Statistical design of an adaptive control chart for linear profile monitoring

Maysa S. De Magalhães[1], Viviany L. Fernandes[2] and Francisco D. Moura Neto[3]

[1]National School of Statistical Sciences, Brazilian Institute of Geography and Statistics (ENCE/IBGE), Rio de Janeiro, Brazil
Email: maysa.magalhaes@ibge.gov.br
[2,3]Polytechnic Institute, State University of Rio de Janeiro (IPRJ/UERJ), Nova Friburgo, RJ, Brazil
[2]Email: vivianylefer@hotmail.com
[3]Email: fmoura@iprj.uerj.br

Abstract: In some production processes the quality characteristics can be represented by profiles or linear functions. We propose an adaptive control chart to monitor the coefficient vector of a simple linear regression model, once fixed parameter control charts are slow in detecting small to moderate shifts in the process parameters, that is, the intercept and the slope. A study on the performance of the proposed control chart was done, considering the adjusted average time until a signal.
Keywords: Linear profile, Adaptive control chart, Markov chain

1 Introduction

In adaptive control charts, one or more design parameters vary in real time during the production process based on recent data obtained from the process. Authors who have been studying this subject have shown that these charts present superior performance when compared to a fixed parameter control chart. Approaches for the design of univariate adaptive control charts have been proposed by several authors, as for example, Reynolds *et al.* (1988), Amin and Miller (1993), Costa (1994, 1997), De Magalhães *et al.* (2002, 2009).

In some processes, however, the simultaneous control of two or more related quality characteristics is necessary, considering that, the design of multivariate fixed parameter and adaptive control charts have been studied by several

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

authors, see for example, Aparisi (1996, 2001), Bersimis *et al.* (2007), Zhang and Shing (2008). Adaptive control schemes have shown better performance than fixed parameter control schemes in detecting small and moderate process shifts.

Nonetheless, some quality characteristics are best represented by a functional relationship between a response variable and one or more explanatory variables, that is, in this case, the quality characteristic is expressed by a function or profile (see, Kang and Albin, 2000; Kim *et al.*, 2003; Zhang and Albin, 2009; Mahmoud *et al.*, 2010; Moura Neto and De Magalhães, 2012). The monitoring of profiles is used to verify the stability of this relationship over time. When the profile does not suffer alteration, it is said that the process is under control. However, if any excessive variation occurs, it is said that the process is out of control, thus, requiring investigation procedures and remedial actions. Some applications of profile monitoring methods include lumber manufacturing (Staudhammer *et al.*, 2007) and calibration of instruments and machines (Stover and Brill, 1998; Kang and Albin, 2000).

Kang and Albin (2000) proposed a fixed parameter chi-square control chart to monitor the intercept and the slope of a linear profile represented by a simple linear regression model.

In this paper, we propose a model for the statistical design of a chi-square control chart with variable sample size and sampling intervals for the monitoring of linear profiles. The performance measure is obtained through a Markov chain approach. The performance of the variable sample size and sampling intervals chi-square chart (VSSI χ^2 control chart) is compared to the fixed parameter chi-square chart (FP χ^2 chart) proposed by Kang and Albin (2000) to monitor the intercept and the slope of a model. Numerical comparisons between these charts are made considering the semiconductor manufacturing process studied in the paper of Kang and Albin.

2. VSSI χ^2 control chart

Based on the studies of Kang and Albin (2000) and Costa (1997), we propose the variable sample size and sampling intervals chi-square chart for monitoring a linear profile. As the chart considered in Kang and Albin, the proposed chart aims to monitor the intercept (β_0) and the slope coefficient (β_1) of a simple linear regression model.

It is considered a production process where the quality of the produced items is evaluated by the value of a measurable characteristic *Y* which is a linear function of an independent variable *x*, that is,

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

where β_0 and β_1 are parameters, ε*'s* are independent random variables and normally distributed with mean zero and variance σ^2 (denoted by, $\varepsilon \sim N(\mu, \sigma^2)$). It is assumed that the parameters $\boldsymbol{\beta} = [\beta_0 \ \beta_1]^T$ and σ^2 of the model, when the process is under control, are known, more specifically, $\boldsymbol{\beta} = \boldsymbol{\beta}_* = [\beta_*^0 \ \beta_*^1]^T$ and $\sigma = \sigma_*$. Then, from the observations, the aim is to verify if the process remains under control, i.e., if the parameters have not changed. Changes or deviations from the parameter vector $\boldsymbol{\beta} = [\beta_0 \, \beta_1]^T$ are analyzed. When the process is out of control, the parameter vector is given by:

$$\widetilde{\boldsymbol{\beta}} = [\beta_*^0 + \delta_0\sigma_* \ \ \beta_*^1 + \delta_1\sigma_*]^T$$

where $\boldsymbol{d} = (\delta_0 \ \ \delta_1)$ represents the vector of the shifts and $\delta_k \ (k = 0, 1)$ is the magnitude of the shifts.

2.1 The statistic used in the monitoring of the process

Consider that the profile *Y* is measured in the values of the independent variable, $x = x_j$, $j = 1,\ldots, n_k$, where $n_k = n_1$ or $n_k = n_2$, depending on the size of the sample that is being used; then, for each sample *i* of size n_k , *k=1*, *2*, the profile monitored is: $Y_i = \beta_{i0} + \beta_{i1}x_j + \varepsilon_{ij}$. For each sample *i* composed by a set data $(x_1 \ y_{i1}), \ (x_2 \ y_{i2}), \ldots, (x_{nk} \ y_{ink})$ the least square estimators for parameters β_0 and β_1 are obtained and the estimator of the vector of parameters is denoted by $\widehat{\boldsymbol{\beta}_i} = [\hat{\beta}_{i0}, \hat{\beta}_{i1}]$. The expressions of $\hat{\beta}_{i0}$ and $\hat{\beta}_{i1}$ are given by,

$$\hat{\beta}_{i0} = \bar{Y} - \hat{\beta}_{i1}\bar{X}$$

$$\hat{\beta}_{i1} = \frac{\sum_{j=1}^{nk}(x_j - \bar{x})(Y_{ij} - \bar{Y})}{\sum_{j=1}^{nk}(x_j - \bar{x})^2}, \qquad k = 1,2$$

For each sample *i* taken from the process, it is assumed that x_j, $1 \leq j \leq n_k$, *k=1, 2, with* $x_1 < x_2 < x_3 < \cdots < x_{n_k}$, are pre-set values for all samples taken.

The proposed control chart monitors the parameters β's, to verify if the process is in control, that is, if the parameters β_0 and β_1 have not shifted. The statistic used in the control chart for monitoring the process is given by,

$$\chi_i^2 = [\widehat{\boldsymbol{\beta}}_{\boldsymbol{i}} - \boldsymbol{\beta}_*]' \, \boldsymbol{\Sigma}^{-1} \, [\widehat{\boldsymbol{\beta}}_{\boldsymbol{i}} - \boldsymbol{\beta}_*]$$

where $\widehat{\boldsymbol{\beta}}_{\boldsymbol{i}} = [\hat{\beta}_{i0}, \hat{\beta}_{i1}]$, the matrix $\boldsymbol{\Sigma}$ and the vector $\boldsymbol{\beta}_*$ are known.

When the process is in control, χ_i^2 has chi-square distribution with two degrees of freedom and the upper control limit is equal to $\chi_{2,\propto}^2$, that is, *UCL* = $\chi_{2,\propto}^2$ where $\chi_{2,\propto}^2$ is the α percentile point of the chi-square distribution. If $\chi_i^2 < \chi_{2,\propto}^2$, it is assumed that the process is in control.

2.2 Surveillance policy

The chi-square control chart with variable sample size and sampling intervals has, besides the upper control limit (*UCL*), a warning limit, *w*, such that $w < UCL$. In contrast to the control chart used by Kang and Albin (2000), which has a fixed sample size (n_0) and sampling interval (h_0), the proposed control chart makes use of two different sample sizes, n_1 and n_2 such that $n_1 < n_0 < n_2$ and two different sampling intervals, h_1 and h_2 such that $h_2 < h_0 < h_1$.

A sample *i* of size n_1 or n_2 is taken randomly and estimates of $\widehat{\boldsymbol{\beta}}_{\boldsymbol{i}}$, the parameter vector of the regression model, are obtained. Then, subsequently, the statistic χ_i^2 is calculated and plotted in the VSSI χ^2 control chart.

Regarding the sample size to be used, if $0 < \chi_{i-1}^2 < w$, the sample *i* will have size n_1 and should be taken after a long time interval, that is h_1, if $w < \chi_{i-1}^2 < UCL$, the sample *i* will have size n_2 and should be taken after a short time interval, that is h_2; finally, if $\chi_{i-1}^2 > UCL$, the process may be out of control. In this case, an investigation should be initiated to verify if there are indeed non-random causes acting in the process, so that corrective action can be undertaken. Otherwise, if an assignable cause is not found the process is considered in control and in this case the signal produced by the chart is a false alarm event.

The probability of a false alarm, that is, the probability of χ_i^2 be greater than *UCL* given that the process is control is $\alpha = P\left(\chi_i^2 \geq UCL \mid \widehat{\boldsymbol{\beta}}_{\boldsymbol{i}} \sim N(\boldsymbol{\beta}_*, \boldsymbol{\Sigma})\right)$.

3. Performance measure

As Costa (1997) to explicitly obtain all the expressions of the statistical model, the production process was represented by a Markov chain with five states:

State 1: if $\chi_i^2 \in [0, w]$ and the process is in control;

State 2: if $\chi_i^2 \in (w, UCL]$ and the process is in control;

State 3: if $\chi_i^2 \in [0, w]$ and the process is out of control;

State 4: if $\chi_i^2 \in (w, UCL]$ and the process is out of control;

State 5 (absorbing state): if $\chi_i^2 \in (UCL, \infty)$.

It is necessary to obtain the transition probabilities to calculate the performance measure. The matrix of transition probabilities between the five states is given by

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} \\ 0 & 0 & p_{33} & p_{34} & p_{35} \\ 0 & 0 & p_{43} & p_{44} & p_{45} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

where p_{lm} denotes the transition probability to go from state l (previous state) to state m (present state).

The transition probabilities between the four transient states are given by

$$p_{11} = P\left[\chi_i^2 < w | \chi_i^2 < UCL\right] P[T > h_1] = \frac{P[\chi_i^2 < w]}{P[\chi_i^2 < UCL]} \cdot e^{-\lambda h_1}$$

$$p_{12} = P\left[w < \chi_i^2 < UCL | \chi_i^2 < UCL\right] P[T > h_1] = \frac{P[w < \chi_i^2 < UCL]}{P[\chi_i^2 < UCL]} \cdot e^{-\lambda h_1}$$

$$p_{21} = P[\chi_i^2 < w | \chi_i^2 < UCL] P[T > h_2] = \frac{P[\chi_i^2 < w]}{P[\chi_i^2 < UCL]} \cdot e^{-\lambda h_2}$$

$$p_{22} = P[w < \chi_i^2 < UCL | \chi_i^2 < UCL] P[T > h_2] = \frac{P[w < \chi_i^2 < UCL]}{P[\chi_i^2 < UCL]} \cdot e^{-\lambda h_2}$$

$$p_{13} = P[\chi_i^2 < w | \chi_i^2 < UCL]\ P[T < h_1] = \frac{P[\chi_i^2 < w]}{P[\chi_i^2 < UCL]} \cdot [1 - e^{-\lambda h_1}]$$

$$p_{14} = P[w < \chi_i^2 < UCL\,|\,\chi_i^2 < UCL]\ P[T < h_1] = \frac{P[w < \chi_i^2 < UCL]}{P[\chi_i^2 < UCL]} \cdot [1 - e^{-\lambda h_1}]$$

$$p_{23} = P[\chi_i^2 < w | \chi_i^2 < UCL]\ P[T < h_2] = \frac{P[\chi_i^2 < w]}{P[\chi_i^2 < UCL]} \cdot [1 - e^{-\lambda h_2}]$$

$$p_{24} = P[w < \chi_i^2 < UCL | \chi_i^2 < UCL]\ P[T < h_2] = \frac{P[w < \chi_i^2 < UCL]}{P[\chi_i^2 < UCL]} \cdot [1 - e^{-\lambda h_2}]$$

$$p_{33} = P\left[\chi_i^2 < \frac{w}{\gamma^2} \,\middle|\, \chi_i^2 \sim \chi_2^2(\tau_1)\ ) \right]$$

$$p_{34} = P\left[\frac{w}{\gamma^2} < \chi_i^2 < \frac{UCL}{\gamma^2} \,\middle|\, \chi_i^2 \sim \chi_2^2(\tau_1) \right]$$

$$p_{43} = P\left[\chi_i^2 < \frac{w}{\gamma^2} \,\middle|\, \chi_i^2 \sim \chi_2^2(\tau_2) \right]$$

$$p_{44} = P\left[\frac{w}{\gamma^2} < \chi_i^2 < \frac{UCL}{\gamma^2} \,\middle|\, \chi_i^2 \sim \chi_2^2(\tau_2) \right]$$

It is assumed that the process starts in control and sometime in the future it goes to out of control and, also, the time that the process remains in control is exponentially distributed with parameter λ.

The performance of the proposed chart, that is, the VSSI χ^2 control chart, was compared to the FP χ^2 control chart proposed by Kang and Albin (2000) for the monitoring of linear profiles. The performance measure employed in this article is the adjusted average time to signal.

3.1 Adjusted average time to signal

The adjusted average time to signal (*AATS*) is the expected time since the instant that the process goes to an out of control state until a signal, that is, until a sample generates a value of statistic χ_i^2 above the *UCL*. When the process is out of control, it is expected to detect this situation rapidly, and then small values of *AATS* are desired. On the other hand, large values of *AATS* are expected when the process is in control. The *AATS* values depend on the magnitude of shift in the process parameters, that is, on the vector $\boldsymbol{d} = (\delta_0 \ \ \delta_1)$ as well as on γ.

The expression for the adjusted average time to signal is given by:

$$AATS = E(TC) - E(T) \text{, then} \quad AATS = E(TC) - \frac{1}{\lambda}$$

where *E*(*TC*) represents the average time of the production cycle, that is, the average time since the beginning of the production process until a signal after an occurrence of a process shift and, *E*(*T*) denotes the time the process remains in control.

The expression for the average time of the production cycle is given by:

$$E(TC) = \begin{bmatrix} P_1^{(0)} & P_2^{(0)} & P_3^{(0)} & P_4^{(0)} \end{bmatrix} \left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix} \right)^{-1} \begin{bmatrix} h_1 \\ h_2 \\ h_1 \\ h_2 \end{bmatrix}$$

where the transition probabilities are given above and $\begin{bmatrix} P_1^{(0)} & P_2^{(0)} & P_3^{(0)} & P_4^{(0)} \end{bmatrix}$ is the vector of initial probabilities with $P_1^{(0)} = P(\chi_i^2 < w | \chi_i^2 < UCL)$, $P_2^{(0)} = 1 - P_1^{(0)} = 1 - P(\chi_i^2 < w | \chi_i^2 < UCL)$, $P_3^{(0)} = 0$ and $P_4^{(0)} = 0$.

The expression for *E*(*TC*) depends on the cumulative probability function of a central chi-square distribution with two degrees of freedom, and a non-central chi square distribution with two degrees of freedom with non-centrality parameter τ.

To compare the performance of the VSSI χ^2 control chart and the FP χ^2 control chart, we use the *AATS*, for a given value of the shift parameters. However, for the comparison to be fair, the same amount of resources/effort spent with

inspections and false alarms, when the process is in control, should be imposed. This is done by the following constraints

$$n_1 P(\chi_i^2 < w | \chi_i^2 < UCL) + n_2 \left\{1 - P(\chi_i^2 < w | \chi_i^2 < UCL)\right\} = n_0$$

$$h_1 P(\chi_i^2 < w | \chi_i^2 < UCL) + h_2 \left\{1 - P(\chi_i^2 < w | \chi_i^2 < UCL)\right\} = h_0$$

4. Comparing charts

In this section, we compare the performance of the VSSI χ^2 and the FP χ^2 control charts for monitoring linear profiles relative to the detection speed of an out-of-control state considering several shifts magnitudes on the parameters.

For comparison purposes, the application of the developed statistical model for the VSSI χ^2 control chart for monitoring linear profiles is shown by the numerical example of Kang and Albin (2000), which consists of a calibration application in a production process of semi-conductors, where several thousand inscriptions of chips need to be provided in a wafer. The critical device in this process is a mass flow controller (*MFC*). The pressure measure in the chamber is approximately a linear function of the mass flux through the *MFC.* In the example presented by Kang and Albin (2000), the chart employed is the FP χ^2 control chart in which a single sample size $(n_0 = 4)$ is used. Moreover, based on the work of Costa (1997) and taking into account the following restrictions $n_1 < n_0 < n_2$, $h_2 < h_0 < h_1$, then the following design parameters were used: n_0=4, h_0=1, α_0=0.005, for the FP χ^2 chart; and n_0=4, n_1=2, n_2=12, h_0=1, h_1=1, 2, h_2=0.2, α_0=0.005, for the VSSI χ^2 chart. Again, based on the work of Costa (1997), we considered $\frac{1}{\lambda} = \frac{1}{0{,}0001}$.

As we are considering the example proposed by Kang and Albin (2000) and also as we are going to compare the chart proposed by them and our proposed chart, the shifts δ_0 in β_*^0 varied from 0.2 to 2.0 in steps of 0.2, the shifts δ_1 in β_*^1 assumed values from 0.025 to 0.250 in steps of 0.025.

Then, with the design parameters considered above, performance measures were calculated for the proposed chart. The results are presented in Tables 1 to 4.

Table 1 presents the values of the *AATS* for both charts with respect to the values of the shift parameter δ_0. Table 2 and Fig.1 present the percentage gain of VSSI chart relative to FP chart as a function of intercept shifts.

Table 1. *AATSs* for both charts with intercept shifts.

χ^2 Chart	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
	δ_0										
FP	200.5	138.24	63.96	28.45	13.69	7.38	4.49	3.08	2.35	1.95	1.73
VSSI	200.5	129.43	40.38	10.46	4.31	2.93	2.43	2.16	1.98	1.86	1.78

Table 2. Percentage gain of VSSI chart relative to FP chart as a function of intercept shifts.

χ^2 Chart	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
	δ_0										
VSSI	0.00	0.06	0.37	0.63	0.69	0.60	0.46	0.30	0.16	0.05	-0.03

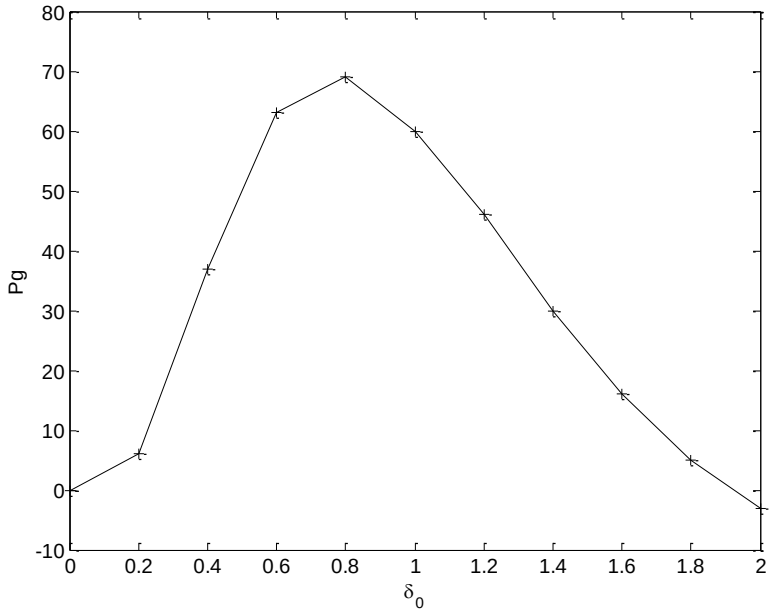


Fig.1. Percentage gain of VSSI chart with respect to FP chart as a function of intercept shifts.

As may be seen in Table 1, when the process is in control the *AATS* is equal to 200.5. It can be observed from this table that from small to moderate shifts in the intercept ($\delta_0 \leq 1.0$), the VSSI χ^2 chart is always quicker than the FP χ^2

control chart, for the design parameters considered. Still the performance of the VSSI χ^2 chart is superior to the FP χ^2 control chart for shifts of magnitude ($\delta_0 \leq 1.8$) for the design parameters considered. In contrast, in the presence of large shifts when ($\delta_0 = 2.0$), the FP χ^2 control chart is more efficient than the VSSI χ^2 chart; although in this case, the average number of samples until a signal is below 2.0 for the VSSI χ^2 chart.

From Table 2 and Fig.1, we can observe when $0.2 \leq \delta_0 \leq 0.8$, the percentage gain varies, approximately, from 6% to 69% and when $1.0 \leq \delta_0 \leq 1.8$, the percentage gain varies, approximately, from 60% to 5%, for the design parameters considered. When $\delta_0 = 2.0$, it is preferable to use the FP χ^2 chart instead of the VSSI χ^2 chart, in the case considered.

Table 3 presents the values of the *AATS* for both charts with respect to the values of the shift parameter δ_1. Table 4 and Fig.2 present the percentage gain of VSSI chart relative to FP chart as a function of slope shifts.

Table 3. *AATSs* for both charts with slope shifts.

	0.000	0.025	0.050	0.075	0.100	0.125	0.150	0.175	0.200	0.225	0.250
χ^2 Chart	δ_1										
FP	200.5	166.50	106.09	61.18	34.98	20.62	12.73	8.30	5.73	4.19	3.24
VSSI	200.5	162.35	87.70	35.39	13.49	6.29	3.98	3.12	2.70	2.43	2.25

Table 4. Percentage gain of VSSI chart relative to FP chart as a function of slope shifts.

	0.000	0.025	0.050	0.075	0.100	0.125	0.150	0.175	0.200	0.225	0.250
χ^2 Chart	δ_1										
VSSI	0.00	0.02	0.17	0.42	0.61	0.69	0.69	0.62	0.53	0.42	0.31

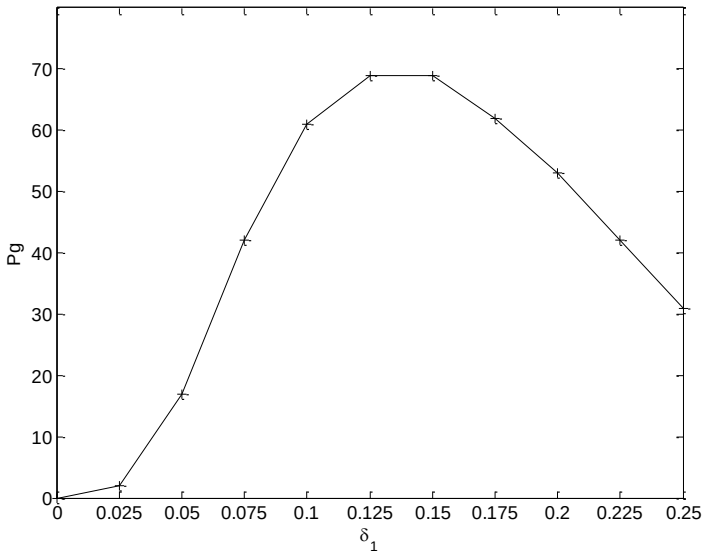


Fig.2. Percentage gain of VSSI chart with respect to FP chart as a function of slope shifts.

As may be seen in Table 3, when the process is in control the *AATS* is equal to 200.5. It can be observed from this table that for all considered shifts in the slope, the VSSI χ^2 chart is always quicker than the FP χ^2 control chart, for the design parameters considered.

From Table 4 and Fig.2, we can observe when $0.025 \leq \delta_1 \leq 0.150$, the percentage gain varies, approximately, from 2% to 69% and when $0.175 \leq \delta_0 \leq 0.250$, the percentage gain varies, approximately, from 62% to 31%, for the design parameters considered.

5. Conclusions

In this article, a model for the statistical design of a chi-square control chart with variable sample size and sampling interval for monitoring a linear profile was developed. This chart contemplates the monitoring of the intercept and the slope coefficient of a linear regression model. The proposed chart was developed based on the fixed parameter chi-square control chart existent in the literature for monitoring a linear profile employed by Kang & Albin (2000). Comparisons between the two charts considered the adjusted average time until a signal (*AATS*). From a numerical example, the performance comparison between the

two charts showed, in general, a better statistical performance for the VSSI chi-square chart.

Acknowledgments

The authors acknowledge the partial financial support from the Brazilian Council for Scientific and Technological Development (CNPq), the State of Rio de Janeiro Research Foundation (FAPERJ), and the Brazilian Council for the Improvement of Higher Education (CAPES).

References

1. R.W. Amin and R.W. Miller. A robustness study of $\overline{X}$ charts with variable sampling intervals. *Journal of Quality Technology*, **25**, 35-44, 1993.
2. F. Aparisi. Hotelling's T^2 control chart with adaptive sample sizes. *International Journal of Production Research*, **34**, 2835-2862, 1996.
3. F. Aparisi and C.L. Haro,. Hotelling's T^2 control chart with sampling intervals, *International Journal of Production Research*, **39**, 3127–3140, 2001.
4. Bersimis, S., Psarakis, S., and Panaretos, J., Multivariate statistical process control charts: an overview, Quality and Reliability Engineering International, **23**, 517–543, 2007.
5. Costa, A.F.B., $\bar{X}$ charts with variable sampling size, *Journal of Quality Technology*, **26**, 155-163, 1994.
6. Costa, A.F.B., $\bar{X}$ charts with variable sample size and sampling intervals, Journal of Quality Technology, **29**, 197–204, 1997.
7. De Magalhães, M.S., Costa, A.F.B., and Epprecht, E.K., Constrained optimization model for the design of an adaptive $\bar{X}$ chart, *International Journal of Production Research*, 40, 3199–3218, 2002.
8. De Magalhães, M. S., Costa, A.F.B., and Moura Neto, F.D., A hierarchy of adaptative $\bar{X}$ control charts, *International Journal of Production Economics*, **119**, 271-283, 2009.
9. Kang, L. and Albin, S.L., On-line monitoring when the process yields a linear profile, *Journal of Quality Technology*, **32**, 418-426, 2000.
10. Kim, K., Mahmoud, M., and Woodall, W.H., On the monitoring of linear profiles, *Journal of Quality Technology*, **35**, 317-328, 2003.
11. Mahmoud, M.A., Morgan, J.P., and Woodall, W.H., The monitoring of simple linear regression profiles with two observations per sample, *Journal of Applied Statistics*, **37**, 1249-1263, 2010.

12. Moura Neto, F.D. and De Magalhães, M.S., A laplacian spectral method in phase analysis of profiles, *Applied Stochastic Models in Business and Industry*, **28**, 251-263, 2012.
13. Reynolds Jr., M.R., Amin, R.W., Nachlas, J.C., $\bar{X}$ charts with variable sampling intervals, *Technometrics*, **30**, 181-192, 1988.
14. Staudhammer, C., Maness, T.C., and Kozac, R.A., Profile charts for monitoring lumber manufacturing using laser range sensor data, *Journal of Quality Technology*, **39**, 224-240, 2007.
15. Stover, F.S. and Brill, R.V., Statistical quality control applied to ion chromatography Calibrations, *Journal of Chromatography A*, **804**, 37-43, 1998.
16. Zhang, H. and Albin, S., Detecting outliers in complex profiles using a χ^2 control chart method, ***IIE*** Transactions, **41**, 335-345, 2009.
17. Zhang, G. and Shing, I., Multivariate EWMA control charts using individual observations for process mean and variance monitoring and diagnosis, *International Journal of Production Research*, **46**, 6855-6881, 2008.

Singular extremals in control problems for wireless sensor networks

Larisa Manita

National Research University Higher School of Economics
Moscow Institute of Electronics and Mathematics
Bolshoy Trehsviatitelskiy Per. 3, 109028 Moscow, Russia
(E-mail: lmanita@hse.ru)

Abstract. Energy-saving optimization is very important for various engineering problems related to modern distributed systems. We consider here a control problem for a wireless sensor network with a single time server node and a large number of client nodes. The problem is to minimize a functional which accumulates clock synchronization errors in the clients nodes and the energy consumption of the server over some time interval $[0, T]$. The control function $u = u(t)$, $0 \le u(t) \le u_1$, corresponds to the power of the server node transmitting synchronization signals to the clients. For all possible parameter values we find the structure of extremal trajectories. We show that for sufficiently large u_1 the extremals contain singular arcs.

Keywords: Pontryagin maximum principle, bilinear control system, singular extremals, wireless sensor network, energy-saving optimization.

1 Model

Power consumption, clock synchronization and optimization are very popular topics in analysis of wireless sensor networks [1]–[7]. In the majority of modern papers their authors discuss and compare communication protocols (see, for example, [4]), network architectures (for example, [3]) and technical designs by using numerical simulations or dynamical programming methods (e.g., [6]). In the present talk we consider a mathematical model related with large scale networks which nodes are equipped with noisy non-perfect clocks [2]. The task of optimal clock sychronization in such networks is reduced to the classical control problem. Its functional is based on the trade-off between energy consumption and mean-square synchronization error. This control problem demonstrates surprisingly deep connections with the theory of singular optimal solutions [8]-[13].

The network consists of a single server node (denoted by 1) and N client nodes (sensors) numbered as $2, \dots, N + 1$.

Let x_i be a state of the node i having the meaning of a local clock value at this node. The network evolves in time $t \in \mathbb{R}_+$ as follows.

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

1) The node 1 is a time server with the perfect clock:

$$\frac{d\,x_1(t)}{dt} = v > 0$$

2) The client nodes are equipped with non-perfect clocks with a random Gaussian noise

$$\frac{d\,x_j(t)}{dt} = v + \sigma dB_j(t) + \text{synchronizing jumps},$$

where $B_j(t)$, $j = 2, \ldots, N+1$, are independent standard Wiener processes, $\sigma > 0$ corresponds to the strength of the noise and "synchronizing jumps" are explained below.

3) At random time moments the server node 1 sends messages to randomly chosen client nodes, u is the *intensity* of the Poissonian message flow issued from the server. The client j, $j = 2, \ldots, N+1$, that receives at time τ a message from the node 1 immediately ajusts its clock to the current value of x_1:

$$\begin{aligned} x_j(\tau + 0) &= x_1(\tau), \\ x_k(\tau + 0) &= x_k(\tau), \qquad k \neq j. \end{aligned}$$

Hence the client clocks $x_j(t)$, $t \geq 0$, are stochastic processes which interact with the time server.

The function

$$R(t) = \mathsf{E}\frac{1}{N}\sum_{j=2}^{N+1} \left(x_j(t) - x_1(t)\right)^2$$

is a cumulative measure of desynchronization between the client and server nodes. Here E stands for the expectation.

It was proved in [2] that the function $R(t)$ satisfies the differential equation

$$\dot{R} = -uR + N\sigma^2$$

2 Optimal control problem

Consider the following optimal control problem

$$\int_0^T (\alpha R(t) + \beta u(t))\, dt \to \inf \tag{1}$$

$$\dot{R}\,(t) = -u(t)R(t) + N\sigma^2 \tag{2}$$

$$R\,(0) = R_0 \tag{3}$$

$$0 \leq u(t) \leq u_1 \tag{4}$$

Here α, β are some positive constants. The control function $u\,(t)$ corresponds to the power of the server node transmitting synchronization signals to the clients. The functional (1) accumulates clock synchronization errors in the

clients nodes and the energy consumption of the server over some time interval $[0, T]$.

The admissible solutions to (1)-(4) are absolutely continuous functions, the admissible controls belong to $L^\infty[0, T]$.

We prove that the problem (1)-(4) has a unique solution. We find a structure of optimal control. We show that optimal solutions may contain singular arcs.

3 Existence of solution

Lemma 1 *For any R_0 and any parameter values T, α, β, N, σ^2, u_1 there exists a solution $(\hat{R}(t), \hat{u}(t))$ to the problem (1)-(4).*

Proof. Let $\mathcal{B}_{R_0}$ denote the set of continuous functions $R : [0, T] \to \mathbb{R}$ such that $R(0) = R_0$. Consider the map $K : L^\infty[0, T] \to \mathcal{B}_{R_0}$ defined as follows:

$$(Ku)(t) = R_0 \exp\left(-\int_0^t u(\xi)d\xi\right) + N\sigma^2 \int_0^t \exp\left(-\int_s^t u(\xi)d\xi\right) ds$$
$$=: A(u, t) + B(u, t). \tag{5}$$

This operator assigns to the control function u the corresponding solution R of (1)-(4).

1. Let $\left\{u^{(n)}(t)\right\}_{n=1}^{\infty}$ be a minimizing sequence for the fuctional

$$\int_0^T (\alpha R(t) + \beta u(t))\, dt,$$

i.e.,

$$\int_0^T \left(\alpha K u^{(n)}(t) + \beta u^{(n)}(t)\right) dt \to \inf_{u \in V} \left\{\int_0^T (\alpha R(t) + \beta u(t))\, dt\right\}, \qquad (n \to \infty),$$

where $V = \{v \in L^\infty[0, T] : \ 0 \le v(t) \le u_1\}$. Recall that the space $L^1[0, T]$ is the adjoint space to $L^\infty[0, T]$. By $\langle \phi, u \rangle$ we denote the value of the functional $\phi \in (L^\infty[0, T])^* \cong L^1[0, T]$ at $u \in L^\infty[0, T]$:

$$\langle \phi, u \rangle = \int_0^T \phi(\xi) u(\xi)\, d\xi\,.$$

Since $u^{(n)}(t) \in [0, u_1]$, one can extract a weakly-$*$ converging in $L^\infty[0, T]$ subsequence $u^{(n_k)}(t)$ by virtue of Banach-Alaoglu theorem. Without loss of generality one can assume that $u^{(n)}$ weakly-$*$ converges to some $\hat{u} \in L^\infty[0, T]$. This means that for each $\rho \in L^1[0, T]$ one has

$$\int_0^T \rho(\xi) u^{(n)}(\xi)\, d\xi \to \int_0^T \rho(\xi)\hat{u}(\xi)\, d\xi, \quad n \to \infty. \tag{6}$$

2. Let us prove that the sequence $R^{(n)}(t) := Ku^{(n)}(t)$ converges pointwise to $\hat{R}(t) := K\hat{u}(t)$ as $n \to \infty$.

Further let $\phi_s^t(\xi) := -\mathbf{1}_{[s,t]}(\xi) = \begin{cases} -1, \xi \in [s,t], \\ 0, \ \xi \notin [s,t]. \end{cases}$ Taking $\rho(\xi) = \phi_0^t(\xi)$ in (6) we obtain

$$\int_0^t u^{(n)}(\xi)\, d\xi \to \int_0^t \hat{u}(\xi)\, d\xi, \quad n \to \infty,$$

hence

$$A(u^{(n)}, t) \to A(\hat{u}, t), \quad n \to \infty$$

for each fixed t. Note that $B(u^{(n)}, t) = N\sigma^2 \int_0^t \exp\left\langle \phi_s^t, u^{(n)} \right\rangle ds$. The functions $\exp\left\langle \phi_s^t, u^{(n)} \right\rangle$ are uniformly bounded and pointwise convergent, hence Lebesgue's dominated theorem yields the convergence

$$B(u^{(n)}, t) \to B(\hat{u}, t), \quad n \to \infty$$

for each fixed t. So we established the required convergence.

3. Let us show that $\hat{R}(t)$ is a solution to (1)–(4).

Obviously $R^{(n)}(t)$ are uniformly bounded (this follows straightforward from the explicit formula (5)). Since they form a pointwise convergent sequence, Lebesgue's dominated theorem yields

$$\int_0^T \alpha R^{(n)}(t)\, dt \to \int_0^T \alpha \hat{R}(t)\, dt, \quad n \to \infty.$$

Moreover, due to weak-$*$ convergence, one has

$$\int_0^T \beta u^{(n)}(t)\, dt = \beta \int_0^T \phi_0^T(t) u^{(n)}(t) dt \to \beta \int_0^T \phi_0^T(t) \hat{u}(t) dt = \beta \int_0^T \hat{u}(t) dt.$$

This yields

$$\int_0^T \left(\alpha R^{(n)}(t) + \beta u^{(n)}(t) \right) dt \to \int_0^T \left(\alpha \hat{R}(t) + \beta \hat{u}(t) \right) dt.$$

Thus $(\hat{R}(t), \hat{u}(t))$ is an optimal solution to (1)–(4). □

4 Pontryagin maximum principle

We will apply Pontryagin Maximum Principle [14] to the problem (1)-(4). Let $\left(\widehat{R}(t), \widehat{u}(t) \right)$ be an optimal solution. Then there exist a constant λ_0 and a continuous function $\psi(t)$ such that for all $t \in (0, T)$ we have

$$H\left(\widehat{R}(t), \psi(t), \widehat{u}(t) \right) = \max_{0 \le u \le u_1} H\left(\widehat{R}(t), \psi(t), u \right) \tag{7}$$

where the Hamiltonian function

$$H(R, \psi, u) = -\lambda_0 (\alpha R + \beta u) + \psi\left(-uR + N\sigma^2 \right)$$

Except at points of discontinuity of $\widehat{u}(t)$

$$\dot{\psi}(t) = -\frac{\partial H\left(\widehat{R}(t), \psi(t), \widehat{u}(t)\right)}{\partial R} = \lambda_0 \alpha + \widehat{u}(t)\psi \tag{8}$$

And ψ satisfies the following transversality condition

$$\psi(T) = 0 \tag{9}$$

The function $\psi(t)$ is called an adjoint function. The condition (7) is called the **maximum condition.**

The dynamics equation (2) and the adjoint equation (8) form a Hamiltonian system

$$\begin{array}{l} \dot{\psi} = \lambda_0 \alpha + \widehat{u}(t)\psi \\ \dot{R} = -\widehat{u}(t) R + N\sigma^2 \end{array} \tag{10}$$

where $\widehat{u}(t)$ satisfies the maximum condition (7). The solutions $(R(t), \psi(t))$ of (10) are called extremals. If $\lambda_0 \neq 0$, we say that $(R(t), \psi(t))$ is normal. One can show [3] that in the problem (1)-(4) every extremal is normal. So we can put $\lambda_0 = 1$.

5 Switching function and singular extremals

Denote

$$H_0(R, \psi) = -\alpha R + \psi N \sigma^2, \quad H_1(R, \psi) = -\beta - R\psi \tag{11}$$

then $H = H_0 + uH_1$. The Hamiltonian H is linear in u. Hence to maximize it over the interval $u \in [0, u_1]$ we need to use boundary values depending on the sign of H_1.

$$\hat{u}(t) = \begin{cases} 0, & H_1(R(t), \psi(t)) < 0 \\ u_1, & H_1(R(t), \psi(t)) > 0 \end{cases} \tag{12}$$

The function H_1 is called a switching function.

Suppose that there exists an interval (t_1, t_2) such that

$$H_1(R(t), \psi(t)) = 0, \quad \forall t \in (t_1, t_2) \tag{13}$$

then the extremal $(R(t), \psi(t)), \quad t \in (t_1, t_2)$, is called a ***singular*** one. In this case we can't find an optimal control from the maximum condition (7). We will differentiate the identity $H_1(R(t), \psi(t)) \equiv 0$ by virtue of the Hamiltonian system (10) until a control u appears with a non-zero coefficient.

We say that a number q is the order of the singular trajectory iff

$$\frac{\partial}{\partial u}\frac{d^k}{dt^k}\bigg|_{(10)} H_1(R, \psi) = 0, \qquad k = 0, \ldots, 2q-1,$$

$$\frac{\partial}{\partial u}\frac{d^{2q}}{dt^{2q}}\bigg|_{(10)} H_1(R, \psi) \neq 0$$

in some open neighborhood of the singular trajectory $(R(t), \psi(t))$.

It is known that q is an integer.

Singular solutions arise frequently in control problems [8]-[12] and are therefore of practical significance. We prove that for suffiently large u_1 a singular control is realised in the problem (1)-(4).

Lemma 2 *Let*

$$\sqrt{\frac{\alpha N\sigma^2}{\beta}} \leq u_1$$

then in the problem (1)-(4) there exists a singular extremal of order 1

$$\hat{R}_s(t) \equiv \sqrt{N\sigma^2\frac{\beta}{\alpha}}, \quad \psi_s(t) \equiv -\sqrt{\frac{\alpha\beta}{N\sigma^2}} \tag{14}$$

and the corresponding singular control is

$$u_s = \sqrt{\frac{\alpha N\sigma^2}{\beta}}$$

Proof. Assume that (13) holds. We will differentiate this identity along the extremal with respect to t:

$$\left.\frac{d}{dt}\right|_{(10)} H_1(R(t), \psi(t)) = 0 \quad \Rightarrow -N\sigma^2\psi(t) - \alpha R(t) = 0 \tag{15}$$

$$\left.\frac{d^2}{dt^2}\right|_{(10)} H_1(R(t), \psi(t)) = 0 \quad \Rightarrow u\left(\alpha R(t) - N\sigma^2\psi(t)\right) - 2\alpha N\sigma^2 = 0 \tag{16}$$

From (13)–(15) we have

$$R(t) = \sqrt{N\sigma^2\frac{\beta}{\alpha}}, \quad \psi(t) = -\sqrt{\frac{\alpha\beta}{N\sigma^2}} \tag{17}$$

Substituting (17) in (16) we obtain

$$2\sqrt{N\sigma^2\alpha\beta}\cdot u - 2\alpha N\sigma^2 = 0$$

Thus

$$R(t) \equiv \sqrt{N\sigma^2\frac{\beta}{\alpha}}, \quad \psi(t) \equiv -\sqrt{\frac{\alpha\beta}{N\sigma^2}}$$

is a singular extremal of order 1 and $u_s = \sqrt{\frac{\alpha N\sigma^2}{\beta}}$ is the corresponding singular control.

Note that if $\sqrt{\frac{\alpha N\sigma^2}{\beta}} > u_1$ then u_s does not satisfy the condition $0 \leq u(t) \leq u_1$ hence optimal solutions to the problem (1)-(4) are nonsingular. □

Recall the well-known generalized Legendre-Clebsch condition [8], the necessary condition for optimality of the singular extremal of order 1:

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}H_1(\widehat{R}(t), \psi(t)) \geq 0$$

We see that this condition holds in our problem. One can show that any concatenation of the singular control with a bang control $u = 0$ or $u = u_1$ satisfies the necessary conditions of the maximum principle [8].

From the transversality condition (9) it is easily seen that on the final time interval the optimal control $\widehat{u}(t)$ in the problem (1)-(4) is nonsingular. Namely, for all initial condition R_0 and for all parameter values α, β, N, σ^2, u_1 we have the following result.

Lemma 3 *There exists $\varepsilon > 0$ such that $\widehat{u}(t) = 0$ for all $t \in (T - \varepsilon, T)$.*

Proof. Using the transversality condition (9) we obtain $H_1(\widehat{R}(T), \psi(T)) = -\beta < 0$. The continuity of the switching function H_1 implies that

$$H_1(\widehat{R}(t), \psi(t)) < 0 \quad \forall t \in (T - \varepsilon, T)$$

for some $\varepsilon > 0$. The maximum condition (7) yields $\widehat{u}(t) = 0, \ t \in (T - \varepsilon, T)$. □

6 The orbits of the Pontryagin maximum principle system

Consider the behaviour of the extremals on the plane (R, ψ). Let Γ be a switching curve, that is, a set of point such that $H_1(R, \psi) = 0$. By (11) we have $\Gamma = \{(R, \psi) | \beta + R\psi = 0\}$. We are interested in the domain $\{(R, \psi) : \ R > 0\}$. Denote

$$\Gamma^+ = \Gamma \cap \{(R, \psi) : \ R > 0\}$$

Above Γ^+ the optimal control $\hat{u}$ equals 0, below Γ^+ the optimal control $\hat{u}$ equals u_1 (see (12)). Let $u = 0$ then the Hamiltonian system (10) has the form

$$\dot{R} = N\sigma^2, \quad \dot{\psi} = \alpha \tag{18}$$

The general solution of (18) is

$$R(t) = N\sigma^2 t + C, \quad \psi(t) = \alpha t + w$$

On the plane (R, ψ) the orbits of the system (18) are straight lines

$$\psi = \frac{\alpha}{N\sigma^2} R + B$$

Let $u = u_1$ than the Hamiltonian system (10) has the form

$$\dot{R} = -u_1 R + N\sigma^2, \quad \dot{\psi} = \alpha + u_1 \psi \tag{19}$$

The general solution of (19) is

$$R(t) = \widetilde{C} e^{-u_1 t} + \frac{N\sigma^2}{u_1}, \quad \psi(t) = \widetilde{w} e^{u_1 t} - \frac{\alpha}{u_1}$$

On the plane (R, ψ) if $\widetilde{C} \neq 0$, $\widetilde{w} \neq 0$, the orbits of the system (19) are hyperbolas

$$|\alpha + \psi u_1| \cdot \left|N\sigma^2 - u_1 R\right| = \omega$$

If $\widetilde{C} = 0$, $\widetilde{w} \neq 0$, the orbit is the straight line $R = \frac{N\sigma^2}{u_1}$, directed upward if $\widetilde{w} > 0$ or downward if $\widetilde{w} < 0$. If $\widetilde{w} = 0$, the orbit is the straight line $\psi = -\frac{\alpha}{u_1}$, directed to the left if $\widetilde{C} > 0$ or to the right if $\widetilde{C} < 0$. If $\widetilde{C} = 0$, $\widetilde{w} = 0$, the point $\left(\frac{N\sigma^2}{u_1}, -\frac{\alpha}{u_1}\right)$ is the stationary orbit.

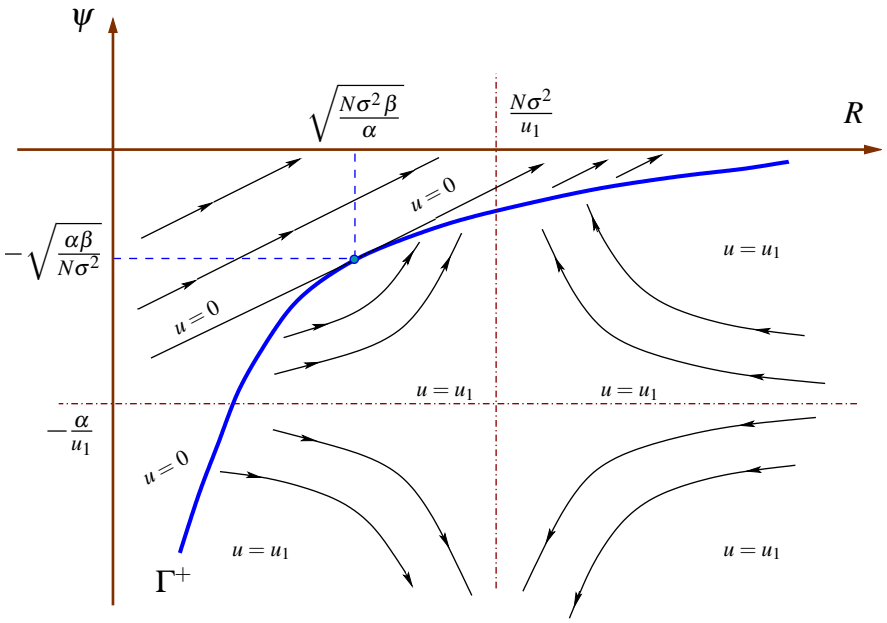


Fig 1. Orbits in the nonsingular case: $\sqrt{\frac{\alpha N\sigma^2}{\beta}} > u_1$

Remark. On Fig. 1 and Fig. 2 we don't show trajectories $(R(t), \psi(t))$ with $\psi(0) > 0$ because they cannot satisfy the transversality condition.

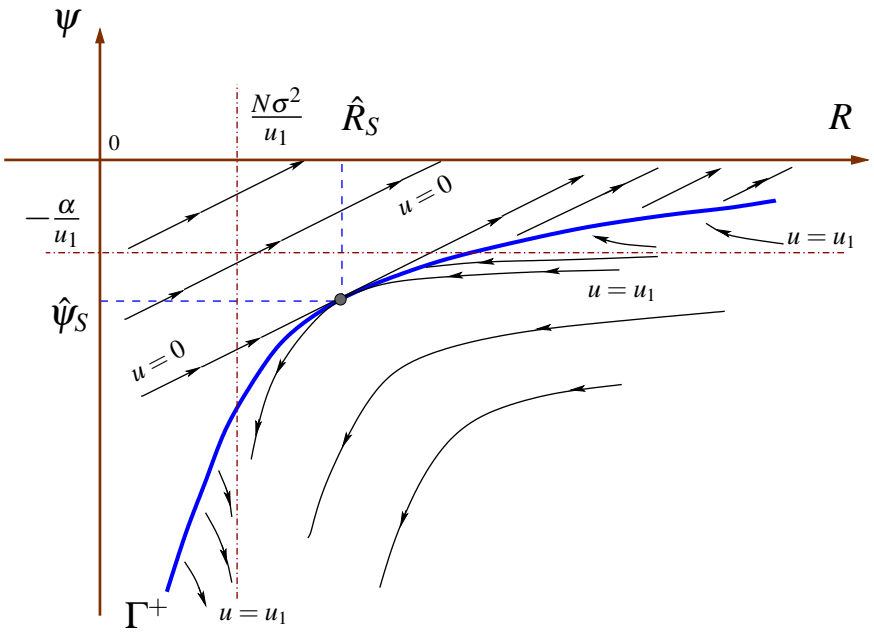


Fig 2. Orbits in the singular case: $\sqrt{\dfrac{\alpha N\sigma^2}{\beta}} \le u_1$

Note that in the case $\sqrt{\alpha N\sigma^2/\beta} \le u_1$ two extremals go out of the singular point $\left(\sqrt{N\sigma^2\frac{\beta}{\alpha}}, -\sqrt{\frac{\alpha\beta}{N\sigma^2}}\right)$ (with $u = 0$ and $u = u_1$). But only one extremal (going of the singular point) satisfies the transversality condition (9).

Thus for any $R_0 \ge 0$ there exists a unique extremal such that $R(0) = R_0$, $\psi(T) = 0$. Since we prove that a solution to problem (1)-(4) exists hence the constructed extremals are optimal.

To summarize the above analysis in the next two sections we consider separately the nonsingular and singular cases. In each case we provide a plot with optimal solutions and state a conclusion on the structure of the optimal control $\hat{u}(t)$ (Theorems 1 and 2). It is interesting also to see how the structure of $\hat{u}(t)$ depends on the parameter R_0 and T. The answer is presented on Figures 4 and 6.

7 Optimal solutions. Nonsingular case

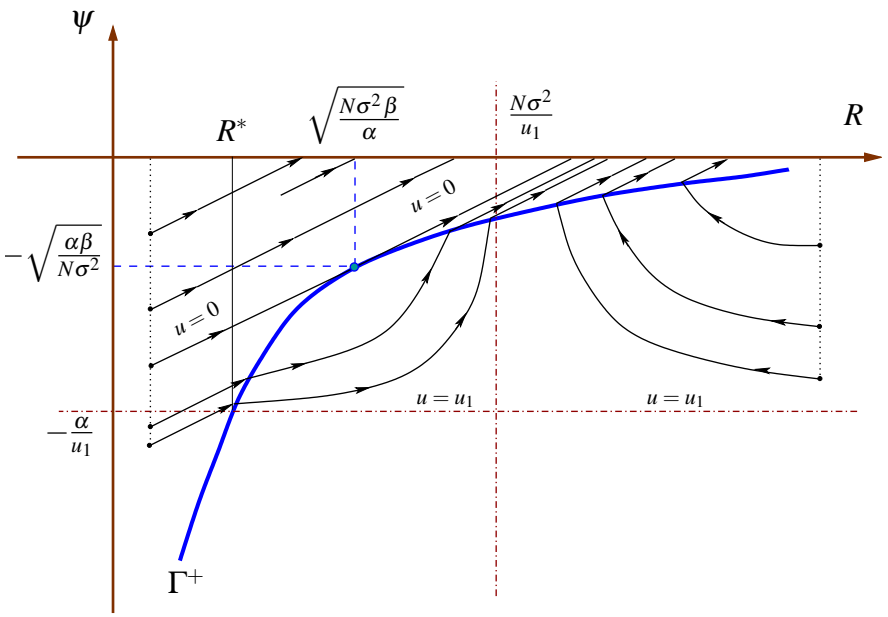


Fig 3. Optimal solutions for different values of the problem parameters. *Nonsingular case.*

Theorem 1 *Let* $\sqrt{\dfrac{\alpha N\sigma^2}{\beta}} > u_1$, *that is, optimal solutions are nonsingular (Lemma 2). Then, depending of values* $R(0)$ *and* T, *the optimal control* $\hat{u}(t)$ *has one of the following forms*

$$1.1. \quad \hat{u}(t) = 0,\ t \in (0,T)$$

$$1.2. \quad \hat{u}(t) = \begin{cases} u_1,\ t \in (0,t_1) \\ 0,\ t \in (t_1,T) \end{cases}$$

$$1.3. \quad \hat{u}(t) = \begin{cases} 0,\ t \in (0,t_1) \\ u_1,\ t \in (t_1,t_2) \\ 0,\ t \in (t_2,T) \end{cases}$$

i.e., the optimal control switches between $u = 0$ *and* $u = u_1$ *and the number of switchings does not exceed 2.*

The Fig. 4 shows how the *structure* of optimal controls $\hat{u} = \hat{u}(t)$, $t \in [0, T]$, depends on T and on the initial value $R(0)$.

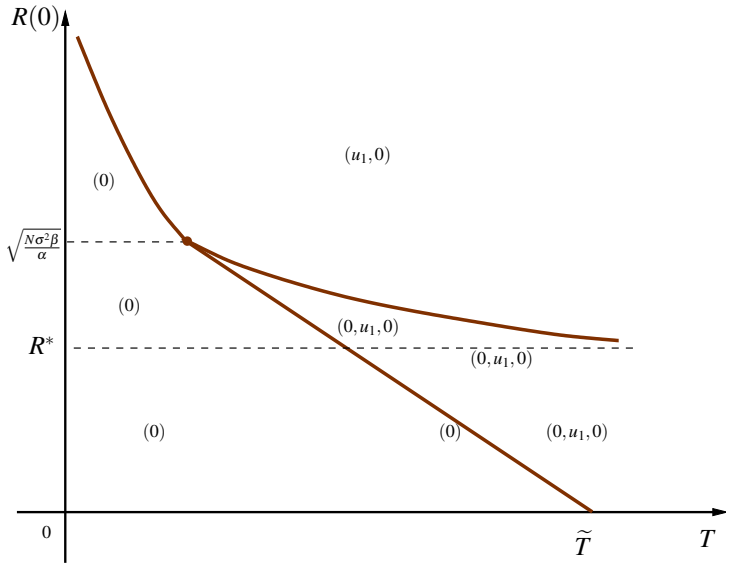


Fig 4.

Let (θ, ρ) be some point on the plane $(T, R(0))$. Assume that (θ, ρ) belongs to a domain labeled, *for example*, by (a, b, c). This means that for the optimal control problem with $T = \theta$ and $R(0) = \rho$ the optimal control function $\hat{u} = \hat{u}(t)$ has the following form

$$\hat{u}(t) = \begin{cases} a, & t \in (0, \tau_1), \\ b, & t \in (\tau_1, \tau_2), \\ c, & t \in (\tau_2, \theta). \end{cases}$$

Here τ_1 and τ_2 are some numbers satisfying the condition $0 < \tau_1 < \tau_2 < \theta$. The numbers τ_1 and τ_2 depend on (θ, ρ) and on all parameters $(\alpha, \beta, N, \sigma)$ of the model. For points (θ, ρ) in the domain labeled by (0) we have $\hat{u}(t) = 0$ for all $t \in [0, T]$.

8 Optimal Solutions. Singular case

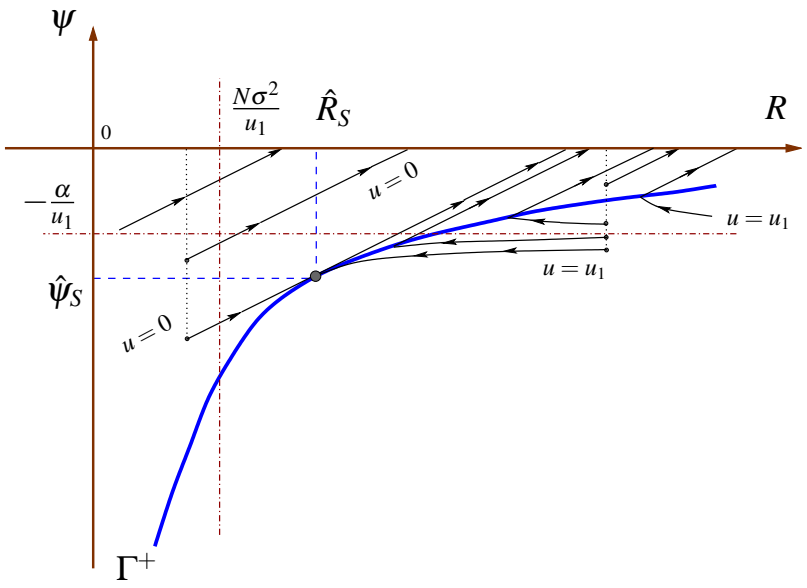


Fig 5. Optimal solutions for different values of the model parameters. *Singular case.*

Theorem 2 *Let* $\sqrt{\frac{\alpha N\sigma^2}{\beta}} \le u_1$. *Then, depending of values* $R(0)$ *and* T, *the optimal control* $\hat{u}(t)$ *has one of the following forms*

$$2.1. \quad \hat{u}(t) = 0,\ t \in (0,T)$$

$$2.2. \quad \hat{u}(t) = \begin{cases} u_1,\ t \in (0,t_1) \\ 0,\ t \in (t_1,T) \end{cases} \qquad 2.3. \quad \hat{u}(t) = \begin{cases} u_s,\ t \in (0,t_1) \\ 0,\ t \in (t_1,T) \end{cases}$$

$$2.4. \quad \hat{u}(t) = \begin{cases} 0,\ t \in (0,t_1) \\ u_s,\ t \in (t_1,t_2) \\ 0,\ t \in (t_2,T) \end{cases}, \qquad 2.5. \quad \hat{u}(t) = \begin{cases} u_1,\ t \in (0,t_1) \\ u_s,\ t \in (t_1,t_2) \\ 0,\ t \in (t_2,T) \end{cases}$$

i.e., the number of control switchings does not exceed 2 and the optimal solutions may contain the singular arcs (cases 2.3-2.5).

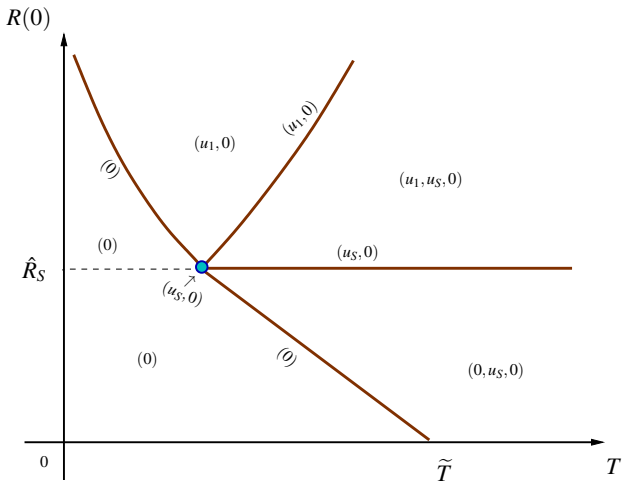


Fig 6.

As it is seen from Fig. 6 in the *singular case* on the plane $(T, R(0))$ we have more domains with different structures of the optimal control $\hat{u} = \hat{u}(t)$. These additional domains are labeled as $(u_S, 0)$ or $(a, u_S, 0)$. Note that on that intervals $t \in \Delta$ where $\hat{u}(t) = u_S$ the function $\hat{R}(t)$ takes the constant value $\hat{R}_S$:

$$\hat{R}(t) = \hat{R}_S, \qquad t \in \Delta.$$

9 Conclusions

We considered the control problem for wireless sensor networks with a single time server node and a large number of client nodes. The cost functional of this control problem accumulates clock synchronization errors in the clients nodes and the energy consumption of the server over some time interval $[0, T]$. For all possible parameter values we found the structure of optimal control function. It was proved that for any optimal solution $\widehat{R}\,(t)$ there exist a time moment τ, $0 \le \tau < T$, such that $\hat{u}(t) = 0$, $t \in [\tau, T]$, i.e., the sending messages at times close to T is not optimal. We showed that for sufficiently large u_1 the optimal solutions contain singular arcs. We found conditions on the model parameters under which different types of the optimal control are realized.

We hope that our study of the energy-saving optimization will also be usefull for analysis of other engineering problems related to modern distributed systems. In future we plan to extend these results to more general models.

References

1. Sundararaman, B., Buy, U., Kshemkalyani, A.D., Clock synchronization for wireless sensor networks: a survey. Ad Hoc Networks, 3, 3, 281–323, 2005
2. Manita A., Clock synchronization in symmetric stochastic networks, Queueing Systems, 76, 2, 149-180, 2014

3. Feistel A., Wiczanowski M., Stanczak S., Optimization of Energy Consumption in Wireless Sensor Networks, Proc. ITG/IEEE International Workshop on Smart Antennas (WSA), 2007, Wien, Austria.
4. Albu R., Labit Y., Gayraud T., Berthou P., An Energy-efficient Clock Synchronization Protocol for Wireless Sensor Networks, Computing Research Repository - CORR , vol. abs/1012.2, 2010
5. Lan Wang, Yang Xiao, Energy Saving Mechanisms in Sensor Networks. Broadband Networks, 2005. BroadNets 2005, 724 - 732, Vol. 1.
6. Xu Ning, Christos G. Cassandras, Dynamic Sleep Time Control in Wireless Sensor Networks, ACM Transactions on Sensor Networks, Vol. 6, No. 3, Article 21, 2010.
7. Moshaddique Al Ameen, S. M. Riazul Islam, Kyungsup Kwak, Energy Saving Mechanisms for MAC Protocols in Wireless Sensor Networks, International Journal of Distributed Sensor Networks Volume 2010, Article ID 163413.
8. Heinz Schattler, Urszula Ledzewicz, Geometric Optimal Control Theory: Methods and Examples. Springer, 2012
9. Volker Michel, Singular Optimal Control: The State of the Art, Berichte der Arbeitsgruppe Technomathematik, V.169, 1996
10. Zelikin M.I., Borisov V.F. Theory of chattering control with applications to Astronautics, Robotics, Economics and Engineering. Boston et al.: Birkhauser, 1994.
11. M.I. Zelikin, L.A. Manita, Optimal control for a Timoshenko beam, C.R. Mécanique 334, Issue 5 (2006) 292-297
12. Manita L., Optimal Chattering Regimes in Nonhomogeneous Bar Model. In C.H. Skiadas, "Theoretical and Applied Issues in Statistics and Demography" (book devoted to ASMDA2013, Barcelona), ISAST, 2014.
13. Powers W. F., On the Order of Singular Optimal Control Problems, J. of Optimization Theory and Applications:V. 32, No, 4, 1980
14. Pontryagin L.S., Boltyanskii V.G., Gamkrelidze R.V., Mishchenko E.F., The Mathematical Theory of Optimal Processes. John Wiley, 1962

Evolution of electoral behavior by principal axes methods

Margarita Marín[1] and Campo Elías Pardo[2]

[1] Universidad Nacional de Colombia
Bogotá, Colombia
(e-mail: mmarinj@unal.edu.cu)
[2] Universidad Nacional de Colombia
Bogotá, Colombia
(e-mail: cepardot@unal.edu.co)

Abstract. This paper study the common voting patterns in Colombian presidential elections between 1986 to 2010. Contingency tables are building with subpartitions on rows and columns, where the rows correspond to the Colombian municipalities, according to their population size and the columns correspond to the votes for candidates in each electoral period. Weighted Intra Blocks Correspondence Analysis (WIBCA) with cluster analysis is develop to study voting patterns, eliminating the variability induced by population differences and election periods. It is possible to conclude that there is an electoral pattern, mainly in the municipalities with population under 20.000, which is more clear before the 2002 election period.
Keywords: WIBCA, Contingency Tables, Cluster Analysis.

1 Introduction

In 1990 Bautista and Pachecho[1] made an study of Colombian presidential election in the period of 1972 to 1990, by the implementation of Principal Component Analysis (PCA) of a dataset that contains the results for all the departments in every period for the Liberal, Conservador and left candidates. They found that the Liberal and Conservador parties have a negative correlated behavior, and that the poll for the left candidates is independent of the results of the others candidates. This work was development before the proclamation of the 1991 new Colombian Political Constitution and the electoral reform in the 90's which laid the groundwork for more flexible rules that allows the entry and exit of new political parties in Colombian. Also, before 1986 the electoral results were reported at the departmental level and the law 136 (CNC[2]) changed the political division of Colombia by created new departments and municipalities.

With this changes in mind, if one considerate this methodology for study the current Colombian presidential election, is possible to find results that do not reflect the reality, since this method does not discount the variation introduce by the change in time caused by the entry and exit the of the new

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

candidate and political parties and the differences in the electoral behaivor of the small municipalities and the big cities.

This work study the Colombian presidential election between 1986 and 2010, excluding the variation introduce by the change of political actors in time and the differences of population size. For this, this paper is divided in five parts including this introduction. In the second part the methodology is explained, then the data and the results are displayed and finally the conclusions are presented.

2 Methodology

2.1 Principal Components Analysis

The Principal Components Analysis (PCA) is a methodology to describe large data sets by the generation of orthogonal variables (known as axes) to the original variables which keeps the most variance (inertia)(UST[3]). This representation allows the study of the relation between rows according to their values of the columns, the relation between the columns and the reduction of dimensionality (Pardo and Cabreras[4]).

Then, from the standardized matrix $\mathbf{X}$ of data, with n rows y p columns is possible to find the row and column geometrical representation of this matrix which correspond to the distance (or metric) $\mathbf{M}$ and $\mathbf{D}$ respectively. This combination of data matrix and metric matrices can be written as $\mathbf{ACP}(\mathbf{X}, \mathbf{M}, \mathbf{D})$ (Escofier and Pagès [5]).

It is possible to demonstrate that the orthogonal axes that maximize the projected inertia corresponds to the eigenvectors associate to the higher eigenvalue of the correlation matrix (Lebart *et al.*[6]).

Then, the rows of the data can be represent as the union of pairs of axes, known as factorial planes, where the plane of the first and second axes (associate whit the first and second eigenvalues and eigenvector) constitute the best projection. In these planes, nearby points indicate similarity between the individuals and distant points indicate dissimilarity. In the case of the columns the representation obtained by crossing pairs of axes allows to get a plane where the points are represented as vectors and the angles formed between the pairs of them indicate the correlation of the columns (Lebart *et al.*[6]).

2.2 Correspondence Analysis with respect to a model

The CA methodology can be used to find the best representation for contingency tables (where the rows and columns represent different variables set) [Benzécri [7], Lebart *et al.*[6]], and can be seen as a Weightes Principal Component Analysis (Pardo *et al.* [8]), denoted as $ACP(\mathbf{X}, \mathbf{M}, \mathbf{D})$.

Escofier[9] generalized the CA to consider it as the relation with a model, which is a matrix that have a relation with $\mathbf{F}$. The best know example of a

model is the independence model that arises by multiplying the marginals of the matrix of frequencies $\mathbf{F}$.

For example, one can consider the $\mathbf{F}$ as the frequency table an $\mathbf{H}$ as the independence model matrix with general term $h_{ik}^{lj} = f_{i.}^{l.} f_{.k}^{.j}$. Then, in the Simple Correspondence Analysis (SCA) which is an $ACP(\mathbf{X}, \mathbf{M}, \mathbf{D})$ where $\mathbf{X}$ has general term $x_{ik}^{lj} = \frac{f_{ik}^{lj} - f_{i.}^{l.} f_{.k}^{.j}}{f_{i.}^{l.} f_{.k}^{.j}}$, $\mathbf{M} = diag(f_{.k}^{.j})$ and $\mathbf{D} = diag(f_{i.}^{l.})$, can also be seen as a $AC(\mathbf{F}, \mathbf{H})$ with respect to the independence model.

2.3 Weighted Intra Blocks Correspondence Analysis

Intra Blocks Correspondence Analysis (IBCA) is a methodology use to represent contingency tables with sub-partitions in rows and columns. In order to facilitate the explanation of the IBCA, the Colombian presidential elections data is presented in the Table 1. In this case, *Ele* represent the year of the election, *Can* the candidate, *Cat* a group of municipalities according to their population size and *Mun* the municipality. One can see that the groups creates to new structures known as band and block.

Table 1: Contingency table with sub-partitions in rows and columns for the presidencial municipality elections

		Ele86			...		**Ele10**	
		Can1	**Can2**	...	...	...	**Can26**	**Can27**
Cat1	**Mun1**							
	Mun2							
	...							
...	...							
	...							
	...							
	...							
Cat7	**Mun960**							
	Mun961							

A band is the partition of the table, created by a group of variables in the rows (row bands) or in the columns (column bands). In the case of the Table 1 an example of row band is the vote for all candidates in all elections for municipalities in category 1, and an example of column band is the voting in all the municipalities and all the categories for 1986 election. A block is create by the intersection of a row band with a column band so, in the Table 1, an example of block is the voting for all the candidates in the 1986 election in all the municipalities in category 1.

Then, the IBCA allows to study the relationship between the municipalities and the candidates, excluding the variation introduce by the size of the populations and the year of elections. This is possible, because this methodology preforms an CA with respect to independence model between the row

and columns bands, which subtract the inertia generate by the bands leaving only the inertia of the variables within the blocks (Pardo[10]).

This implies that the IBCA can be seen as a $PCA(\mathbf{X}, \mathbf{D}, \mathbf{M})$ or an $CA(\mathbf{F}, \mathbf{B})$, where the general term of each matrix is presented in the Table 2 [Pardo[10], Pardo[8]].

Table 2: General terms in the IBCA matrix

Method	**Matrix** X	**Matrix D**	**Matrix M**	**Modelo**
IBCA	$x_{ik}^{lj} = \frac{f_{ik}^{lj} - \frac{f_{i.}^{lj} f_{.k}^{lj}}{f_{..}^{lj}}}{f_{i.}^{l.} f_{.k}^{.j}}$	$diag(f_{i.}^{l.})$	$diag(f_{.k}^{.j})$	$b_{ik}^{lj} = \frac{f_{i.}^{lj} f_{.k}^{lj}}{f_{..}^{lj}}$

However, the IBCA great limitation is that can be influenced by oversized bands (bands whit a lot of variables or weight). Taking this into account, Pardo[10] propose the Weighted Intra Blocks Correspondence Analysis (WIBCA) (as an extension of the Multiple Factorial Analysis for Contingency Table (MFACT) presented by Bécue-Bartaut and Pagès[11] in which is possible to introduce simultaneously weights to $\mathbf{M}$ y $\mathbf{D}$, in order to eliminate the effect of the oversized bands. Pardo[10] demonstrate that this weighted matrix are $\mathbf{M} = diag(\alpha_j f_{.j}^{.k})$ and $\mathbf{D} = diag(\beta_l f_{l.}^{i.})$, where α_j y β_l are the weights, which have to be estimated by iterative process.

2.4 Clustering strategy

In addition to the previous methodology, this papers implements clustering strategies for the interpretation of the results at the municipality level. This is necessary since the amount of municipalities complicates the individual analysis for the rows.

Having this in mind, in this work the mix algorithm for the classification of the individuals is used. This algorithm implement the Ward algorithm, for hierarchical classification, in order to choose the number of clusters, the gravity centres and an initial classification. Then, the results are optimized by the K-means algorithm (Lebart *et al.* [6]).

3 The data

This paper study the relations between Colombian municipalities and votes for the principals presidential candidates en each election from 1986 to 2010, according with the configuration present in the Table 1.

In Colombia the presidential term has a duration of 4 year, that means that in the period of interest seven presidential election took place. Also, this

Table 3: Presidential candidates and year of participation

Election year	Candidate Name	Candidate-year code
1986	Virgilio Barco	Bar86
	Alvaro Gomez	Gom86
	Jaime Pardo	Par86
1990	Cesar Gaviria	Gav90
	Alvaro Gomez	Gom90
	Rodrigo Lloreda	Llo90
	Antonio Navarro	Nav90
1994	Antonio Navarro	Nav94
	Andres Pastrana	Pas94
	Ernesto Samper	Sam94
1998	Harold Bedoya	Bed98
	Andres Pastrana	Pas98
	Noemi Sanin	San98
	Horacio Serpa	Ser98
2002	Luis Garzon	Gar2
	Noemi Sanin	San2
	Horacio Serpa	Ser2
	Alvaro Uribe	Uri2
2006	Carlos Gaviria	Gav6
	Horacio Serpa	Ser6
	Alvaro Uribe	Uri6
2010	German Vargas Lleras	Lle10
	Antanas Mockus	Moc10
	Rafael Pardo	Par10
	Gustavo Petro	Pet10
	Noemi Sanin	San10
	Juan Manuel Santos	Sant10

paper only considers the 27 candidates who obtained a total number of votes greater than the blank vote. The Table 3 shows the candidates included in the analysis and the year of participation.

Table 4: Classification of Colombian municipality according to population size

Category	Minimum	Maximum	Number of municipalities
Cat1	500.001	-	9
Cat2	100.001	500.000	51
Cat3	50.001	10.0000	60
Cat4	30.001	50.000	107
Cat5	20.001	30.000	134
Cat6	10.001	20.000	317
Cat7	0	10.000	441

Also, this work only takes into account the 961 municipalities, and not the 1120 municipalities that currently exist, with voting between 1986 and 2010. The absence of voting in the other 159 municipalities can respond to various reasons such as lack of the municipality, inability to install polling stations because of armed conflict, among other reasons.

The municipality classification, for the creation of the bands, is made according to the parameters established in the law 136 (CNC[2]) which is presented in the Table 4.

4 Application

This section present the principal results for the application of the WIBCA in the Colombian presidential elections data. For the implementation of the WIBCA the R-package *pamctd* is used (Pardo[13]) and for the cluster classification the R-package FactoClass (Pardo and Del Campo[12])are employed. In some cases was necessary to modify the functions to make them compatible.

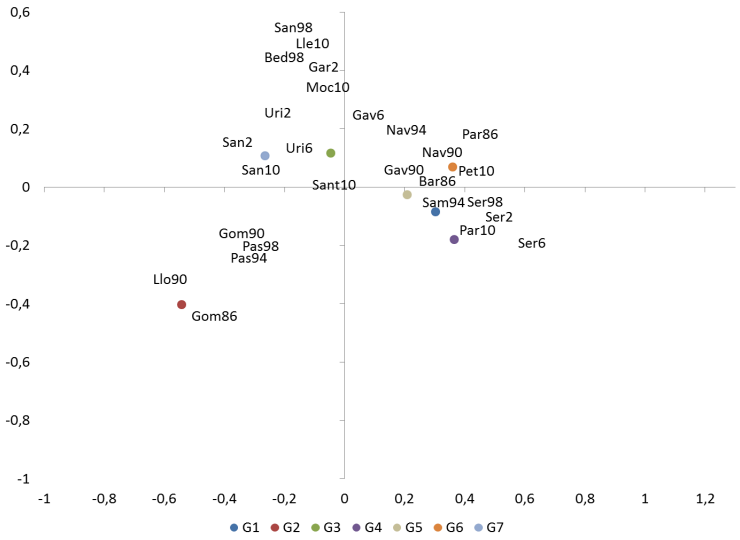


Fig. 1: WIBCA for municipalities presidential elections between 1986 and 2010

The inertia analysis and the Figure 1 (which represent the first two axes and the centres of the cluster analysis) shows the candidates with the higher percent of votes. The first and second axes explain the 51% of the inertia (31% and 20% respectively) and identify the candidates of the Liberal and Conservador parties like Serpa, Pastrana, Samper and Barco. The second axis is also associate with candidates who do not belong to traditional Colombian parties like Lleras, Uribe in 2002 and Mockus.

Leftist candidates like Petro, Gaviria, Jaime Pardo and Navarro are characterized by the third and fourth axes (11% and 9% of the inertia respectively) which means that this candidates do not have as many percentage of votes as

the previous ones but they receive voting from a different set of municipalities that the previous candidates.

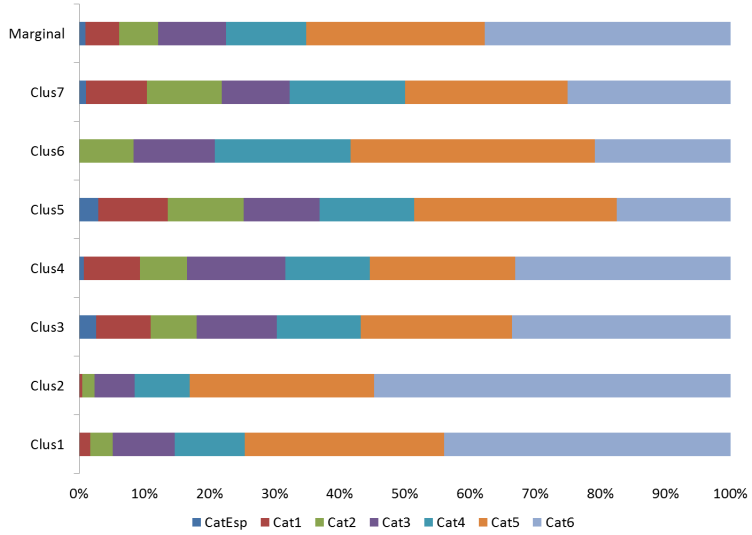


Fig. 2: Relationship between clusters and categories of municipalities

Finally, candidates like Santos and Uribe in 2006 are represented by all the axes. This could mean that this candidates get votes from all the municipalities and not only an specific category of municipality.

The Tables 5 and 6 has the cluster characterization of the WIBCA that is also presented in the Figure 1 and represented in the Figures 2, 3 and 4. In the first group the candidates Barco, Jaime Pardo, Cesar Gaviria, Samper, Serpa, Rafael Pardo and Santos present a higher percentage of voting, comparing with their national result. Except for Jaime Pardo and Santos, these candidates are affiliate with the Liberal party. This cluster has 10% of the voting, 232 municipalities and around the 75% of this municipalities (Figure 2) belong to categories 6 and 7.

In the second group the candidates Gomez, Lloreda, Pastrana, Sanin, Uribe and Santos have a higher percentage of voting, comparing with their national percentage. Except for Uribe and Santos, these candidate represent the Conservador party. This cluster has 8% of the voting, 212 municipalities and around the 80% of this municipalities (Figure 4) belong to categories 6 and 7.

The third group presents the most similar percentage of the vote compare with the national level. This cluster has 40% of the voting, 155 municipalities and and has not a dominant category.

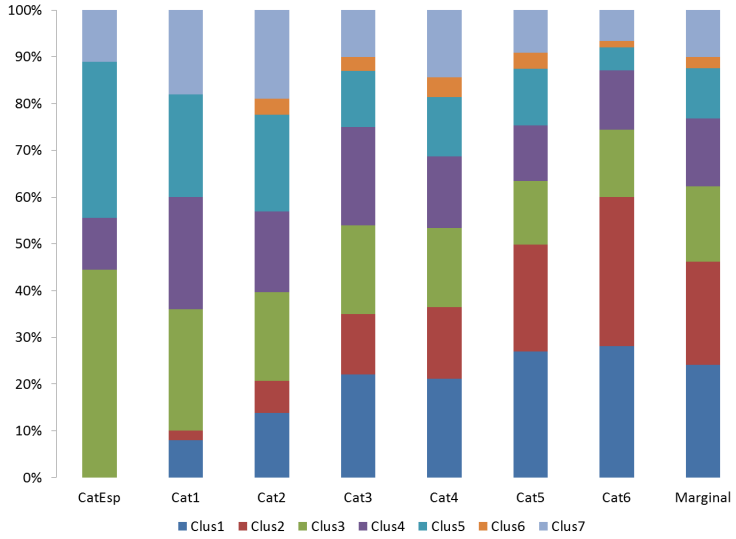


Fig. 3: Relationship between categories and cluster of municipalities

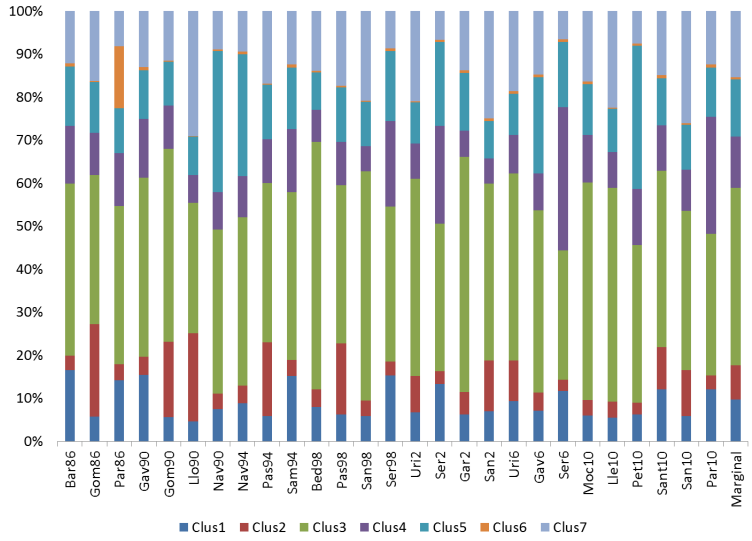


Fig. 4: Relationship between clusters and candidats

In the fourth group the candidates Barco, Cesar Gaviria, Samper, Serpa, Rafael Pardo and Petro show a higher percentage of voting, comparing with their national result. Except for Petro, these candidates are affiliate with the

Liberal party. This cluster has 12% of the voting, 139 municipalities and has not a dominant category.

Table 5: Cluster characterization for the presidential elections between 1986 and 2010: groups one to four

Candidate	Group 1		Group 2		Group 3		Group 4		Mean
	Clas/Cat	Cat/Clas	Clas/Cat	Cat/Clas	Clas/Cat	Cat/Clas	Clas/Cat	Cat/Clas	
Bar86	16,5	11,2	3,4	2,8	40	6,4	13,4	7,4	6,6
Gom86	5,8	2,4	21,4	10,9	34,7	3,4	9,8	3,3	4
Par86	14,2	0,7	3,6	0,2	36,8	0,4	12,3	0,5	0,5
Gav90	15,4	7,1	4,2	2,4	41,7	4,5	13,7	5,1	4,5
Gom90	5,6	1,3	17,5	4,9	44,8	2,4	10,1	1,9	2,2
Llo90	4,6	0,5	20,5	2,9	30,3	0,8	6,5	0,6	1,1
Nav90	7,5	0,9	3,6	0,5	38,2	1,1	8,7	0,9	1,2
Nav94	8,8	0,3	4,1	0,2	39,2	0,3	9,5	0,3	0,3
Pas94	5,9	2,4	17,1	8,7	36,9	3,6	10,3	3,5	4
Sam94	15,1	6,3	3,7	1,9	39	3,8	14,7	5	4,1
Bed98	8	0,2	4,1	0,2	57,5	0,4	7,4	0,2	0,3
Pas98	6,2	3,6	16,6	11,7	36,8	5	10,1	4,7	5,6
San98	5,9	2,7	3,6	2	53,3	5,8	5,8	2,1	4,4
Ser98	15,3	9	3,1	2,2	36,1	5	19,8	9,4	5,7
Gar2	6,2	0,7	5,2	0,7	54,6	1,4	6,1	0,5	1,1
San2	7	0,7	11,8	1,5			5,9	0,5	1
Ser2	13,3	7,3	3	2	34,4	4,5	22,7	10,2	5,4
Uri2	6,7	6,1	8,5	9,5	45,9	9,9	8,2	6,1	8,9
Gav6	7,1	2,9	4,3	2,1	42,3	4,1	8,7	2,9	4
Ser6	11,7	2,6	2,7	0,7	30,1	1,6	33,3	5,9	2,1
Uri6	9,4	10,8	9,4	13,2	43,5	11,8	9	8,4	11,2
Lle10	5,5	1,3	3,7	1,1	49,7	2,8	8,3	1,6	2,3
Moc10	5,9	2,9	3,6	2,2	50,6	5,9	11,1	4,5	4,8
Par10	12	1,2	3,3	0,4	32,9	0,8	27,2	2,2	1
Pet10	6,3	1,3	2,7	0,7	36,7	1,8	13	2,2	2
San10	5,8	0,8	10,7	1,9	37	1,2	9,7	1,1	1,4
Sant10	12	12,6	9,9	12,6	40,9	10,1	10,6	9	10,1

In the fifth group the candidates Barco, Navarro, Samper, Serpa, Carlos Gaviria and Petro have a higher percentage of voting, comparing with their national result. This candidates can be associate with leftist politics. This cluster has 13% of the voting, 103 municipalities and has not a dominant category.

In the sixth group the candidates Barco, Jaime Pardo, Cesar Gaviria, Samper, Gaviria, Pardo and Santos present a higher percentage of voting, comparing with their national result. Except for Santos, this candidates are associate whit softer leftist politics that the ones in the fifth group. This cluster has 1% of the voting, 24 municipalities and and has not a dominant category.

In the seven group the candidate Gomez, Lloreda, Pastrana, Sanin, Uribe, Lleras and Mockus show a higher percentage of voting, comparing with their national result. The majority of this candidates are associate whit right policies. This cluster has 15% of the voting, 96 municipalities and and has not a dominant category.

Table 6: Cluster characterization for the presidential elections between 1986 and 2010: groups five to seven

Candidate	Group 5		Group 6		Group 7		Mean
	Clas/Cat	Cat/Clas	Clas/Cat	Cat/Clas	Clas/Cat	Cat/Clas	
Bar86	13,9	6,9	0,7	7,7	12,2	5,2	6,6
Gom86	11,9	3,6	0,2	1,5	16,2	4,3	4
Par86	10,5	0,4	14,4	12,5	8,2	0,3	0,5
Gav90	11,3	3,8	0,7	5,6	13	3,8	4,5
Gom90	10,1	1,7	0,3	1,3	11,5	1,7	2,2
Llo90	8,9	0,8	0,2	0,4	29	2,2	1,1
Nav90	32,8	2,9	0,3	0,6	9	0,7	1,2
Nav94	28,4	0,7			9,5	0,2	0,3
Pas94	12,6	3,8	0,3	2,2	16,8	4,4	4
Sam94	14,3	4,4	0,7	5	12,4	3,3	4,1
Bed98	8,7	0,2	0,3	0,2	13,9	0,3	0,3
Pas98	12,7	5,4	0,3	2,8	17,4	6,4	5,6
San98	10,4	3,5	0,3	2	20,8	6	4,4
Ser98	16,4	7	0,5	5,2	8,7	3,2	5,7
Gar2	13,5	1,1			13,8	0,9	1,1
San2	8,7	0,7	0,5	0,9	25	1,6	1
Ser2	19,5	7,9	0,5	4,3	6,7	2,3	5,4
Uri2	9,5	6,4	0,3	4,4	21	12,2	8,9
Gav6	22,3	6,7	0,7	5,1	14,7	3,9	4
Ser6	15,2	2,4	0,5	2	6,6	0,9	2,1
Uri6	9,6	8,1	0,6	10,9	18,7	13,6	11,2
Lle10	10,2	1,8	0,2	0,8	22,4	3,4	2,3
Moc10	11,8	4,3	0,5	4,6	16,4	5,1	4,8
Par10	11,4	0,8	0,7	1,2	12,4	0,8	1
Pet10	33,3	5,1	0,5	1,9	7,5	1	2
San10	10,4	1,1	0,4	1	26,1	2,3	1,4
Sant10	10,9	8,3	0,8	14,4	14,8	9,8	10,1

5 Conclusions

This paper analyse the relation between municipalities and the results of presidential elections between 1986 and 2010, excluding the variation introduce by the size of the populations and the year of elections. For this a Weighted Intra Blocks Correspondence Analysis (WIBCA) and a mix algorithm of classification is used.

The first plan and inertia analysis show that the candidates with the higher percent of votes are the best represented in this two axes, specially the candidates Serpa, Pastrana, Samper, Barco, Lleras, Uribe in 2002 and Mockus. In the other hand, leftist candidates like Petro, Gaviria, Jaime Pardo and Navarro are characterized by the third and fourth axes, which means that this candidates do not have as many percentage of votes which means that they receive voting from a different set of municipalities as the previous candidates. Finally, candidates like Santos and Uribe in 2006 are represented by all the axes, because this candidates get votes from all the municipalities and not only an specific type.

The cluster analysis of this results shows the existence of a electoral patron in the small population size municipalities. One group of this municipalities vote for candidates which can be associate with the Liberal party and the

other group vote for candidates close to the Conservador party. However, in the 2006 election, this patron is less clear, because of the tendency of Santos and Uribe to get votes from all the municipalities.

References

1. L. Bautista and P. Pacheco, Análisis de la evolución del comportamiento electoral departamental en los últimos años: aplicación de los métodos factoriales al estudio de series temporales cortas. *Revista Colombiana de Estadística*, vol. 19, pp. 94-112, 1989.
2. CNC, *Ley 136*. Congreso Nacional de Colombia, 1994
3. USTA-OCHA.,Índice de riesgo en situación humanitaria, *Universidad Santo Tomás (USTA) y Oficina para la coordinación de Asuntos Humanitarios (OCHA)*, Bogotá, 2009.
4. C. Pardo and G. Cabarcas, Métodos estadísticos multivariados en investigación social, *Simposio de Estadística. Universidad Nacional de Colombia*, Santa Marta, 2001.
5. B. Escofier and J. Pagès, Análisis factoriales simples y múltiples: objetivos, métodos e interpretaciones. *Servicio Editorial Universidad del País Vasco*, 1992.
6. L. Lebart, M. Piron, and A. Morineau, Statisitique exploratoire multidimensionnelle. Visualisation et inférence en fouilles de données. *Dunod*, Paris, 2006.
7. J. Benzécri, Statistical analysis as a tool to make patterns emerge from the data in methodologies of pattern recognition. *Academic Press*, 1969.
8. C. E. Pardo, M. Bécue-Bertaut, and J. Ortiz, Análisis de correspondencias de tablas de contingencias con subparticiones en filas y columnas, *Revista Colombiana de Estadística*, vol. 36, pp. 115-144, 2013.
9. B. Escofier, Analyse factorielle en reference a un modele. application a l'analyse de tableaux d'echanges, *Revue de Statistique Appliquée*, vol. 32, no. 4, pp. 25-36, 1984.
10. C. E. Pardo, Métodos en ejes principales para tablas de contingencia con estructuras de partición en filas y columnas. *PhD thesis, Universidad Nacional de Colombia. Facultad de Ciencias*, Bogotá, 2011.
11. M. Bécue-Bertaut and J. Pagès, A principal axes method for comparing contingency tables: MFACT, *Computational Statistics and Data Analysis*, vol. 45, pp. 481-503, Apr. 2004.
12. C. Pardo and P. DelCampo, Combinación de métodos factoriales y de análisis de conglomerados en R: el paquete FactoClass, *Revista Colombiana de Estadística*, vol. 30, no. 2, pp. 231-245, 2007.
13. C. Pardo, pamctdp: Principal Axes Methods for Contingency Tables with Partition Structures on Rows and Columns. R, 2013.

The Coxian Phase-type distribution with a Hidden Node

Hannah Mitchell[1], Adele H. Marshall[1], and Mariangela Zenga[2]

[1] Queen's University Belfast Department of Mathematics and Physics, Belfast, United Kingdom
(E-mail: hmitchell03@qub.ac.uk a.h.marshall@qub.ac.uk)
[2] University of Milan-Bicocca, Department of Statistics and Quantitative Methods, Milan, Italy
(E-mail: mariangela.zenga@unimib.it)

Abstract. Healthcare providers are under increased pressure to ensure that the quality of care delivered to patients are off the highest standard. Modelling quality of care is difficult due to the many ways of defining it. This paper introduces a potential model which could be used to take quality of care into account when modelling length of stay. The Coxian phase-type distribution is used to model length of stay and quality of care incorporated into this using a Hidden Markov model. This model is then applied to a simulation dataset as well as patient data from the Lombardy region of Italy
Keywords: Coxian phase-type distribution, Hidden Markov model, Quality of Care.

1 Introduction

Healthcare systems across Europe and further afield have come under increased scrutiny in recent years. Many European countries are encountering a growing older population and this comes with many problems for not only governments but also healthcare providers. Elderly individuals tend to spend longer in care than the rest of the population due to complex and time consuming medical conditions and rehabilitation. This in-turn puts a strain on the hospitals budget, with healthcare managers coming under increased pressure to make sure that the hospitals deliver the best quality of care available but at the same time effectively and efficiently managing an already stretched budget [1].

Quality of care can be defined in many ways making it difficult to measure and use within a scientific study. The purpose of this paper is to develop a model which incorporates the concept of quality of care within it. An outline of how the Hidden Markov model (HMM) could potentially incorporate quality of care into a model is given. The Hidden Markov model has an underlying hidden stochastic process which potentially could represent quality of care. Initially developed as extensions for measurement errors of the standard Markov chain

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

model, the HMM but have been used in many areas of research including signal processing, in particular speech processing, medical applications and economics [4], [2], [6].

Modelling patient flow in healthcare systems is considered vital in understanding the system's activity. The Coxian phase-type distribution has successfully managed to achieve this where the distribution describes the time to absorption of a finite Markov chain in continuous time where there is a single absorbing state and the stochastic process starts in the first transient state [8]. Using the Coxian phase-type distribution has proved to be not only useful in healthcare modelling [22] but also in modelling the length of stay in Italian and Greek Universities [3] and the length of time taken for a component to fail [11].

The aim of this paper is to combine the Hidden Markov Model and the Coxian phase-type distribution with the intention of encapsulating quality of care. The Coxian phase-type distribution with a hidden node will attempt to model patient flow in healthcare with the quality of care delivered by the hospital incorporated into it via the hidden layer. The quality of care delivered by hospitals has previously been modelled using measures such as the nurse-staffing levels and the number of deaths [19], number of readmissions [13], and patient length of stay in hospital [18].

The Coxian pahse-type distribution with a hidden node will be applied to healthcare data from the Lombardy region of Italy between 2008-2011. The model will also be applied to a simulated dataset from a known Coxian phase-type distribution as a means of demonstrating the hidden node representing quality of care and how it can affect the length of stay of individuals. This model will subsequently highlight when the quality of care delivered by the hospital has changed and how quality of care affects a patients length of stay.

1.1 Hidden Markov Models

Hidden Markov models were developed by Baum in 1960 [2]. They are used extensively in signal progressing particularity in speech recognition, but their application has since been expanded upon and they have now been used in healthcare [4], financial [5] and economic applications [6].

The Hidden Markov model (HMM) is used to describe a stochastic system and is made up of a finite set of states. Each state generates an observation as well as being associated with a probability distribution. The state of the underlying Markov process cannot be observed directly but can be inferred from observations. The HMM was first developed with discrete outcomes with the Poisson distribution proving to be a popular probability distribution for this. Some applications however have observations that have continuous signals or outputs and so it would be advantageous to develop the model so that continuous densities can be used. The model has since then been developed further to allow for continuous outcomes with the mixed Guassian/Guassian distribution

being used for those applications [23].

Figure [1] shows the HMM which is a doubly embedded stochastic process,

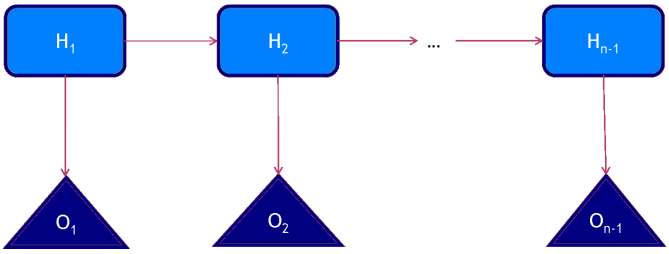


Fig. 1. The Hidden Markov Model

where the outcome (probability distribution) is observed with an underlying hidden stochastic process.

The formal definition of a Hidden Markov Model is as follows [2],

$$\lambda = (A, B, \pi) \tag{1}$$

A is a transition array, storing the probability of state j following state i.

$$A = [a_{ij}], \quad a_{ij} = P(q_t = s_j |\ q_{t-1} = s_i)$$

where $1 \leq i,\ j \leq N$

B is the observation array, storing the probability of observation k being produced from the state j, independent of t:

$$B = [b_i(k)], \quad b_i(k) = P(x_t = v_k | q_t = s_i)$$

S is the state alphabet set, and V is the observation alphabet set:

$$S = (s_1, s_2, ..., s_N)$$

Where N is the number of states in the model with the individual sates s

$$V = (v_1, v_2, ..., v_M)$$

M is the number of distinct observation symbols per state. The observation symbols correspond to the physical output of the system being modelled. The individual symbols are denoted by v.

π is the initial probability array:

$$\pi = [\pi_i], \quad \pi_i = P(q_1 = s_i)$$

The model makes two assumptions. The first is that the current state is dependent only on the previous state, which is known as the Markov property. The second is that the output observation at time t is dependent only on the current state, it is independent of previous observations and states.

The Baum-Welch algorithm which incorporates the EM-algorithm is used to estimate the parameters of the HMM. It is an iterative procedure which adjusts the HMM parameters to obtain the maximum probability of obtaining the observation sequence. Details of this algorithm can be seen in Rabiner [2].

1.2 Coxian Phase-type Distribution

Past investigations of modelling length of stay led to the discovery that a two-term mixed exponential model produces a good representation of patient survival [24]. Since then further research has endeavoured to improve the mixed exponential models with the incorporation of more complex compartmental systems and more sophisticated stochastic models such as the Coxian phase-type distribution.

The Coxian phase-type distribution is a special type of stochastic model that represents the time to absorption of a finite Markov chain in continuous time where there is a single absorbing state and the stochastic process starts in a transient state. They are a subset of the phase-type distributions introduced by Neuts in 1975 [8], having the benefit of overcoming the problem of generality within phase-type distributions.

A Coxian phase-type distribution $(X(T); t \geq 0)$ may be defined as a (latent) Markov chain in continuous time with states 1,2,...,n,n+1, $X(0)$=1,

and for i=1,2,...,n-1,

$$\text{prob}\{X(t+\delta t) = i+1 | X(t) = i\} = \lambda_i \delta t + 0(\delta t) \tag{2}$$

and for i=1,2,...,n

$$\text{prob}\{X(t+\delta t) = n+1 | X(t) = i\} = \mu_i \delta t + 0(\delta t). \tag{3}$$

here states 1,2,...,n are latent (transient) states of the process and state n+1 is the absorbing state. λ_i represents the transition from state i to state (i+1) and μ_i the transition from state i to the absorbing state (n+1). The Coxian phase-type distribution is defined as having a transition matrix $\mathbf{T}$ of the following form.

$$T = \begin{pmatrix} -(\lambda_1+\mu_1) & \lambda_1 & 0 & ... & 0 & 0 \\ 0 & -(\lambda_2+\mu_2) & \lambda_2 & ... & 0 & ... \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & -(\lambda_{n-1}+\mu_{n-1}) & \lambda_{n-1} \\ 0 & 0 & 0 & ... & 0 & -\mu_n \end{pmatrix} \tag{4}$$

The survival probability that $X(t)$=1,2,...,n is given by

$$S(x) = \alpha \exp(\mathbf{T}x)\mathbf{e} \tag{5}$$

where α=(1, 0, 0, ..., 0, 0), $\mathbf{e}$ is a column vector of 1's and $\mathbf{T}$ is the transition matrix.

The probability density function of X is

$$f(x) = \mathbf{p} \exp(\mathbf{T}x)\mathbf{q} \tag{6}$$

where $\mathbf{q} = -\mathbf{T}\mathbf{e} = (\mu_1, \mu_2, ..., \mu_n)^T$ and p=α=(1, 0, 0, ..., 0, 0). An illustration

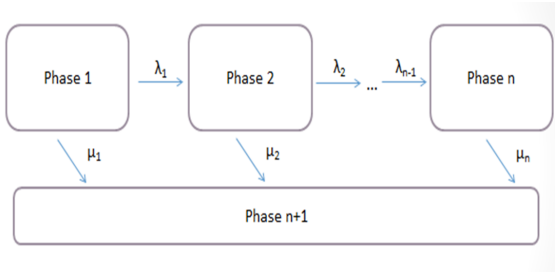


Fig. 2. Phase diagram of the Coxian Distribution

of the Coxian phase-type distribution can be seen in Figure (2). From this figure it can be seen that the process starts in the first phase and sequentially moves either through each transient state or into the absorbing state at any stage, because of this the Coxian phase-type distribution can be thought of as having some real world meaning. For example the distribution could be thought of in a hospital scenario with each phase seen as some progression of treatment. The first phase could be admittance followed by treatment and rehabilitation, with the individual being able to leave the hospital during any phase due to discharge, transfer or death.

Coxian phase-type distributions have been used in a variety of settings from component failure data [14] to prisoner remand times [15]. Marshall et al. [3] used the Coxian phase-type distribution to model career progression of students at university. Most applications of the Coxian phase-type distribution have been in modelling the length of time spent in hospital in particular McClean et al. [22] showed that the distribution was appropriate for describing

the length of time of geriatric patients in hospital.

The parameters of the Coxian phase-type distribution can be estimated in a variety of ways using a range of computer programmes or packages. Payne et al. [9] investigated the efficiency of fitting the Coxian phase-type distribution to healthcare data using SAS, R, Matlab and EMpht. EMpht which is a programme written in C uses the EM algorithm as its optimisation function and it was shown to have consistently high rates of convergence. This approach to fitting the Coxian phase-type distribution has been coded in Matlab.

1.3 Quality of Care

Quality of Care is a multifaceted concept, whose incorporation into scientific study as a result is deemed difficult but one of great importance and interest [16]. In 1855 (during the Crimean War), Florence Nightingale noticed that soldiers operated in large hospitals were more likely to die than those operated on in smaller hospitals. She identified that poor sanitation and the rapid spread of infection from patient to patient in large hospitals was the cause, so she set about doing something to improve the sanitary conditions in English field hospitals. More than a century later, there is still great interest in characterising hospitals that provide better or worse care with the aim of improving the quality delivered [17].

Until recently, to ensure that patients are receiving high-quality medical care, professional judgement was relied upon [21]. Hospitals routinely monitored poor outcomes, such as deaths or infections to identify ways to improve the quality of care. In general the monitoring of and improvement of quality was left to the clinician.

Quality of care can be defined in many ways and is not amenable to a single performance measure. In general it is defined as having the following six key domains [20]:

1) **Effectiveness**: This refers to the extent to which an intervention produces its intended result, and the concept of appropriateness; concentrating on whether interventions or services are provided to those who would benefit from them and withheld from those who would not.

2) **Access**: Access monitors waiting times, with lower waiting times for patient procedures being more beneficial.

3) **Capacity**: This takes into account the number of medical staff, bed numbers, along with how well equipped the hospitals and or surgeries are as well as the budget allocated/available to each provider.

4)**Safety**: Safety is concerned with infection control while the patient is in hospital and the elimination of unnecessary risk of harm to patients.

5) **Patient Centredness**: This measures how patients rate the quality that they are receiving whilst in hospital.

6) **Equity**: Equity is concerned with significant inequalities in life expectancy and mortality from major diseases between the least and most deprived groups.

Quality of Care has previously been incorporated into studies using the number of deaths [19], readmission [13] and length of stay of patients [18] as common proxies for the measurement of it.

Each of these proxies do have limitations when using them as a measurement for quality of care. Using death rate might not be the best measure of quality of care as it is not necessarily the quality that the hospital offers that causes the outcome but rather the disease or injury endured by the patient. Given that it is a geriatric ward that this paper will look at, death is more likely amongst this group of individuals than any other due to old age or the lowered ability to recover from disease and infection. Readmissions also could have the limitation that the data is not available for use (as is this case with the data used here) or indeed the individual could be readmitted into the same hospital but with a different complaint. Elderly patients tend to be admitted into hospital due to one condition but in fact they may be suffering from several other aliments. This also has an effect on length of stay. Elderly individuals tend to spend longer in hospital due to the multitude and variety of illness that many of them suffer from at any one time. Therefore to use length of stay data as a proxy, very long length of stays as well as short length of stays (which are mainly attributed to patients who die) could serve as flagging up potentially poorer quality of care delivered to these patients [10]. Due to length of stay being related to the outcome of patients, and that the data used does not have readmission information, length of stay was used and quality of care inferred from it.

Quality of care is difficult to measure due to its many factors, internal and external. One potential way of measuring quality is by treating it as a hidden layer. Length of stay has been seen as an indicator of quality of care and it has been shown that the Coxian phase-type distribution gives a good representation of length of stay. In this paper the Coxian phase-type distribution has been combined with the HMM, thus giving the effect of a hidden node (representing quality) being incorporated into the Coxian phase-type distribution.

1.4 Coxian phase-type distribution with a hidden node

The Hidden Markov model with continuous outcomes has only ever used the Gaussian or the mixed Gaussian distribution as its probability density function. The model was expanded upon so that the outcomes which are best represented using a Coxian phase-type distribution could be used, by letting this distribution be the probability density function.

The Coxian phase-type distribution with a hidden node was developed, with the hidden node representing quality. This model has the Markov property within the hidden layer and the probability density function. Given quality of care is being represented as the hidden layer the Markov property could infer that the quality of care delivered by the hospital for example does not depend on the past but from the previous quality delivered. This assumes that if poor quality of care was obtained during a measurement that the hospital would rectify the problem immediately so that the next patient or time measured would receive the newly modified care system. Figure [3] shows a representation of

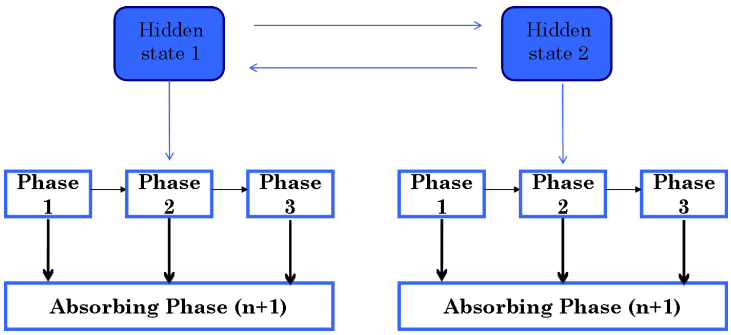


Fig. 3. Phase diagram of the Coxian phase-type distribution with a hidden node

the Coxian phase-type distribution with the hidden node. The hidden Markov model was given two hidden states, to represent good and poor quality of care, and from these two states a Coxian phase-type distribution produced. The formal definition of the HMM is given by equation (1), where λ is the parameter estimates of the model.

For the HMM the general Q function for the complete-data log-likelihood is given as

$$Q(\lambda, \lambda') = \Sigma_{q \epsilon Q} \log P(O, q|\lambda) P(O, q|\lambda') \tag{7}$$

where λ' is the initial/previous parameter estimates, O=$(o_1, ..., o_T)$ are the observed data and q=$(q_1, ..., q_T)$ is the underlying hidden state sequence.

Given a particular state sequence q, representing $P(O, q|\lambda')$ that is,

$$P(O, q|\lambda') = \pi_{q0} \Pi_{t=1}^{T} a_{q_{t-1} q_t} b_{q_t}(o_t) \tag{8}$$

where π_{q0} is the probability of initially being in state q, $a_{q_{t-1} q_t}$ is the probability of moving between the hidden states and $b_{q_t}(o_t)$ is the probability of a particular observation vector at a particular time t for state q_t. The Q function then becomes:

$$Q(\lambda, \lambda') = \Sigma_{q \epsilon Q} \log \pi_{q0} P(O, q|\lambda') + \Sigma_{q \epsilon Q} \{\Sigma_{t=1}^{T} \log a_{q_{t-1} q_t}\} p(O, q|\lambda') + \tag{9}$$

$\Sigma_{q\epsilon Q}\{\Sigma_{t+1}^{T}\log b_{q_t}(o_t)\}P(O,q|\lambda')$

The parameters that require optimisation are in three independent terms and can thus be optimised individually. The Coxian phase-type distribution being the output of the Hidden Markov model, $b_{q_t}(o_t)$ in equation [9] can be replaced by the probability density function of the Coxian phase-type distribution, equation [6]. The model is implemented in Matlab with the EM and Runge-kutta algorithms used to for fitting the Coxian phase-type distribution and the Baum-Welch used to fit the HMM.

The Viterbi algorithm [2] finds the best state sequence for the observations of the hidden states. It is also coded in Matlab so that the state which best represents the quality of care at each time point can be produced. This will also highlight when a change of state is most likely to occur.

1.5 Simulation Study

The actuar package in R was used to simulate length of stay data from a three phase Coxian phase-type distribution.The simulated dataset consists of 100 data points ranging from 0.19 to 84.65 days with an average of 23 days. Table [1] displays the three phase Coxian distribution parameters used to simulate the data. The proposed Coxian HMM was applied to the dataset using Matlab.

Table 1. Coxian phase-type distribution for simulated data

Phase	Transition Rates
3	μ_1 =0.0462, μ_2 =0.0011, μ_3 =0.0779, λ_1 =0.0589, λ_2 =0.0779

The model probability of initially being in either state is

$$\pi = (\,0.2367\ 0.7633\,) \tag{10}$$

The initial probability suggests that the hospital quality of care is initially in state 2. The transition matrix for the Hidden Markov model was given as

$$A = \begin{pmatrix} 0.5091 & 0.4909 \\ 0.5107 & 0.4893 \end{pmatrix} \tag{11}$$

The hidden transition matrix and initial probabilities of the model suggest that the hospital starts in state 2 with a probability of 0.7633. The hidden transition matrix shows that the probability of being in state 1 and moving to state 2 at the next time point is 0.4904 and the probability of being in state 2 and moving to state 1 is 0.5107. This suggests that although the hospital care is initially in state 2 there is a slightly higher probability that the hospital quality of care will transition to state 1 and a slightly higher probability of staying in state 1 than transitioning back to state 2.

The parameters for the Coxian phase-type distributions given by each state are provided below. Table [2] shows the parameters for state 1 and state 2. The Coxian phase-type distribution for state 1 and state 2 both suggest that

Table 2. State 1 and 2 Coxian Phase-type distribution parameter estimates

State	Parameter Estimates
1	$\mu_1 = 0.0000, \mu_2 = 3.3570, \mu_3 = 0.0608$ $\lambda_1 = 0.0608, \lambda_2 = 2.0900$
2	$\mu_1 = 0.0000, \mu_2 = 2.7302, \mu_3 = 0.0609$ $\lambda_1 = 0.0583, \lambda_2 = 2.2430$

no patients left the system in the first phase but all transitioned through to the next stage of treatment. In state 1 the patients left the second phase of the Coxian phase-type distribution at a faster rate than in state 2. They exited into the absorbing state at a transition rate of 3.357 in comparison to state 2 which had a transition rate of 2.7302. The original Coxian phase-type distribution showed that no one left at the second phase and the transition through to the second phase for the original distribution was similar to that of the two Coxian phase-type distributions with hidden nodes. The Coxian phase-type distributions between the two states after phase 2 are very similar with state 2 showing larger transition rates. State 2 suggests that patients go through the system slightly quicker in the final part of the Coxian phase-type distribution than in state 1.

The interpretation of the states are difficult. In this context they are thought of as quality of care. In general the interpretation of the states are taken by the average of the outcomes of the HMM. The Viterbi algorithm gives the best state that each observation belongs to, from this the average length of stay of each state was found. The outcome of the Viterbi algorithm also gives how the states change over time, with lower length of stays changing the state from state 1 to state 2. Those observations which the Viterbi algorithm showed to be in state 1 had an average length of stay of 42.62 days, whereas the average length of stay for the observations in state 2 was 9.15 days. This would suggest that given the individuals are in the system longer that state 1 shows a better quality of care than state 2.

In general patients who leave the system at the start tend to leave due to death in-comparison to those individuals who are transferred to another hospital or home. This suggests given the average length of stay of each of the states that state 1 suggests better quality than state 2. However state 1 shows that individuals leave phase 2 at a faster rate than state 2.

The phases in the 3 phase Coxian distribution could be thought of as acute care for the first phase, further treatment for the second and rehabilitation for the final phase. Given both outcomes suggest that no one left the first

phase through the absorbing state, patients left at a slightly faster rate in state 1 and along with the average length of stay being a lot greater than state 2 would suggest possibly that the hospital is potentially ill equipped to cope with certain individuals aliments and are waiting for them to be transferred to a different hospital or care-home. This may fall under the quality of care domain of Capacity if the hospital is not well equipped, and hence the quality of care delivered to some individuals is of a poorer standard.

From previous research small and long length of stays both are possibly highlighting decreased quality of care [10]. However with this dataset the model has partitioned the small and large length of stays as two different quality of care states. If more detail is available, for example the outcome of the patients along with the aliments/problems of the patients, this may also help with the interpretation of the states.

1.6 Italian dataset

The proposed Coxian HMM is applied to a large Italian administrative dataset, which contains all of the geriatric wards in the Lombardy Region. There were no day hospital cases with all patients being ordinary admission. The data consists of length of stay information for 2174 patients aged 65 years of age or older that were admitted into a geriatric ward between 2008-2011. The hospital considered for this illustration is a public hospital with an average length of stay of 17.83 days.
A coxian phase-type distribution was fitted to the data resulting in a four Coxian as the best fit, determined so by comparing the AIC values. The parameter estimates of this Coxian phase-type distribution can be seen in table (3) The

Table 3. State 1

Phases	Transition rates
4	μ_1 =0.0000, μ_2 =0.0000, μ_3 =0.0000, μ_4 =0.220772
	λ_1 =0.220772, λ_2 =0.220772, λ_3 =0.220772

Coxian phase-type distribution with a hidden node was then applied to this data. The results show that the initial probability of being in each state is

$$\pi = (\,0.4984\ 0.5016\,) \tag{12}$$

The transition matrix is

$$A = \begin{pmatrix} 0.4986 & 0.5014 \\ 0.4986 & 0.5014 \end{pmatrix} \tag{13}$$

The initial probabilities show that the hospital is initially in state 2 with a probability of 0.5016. The A matrix shows that the hospitals hidden state has a small probability of changing if it is in state 1. The probability of it transitioning to state 2 is 0.5014 and the probability of remaining in state 2 if it is in state 2 already is also 0.5014. Given the probabilities for the hidden states

show that the observations are equally as likely to be in state 1 as in state 2 this potentially shows that the observations each month have not changed significantly to warrant a change in hidden state and therefore quality of care.

The Coxian phase-type distribution for each of the hidden states is displayed in table [4]. From table [4] the transition rates for the two states of the Coxian

Table 4. Coxian phase-type distributions for state 1 and state 2

States	Parameter Estimates
1	μ_1 =0.000000, μ_2 =0.000000, μ_3 =0.000000, μ_4 =0.220756 λ_1 =0.220756, λ_2 =0.220756, λ_3 =0.220756
2	μ_1 =0.000000, μ_2 =0.000000, μ_3 =0.000000, μ_4 =0.220774 λ_1 =0.220774, λ_2 =0.220774, λ_3 =0.220774

phase-type distribution with a hidden node are very similar. They are also similar to the original Coxian phase-type distribution parameters table [3]. This could suggest that there is barely any difference between the two states suggesting that the quality of care has not changed dramatically over time. Given that the Hidden Markov model with the Coxian phase-type distribution has given parameters very similar to the Coxian phase-type distribution perhaps this model "fine tunes" the Coxian phase-type distribution in allowing a hidden layer of unobserved factors to be taken into account.

When the Viterbi algorithm is applied, it shows that the best state for each of the data points were the same. This suggests that the quality of care delivered has remained the same throughout the months. Looking at the average length of stay for each month the range is small suggesting no real difference in length of stay over the years.

The Hidden transition matrix, the Coxian phase-type distributions for the two hidden states and the Viterbi algorithm all suggest that the quality if care in Hospital A has remained the same over the 4 year period.

A more in-depth approach could be suggested by looking at the length of stay data over each week or to use data manipulation so it could be used for each day if there is missing time points. This would then go into more detail if there had been any changes of quality of care on a day to day basis. To get a better picture of quality of care within this hospital, covariates could be incorporated into the model. If the hospital is a public or private hospital, if its small or large, the number of staff that they have working each day and the number of beds that are available each day may give a better picture as to the quality delivered by the hospital.

1.7 Conclusions/Further Work

This paper introduces the Coxian phase-type distribution with a hidden node. It expanded the Hidden Markov model to allow the Coxian phase-type distribution as the probability density function to represent the observations. This model was then applied to a simulated dataset and a real hospital dataset from the Lombardy region of Italy. The model was used to measure quality of care and how it affects the Coxian phase-type distribution as well as giving how the quality of care changes over time and the probabilities of a change of quality of care happening. The results show how interpreting the states of the hidden Markov model is difficult, and that taking the average outcome of each state requires further refinement.

Further work includes incorporating covariates into the transition matrix of the Hidden Markov model with the Coxian phase-type distribution as the output. This will show how quality of care changes or effects the length of stay of patients when for example the number of beds, staff levels etc change. This has the potential to be used therefore by hospital managers for planning and efficiently running the healthcare system. Other future work includes incorporating covariates into the Coxian phase-type distribution to show how certain they affect patient length of stay. This has previously been incorporated into a Coxian (without the HMM). Cost is factor in the running of the hospital and the quality of care delivered. Quality of care and cost potentially go hand in hand thus further work will investigate the inclusion of cost into the model. The Coxian phase-type distribution has been previously developed to estimate costs therefore there is potential to incorporate this theory into the Hidden model. Having Cost in this model will also benefit the healthcare managers so they can plan and evaluate the possible benefits or problems associated with reducing the amount of money, staff, beds etc when trying to run a hospital which delivers high quality of care to all patients within a tight and stringent budget.

References

1. Janssen. Economist Intelligence Unit, The future of healthcare in Europe. *The Economist*, Janssen, 2011
2. L. R. Rabiner. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77, **2**, 257-286, 1989.
3. A.H.Marshall, M. Zenga, S. Giordano. Modelling Students' Length of Stay at University Using Coxian Phase-type Distributions. *International Journal of Statistics.* 73-89, 2013
4. B. Cooper and M. Lipsitch. The analysis of hospital infection data using hidden Markov models. *Biostatistics.* 2, **2**, 223-237, 2004.
5. A. Gupta and B. Dhingra. Stock Market Prediction Using Hidden Markov models. *IEEE.* 2012.
6. R. Bhar and S. Hamori. Hidden Markov Models: Applications to Financial Economics. *Advanced Studies in Theoretical and Applied Econometrics.* Springer, 2004.

7. S. Asmussen, O. Nerman and M. Olsson. Fitting Phase-type Distributions via the EM Algorithm. *Scandinavian Journal of Statistics*, 23, **4**,419-441, 1996.
8. M.F. Neuts, Matrix-Geometric Solutions in Stochastic Models. *John Hopkins University Press*, 1981.
9. K. Payne, A.H. Marshall and K.J. Cairns. Investigating the efficiency of fitting Coxian phase-type distributions to health care data. *IMA Journal of Management Mathematics*, 2011.
10. J.W. Thomas, K. E. Guire, G.G. Horvat. Is patient length of stay related to quality of care. *Hospital Health Service.* 42, **4**, 489-507, 1997.
11. M.J. Faddy. Phase-type distributions for Failure Time. *Mathematical computer Modelling*, 22, **10-12**, 63-70, 1995
12. S.I. McClean, P. Millard. Modelling In-Patient Bed Usage Behaviour in a Department of Geriatric Medicine *Methods of Information in Medicine*, 32, 79-81, 1993.
13. A. Clarke. Readmission to Hospital: a measure of quality or outcome? *The international journal of healthcare improvement.* 13, 10-11,2004.
14. M. J. Faddy. Phase-type distributions for Failure Time. *Mathematical computer Modelling*, 22, 247-255,1994.
15. M. J. Faddy. Examples of fitting structured phase-type distributions. *Applied Stochastic models and data analysis.*247-255, 1994.
16. L. Fallowfield. What is quality of life? *Health economics*,2009.
17. E. B. Keeler, L. V. Rubenstein, K.L.Kahn, D. Draper, E.R. Harrison, M.J. McGinty, W.H. Rogers, R.H. Brook. Hospital Characteristics and Quality of Care. *The Journal of the American Medical Associatio.* 268,**13**, 1709-1714, 1992.
18. T. P. Meehan, M.J. Fine, H.M. Krumholz, J.D. Scinto, D.H. Galusha, J.T. Mockalis, G.F. Weber, M.K. petrillo, P.M. Houck, J.M. Fine. Quality of care, Process, and Outcomes in Elderly Patients With Pneumonia. *The Journal of the American Medical Association.* 278, **23**, 2080-2084,1997.
19. J. Needleman, P. Buerhaus, S. Mattke, M. Stewart. Nurse-Staffing levels and the Quality of Care in Hospitals. *The New England Journal of Medicine.* 346, **22**, 2002.
20. K. Sutherland, N. Coyle. Quality in Healthcare in England, Wales, Scotland, Northern Ireland: an intra-UK chartbook. *The Health Foundation*, 2008.
21. D. Blumenthal. Quality of Health Care, Part 2: Measuring Quality of Care. *The New England Journal of Medicine.*, 1996.
22. S.I. McClean, P. Millard. Patterns of Length of Stay after Admission in Geriatric Medicine: An Event History Approach. *The Statistician*, 263-274, 1993.
23. J.A. Bilmes. A Gentle Tutorial of the EM Algorithm and its Applications to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. *International Computer Science Institute.* Berkeley, 1998.
24. P.H. Millard. Throughput in a department of geriatric medicine: a problem of time, space and behaviour. *Health trends.* 24, 20-24, 1991.

Spectral Theory of Convolution Operators with Multi-point Perturbation and its Applications to Population Dynamics

Stanislav Molchanov[1] and Elena Yarovaya[2]

[1] Department of Mathematics, University of North Carolina, Charlotte, USA
(E-mail: smolchan@uncc.edu)
[2] Department of Probability Theory, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Moscow, Russia
(E-mail: yarovaya@mech.math.msu.su)

Abstract. Spectral properties of linear operators play an important role in the theory of branching random walks. Resolvent analysis of bounded symmetric operators with multi-point potential generating continuous-time branching random walks on d-dimensional lattices with a finite set of branching sources has allowed to study large deviations for branching random walks in a number of works of the authors. Using these results the limit structure of the population inside of the propagating front is investigated. A special attention is paid to the case when the spectrum of an evolution operator of mean numbers of particles contains only one positive isolated eigenvalue.
Keywords: Branching Random Walks, Evolution Operator, Green Functions, Large Deviations, Front.

1 Introduction

We review the recent studies of the branching random walk with a compactly supported birth rate potential. To simplify the exposition, we formulate only the most principal results and ideas skipping technical details and possible generalisations.

Let us formulate the problem. On the lattice $\mathbb{Z}^d$ we consider the branching random walk, similar to the Kolmogorov-Petrovsky-Piskunov (KPP) model [7]. The central object here is the point field $n(t,x,y)$, $t \geq 0$, $y \in \mathbb{Z}^d$, where $n(t,x,y)$, at any given time $t \geq 0$, is the number of particles at the point $y \in \mathbb{Z}^d$, provided that at the start a single particle is located at the point $x \in \mathbb{Z}^d$, that is, $n(0,x,y) = \delta(x-y)$.

The initial particle performs the symmetric random walk $x(t)$ with the generator

$$\mathcal{L}_x f = (\mathcal{L}f)(x) = \sum_{z \neq 0} (f(x+z) - f(x))\, a(z),$$

that acts on the space $l^p(\mathbb{Z}^d)$, $1 \leq p \leq \infty$, and has the following properties: $a(z) = a(-z)$ (symmetry: i.e., $\mathcal{L} = \mathcal{L}^*$ in $l^2(\mathbb{Z}^d)$); $\sum_{z \neq 0} a(z) = -a(0) = 1$

3^{rd} *SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal*
C. H. Skiadas (Ed)

(normalization: the total intensity of jumps is 1); for every $z \in \mathbb{Z}^d$ there exists a set of vectors $z_1, z_2, \dots, z_k \in \mathbb{Z}^d$, such that $z = \sum_{i=1}^k z_i$ and $a(z_i) > 0$ for $i = 1, 2, \dots, k$ (irreducibility), for details see Yarovaya [15].

For the random walk $x(t)$ with the generator $\mathcal{L}$, the time τ spent by a particle at a point x is exponentially distributed with parameter 1; at time $\tau + 0$ the particle jumps to a point $x + z$ with intensity $a(z)$, which determines the distribution of the jump of the process.

In addition, for any time interval $(t, t + dt)$ each particle at $x \in \mathbb{Z}^d$ in the population independently of others can split in two particles, located at the same point. Later on these two particles (the parental one and the offspring) evolve independently of each other according to the same law as the initial particle.

The rate of splitting is presented in the form $V(x)dt = \beta V_0(x)dt$, where β is the coupling constant, and $V_0(\cdot)$ is a function subjected to the normalization

$$\max_{x \in \mathbb{Z}^d} V_0(x) = 1.$$

The central assumption is the finiteness of the support $\operatorname{supp} V$ of the function V: there are finitely many points $x_1, x_2, \dots, x_N \in \mathbb{Z}^d$, where $V_0(x) > 0$ and, say, $V_0(x_1) = V_0(x_2) = \dots = V_0(x_m) = 1$, $m \le N$. They form $\operatorname{supp} V$, whereas $V_0(x) \equiv 0$ if $x \notin \operatorname{supp} V$.

One can consider more general schemes of the branching (more than one offspring or non-local but fast decreasing potential $V_0(\cdot)$, etc.) but we will concentrate on the simplest case. The theory also can include the mortality rate, as in [14,15], but again for transparency we exclude it. More details can be found in [9–12].

Consider the generating function

$$u := u_z = u_{z_1,\dots,z_l}(t, x, y_1, \dots, y_l) = \mathbf{E}_x z_1^{n(t,x,y_1)} \cdots z_l^{n(t,x,y_l)},$$

where $y_1, \dots, y_l$ are different points of the lattice $\mathbb{Z}^d$ and $z_1, \dots, z_l$ are complex variables. For the evolution of u the standard calculation gives the KPP-type equation

$$\partial_t u = \mathcal{L}_x u + \beta V_0(x)(u^2 - u),$$

where

$$u_{z_1,\dots,z_l}(0, x, y_1, \dots, y_l) = \begin{cases} z_i, & x = y_i,\ i = 1, \dots, l, \\ 1, & x \ne y_1, \dots, y_l. \end{cases}$$

This case differs from the classical KPP-situation in two ways: we deal with the lattice $\mathbb{Z}^d$ instead of the continuum $\mathbb{R}^d$ and with the convolution bounded symmetric operator $\mathcal{L}$ instead of the Laplacian Δ.

Differentiating $u_{z_1,\dots,z_l}(t, x, y_1, \dots, y_l)$ in variables z_j, $j = 1, \dots, l$, we get the moment equations. The simplest one is the equation for the first moment. If $m_1(t, x, y) = \mathsf{E}_x n(t, x, y)$ then

$$\begin{cases} \partial_t m_1 & = \mathcal{H}_\beta m_1, \\ m_1(0, x, y) & = \delta(x - y), \end{cases} \tag{1}$$

where the Hamiltonian $H_\beta = \mathcal{L}_x + \beta V_0(\cdot) I$ is a bounded self-adjoint operator on $l^2(\mathbb{Z}^d)$. For more details about the properties of the operator H_β see [12,14,16]. For the mixed second moment

$$m_2(t,x,y_1,y_2) = \mathsf{E}_x n(t,x,y_1) n(t,x,y_2), \quad y_1 \neq y_2,$$

the relevant equation has the form

$$\begin{cases} \partial_t m_2(t,x,y_1,y_2) = \mathcal{H}_\beta m_2 + 2\beta V_0(x) m_1(t,x,y_1) m_1(t,x,y_2), \\ m_2(0,x,y_1,y_2) \quad = \delta(x-y_1) + \delta(x-y_2). \end{cases}$$

Equations for all higher order moments can be obtained in a similar way, see, e.g., [14]. For any moment $m_k(t,x,y_1,\dots,y_l) = \mathsf{E}_x n^{j_1}(t,y_1)\dots n^{j_l}(t,y_l)$ of order k depending on several points $y_1,\dots,y_l$, where $j_1+\dots+j_l = k$, the related equation has the form

$$\begin{cases} \partial_t m_k(t,x,y_1,\dots,y_l) = \mathcal{H}_\beta m_k + g_k(m_1,\dots,m_{k-1}), \qquad k \geq 2, \\ m_k(0,x,y_1,\dots,y_l) \quad = \delta(x-y_1) + \dots + \delta(x-y_l), \end{cases}$$

where g_k is a polynomial of order k depending on the moments m_j, $j \leq k-1$. Here all the moment equations include the Hamiltonian H_β.

Let us consider the fundamental solutions of two closely related parabolic problems:

$$\begin{cases} \partial_t p(t,x,y) = \mathcal{L}_x p, \\ p(0,x,y) = \delta(x-y). \end{cases}$$

Here $p(t,x,y) = \mathsf{P}_x(x(t)=y) = p(t,0,y-x) = p(t,0,x-y)$ is the transition probability of the underlying random walk $x(t)$.

The second Schrödinger parabolic problem contains information about the first moment.

$$\begin{cases} \partial_t m_1 \quad\quad = \mathcal{H}_\beta m_1, \\ m_1(0,x,y) = \delta(x-y), \end{cases}$$

Due to the Feynman-Kac formula one has

$$m_1(t,x,y) = p(t,x,y)\mathsf{E}_x\left[e^{\beta\int_0^t V(x_s)\,ds} \mid x(t)=y\right],$$

and then $m_1(t,x,y) \geq p(t,x,y)$. Of course, the fundamental solution $m_1(t,x,y)$ is not translation invariant. For fixed t, x and very large $|y|$ one can expect that $p(t,x,y) \sim m_1(t,x,y)$. An asymptotic analysis of $p(t,x,y)$ and $m_1(t,x,y)$ will be given below.

Let us note that using these results and Duhamel's formula for semigroups

$$\mathsf{P}_t = e^{t\mathcal{L}}, \quad \mathsf{M}_t = e^{t\mathcal{H}_\beta},$$

one can calculate moments. Let $n(t,x,\Gamma) = \sum_{y\in\Gamma} n(t,x,y)$, $\Gamma \in \mathbb{Z}^d$, then for $m_1(t,x,\Gamma)$ we get

$$m_1(t,x,\Gamma) = (\mathsf{P}_t \boldsymbol{I}_\tau)(x) = \sum_{y\in\Gamma} p(t,x,y).$$

For the equation

$$\partial_t u = \mathcal{L}_x u + f(t,x), \quad u(0,x) = 0,$$

we have by Duhamel's formula

$$u(t,x) = \int_0^t ds(\mathsf{P}_{t-s} f)(s,x).$$

Similarly for

$$\partial_t u = \mathcal{H}_\beta u + f(t,x), \quad u(0,x) = 0,$$

we have

$$u(t,x) = \int_0^t (\mathsf{M}_{t-s} f)(s,x) ds.$$

2 Spectral theory of the operator $\mathcal{H}_\beta$

The operator $\mathcal{H}_\beta$ is a bounded self-adjoint operator on $l^2(\mathbb{Z}^d)$. Its spectrum consists of two components: the absolutely continuous spectrum located in $[-\min_\kappa \widehat{a}(\kappa), 0]$ where $\widehat{a}(\kappa) = \sum_{x \neq 0} a(x) \cos(\kappa, x)$ for $\kappa \in [-\pi, \pi]^d$, and the finite discrete spectrum $\sigma_d(\mathcal{H}_\beta)$ belonging to the positive part of the λ-axis. The detailed analysis of the properties of the operator $\mathcal{H}_\beta$ for different types of branching random walks can be found in [16,9] .

The absolutely continuous spectrum can be described in terms of the general scattering theory [13] which includes the irradiation conditions on the infinity. For the applications to the population dynamics, the most important part of the spectrum of the operator $\mathcal{H}_\beta$ is its discrete spectrum $\sigma_d(\mathcal{H}_\beta)$, see Cranston *et al* [2] and Molchanov and Yarovaya [17]. It contains no more than N positive eigenvalues λ_j, $j \geq 0$, since $\beta V_0(x)$ is the rank N perturbation of the operator $\mathcal{L}_x$ with a purely absolutely continuous spectrum, see Yarovaya [16]. If β is very large then $\sigma_d(\mathcal{H}_\beta)$ consists of N nonnegative eigenvalues $\lambda_0 > \cdots > \lambda_{N-1} > 0$. It is true for instance, if $\beta \min_{\mathrm{supp}\, V} V_0(x) > \|\Delta\|_2$. The situation of small β is more interesting.

Theorem 1. *If the underlying random walk $x(t)$ is recurrent then $\sigma_d(\mathcal{H}_\beta) \neq \emptyset$ for any $\beta > 0$. In particular, there exists a simple leading eigenvalue $\lambda_0(\beta, V_0) > 0$ for which the corresponding eigenfunction $\psi_0(x, \beta)$ is strictly positive.*

If the underlying random walk $x(t)$ is transient then there exists a value $\beta_{cr} > 0$, such that for

- *$\beta \leq \beta_{cr}$ the positive discrete spectrum is empty: $\sigma_d(\mathcal{H}_\beta) = \emptyset$;*
- *$\beta > \beta_{cr}$ there exists at least one leading eigenvalue $\lambda_0(\beta, V_0) > 0$ corresponding to eigenfunction $\psi_0(x, \beta) > 0$.*

Remark 1. If the underlying random walk $x(t)$ on $\mathbb{Z}^d$ is transient and $\beta = \beta_{cr}$ then there may exist an eigenvalue $\lambda_0(\beta_{cr}, V_0)$, which is equal to 0. As was shown in [14], for branching random walks with a finite variance of jumps such an eigenvalue $\lambda_0(\beta_{cr}, V_0) = 0$ exists if and only if $d \geq 5$.

Let us present an algorithm for calculating the eigenvalue $\lambda_0(\beta, V_0)$ in the transient case.

By definition, for a number $\lambda > 0$ to be an eigenvalue of the operator $\mathcal{H}_\beta$, it is necessary and sufficient that there exists a nonzero element $\psi \in l^2(\mathbb{Z}^d)$ satisfying the equation

$$\mathcal{H}_\beta \psi(x) = \mathcal{L}_x \psi(x) + \beta V_0(x)\psi(x) = \lambda \psi(x).$$

Rewrite the last equation in the form

$$(\mathcal{L}_x - \lambda I)\psi(x) = -\beta \sum_{j=1}^{N} V_0(x_j)\delta(x - x_j)\psi(x_j). \tag{2}$$

Let us recall that the solution of the equation

$$(\mathcal{L}_x - \lambda I)\psi(x) = -\delta(x - y),$$

is called the Green function

$$G_\lambda(x, y) = \int_0^\infty e^{-\lambda t} p(t, x, y)\, dt.$$

Then from (2) one can deduce that

$$\psi(x) = \beta \sum_{j=1}^{N} V_0(x_j) G_\lambda(x, x_j)\psi(x_j),$$

and by taking the vector x from $\operatorname{supp} V$, i.e. $x = x_1, \ldots, x_N$, we get the linear system

$$\psi(x_i) = \beta \sum_{j=1}^{N} V_0(x_j) G_\lambda(x_i, x_j)\psi(x_j) = (\mathcal{A}(\lambda, \beta)\psi)(x_i).$$

Consider now a square $N \times N$ matrix

$$\mathcal{A}(\lambda, \beta) = [a_{ij}] = [\beta V_0(x_j) G_\lambda(x_i, x_j)],$$

which has strictly positive elements. Then the Perron-Frobenius theorem ensures the existence of a strictly positive simple eigenvalue $\mu_0(\lambda, \beta)$ with the corresponding positive eigenvector $\psi_0(x_j)$, $j = 1, \ldots, N$. These objects are analytically depend on β and λ, see Kato [6]. Since $G_\lambda(\cdot, \cdot) \le \frac{1}{\beta}$ then the eigenvalue $\mu_0(\lambda, \beta)$ decreases in the first variable and increases in the second one. It is also clear that $\mu_0(\lambda, \beta) \to 0$, as $\lambda \to \infty$, for any fixed $\beta > 0$.

Let us note that for a fixed β the leading eigenvalue $\lambda_0(\beta, V_0)$ of the operator $\mathcal{H}_\beta$ is a root of the equation

$$\mu_0(\lambda, \beta) = 1$$

or of the equation

$$\det(\mathcal{A}(\lambda, \beta) - I) = 0,$$

which can be written in the form

$$\det\begin{bmatrix} \beta V(x_1)G_\lambda(x_1,x_1)-1 & \cdots & \beta V(x_1)G_\lambda(x_1,x_N) \\ \beta V(x_2)G_\lambda(x_2,x_1) & \cdots & \beta V(x_2)G_\lambda(x_2,x_N) \\ \cdots & \cdots & \cdots \\ \beta V(x_N)G_\lambda(x_N,x_1) & \cdots & \beta V(x_N)G_\lambda(x_N,x_N)-1 \end{bmatrix} = 0$$

or

$$\det\begin{bmatrix} G_\lambda(x_1,x_1)-\frac{1}{\beta V_0(x_1)} & \cdots & G_\lambda(x_1,x_N) \\ G_\lambda(x_2,x_1) & \cdots & G_\lambda(x_2,x_N) \\ \cdots & \cdots & \cdots \\ G_\lambda(x_N,x_1) & \cdots & G_\lambda(x_N,x_N)-\frac{1}{\beta V_0(x_N)} \end{bmatrix} = 0,$$

where $G_\lambda(x_i,x_i)=G_\lambda(0,0)$, $i=1,2,\dots,N$.

It is not difficult to prove that in the case $\beta>\beta_{cr}$ the number $N=N(\beta)$ of the positive eigenvalues of the operator $\mathcal{H}_\beta$ grows when the parameter β increases.

Let us denote by β_1 the critical value of the parameter β such that for $\beta_{cr}<\beta<\beta_1$ the operator $\mathcal{H}_\beta$ has only one eigenvalue (the ground state energy). Here $\beta_{cr}=0$ if $x(t)$ is recurrent and $\beta_{cr}>0$ for the transient $x(t)$.

Consider the case when $\beta_{cr}<\beta<\beta_1$, and solve the first moment equation (1) for $m_1(t,x,\Gamma)=\mathsf{E}_x n(t,x,\Gamma)$ with the initial conditions $m_1(0,x,\Gamma)=\mathbf{I}_\Gamma(x)$. Then we have

$$m_1(t,x,\Gamma)=(\boldsymbol{I}_\Gamma,\psi_0)\psi_0(x)e^{\lambda_0(\beta)t}+\tilde{m}_1(t,x,\Gamma).$$

Here $\tilde{m}_1$ is the projection of m_1 on the invariant spectral subspace corresponding to the absolutely continuous part of the spectrum of $\mathcal{H}_\beta$. Due to scattering theory for the operator $\mathcal{H}_\beta$ [13] we have that $\|\tilde{m}_1(t,x,\Gamma)\|_\infty=O(1)$ uniformly in Γ, i.e., main contribution to $m_1(t,x,\Gamma)$ for large sets Γ and not very large x give the first term:

$$m_1(t,x,\Gamma)=(\boldsymbol{I}_\Gamma,\psi_0)\psi_0(x)e^{\lambda_0(\beta)t}.$$

The asymptotics of $m_1(t,0,y)$ depends mainly on the structure of $\psi_0(y)$ when $|y|\to\infty$. It is not difficult to prove that

$$\psi_0(y)\asymp G_\lambda(0,y),\quad |y|\to\infty.$$

where we are writing $f(x)\asymp g(x)$ if $0<c_1\le\frac{f}{g}\le c_2<\infty$ for some positive constants c_1,c_2. Moreover, in many cases there exists a constant $C>0$ such that

$$\psi_0(y)\sim C\cdot G_\lambda(0,y),\quad |y|\to\infty.$$

3 Asymptotic behavior of the Green functions

The asymptotics of the Green function $G_\lambda(0,y)$ is essentially depends on the tails of the random walk distribution, see [18]. In the present publication we consider only a very important case for branching random walks in which the

underlying random walks have light tails, see details in [10,12]. In this case the characteristic function

$$\widehat{a}(\kappa) = \sum_{x\in\mathbb{Z}^d} e^{i(\kappa,x)}a(x)$$

is the entire function of d complex variables $\kappa_1,\dots,\kappa_d$. As was shown in [12] instead of $\widehat{a}(\kappa)$, for our purposes, we may consider the moment generating function

$$\widehat{A}(\kappa) = \sum_{x\in\mathbb{Z}^d} e^{(\kappa,x)}a(x).$$

Put $H(\kappa) = \widehat{A}(\kappa) - 1$. This is the convex function and one can define its Legendre transform

$$H_*(y) = \max_{\kappa\in\mathbb{R}^d}\{(y,\kappa) - H(\kappa)\}.$$

In terms of $H(\kappa)$ and $H_*(\kappa)$ by using the classical Cramer's approach one can prove the central limit theorem for $x(t)$, which covers the area of arbitrary large deviations. Applying the Laplace transform to the transition probabilities of $x(t)$ one can prove that for any fixed $\lambda > 0$

$$\ln G_\lambda(0,y) \sim -|y|\nu\left(\lambda, \frac{y}{|y|}\right), \quad |y|\to\infty.$$

Here the factor $\nu\left(\lambda, \frac{y}{|y|}\right)$ is strictly positive and can be explicitly expressed in terms of $H(\kappa)$ and $H_*(\kappa)$ as in [12]. For small λ

$$\nu\left(\lambda, \frac{y}{|y|}\right) = O(\sqrt{\lambda}).$$

Let us recall that for a Brownian motion $b(t)$ with the generator Δ

$$G_\lambda(0,y) \sim \frac{e^{-\sqrt{\lambda}|y|}}{|y|^{\frac{d-1}{2}}},$$

i.e.,

$$\ln G_\lambda(0,y) \sim -\sqrt{\lambda}|y|.$$

Situation of very light tails is similar to the case of the Brownian motion considered in [1].

Now we define the front $F(t)$ of the population in the spirit of the classical KPP model. Let us assume that $x = 0$, that is the population starts from the single particle at the origin. Then for the density $m_1(t,0,y)$ we have the asymptotics

$$m_1(t,0,y) \sim \psi_0(0)\psi_0(y)e^{-\lambda_0(\beta)t}$$

or, taking into account that $\psi_0(y) \asymp G_\lambda(0,y)$,

$$m_1(t,0,y) \asymp G_\lambda(0,y)e^{-\lambda_0(\beta)t}.$$

Like in the case of the KPP model we define the particle propagation front $F(t)$ by the relation $m_1(t,0,y) \asymp \text{const}$, i.e., with the log accuracy

$$\lambda_0(\beta)t - \ln G_\lambda(0,y) = O(1).$$

Hence in the case of very light tails the front propagates linearly in t and its shape depends on level lines of the functions H and H_*.

Our next goal is the analysis of the population inside the propagating front.

4 Non-intermittency of the population inside the propagating front

The concept of intermittency for the study of the Solar magnetic field on the temperature field of the Earth ocean was proposed by Ya. Zeldovich and introduced in mathematics in the semi-physical review [19]. It was developed later in the numerous pure mathematical works [1,3–5]. At the physical level, say, the magnetic field is intermittent if almost all its energy is concentrated on the set of very low density, or, as in a population dynamic case, almost all particles are concentrated inside compactly supported clusters (patches, spots) and between the clusters we have small amount of particles.

Mathematically, definition of intermittency must include the passing to the limit: not simply very small ε, but $\varepsilon \to 0$. The formal definition: consider a family of non-negative homogeneous and continuous in space and time ergodic random fields $X(t,x)$, $t \geq 0$, $x \in \mathbb{R}^d$. We say that this family is *intermittent* asymptotically for $t \to \infty$ if

$$\mathsf{E}X^2(t,x) = m_2(t) \gg (\mathsf{E}X(t,x))^2 = m_1^2(t),$$

i.e.,

$$\frac{m_2(t)}{m_1^2(t)} \to \infty.$$

If $m_2(t) = o(m_1^2(t))$, but one can find integer $k \geq 3$ such that

$$\mathsf{E}X^k(t,x) = m_k(t) \gg (\mathsf{E}X(t,x))^k = m_1^k(t).$$

we say that the family $X(t,x)$ is *weakly intermittent*. In the case of the lattice population dynamics $X(t,y) = n(t,\cdot,y)$.

We will use the same definitions also in the case of non-homogeneous fields. In our model we will apply the concept of the intermittency deep enough inside the front of the propagation. The discussion of the phenomenon of intermittency including the explanation why the progressive growth of the moments implies the high irregularity of the field $n(t,x,y)$: clusterization, patches etc., can be found in the above mentioned publications [19,1,3,4].

Recently in [8] the intermittency of the point field $n(t,x,\Gamma)$ was proven for the classical KPP model in $\mathbb{R}^d$, which is based on the non-linear equation

$$\partial_t u_z = \Delta u_z + \beta(u_z^2 - u_z),$$

where

$$u_z(0,x,\Gamma)=\begin{cases} z, & x\in\Gamma,\\ 1, & x\neq\Gamma,\end{cases}$$

and $x\in\mathbb{R}^d$ is the location of the initial particle. Of course, we can not apply the definition of the intermittency directly to the generalized field $n(t,x,\Gamma)$, but one can use the natural averaging: $\tilde{n}(t,x,B_1(x))$ is the number of particles at the moment t in the unite ball $B_1(x)$ centered at $x\in\mathbb{R}^d$. In the KPP model the front propagates linearly in t which is seen, for $x=0$, from the relation

$$\frac{e^{-y^2/4t}}{(4\pi t)^{d/2}}e^{\beta t}\approx 1.$$

It gives for $F(t)$ the equation

$$|y|=2\sqrt{\beta}t=k_0t,\quad k_0=2\sqrt{\beta}.$$

Speed of the front is equal to $2\sqrt{\beta}$. According to [8] there exists the constant $k_1\in(0,k_0)$, such that for

$$k_1t<|y|<k_0t$$

the field $\tilde{n}(t,x)=n(t,x,B_1(x))$ is intermittent. If $|y|\geq k_0t$ the field is also intermittent, but in the weak sense.

The fundamental difference between the KPP model, where $V(x)\equiv\beta>0$, and our model with compactly supported potential $V(x)=\beta V_0(x)$, where $\operatorname{supp}V_0=B_{\mathbb{R}}(0)$, is related to the absence of intermittency in our model.

Theorem 2. *For any $\varepsilon>0$ inside the ε-neighbourhood of the front, i.e., for*

$$|y|\leq(1-\varepsilon)F\left(t,\frac{y}{|y|}\right),$$

we have for $t\to\infty$

$$\mathsf{E}\tilde{n}^2(t,y,B_1(y))=O(\mathsf{E}\tilde{n}(t,y,B_1(y))^2.$$

The proof is based on direct calculations. To solve the problem

$$\begin{cases}\partial_t m_2 &= \mathcal{H}_\beta m_2+2\beta m_1,\\ m_2(0,\cdot)=\boldsymbol{I}_{B_1(y)}.\end{cases}$$

one can project it on two invariant subspaces: $L_2^{(0)}=\{f:\operatorname{span}\{\psi_0\}\}$ and $L_2^{(0)\perp}=\{f:(\psi_0,f)=0)\}$. In $L_2^{(0)}$ we get the differential equation which can be solved explicitly, the part of the solution in $L_2^{(0)\perp}$ is growing much slower. We are not presenting the pure analytical proof due to paper volume restriction. The same approach is working for the higher moments of the order more than or equal to 3 and leads to the following limit theorem.

Theorem 3. *Assume that $t\to\infty$, $|y|=|y(t)|\in(1-\varepsilon)F(t,y/|y|)$. Then for the random variables $n^*=\frac{n(t,0,y)}{\mathbf{E}_0n(t,0,y)}$ there exists the limiting distribution. The moments of the distribution can be calculated successively, starting from the second moment.*

At the end of this section let us discuss another property of the random walk with the finite number of the generating sites. Consider the event: $A_\infty = \{n(t,x,\mathbb{Z}^d) \le C\}$ for any $t < \infty$. It means, that the population is uniformly bounded for any $t > 0$. Since the model does not contain the mortality then $n(t,x,\mathbb{Z}^d) \ge n(0,x,\mathbb{Z}^d) = 1$.

Theorem 4. *If the underlying random walk $x(t)$ is recurrent then $\mathsf{P}(A_\infty) = 0$, i.e., the population exponentially increases P-almost sure for $t \to \infty$.*

The proof of this result is simple: the initial particle returns to the support of the function $V(x)$ infinitely many times, i.e., it produces infinitely many offspings since at any visit to the support of the function $V(x)$ it produces the offsprings with uniformly positive probability.

In other words it means that the condition $\lambda_0(\beta) > 0$ for any $\beta > 0$ implies $\mathsf{P}(A_\infty) = 0$, and the population is exponentially growing P-almost sure.

For the transient process $x(t)$ the situation is different. It is obvious that x does not belong to the support of the function $V(x)$ with positive probability because the random walk of the initial particle started from $x \in \mathbb{Z}^d$ never visits the support of the function $V(x)$, that is $n(t,x,\mathbb{Z}^d) \equiv 1$ for $t > 0$.

Let us introduce the event

$$A_1 = \{n(t,x,\mathbb{Z}^d) \equiv 1\}, \quad t \ge 0,$$

and the function

$$\pi_1(x) = \mathsf{E}_x \boldsymbol{I}_{A_1} = \mathsf{P}_x(A_1).$$

Similarly, for every $k \in \mathbb{N}$ and $t \to \infty$ we define

$$A_k = \{n(t,x,\mathbb{Z}^d) \to k\} = \{\exists \tau_k : n(t,x,\mathbb{Z}^d) \equiv k,\ t \ge \tau_k\},$$

and

$$\pi_k(x) = \mathsf{E}_x \boldsymbol{I}_{A_k} = \mathsf{P}_x(A_k),$$

and also

$$A_\infty = \{n(t,x,\mathbb{Z}^d) \to \mathbf{C} < \infty\}, \quad \mathbf{C} > 0,$$

and

$$\pi_\infty(x) = \mathsf{E}_x \boldsymbol{I}_{A_\infty} = \mathsf{P}_x(A_\infty).$$

All these events are the elements of the final σ-algebra of our branching random walk [11]. Clearly,

$$\pi_\infty(x) = \sum_{k=1}^{\infty} \pi_k(x).$$

All the functions $\pi_k(x)$, $k \ge 1$, can be calculated directly as solutions of appropriate equations [11]. However, for our purpose it is more convenient to use the generating functions like in the classical Galton-Watson theory. Let us point out the difference between our case and the Galton-Watson branching processes. In the latter case the population in the supercritical regime can degenerate if the rate of mortality b_0 is positive. In the situation of a branching

random walk with $b_0 = 0$ the population remains bounded due to the fact that initial particle or its offsprings leave forever the support of the function V.

Put $u_z^*(x) = \lim_{t\to\infty} u_z(t,x,\mathbb{Z}^d) = \sum_{k=1}^{\infty} \pi_k(x) z^k$, $|z| < 1$. Note that $z^{n(t,x,\mathbb{Z}^d)} \to 0$ if $n(t,x,\mathbb{Z}^d) \to \infty$. The limit exists since $n(t,x,\mathbb{Z}^d)$ increases when $t \to \infty$. Passing to the limit in the initial equation we get

$$\mathcal{L}_x u_z^* + \beta V_0(x)((u_z^*)^2 - u_z^*) = 0$$

with the constant boundary conditions at infinity. It is clear that $\pi_1(x) \to 1$ as $|x| \to \infty$, $\pi_k(x) \to 0$ as $|x| \to \infty$ and $k \geq 2$, i.e., $u_z^*(x) \to z$ as $|x| \to \infty$.

To find $u_z^*(x)$ let us introduce $v_z(x) = z - u_z^*(x)$ where $v_z(x) \to 0$ as $|x| \to \infty$. Then

$$\mathcal{L}_x v_z + \beta V_0(x)((z - v_z)^2 - (z - v_z)) = 0,$$

that is

$$\mathcal{L}_x v_z = -\beta V_0(x) \Phi_z(x),$$

where $\Phi_z = ((z - v_z)^2 - (z - v_z))$. For every $x = y_i$ belonging to the support of the function V this last equation can be treated as the quadratic system

$$v_z(y_i) = \beta \sum_{y_j \in \operatorname{supp} V} G_0(y_i, y_j) V_0(y_j) \Phi_z(y_j).$$

In some cases, e.g., for systems with a single generating centre, several centres with high level of symmetry and condition $V(x) \equiv C$ on $\operatorname{supp} V$ where C is a constant, and so on, this system can be solved explicitly, see [11] for additional information.

5 Acknowledgement

The research is supported by the RSF grant 14-21-00162.

References

1. R. Carmona and S. Molchanov. Parabolic Anderson model and intermittency. *Memoirs of American Math Soc.*, **518**, 1994.
2. M. Cranston, L. Koralov, S. Molchanov S. and B. Vainberg. Continuous model for homopolymers. *J. Funct. Anal.*, **256**, 8, 2656-2696, 2009.
3. J. Gartner and S. Molchanov. Parabolic problems for the Anderson model. I. Intermittency and related topics. *Comm. Math. Phys.*, **132**, 3, 613-655, 1990.
4. J. Gartner and S. Molchanov. Parabolic problems for the Anderson model. II. Second-order asymptotics and structure of high peaks. *Probab. Theory Related Fields*, **111**, 1, 17–55, 1998.
5. J. Gartner, W. Konig and S. Molchanov. Geometric characterization of intermittency in the parabolic Anderson model. *Ann. Probab.*, **35**, 2, 439–499, 2007.
6. T. Kato. *Perturbation Theory for Linear Operators*, Springer-Verlag, Heidelberg, 1966; Mir, Moscow, 1972.

7. A. Kolmogorov, I.Petrovskii and N. Piskunov. A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem. *Byul. Mosk. Gos. Univ.*, Mat. Mekh. **1**, 6, 1-26, 1937.
8. L. Koralov and S. Molchanov. Structure of the population inside the propagating front. *J. Math. Sci.*, **189**, 4, 637–659, 2013.
9. S. Molchanov and E. Yarovaya. Branching processes with lattice spatial dynamics and a finite set of particle generation centers. *Dokl. Math.*, **86**, 2, 638-641, 2012.
10. S. Molchanov and E. Yarovaya. Limit theorems for the Green function of the lattice Laplacian under large deviations of the random walk. *Izv. Math.*, **76**, 1190-1217, 2012.
11. S. Molchanov and E. Yarovaya. Population structure inside the propagation front of a branching random walk with finitely many centers of particle generation. *Dokl. Math.*, **86**, 3, 787-790, 2012.
12. S. Molchanov and E. Yarovaya. Large Deviations for a Symmetric Branching Random Walk on a Multidimensional Lattice. *Proceedings of the Steklov Institute of Mathematics*, **282**, 186-201, 2013.
13. W. Shaban and B. Vainberg. Radiation conditions for the difference Schrödinger operators. *Appl. Anal.*, **80**, 3-4, 525–556, 2001.
14. E. Yarovaya. *Branching Random Walks in Nonhomogeneous Environment*, Tsentr Prikl. Issled. Mekh.Mat. Fak. MGU, Moscow, 2007. [in Russian].
15. E. Yarovaya. Criteria of exponential growth for the numbers of particles in models of branching random walks. *Theory Probab. Appl.*, **55**, 661-682, 2011.
16. E. Yarovaya. Spectral properties of evolutionary operators in branching random walk models. *Math. Notes*, **92**, 115-131, 2012.
17. E. Yarovaya. Branching Random Walks with Several Sources. *Mathematical Population Studies*, **20**, 14–26, 2013.
18. E. Yarovaya. Criteria for transient behavior of symmetric branching random walks on $\mathbb{Z}$ and $\mathbb{Z}^2$. *New Perspectives on Stochastic Modeling and Data Analysis*, ISAST: Athens, Greece, 283-294, 2014.
19. Y. Zeldovich, S. Molchanov, A. Ruzmakin and D. Sokoloff. Intermittency, diffusion and generation in a nonstationary random medium. *Mathematical physics reviews*, Chur: Harwood Academic Publ., **7**, 3–110, 1988.

Sommerfeld's Integrals and Hallén's Integral Equation in Data Analysis for Horizontal Dipole Antenna above Real Ground

Farid Monsefi[1], Milica Rančić[1,2], Sergei Silvestrov[1], and Slavoljub Aleksić[2]

[1] Division of Applied Mathematics, UKK, Mälardalen University, MDH Västerås, Sweden
(e-mail: farid.monsefi, milica.rancic, sergei.silvestrov@mdh.se)
[2] Dept. of Theoretical Electrical Engineering, ELFAK, University of Niš Niš, Serbia
(e-mail: slavoljub.aleksic@elfak.ni.ac.rs)

Abstract. High frequency (HF) analysis of the horizontal dipole antenna above real ground, which is employed in this paper, is based on the electric-field integral equation method and formulation of the Hallén's integral equation solved for the current using the point-matching method. The Sommerfeld's integrals, which express the influence of the real ground parameters, are solved approximately. Influence of different parameters of the geometry and ground on current distribution and input admittance is investigated. Furthermore, the method validation is done by comparison to the full-wave theory based exact model, and available measured data.
Keywords: Horizontal dipole antenna, Hallén's integral equation, Point-matching method, Polynomial current approximation, Real ground, Sommerfeld's integrals.

1 Introduction

Increase of the radiation power in different frequency bands during the last decades, has called for a study of harmful effects of the radio frequency energy on the living organisms and electronic equipment. An accurate determination of the near field strength in the vicinity of higher-power transmitting antennas is necessary for assessing any possible radiation hazards. In that sense, it is of great importance to account for the influence of the finite ground conductivity on the electromagnetic field structure in the surroundings of these emitters. The estimation of this influence has been intensively studied by Wait and Spies[1], Popović[2], Bannister[3], Popović and Djurdjević[4], Popović and Petrović[5], Rančić and Rančić[7], [8], Rančić and Aleksić[9], [11], Rančić[10], Arnautovski-Toseva *et al.*[12], [13], Nicol and Ridd[14], and a number of approaches has been applied in that sense, ranging from the exact full-wave based ones (Popović and Djurdjević[4], Arnautovski-Toseva *et al.*[12], [13]) to different forms of approximate, less time-consuming ones (Wait and Spies[1], Popović[2],

3^{rd} *SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal*
C. H. Skiadas (Ed)

Bannister[3], Popović and Petrović[5], Rančić and Rančić[7], [8], Rančić and Aleksić[9], [11], Rančić[10]). Although the approximate methods introduce a certain level of calculation error, their simplicity is of interest in the electromagnetic compatibility (EMC) studies. For that reason, finding an approximate, but satisfyingly accurate method applicable to wide range of parameters is often a goal of researches done in this field.

In this paper, the authors perform analysis of a thin horizontal dipole antenna (HDA) above lossy half-space (LHS) of known electrical parameters. The approach is based on the electric-field integral equation method, and formulation of the Hallén's integral equation (HIE), Balanis[6]. This equation is then solved for the current, which is assumed in a polynomial form Popović[2], using the point-matching method (PMM) (Balanis[6]). This way obtained system of linear equations involves improper Sommerfeld's integrals, which express the influence of the real ground, and are here solved approximately using simple, so-called OIA and TIA, approximations (Rančić and Rančić[7], [8], Rančić and Aleksić[9], [11], Rančić[10]). Both types of approximations are in an exponential form, and therefore, are similar to those obtained applying the method of images. It should be kept in mind that the goal of this approach is to develop approximations that have a simple form, whose application yields satisfyingly accurate calculations of the Sommerfeld's type of integrals, and are widely applicable, i.e. their employment is not restricted by the values of electrical parameters of the ground, or the geometry, Rančić and Rančić[7], [8], Rančić and Aleksić[9], [11], Rančić[10].

Thorough analysis is performed in order to observe the influence of different parameters of the geometry, and the ground, on current distribution and the input impedance/admittance of the HDA. Furthermore, the verification of the method is done by comparison to the exact model based on the full-wave theory (Arnautovski-Toseva *et al.*[12], [13]), and experimental data from Nicol and Ridd[14]. Obtained results indicate a possibility of applying the described methodology to inverse problems involving evaluation of electrical parameters of the ground (or detection of ground type change) based on measured input antenna impedance/admittance.

2 Theory

Considered HDA is positioned in the air (conductivity $\sigma_0 = 0$, permittivity ϵ_0, permeability μ_0) at height h above semi-conducting ground that can be considered a homogeneous and isotropic medium of known electrical parameters. Antenna conductors are of equal lenght $l_1 = l_2 = l$ and cross-section radius $a_1 = a_2 = a$ ($a \ll l$ and $a \ll \lambda_0$, λ_0 – wavelength in the air). The HDA is fed by an ideal voltage generator of voltage U and frequency f, and is oriented along the x-axis.

For such antenna structure, the Hertz's vector potential has two components, i.e. $\mathbf{\Pi}_{00} = \Pi_{x00}\ \hat{x} + \Pi_{zx00}\ \hat{z}$, which are described, at an the arbitrary field point $\mathrm{M}_0(x, y, z)$, by the following expressions:

$$\Pi_{x00} = \frac{1}{4\pi\underline{\sigma}_0}\int_{-l}^{l} I(x')\left[K_0(r_{1k}) + S_{00}^{h}(r_{2k})\right]\mathrm{d}x', \tag{1}$$

$$\Pi_{zx00} = \frac{1}{4\pi\underline{\sigma}_0}\frac{\partial}{\partial x}\int_{-l}^{l} I(x')\int_{\alpha=0}^{\infty}\left[-\underline{n}^{-2}\tilde{T}_{z10}(\alpha) + \tilde{T}_{\eta 10}(\alpha)\right]\frac{\tilde{K}_{00}(\alpha, r_{2k})}{u_0}\,\mathrm{d}\alpha\,\mathrm{d}x'. \tag{2}$$

where $I(x')$ - current distribution (x' - axis assigned to the HDA); $\underline{\gamma}_i$ - propagation constant and $\underline{\sigma}_i$ - equivalent complex conductivity of the i-th medium ($i = 0$ for the air, and $i = 1$ for the lossy ground); $\underline{n} = \underline{\gamma}_1/\underline{\gamma}_0 = \sqrt{\underline{\epsilon}_{r1}}$ - complex refractive index ($\underline{\gamma}_0 = \mathrm{j}\beta_0$ in the air); $\underline{\epsilon}_{r1} \approx \epsilon_{r1} - \mathrm{j}60\sigma_1\lambda_0$ - complex relative permittivity; α - continual variable over which the integration is done; $\tilde{K}_{00}(\alpha, r_{2k})$ - spectral form of the potential kernel, $K_0(r_{ik}) = e^{-\underline{\gamma}_0 r_{ik}}/r_{ik}$ - standard potential kernel, $i = 1, 2$; $S_{00}^{h}(r_{2k})$ - a type of the Sommerfeld's integral; $\tilde{T}_{z10}(\alpha)$ and $\tilde{T}_{\eta 10}(\alpha)$ - spectral transmission coefficients; $r_{1k} = \sqrt{\rho_k'^2 + (z-h)^2}$, $r_{2k} = \sqrt{\rho_k'^2 + (z+h)^2}$, $\rho_k'^2 = (x - x_k')^2 + (y - y_k')^2$, $k = 1, 2$; $u_0 = \sqrt{\alpha^2 + \underline{\gamma}_0^2}$, x_k' and y_k' - coordinates of the k-th current source element.

Boundary condition for the total tangential component of the electric field vector must be satisfied at any given point on the antenna surface, i.e.:

$$E_x + U\delta(x) = 0, \;\; -l \le x \le l, \; y = a, \; z = h, \tag{3}$$

where E_x - x-component (tangential one) of the electric field vector $\boldsymbol{E}$

$$E_x = \boldsymbol{E}\hat{x} = \left[\mathrm{graddiv}\;\boldsymbol{\Pi}_{00} - \underline{\gamma}_0^2\,\boldsymbol{\Pi}_{00}\right]\hat{x} = \frac{\partial^2\Pi_{x00}}{\partial x^2} + \frac{\partial^2\Pi_{zx00}}{\partial x\partial z} - \underline{\gamma}_0^2\,\Pi_{x00}. \tag{4}$$

The second term in (4) can be written in the following manner:

$$\frac{\partial^2\Pi_{zx00}}{\partial x\partial z} = \frac{\partial^2\Pi_{zx00}^{*}}{\partial x^2}, \tag{5}$$

where Π_{zx00}^{*} denotes the modified z-component of the Hertz's vector potential

$$\Pi_{zx00}^{*} = \frac{-1}{4\pi\underline{\sigma}_0}\int_{-l}^{l} I(x')\int_{\alpha=0}^{\infty}\left[-\underline{n}^{-2}\tilde{T}_{z10}(\alpha) + \tilde{T}_{\eta 10}(\alpha)\right]\tilde{K}_{00}(\alpha, r_{2k})\,\mathrm{d}\alpha\,\mathrm{d}x' =$$

$$= \frac{-1}{4\pi\underline{\sigma}_0}\int_{-l}^{l} I(x')\left[(1 - \underline{n}^{-2})K_0(r_{2k}) - \underline{n}^{-2}S_{00}^{v}(r_{2k}) + S_{00}^{h}(r_{2k})\right]\mathrm{d}x'. \tag{6}$$

where $S_{00}^{v}(r_{2k})$ - another type of the Sommerfeld's integral. Substituting (4) into (3) and adopting (5), the boundary condition (3) becomes:

$$\underline{\gamma}_0^2\,\Pi_{x00}^{*} - \frac{\partial^2\Pi_{x00}^{*}}{\partial x^2} = \underline{\gamma}_0^2\,\Pi_{zx00}^{*} + U\delta(x), \;\; -l \le x \le l, \; y = a, \; z = h, \tag{7}$$

where Π_{x00}^{*} denotes the *modified* x-component of the Hertz's vector potential

$$\Pi_{x00}^{*} = \Pi_{x00} + \Pi_{zx00}^{*} =$$

$$= \frac{1}{4\pi\underline{\sigma}_0}\int_{-l}^{l} I(x')\left[K_0(r_{1k}) + (\underline{n}^{-2}-1)K_0(r_{2k}) + \underline{n}^{-2}S_{00}^{v}(r_{2k})\right]\mathrm{d}x'. \quad (8)$$

Equation (7) presents the second order nonhomogeneus partial differential equation whose solution can be expressed as:

$$\Pi_{x00}^{*} = C_1'\cos\beta_0 x + C_2'\sin\beta_0 x -$$

$$-\frac{1}{\beta_0}\int_{s=0}^{x}\left[\underline{\gamma}_0^2\ \Pi_{zx00}^{*} + U\delta(x)\right]_{\substack{x=s\\y=a\\z=h}}\sin\beta_0(x-s)\mathrm{d}s, \quad (9)$$

i.e.

$$4\pi\underline{\sigma}_0\Pi_{x00}^{*} = C_1\cos\beta_0 x + C_2\sin\beta_0 x +$$

$$+\mathrm{j}\underline{\gamma}_0\int_{-l}^{l} I(x')\int_{s=0}^{x}\left[\begin{matrix}(1-\underline{n}^{-2})K_0(r_{2k})-\\ -\underline{n}^{-2}S_{00}^{v}(r_{2k})+\\ +S_{00}^{h}(r_{2k})\end{matrix}\right]_{\substack{x=s\\y=a\\z=h}}\sin\beta_0(x-s)\mathrm{d}s\ \mathrm{d}x', \quad (10)$$

where $C_1 = 4\pi\underline{\sigma}_0 C_1'$, and $C_2 = 4\pi\underline{\sigma}_0(C_2' - \mathrm{j}U/\underline{\gamma}_0)$ is a constant that will be obtained from the potential gap condition $\varphi_{00}(x=0^+) - \varphi_{00}(x=0^-) = U$ at feeding points. The electric scalar potential can be expressed as:

$$\varphi_{00} = -\mathrm{div}\mathbf{\Pi}_{00} = -\frac{\partial\Pi_{x00}}{\partial x} - \frac{\partial\Pi_{zx00}}{\partial z} = -\frac{\partial\Pi_{x00}}{\partial x} - \frac{\partial\Pi_{zx00}^{*}}{\partial x} = -\frac{\partial\Pi_{x00}^{*}}{\partial x}, \quad (11)$$

and substituting (10) in (11) we get

$$\varphi_{00} = -\mathrm{j}30C_1\sin\beta_0 x + \frac{U}{2}\cos\beta_0 x -$$

$$-\mathrm{j}30\frac{\partial}{\partial x}\int_{-l}^{l} I(x')\int_{s=0}^{x}\left[\begin{matrix}(1-\underline{n}^{-2})K_0(r_{2k})-\\ -\underline{n}^{-2}S_{00}^{v}(r_{2k})+\\ +S_{00}^{h}(r_{2k})\end{matrix}\right]_{\substack{x=s\\y=a\\z=h}}\sin\beta_0(x-s)\mathrm{d}s\ \mathrm{d}x'. \quad (12)$$

Knowing (12), the potential gap condition yields $C_2 = -\mathrm{j}U/60$. Finally (10) becomes:

$$4\pi\underline{\sigma}_0\Pi_{x00}^{*} = C_1\cos\beta_0 x - \mathrm{j}\frac{U}{60}\sin\beta_0 x +$$

$$+\mathrm{j}\underline{\gamma}_0\int_{-l}^{l} I(x')\int_{s=0}^{x}\left[\begin{matrix}(1-\underline{n}^{-2})K_0(r_{2k})-\\ -\underline{n}^{-2}S_{00}^{v}(r_{2k})+\\ +S_{00}^{h}(r_{2k})\end{matrix}\right]_{\substack{x=s\\y=a\\z=h}}\sin\beta_0(x-s)\mathrm{d}s\ \mathrm{d}x'. \quad (13)$$

Expression (13) presents the Hallén's integral equation (HIE) (Balanis[6]), having the current distribution $I(x')$ and the integration constant C_1 as unknowns. With a suitable function chosen to approximate the current distribution, HIE (13) is transformed to a system of linear equations appying the point-matching method at so-called matching points along the antenna.

It is of great importance to select an appropriate approximation for the current distribution since it will affect the calculation accuracy of both the near- and the far-field characteristics. There is a variety of proposed functions in the literature, but the polynomial current approximation proposed in

Popović[2] was repeatedly proven as a very accurate one when analysing different wire antenna structures, Popović[2], Popović and Djurdjević[4], Popović and Petrović[5], Rančić and Rančić[7], [8], Rančić[10], Rančić and Aleksić[9], [11]. The form that will be used in this paper is as follows:

$$I(x') = \sum_{m=0}^{M} I_m \left(\frac{x'}{l}\right)^m, \tag{14}$$

where I_m, $m = 0, 1, 2, \cdots, M$, present unknown complex current coefficients.

Adopting (14), HIE (13) becomes:

$$\sum_{m=0}^{M} I_m \int_{-l}^{l} \left(\frac{x'}{l}\right)^m \left[\begin{array}{c} K_0(r_{1k})+(\underline{n}^{-2}-1)K_0(r_{2k})+\underline{n}^{-2}S_{00}^{v}(r_{2k})- \\ -\mathrm{j}\underline{\gamma}_0 \int_{s=0}^{x} \left[\begin{array}{c}(1-\underline{n}^{-2})K_0(r_{2k})- \\ -\underline{n}^{-2}S_{00}^{v}(r_{2k})+ \\ +S_{00}^{h}(r_{2k})\end{array}\right]_{\substack{x=s \\ y=a \\ z=h}} \sin\beta_0(x-s)\mathrm{d}s \end{array}\right] \mathrm{d}x' -$$

$$-C_1 \cos\beta_0 x = -\mathrm{j}\frac{U}{60}\sin\beta_0 x. \tag{15}$$

Unknown complex current coefficients I_m, $m = 0, 1, 2, \cdots, M$, are determined from the system of linear equations obtained matching (15) at points:

$$x_i = \frac{i}{M} l, \; i = 0, 1, 2, \cdots, M. \tag{16}$$

This way, system of $(M + 1)$ linear equations is formed, lacking one additional equation to account for the unknown integration constant C_1. This remaining linear equation is obtained applying the condition for the current at the conductor's end. Standardly, the vanishing of the current is assumed at the end of antenna arm (Popović[2], Popović and Djurdjević[4], Popović and Petrović[5], Rančić and Rančić[7], [8], Rančić and Aleksić[9], [11], Rančić[10]), which corresponds to $I(-l) = I(l) = 0$, i.e. based on (14) to

$$\sum_{m=0}^{M} I_m = 0. \tag{17}$$

(Note: A more realistic condition for the current at the conductor's ending, derived satisfying the continuity equation at the end of an antenna arm, can also be used.)

This way, the system of equations needed for computing the current distribution of the observed antenna is formed. Based on that, for the given generator voltage U, the input admittance is determined from $Y_{in} = I_0/U$, where $I_0 = I_m|_{m=0}$.

Remaining problem are two Sommerfeld's integrals appearing in (15) expressed by

$$S_{00}^{v}(r_{2k}) = \int_{\alpha=0}^{\infty} \tilde{R}_{z10}\tilde{K}_{00}(\alpha, r_{2k})\mathrm{d}\alpha, \tag{18}$$

$$S_{00}^{h}(r_{2k}) = \int_{\alpha=0}^{\infty} \tilde{R}_{\eta 10}\tilde{K}_{00}(\alpha, r_{2k})\mathrm{d}\alpha, \tag{19}$$

where the first terms in both integrands represent spectral reflection coefficients (SRCs):

$$\tilde{R}_{z10}(\alpha) = \frac{\underline{n}^2 u_0 - u_1}{\underline{n}^2 u_0 + u_1},\ u_i = \sqrt{\alpha^2 + \underline{\gamma}_i^2},\ i = 0, 1, \tag{20}$$

$$\tilde{R}_{\eta 10}(\alpha) = \frac{u_0 - u_1}{u_0 + u_1},\ u_i = \sqrt{\alpha^2 + \underline{\gamma}_i^2},\ i = 0, 1. \tag{21}$$

In order to solve the type of Sommerfeld's integral given by (18) the methodology proposed in Rančić and Rančić[7] will be applied. Let us assume the SRC (20) in a so-called - TIA (*two-image approximation*) form:

$$\tilde{R}_{z10}(u_0) \cong B_v + A_{1v} e^{-(u_0 - \underline{\gamma}_0)\underline{d}_v}, \tag{22}$$

where B_v, A_{1v} and $\underline{d}_v$ are unknown complex constants. When (22) is substituted into (18), the following general TIA approximation is obtained:

$$S_{00}^v(r_{2k}) \cong B_v K_0(r_{2k}) + A_v K_0(r_{2kv}), \tag{23}$$

where $r_{2kv} = \sqrt{\rho_k'^2 + (z + h + \underline{d}_v)^2}$, presents the distance between the second image and the observation point M_0, and $A_v = A_{1v} \exp(\underline{\gamma}_0 \underline{d}_v)$. Now, matching expressions (20) and (22) at $u_0 \to \infty$ and $u_0 = \underline{\gamma}_0$, and the first derivative of the same expressions at $u_0 = \underline{\gamma}_0$, the following values for the unknown complex constants in (22) are obtained:

$$B_v = R_\infty, A_{1v} = R_0 - R_\infty, \underline{d}_v = (1 + \underline{n}^{-2})/\underline{\gamma}_0, \tag{24}$$

where: $R_\infty = \tilde{R}_{z10}(u_0 \to \infty) = (\underline{n}^2 - 1)/(\underline{n}^2 + 1)$ and $R_0 = (\underline{n} - 1)/(\underline{n} + 1)$.

Substituting (24) into (23), the following TIA form of (18) is obtained:

$$S_{00}^v(r_{2k}) \cong R_\infty K_0(r_{2k}) + (R_0 - R_\infty) e^{\underline{\gamma}_0 \underline{d}_v} K_0(r_{2kv}). \tag{25}$$

Similarly, we can assume (21) in the following form (Rančić and Rančić[8], Rančić and Aleksić[9], [11], Rančić[10]):

$$\tilde{R}_{\eta 10}(u_0) \cong B_h + A_{1h} e^{-(u_0 - \underline{\gamma}_0)\underline{d}_h}, \tag{26}$$

where B_h, A_{1h} and $\underline{d}_h$ - unknown complex constants. Substituting (26) into (19), the following general approximation is obtained:

$$S_{00}^h(r_{2k}) \cong B_h K_0(r_{2k}) + A_h K_0(r_{2kh}), \tag{27}$$

where $A_h = A_{1h} \exp(\underline{\gamma}_0 \underline{d}_h)$, and $r_{2kh} = \sqrt{\rho_k'^2 + (z + h + \underline{d}_h)^2}$.

After matching (21) and (26) at points $u_0 \to \infty$ and $u_0 = \underline{\gamma}_0$, and their first derivatives at $u_0 = \underline{\gamma}_0$, we get values $B_h = 0$, $A_{1h} = -R_0$, and $\underline{d}_h = 2/(\underline{\gamma}_0 \underline{n})$, i.e. (27) gets the OIA (*one-image approximation*) form, Rančić and Aleksić[9], [11], Rančić[10]:

$$S_{00}^h(r_{2k}) \cong -R_0 e^{\underline{\gamma}_0 \underline{d}_h} K_0(r_{2kh}). \tag{28}$$

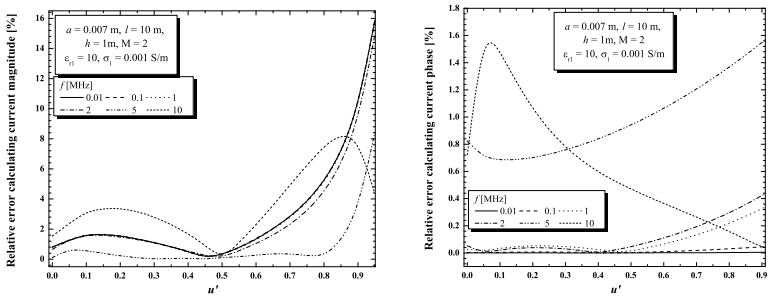


Fig. 1. Relative error of the current magnitude (left) and phase (right) along the HDA arm.

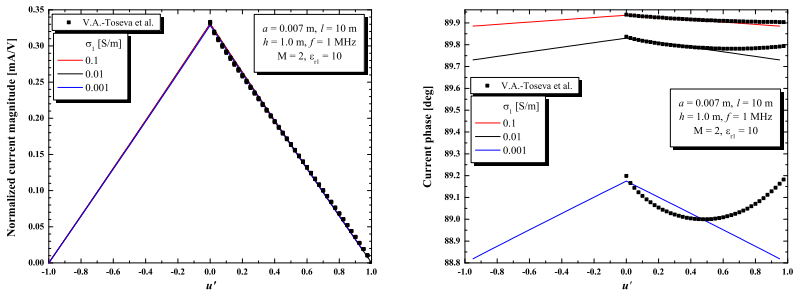


Fig. 2. Current magnitude (left) and phase (right) along the HDA for different ground conductivities.

3 Numerical results

Described numerical procedure is applied to near-field analysis of the symmetrical HDA fed by an ideal voltage generator of voltage U.

Firstly, results of the relative error of current distribution calculation are given in Figure 1. The conductor is $2l = 20$ m long with the cross-section radius of $a = 0.007$ m, and it is placed at $h = 1.0$ m above lossy ground with electrical permittivity $\epsilon_{r1} = 10$. In this case, the variable parameter is the frequency that takes values from a wide range (10 kHz to 10 MHz). The relative error is shown separately for the current magnitude and phase along the HDA arm for the case of the specific conductivity of $\sigma_1 = 0.001$ S/m. As a reference set of data, those from Arnautovski-Toseva *et al.*[12], [13] are taken.

Current distribution's magnitude and phase at 1 MHz, can be observed from Figure 2. The HDA has the same dimensions as previously, and it is placed at $h = 1.0$ m above lossy ground with electrical permittivity $\epsilon_{r1} = 10$. The value of the specific conductivity has been taken as a parameter: $\sigma_1 = 0.001, 0.01, 0.1$ S/m. Comparison has been done with the results from Arnautovski-Toseva *et al.*[12], [13].

Further, the influence of the conductor's position on the current distribution has been analysed. The results are graphically illustrated in Figure 3 together

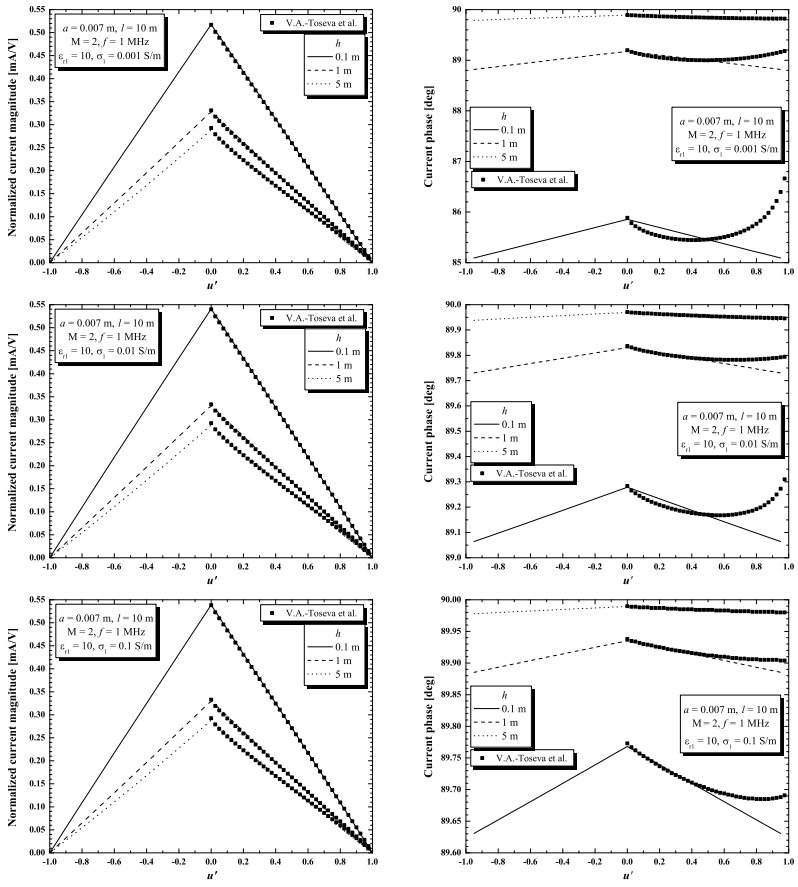


Fig. 3. Current magnitude (left) and phase (right) along the HDA above LHS at different heights.

with the ones from Arnautovski-Toseva *et al.*[12], [13]. Three cases were observed that correspond to heights $h = 0.1, 1.0, 5.0$ m. The current has been calculated at frequency of 1 MHz, and analysis has been done for the following values of the specific ground conductivity: $\sigma_1 = 0.001, 0.01, 0.1$ S/m. HDA dimensions are the same as previously.

Next example explores the dependence of the current (its magnitude and phase) on different ground conductivities calculated at the feeding point A($l =$ 0 m), which can be observed from Figure 4. Two cases are considered: solid line represents the value of $\sigma_1 = 0.001$ S/m, and the dashed one corresponds to $\sigma_1 = 0.1$ S/m. The first row of Figure 4 corresponds to HDA height of $h = 2.5$ m, and the second one to $h = 5.0$ m. The same influence for height $h = 0.5$ m is given in Rančić and Aleksić[11].

Similarly, the dependence of the current (its magnitude and phase) at specific points along the HDA arm in the frequency range from 10 kHz to 10 MHz,

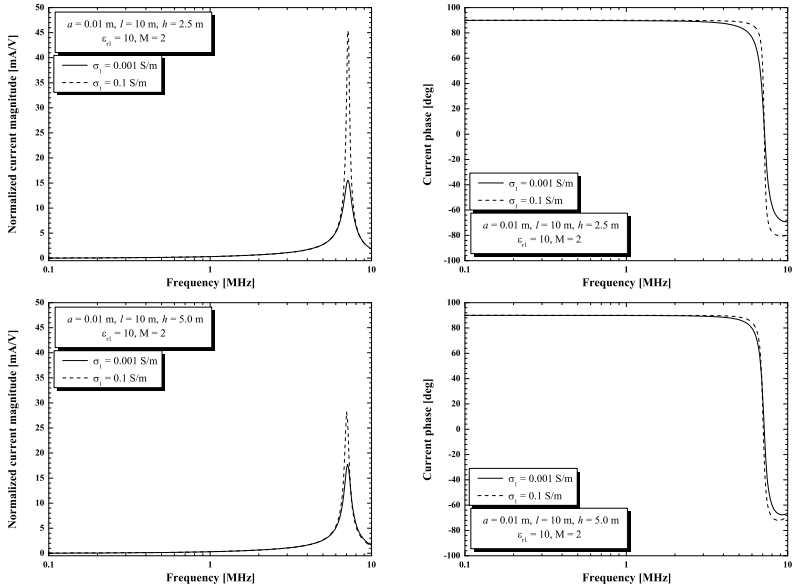


Fig. 4. HDA current magnitude (left) and phase (right) at point A for different ground conductivities.

is presented in Figure 5. The antenna is $2l = 20$ m long with a cross-section radius of $a = 0.01$ m, and considered heights are: $h = 0.5, 2.5, 5.0$ m. Electrical parameters' values of the ground are: electrical permittivity $\epsilon_{r1} = 10$, and specific conductivity $\sigma_1 = 0.1$ S/m. Current is calculated at points: A($l = 0$ m), B($l = 2.5$ m), C($l = 5.0$ m), and D($l = 7.5$ m). This example for $\sigma_1 = 0.001$ S/m and $h = 0.5$ m is given in Rančić and Aleksić[11].

Finally, Figure 6 shows comparison between theoretical calculations performed using the methodology described in this paper, and the results of the admittance measurements for the frequency range of $7 - 12$ MHz (Nicol and Ridd[14]). Observed HDA is 15 m long suspended at height of 0.3 m above the LHS. Two boundary cases of the ground are observed: a perfect dielectric (blue data), and a highly conducting plane (black data). Corresponding results obtained by the method of images are also shown (open circles). It can be observed that the better accordance is achieved using the method described here, which was expected since the observed antenna is very close to the ground (for the frequency of 10 MHz, height of 0.3 m corresponds to $0.01\lambda_0$), and the accuracy of the method of images decreases when the antenna is at height less than $h/\lambda_0 = 0.025$ (Popović and Petrović[5]).

4 Conclusions

Approximate method for the analysis of horizontal dipole antenna has been applied in this paper for the purpose of the current distribution and input admittance evaluation for the HDA positioned in the air at arbitrary height

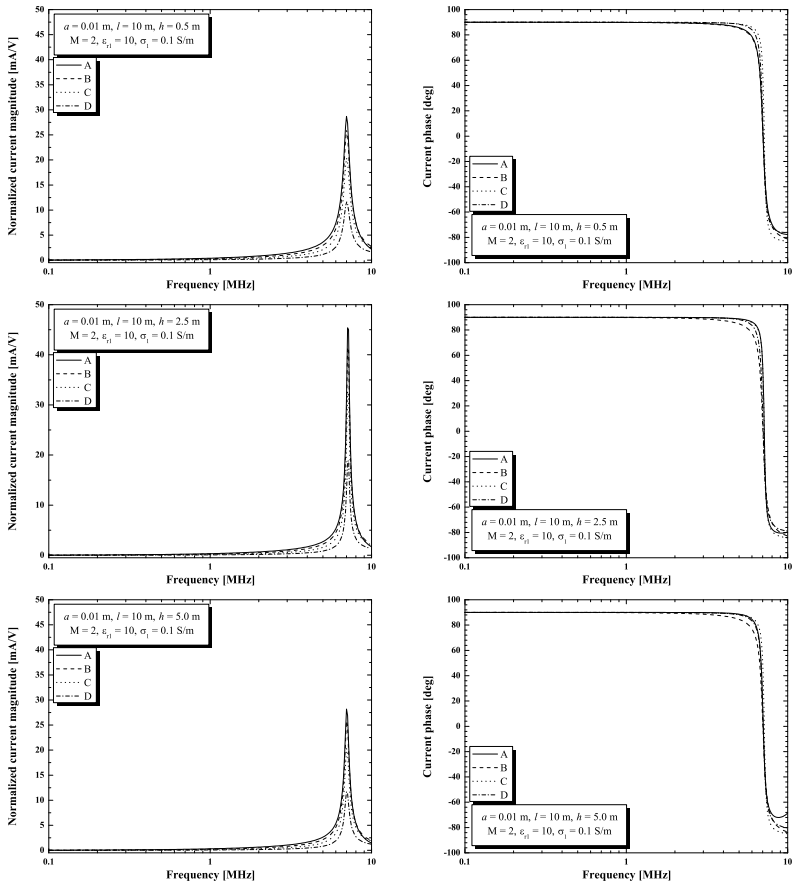


Fig. 5. HDA current magnitude (left) and phase (right) at different points along the antenna.

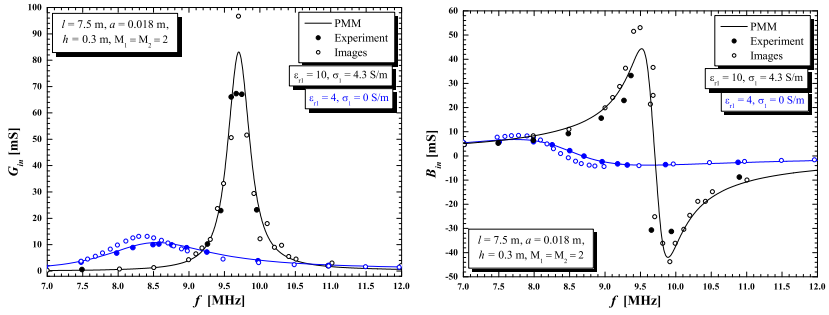


Fig. 6. HDA input conductance (left) and susceptance (right) versus frequency.

above LHS, which is considered a homogenous medium. The aim of the paper was to validite the applied method for the cases of interest in the EMC studies.

The analysis has been performed in a wide frequency range, and for different positions of the antenna, as well as for various values of the LHS's conductivity. It has been proven, based on the comparison with the exact model from Arnautovski-Toseva *et al.*[12], [13], that the methodology used here yields very accure results in the observed parameters' ranges. This indicates a possibilty of applying this method for analysis of different wire structures in the air above LHS, and more importantly, very close to the ground where the finite conductivity's influence is the greatest.

5 Acknowledgement

This work is partly supported by the RALF3 project funded by the Swedish Foundation for Strategic Research (SSF), and the EUROWEB Project funded by the Erasmus Mundus Action II programme of the European Commission.

The second author would like to thank members of the Division of Applied Mathematics at the MDH University, Sweden for inspiring and fruitful collaboration.

References

1. Wait, J. R. and Spies, K. P., "On the Image Representation of the Quasi-Static Fields of a Line Current Source above the Ground", *Can. J. Phys.* 47, 2731–2733 (1969).
2. Popović, B. D., "Polynomial Approximation of Current along thin Symmetrical Cylindrical Dipoles", *Proc. Inst. Elec. Eng.* 117, 5, 873–878 (1970).
3. Bannister, P. R., "Extension of Quasi-Static Range of Finitely Conducting Earth Image Theory Technique to other Ranges", *IEEE Trans. on AP* 26, 3, 507-508 (1978).
4. Popović, B. D. and Djurdjević, D. Ž., "Entire-Domain Analysis of Thin-Wire Antennas near or in Lossy Ground", *IEE Proc., Microw. Antennas Propag.* 142, 213-219 (1995).
5. Popović, B. D. and Petrović, V. V., "Horizontal Wire Antenna above Lossy Half-Space: Simple Accurate Image Solution", *International journal of numerical modelling: Electronic networks, devices and fields* 9, 194-199 (1996).
6. Balanis, C. A., *Antenna Theory: Analysis and Design, Chapter 8*, 3rd Edition: J. Wiley and Sons, Inc., Hoboken, New Jersey (2005).
7. Rančić, M. P. and Rančić, P. D., "Vertical Dipole Antenna above a Lossy Half-Space: Efficient and Accurate Two-Image Approximation for the Sommerfeld's Integral", in Proc. of *EuCAP 2006* Nice, France, paper No121 (2006).
8. Rančić, M. and Rančić, P., "Horizontal Linear Antennas above a Lossy Half-Space: A New Model for the Sommerfeld's Integral Kernel", *Int. J. El. Commun. AEU* 65, 879-887 (2011).
9. Rančić, M. and Aleksić, S., "Horizontal Dipole Antenna very Close to Lossy Half-Space Surface", *Electrical Review* 7b, 82-85 (2012).
10. Rančić, M., *Analysis of Wire Antenna Structures in the Presence of Semi-Conducting Ground*, Ph.D dissertation, Faculty of electronic engineering, University of Niš, Niš, Serbia (2012).

11. Rančić, M. and Aleksić, S., "Analysis of Wire Antenna Structures above Lossy Homogeneous Soil", in Proc. of *21st Telecommunications Forum (TELFOR)* Belgrade, Serbia, 640-647 (2013).
12. Arnautovski-Toseva, V., Khamlichi Drissi, K. El, and Kerroum, K., "Comparison of Approximate Models of Horizontal Wire Conductor above Homogeneous Ground", in Proc. of *EuCAP 2012* Prague, Czech Republic, 678-682 (2012).
13. Arnautovski-Toseva, V., Khamlichi Drissi, K. El, Kerroum, K., Grceva, S. and Grcev, L., "Comparison of Image and Transmission Line Models of Energized Horizontal Wire above Two-Layer Soil", *Automatika* 53, 38-48 (2012).
14. Nicol, J.L. and Ridd, P.V., "Antenna Input Impedance: Experimental Confirmation and Gological Application", *Can. J. Phys.* 66, 818-823 (1988).

Building-type classification based on measurements of energy consumption data

Ying Ni[1], Christopher Engström[1], Anatoliy Malyarenko[1], and Fredrik Wallin[2]

[1] School of Education, Culture and Communication, Division of Applied Mathematics, Mälardalen University, Västerås, Sweden (e-mail: ying.ni@mdh.se, christopher.engstrom@mdh.se, anatoliy.malyarenko@mdh.se)

[2] Department of Energy, Building and Environment, Mälardalen University, Västerås, Sweden (e-mail: fredrik.wallin@mdh.se)

Abstract. In this paper we apply data-mining techniques to a classification problem on actual electricity consumption data from 350 Swedish households. More specifically we use measurements of hourly electricity consumption during one-month and fit classification models to the given data. The goal is to classify and later predict whether the building type of a specific household is an apartment or a detached house. This classification/prediction problem becomes important if one has a consumption time series for a household with unknown building type. To characterise each household, we compute from the data some selected statistical attributes and also the load profile throughout the day for that household. The most important task here is to select a good representative set of feature variables, which is solved by ranking the variable importance using technique of random forest. We then classify the data using classification tree method and linear discriminant analysis. The predictive power of the chosen classification models is plausible.

Keywords: data-mining, energy consumption data, classification of energy customers, clustering of energy customers.

1 Introduction

In this paper we consider a classification problem on a large data set of actual energy (electricity) consumption measurements for 350 Swedish households. The goal is to classify the energy-using households into two categories, one being apartments and the other being detached houses. The motivation is to both discover knowledge from the real-world data and also make predictions later when one needs to determine the source of a given time series of energy measurements. To solve this classification problem, we apply data-mining techniques including random forests, classification tree and discriminant analysis. The technique of random forests is used for ranking the set of the candidate feature variables and the classification tree and discriminant analysis are used as classification models.

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

There are several reviews summarizing the methods used when classifying and clustering different electric loads (Zhou et. al [12]) and consumption patterns (Chicco [3]). Both of these discuss the importance of load data preparation, pre-clustering, clustering implementation, cluster analysis and finally determine applications for the clusters. Some key methodologies used are: K-Means, hierarchical clustering, fuzzy clustering and self-organization mapping (López et al. [7]; Zhou et. al. [12]). Jota et al. [6] shows on the possibilities to use load curves to obtain a deeper understanding of the building energy usage in various types of commercial buildings and hospitals. In several papers clustering techniques are used in order to provide electric customer segmentation (López et al.[7]; Chicco [4]) and same principles are valid to distinguish different types of domestic consumers such as apartments and houses.

In the present work, the most crucial task for the classification problem is to find good predictors, i.e. to select a set of feature variables that can well explain the differences in terms of energy consumption between an apartment and a detached house. Some standard statistical attributes like mean, standard deviation, skewness, kurtosis and quantiles are used. In searching for the best predictors we also find it interesting to use some heuristic variables like the mean of daily maximum/minimum. Moreover the daily load profile, defined here as the hourly consumption at clock hour 00:00, 01:00, ... 23:00 averaged over all working days during the month, are included in the set of attributes. Daily load profiles are used to characterise the consumption patterns of households in for example Rodrigues et.al.[8].

2 The data

Our study is on a data set of electricity consumption measurements provided by the Swedish Energy Agency. The data set consists of 350 Swedish households with each household observed for one month during the period from year 2005 to 2008 . The original data contains electricity consumption measurements and its seasonally-adjusted counterparts on each ten-minute time interval for each light source and each electrical appliance like refrigerator and television. For more details regarding to this data set we refer to the report from the Swedish end-use metering campaign (Zimmerman [13]) and the description of its project "Improved energy statistics in buildings and industry" (Swedish Energy Agency [9]). The present study uses one-hour data which is the total electricity consumption aggregated in two steps. The first step is to obtain electricity consumption for each ten-minute time interval by aggregating measurements from all sources, i.e. including all light sources and all appliance sources. The second step is to sum up all ten-minute intervals within a specific clock hour.

The original data contains measurement for 389 households and we use a subset of the data here.

Seasonal correction is only done on some categories of the electrical equipments, we will however treat these values as the seasonally-adjusted ones and use these values in the data analysis.

There are problems of missing values. Oftentimes the problem is that the recorded consumption values exist for electrical appliances but not for the light sources. Such problems are treated in the ten-minute measurements level. The treatments are different for two different cases. The more frequent case is that those missing values for light sources appear at the beginning or the end of the observation time period, which is handled by simply removing the existing extra values for appliances before aggregation. In some cases the missing values happen in the middle of the observation period, these missing values are approximated by the average value at that time point for the same weekdays.

3 Feature selection

The processed data contains information on building type, namely apartments or detached houses, and hourly electricity consumption values which we will use to perform the classification task. Take the hourly electricity consumption observations for a specific household as a sample. Since we have in total 350 households monitored for approximately one month, we have then 350 samples each contains approximately $30 \times 24 = 720$ observation values. To fit a classifier on this data, we need to select feature variables that are relevant for this classification. For example, for each sample (household), we can compute statistical quantities like the mean hourly consumption and use it as a candidate feature variable for that sample. Inspection of the data indeed indicates that the mean hourly consumption for apartment-households tend to be much smaller than the corresponding means for detached house-households, which is of course not surprising.

To confirm that the building type (apartment or detached house) does have an influence on the distribution or character of the hourly electricity consumption, we can build two separate histograms on all samples that have building type as apartment and detached house respectively, as shown in Figure 1 below.

The histograms are not very similar which lead us to conclude that values of hourly consumption are indeed influenced by the building type. For examples it suggests that the empirical distributions of apartments (called flats in Figure 1 for brevity) and houses differ obviously in mean and the tail behaviour of both left and right tails. Therefore we include into our set of candidate features nine standard statistics such as mean (of hourly consumption for each household), standard deviation, skewness, kurtosis, median, 25% and 75% quantiles, "gMaxima"/"gMinima" (the maximum /minimum value over all observed hourly consumption values for that household). We shall also include two more heuristic variables, namely the average/mean of daily maximum (hourly) consumption values (shortened as "dailyMaxMean") and the average/mean of daily minimum (hourly) consumption values (shortened as "dailyMinMean"). For illustration, variable"dailyMinMean" is calculated in two steps, first compute the minimum hourly consumption value for each day then average over all 30 observation days. The intuition to "dailyMaxMean" is obvious, apartments should in general have lower hourly peak than houses. The motivation of using "dailyMinMean" is that apartments tend to have less flexibility in minimising the electricity consumption than a detached house. For

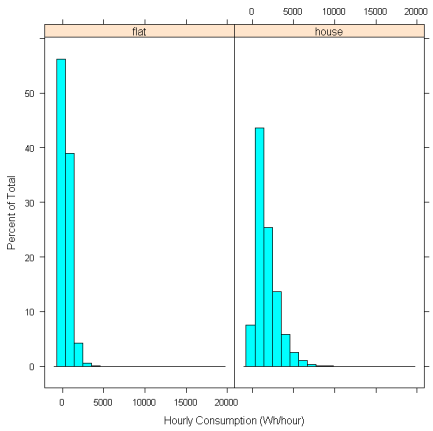


Fig. 1. Histograms of hourly energy consumption values for apartments (left) and houses (right)

each sample/household, we can compute a single value of each of the above-mentioned 11 variables which are then used to characterise the corresponding sample/household.

We run a *random forest* in R to rank the importance of these variables. The random forest technique is an ensemble method formed by a large set of tree-based models each of which are grown with random subspace method (see Breiman [1] for details). As shown in Figure 2 (left), the random forest technique ranks the variable "25% quantile" as the most explanatory variable for predicting the building type, followed by "dailyMinMean", "mean" and so on. This is an interesting result since one might think the mean hourly consumption is the most useful variable in predicting whether the corresponding household is a detached house or an apartment.

Let us define the load profile of a household as the electricity consumption for each specific clock hour during an average working day. For instance, suppose that a household is observed for one month, then its load profile at 8 am is the hourly consumption occurred during 8:00 and 9:00 o'clock (or between 7:50-8:50, depends on how the original data is measured) averaged over all working days (i.e. about 20 working days). It might be interesting to include the load profile into the set of feature variables since houses and apartments might have different patterns in terms of its load profile throughout the day. Figure 2 (right) gives the ranking by the random forest technique when we include the 24 load values at clock hour 00:00, 01:00, 02:00, ... 23:00, namely, "X0", ..., "X23", averaged over all working days during the month. It confirms that "25% quantile" and "dailyMinMean" are important variables, but also ranks the load values at 06:00 and 07:00 clock as highly relevant for classifying houses and apartments.

The horizontal axis in Figures 2 gives the "MeanDecreaseAccuracy" of each variable which is the percentage increase in the error of the forest if we remove that variable. This percentage increase in error is calculated as the average

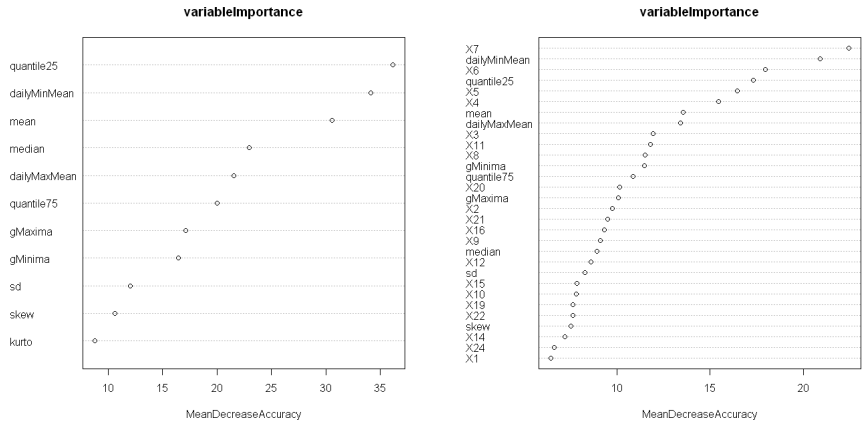


Fig. 2. Variable importance ranking on statistical feature variables (left) and on all feature variables (right)

increase in the mean squared error of each classification tree grown in the forest. To select a subset of the feature variables, we decide on the threshold value of 10 for "MeanDecreaseAccuracy" in Figures 2 (right) which gives the following set of attributes: "quantile25","dailyMinMean", "mean", "dailyMaxMean", "gMinima", "quantile75" , "gMaxima", and the load values Xi with $i = 3 - 8, 11, 20$. To summarize, our training data consists $N = 350$ observations $(x_i, y_i), i = 1, 2, \ldots, N$. Each observation contains $p = 15$ features variables, i.e. $x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})$ and the target y_i which is the classification outcome $k, k = 1, 2$ (apartment or house). We will fit the data using classification tree and the linear discriminant analysis in the following sections, using both the whole set of 15 feature variables or some selected subsets of it.

4 A classification tree

A classification tree uses recursive binary splitting of feature variables $x = (x_1, x_2, \ldots, x_p)$ to partition the feature space into a set of regions, say M regions $R_1, R_2, \ldots, R_M$ and then model the classification outcome for each region by k_m, i.e.

$$\hat{f}(x) = \sum_{m=1}^{M} k_m I(x \in R_m),$$

where I is the indicator function. The estimator for k_m can be for example the majority class in region R_m.

Some error minimization criterion, for instance minimizing the impurity measures the Gini index or the entropy, is adopted to decide on the splitting variables and splitting points and in general the shape of the tree. For a detailed discussion on the classification tree method we refer to Breiman et. al.[2].

We use `rpart` package (Therneau and Atkinson [10]) in R to fit a classification tree to our training data. To make a pretty graphical representation of the tree we use the `DMwR` package (Torgo [11]).

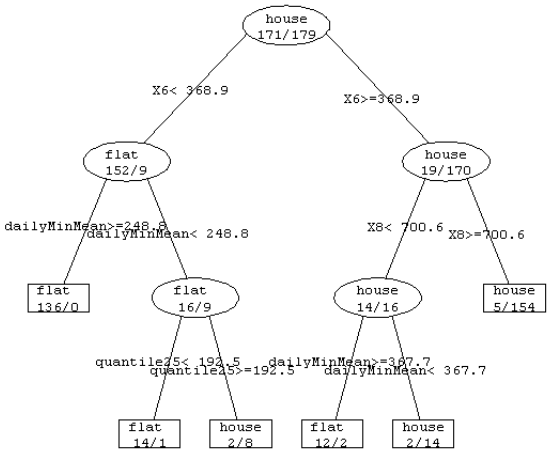


Fig. 3. The default classification tree (CP 0.01)

The tree in Figure 3 uses the whole set of the feature variables. It partitions the feature space into six regions $R_i, i = 1, 2, \ldots$ corresponding to the six terminal nodes/leaves of the tree. For example region R_1 is determined by satisfying $X6 < 368.9$ and dailyMinMean ≥ 248.8. There're 136 apartments (called flats in the tree) and 0 house in this region, the classification outcome k_1 for region R_1 is then modeled as apartment which is the majority class in this region.

When we build a classification tree, we need to consider the trade-off between the benefit of an accurate prediction and the risk of an oversized and over-fitted tree. To reach the best balance between these two we can check the corresponding cost complexity parameter (cp) values and relative errors, i.e. to perform a so-called a *cost complexity pruning* (Breiman et. al.[2]) of the tree.

For each sub-tree of the default tree given in Figure 3, we have the following information on the corresponding cost complexity parameter (`CP`), number of splits (`Nsplit`), relative error (`Rel error`) on the training data itself, the ten-fold cross validation relative error (`Rel xerror`) and its standard deviation (`std`) with prefix "`x`" in `xerror` standing for "cross-validation", as presented in Table 1.

The most useful estimate on the predictive performance is `Rel xerror` which refers to the average relative error computed from ten-fold cross validation procedures. This validation procedure randomly partitions the data set into 10 subsets, use 9 of them as the training data to fit a classifier on, the remaining 1 subset is used for evaluating the classifier. Note that the relative

Table 1. Errors on the training set and 10-fold cross validation errors for sub-trees of the default tree (Root node error: 171/350 = 0.48857)

CP	Nsplit	Rel error	Rel xerror (std)	Abs error	Abs xerror (std)
1. 0.83626	0	1.00000	1.00000 (0.05469)	0.48857	0.48857 (0.02672)
2. 0.02924	1	0.16374	0.20468 (0.03282)	0,08000	0.10000 (0.01604)
3. 0.01754	3	0.10526	0.17544 (0.03063)	0,05143	0.08571 (0.01496)
4. 0.01000	5	0.07018	0.16374 (0.02968)	0.03429	0.08000 (0.01450)

error from a ten-fold cross validation procedure is a random variable with estimated mean `Rel xerror` and standard deviation `std` due to the randomness of this validation method.

It is worth mentioning that both `Rel error` and `Rel xerror` are errors relative to the root node error (0.48857). The root node error is simply the misclassification error if we use the majority class to predict all households before any splitting. Since we have 179 houses and 171 apartments in the data sets, by classifying all households as houses lead to a root node error of 171/350 = 0.48857. The corresponding absolute (misclassification) errors on the training data, denoted by `Abs error`, and the average absolute error for ten-fold cross validation method, denoted by `Abs xerror`, are smaller and equal to `Rel error` and `Rel xerror` multiplied with the root node error (0.48857) respectively. For convenience we present also `Abs error` and `Abs xerror` together with their standard deviations in Table 1 since when we will only use absolute errors for the linear discriminant analysis conducted in the coming section.

The default tree in Figure 3 is tree number 4 with 5 tests and an average relative error of 0.16374 with standard deviation of 0.029681 using ten-fold cross validation. However the simpler tree number 3 has a value of 0.17544 which is within one standard deviation of `Rel xerror` for tree number 3, i.e. $0.17544 \in (0.16374 - 0.029681, 0.16374 + 0.029681)$. Using the so-called "1-SE" selection approach, we may choose tree number 3 as our classification model. The corresponding absolute (misclassification) cross-validation error is 8.57%.

Experiments have been made by fitting a classification tree on various subsets of the feature variables. The error measure `Rel xerror` (and hence `Abs xerror`) is slightly different each time one runs ten-fold cross validation due to the randomness of this validation procedure. This measure changes also when one fits a tree with another subset of the features. However the measure `Rel xerror` in general is moderately low with the worst score being about 21%. To conclude one can note that variables "25% quantile", "dailyMinMean" and also the load values during morning clock hours like 6 am and 8 am are the most relevant ones for the classification task. These quantities are very easy to obtain given a time series of the energy consumption data. The classification tree built on these easily-attainable variables also have plausible prediction performance, given a low average relative error (17.54%) and hence a low absolute error (8.57%) of ten-fold validation process in the example tree above.

To illustrate the explanatory power of these variables, several scatter plots on these two variables are given below in Figure 4, with houses plotted in red plus signs and apartments in green circles.

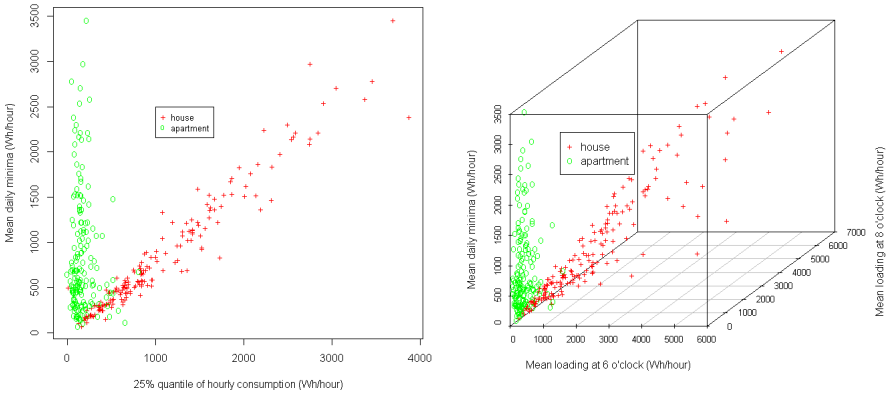


Fig. 4. A scatter plot (left) and a 3D scatter plot on selected variables (right)

5 Linear discriminant analysis

Looking at the scatter plots in Figure 4 we can clearly make out the two classes (houses and apartments). As an alternative to the tree classification method we also classify the data using linear discriminant analysis as described in Johnson and Wichern [5]. The analysis will be done using all the feature variables as well as a subset of the "best" feature variables.

First a multivariate Gaussian distribution is fitted to each class $k, k = 1, 2$ with mean μ_k, the mean is estimated from the N observations using the sample mean:

$$\hat{\mu}_k = \frac{\sum_{n=1}^{N} M_{nk} x_n}{\sum_{n=1}^{N} M_{nk}},$$

where $M_{nk} = 1$, if obeservation n belongs to class k and $M_{nk} = 0$ otherwise.

Both classes are assumed to have the same covariance matrix. The covariance matrices are estimated using the sample covariance where we first subtract the sample mean of each class from the observations of each class. The elements q_{nm} of the sample covariance matrix Q can be written as:

$$q_{nm} = \frac{1}{N-1} \sum_{i=1}^{N} (x_{in} - \hat{\mu}_{nk})(x_{im} - \hat{\mu}_{mk})^{\top}$$

where $\hat{\mu}_{nk} = \sum_{k=1}^{K} M_{nk}\hat{\mu}_k$ where K is the number of classes ($K = 2$ in this case). To classify an observation we start by calculating the posterior probability that the observation belongs to every class and find the class it is most likely to come from. With 2 classes, using equal misclassification cost and prior probability (we have approximately the same number of houses as apartments in our data) we can classify an observation x_i by calculating:

$$\hat{y}_i = (\hat{\mu}_1 - \hat{\mu}_2)Q^{-1}x_i$$

$$\hat{m} = \frac{1}{2}(\hat{\mu}_1 - \hat{\mu}_2)Q^{-1}\hat{\mu}_1 + \hat{\mu}_2.$$

If $\hat{y}_i > \hat{m}$ we assign x_i to class 1 and to class 2 otherwise. $\hat{m}$ can be seen as the "midpoint" between the two classes where we simply check on which side an observation lies. This is also equal to Fisher's discriminant function since we have equal misclassification and prior probabilities (Johnson and Wichern [5]).

To do the classification we use the "ClassificationDiscriminant" object and the related functions in MATLAB (Statistics Toolbox). We perform linear discriminant analysis for multiple sets of predictors, all feature variables excluding load values (hereafter shortened as "all variables") or all load values, all variables + load values or a selection of only a few variables. The resulting absolute misclassification errors on the training set and from a 10-fold cross validation procedure can be found in Table 2. A number of other combinations

Table 2. Error on training set and 10-fold cross validation error for different predictor variables

Predictors	**Abs error**	**Abs xerror (std)**
all variables + all load values	0.0971	0.1057 (0.0487)
all variables	0.0857	0.0886 (0.0494)
all load values	0.1400	0.1600 (0.0715)
dailyMinMean + 25% quantile	0.0600	0.0686 (0.0335)
load values X6 +dailyMinMean	0.0800	0.0800 (0.0295)
load values X7 +dailyMinMean	0.1057	0.1057 (0.0674)

of $2 - 4$ variables have also been experimented with, typically resulting in an absolute cross-validation error around 0.09. While some of the load values by themselves give useful information as indicated by the *random forest*, the high ranking load values have a very high correlation with each other as well as the "25% quantile". For example "25% quantile" and load value X7 have a correlation of 0.89. In our experiments the "dailyMinMean" together with "25% quantile" give the best result, replacing 25% quantile with one of the load values gave worse result even though they have a higher ranking from the *random forest*. Using the "dailyMinMean" together with "25% quantile" we obtain a very low error rate as seen in Table 2, both in terms of absolute error on the training data set as well as the absolute error from a 10-fold cross-validation

procedure. A scatterplot of these two variables as well as the line separating the two classes made from our analysis can be seen in Fig. 5.

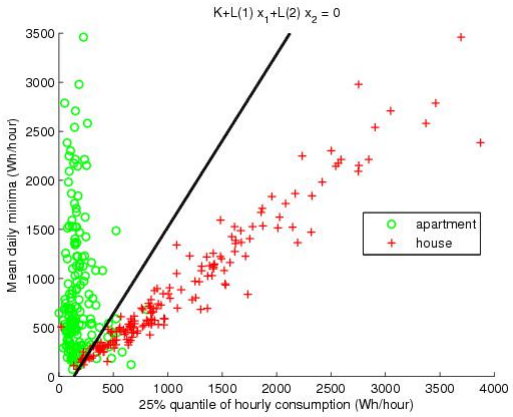


Fig. 5. A scatter plot of "dailyMinMean" and "25% quantile" and a plot of the classification boundary. $K = -241,58$, $L(1) = 1.76$ and $L(2) = 1$.

6 Conclusion and future work

We have observed different patterns in terms of energy consumption for different building types as house and apartments on a data set of energy consumption values for 350 Swedish households. Motivated by this observation, we conduct a classifications task using techniques as random forest, classification tree and linear discriminant analysis method. We note that statistics like "25% quantile" and the mean of daily minimum (hourly) consumption values (shortened as "dailyMinMean") and a couple of other variables ranked high by the random forest technique are the most relevant predictors in classifying building types (house or apartment). This result is further confirmed by the fitted classification models, namely the classification tree and linear discriminant analysis method. In addition, load profile at certain morning clock hours can also be useful in the classifying task, given the fact that they are highly correlated to the feature variable "25% quantile". The results of the classification models are plausible using the selected feature variables. Given that the 10-fold cross validation absolute/misclassfication error being very low in both the classification tree model and the linear discriminant analysis. In particular, the chosen tree model, i.e. tree number 3 in Table 1, gives a low 10-fold cross validation (absolute) error of 8.57%. When the best predictors as "25% quantile" and "dailyMinMean" are used for the linear discriminant analysis, we obtain only 6.29% 10-fold cross validation (absolute) error. These quantities can be easily obtained given a time series of energy consumption data with unknown building type.

The classification has been done with hourly energy consumption values aggregated from the original data of consumption values for each ten-minute. A natural question arises as to whether the high-frequency ten-minute data provides more information in the classification task or a customer clustering task than the lower-frequency one-hour data. These are some questions which we aim to address in our future work.

7 Acknowledgements

The authors are grateful to the Swedish Energy Agency for the financial support under research grant 33707-1. Appreciations also go to professor Sergei Silvestrov who has been very helpful in facilitating the research project and company Revolution Analytics for providing the useful software Revolution R Enterprise 6.0 for academic users.

References

1. Breiman, L., Random forests. *Machine Learning*, 45,**1**, 5–32, 2001.
2. Breiman, L., Friedman, J. Olshen, R., and Stone, C.,*Classification and regression trees*. Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, 1984.
3. Chicco, G. Overview and performance assessment of the clustering methods for electrical load pattern grouping. *Energy*, Vol. 42, **1**, pp. 68–80, 2012.
4. Chicco, G., Napoli, R., Piglione, F. Application of clustering algorithms and self organising maps to classify electricity customers. *Proc. of IEEE Power Tech Conference*, Bologna, 23–26 June, 2003.
5. Johnson, R. A. and Wichern, D. W. *Applied multivariate statistical analysis,* 5^{th}ed. Prentice Hall, 2002.
6. Jota, P.R.S., Silva, V.R.B., Jota, F.G. Building load management using cluster and statistical analyses. *International Journal of Electrical Power and Energy Systems*, Vol. 33, **8**, pp. 1498–1505, 2011.
7. López, J.J., Aguado, José A, Martín, F., Muñoz, F., Rodríguez, A., Ruiz, José E. Hopfield-K-Means clustering algorithm: A proposal for the segmentation of electricity customers. *Electric Power Systems Research*, Vol. 81, pp. 716–724, 2011.
8. Rodrigues,F., Duarte, J.,Figueiredo, V., Vale, Z.and Cordeiro, M., A comparative Analysis of Clustering Algorithms Applied to Load Profiling, *Machine Learning and Data Mining in Pattern Recognition,* ser. Lecture Notes in Computer Science, P. Perner and A. Rosenfeld, Eds. Springer Berlin Heidelberg, vol. 2734, pp. 73–85, 2003.
9. Swedish Energy Agency, http://www.energimyndigheten.se/Statistik/FESTIS/ (Accessed 2014-03-20).
10. Therneau, T. M. and Atkinson, B. R port by Brian Ripley. *rpart: Recursive Partitioning*. R package version 3.1-46, 2010.
11. Torgo, L., *Data Mining with R, learning with case studies*, Chapman and Hall/CRC. URL: http://www.dcc.fc.up.pt/ ltorgo/DataMiningWithR, 2010.
12. Zhou, KL., Yang, SL., Shen, C. A review of electric load classification in smart grid environment. *Renewable and Sustainable Energy Reviews*, Vol. 24, pp. 103–110, 2013.

13. Zimmerman, J.P. End-use metering campaign in 400 households in Sweden. September 2009, http://www.energimyndigheten.se/Global/Statistik/F%C3%B6rb%C3%A4ttrad%20energistatistik/Festis/Final_report.pdf (Accessed 2014-03-20)

Vandermonde matrices, extreme points and orthogonal polynomials

Jonas Österberg, Karl Lundengård, and Sergei Silvestrov

Division of Applied Mathematics, School of Education, Culture and Communication, Mälardalen University, Box 883, SE-721 23 Västerås, Sweden (e-mail: jonas.osterberg@mdh.se, karl.lundengard@mdh.se, sergei.silvestrov@mdh.se)

Abstract. Vandermonde matrices and their determinant appear in many different problems, including interpolation of data, moment matching in stochastic processes with applications to computer-aided decision support and in various types of numerical analysis. Motivated by these and other applications the values of the determinant of Vandermonde matrices are analyzed both visually and analytically over the unit sphere in various dimensions and under various norms. The extreme points of the determinant over these surfaces are related to polynomials and classical orthogonal polynomials in particular. Some of the polynomials are identified and recursive definitions are provided.
Keywords: Vandermonde matrix, Determinant, Extreme points, Orthogonal polynomials, Moment matching.

1 The Vandermonde matrix

A rectangular Vandermonde matrix of size $m \times n$ is determined by n values $\boldsymbol{x}_n = (x_1, \cdots, x_n)$ and is defined by

$$V_{mn}(\boldsymbol{x}_n) = \left[x_j^{i-1}\right]_{mn} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{m-1} & x_2^{m-1} & \cdots & x_n^{m-1} \end{bmatrix}. \tag{1}$$

Note that some authors use the transpose of this as the definition and possibly also let indices run from 0. All entries in the first row of Vandermonde matrices are ones and by considering $0^0 = 1$ this is true even when some x_j is zero. We have the following well known theorem.

Theorem 1. *The determinant of (square) Vandermonde matrices has the well known form*

$$\det V_n(\boldsymbol{x}_n) \equiv v_n(\boldsymbol{x}_n) = \prod_{1 \le i < j \le n} (x_j - x_i). \tag{2}$$

3^{rd} *SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal*
C. H. Skiadas (Ed)

This determinant is also simply referred to as the *Vandermonde determinant*.

Perhaps the most well known application of Vandermonde matrices is polynomial interpolation. Given a set of points $(x_i, y_i) \in \mathbb{C}^2$, $1 \leq i \leq n$, define the two row vectors $\boldsymbol{x}_n = (x_1, \cdots, x_n)$ and $\boldsymbol{y}_n = (y_1, \cdots, y_n)$ and let all x_i be pairwise distinct. Then the polynomial of minimal degree that interpolates these points has coefficients $\boldsymbol{c}_n$ where $\boldsymbol{c}_n V_n(\boldsymbol{x}_n) = \boldsymbol{y}_n$, that is $\boldsymbol{c}_n = \boldsymbol{y}_n V_n(\boldsymbol{x}_n)^{-1}$. Note that since the x_i are pairwise distinct the determinant $v_n(\boldsymbol{x}_n)$ is nonzero by Theorem 1 and so the solution is unique. More information and an expansion on this can be found in El-Mikkawy[1] and references therein.

There are other applications as well, for instance in Lundengård *et al.*[4] we find an application of the Vandermonde matrix to moment matching of random variables with a discrete random variable. This has applications in pricing of Asian options and other areas.

In the article Klein and Spreij[3] we find an application of Vandermonde matrices to the important area of time-series analysis.

The work done here recaptures the work presented in Szegő[7]. A more explicit treatment is provided together with some slight generalization of the ideas.

2 Optimizing the determinant

Consider the unit $(n-1)$-sphere under the p-norm ($p \in \mathbb{Z},\ p \geq 1$), we define

$$\mathcal{S}_p^{n-1} = \{\boldsymbol{x}_n \in \mathbb{R}^n \ : \ |x_1|^p + \cdots + |x_n|^p = 1\}\,.$$

When $p = 2$ we have the Euclidean norm and thus the usual $(n-1)$-sphere, when $p = \infty$ we have the cube defined by the boundary of $[-1, 1]^n$. The 2-sphere for some different norms are depicted in Figure 1.

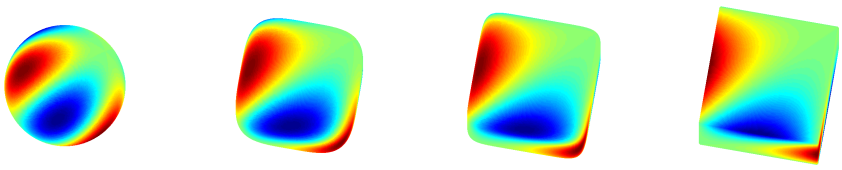


Fig. 1. Value of $v_3(\boldsymbol{x}_3)$ over: $\mathcal{S}_2^2$ (left), $\mathcal{S}_4^2$ (middle left), $\mathcal{S}_8^2$ (middle right) and $\mathcal{S}_\infty^2$ (right).

The four separate plots in Figure 1 are all rotated slightly by the same transformation for visual clarity and the mappings of color to value are slightly different between the figures. The positive maxima can be seen to lie in the dark red regions and the minima can be found in the dark blue regions. As we soon will see the extreme points on $\mathcal{S}_2^2$ are $(-1/\sqrt{2}, 0, 1/\sqrt{2})$ and the vectors constructed by permutating the coordinates of this vector, making a total of $6 = 3!$ extreme points. Similarily, for the cube we have the vectors formed by the six

permutations of the coordinates in $(-1, 0, 1)$. The two stated vectors are both maxima and are shown in the top left maxima in the figures, odd permutations of these vectors will give minima and even permutations will give again maxima. This follows directly from the anti-symmetry of the Vandermonde determinant, given a permutation $\sigma \in S_n$ we have $v_n(x_{\sigma_1}, \cdots, x_{\sigma_n}) = \operatorname{sgn}(\sigma) v_n(x_1, \cdots, x_n)$.

We are thus faced with the problem of maximizing $|v_n(\boldsymbol{x}_n)|$ over $\mathcal{S}_p^{n-1}$. By the symmetry of $|v_n|$ we do not loose any generality by assuming that the coordinates are ordered and pairwise distinct, $x_1 < \cdots < x_n$.

Theorem 2. *The coordinates $x_1 < \cdots < x_n$, $n \geq 2$, that maximize the absolute value of the Vandermonde determinant over the unit sphere in $\mathbb{R}^n$, under the p-norm, that is $\mathcal{S}_p^{n-1}$, are unique. Furthermore, the coordinates are symmetric so that $x_i = -x_{n-i+1}$ for $1 \leq i \leq n$ and the total number of maxima, counting permutations, are $n!$.*

This is an extension of a result presented by Szegő[7, p.140].

Proof. Suppose that we have two different sequences of coordinates $\boldsymbol{x}_n$ and $\boldsymbol{x}'_n$:

$$x_1 < \cdots < x_n,$$

$$x'_1 < \cdots < x'_n,$$

both maximizing $|v_n|$ on $\mathcal{S}_p^{n-1}$, so $|v_n(x_1, \cdots, x_n)| = |v_n(x'_1, \cdots, x'_n)| = v_n^{max}$ is the maximum value. Now define a new configuration $\boldsymbol{z}_n$ defined by,

$$z_i = \frac{x_i + x'_i}{2\sigma}, \quad 1 \leq i \leq n,$$

where $\sigma > 0$ is a normalization constant to assure that $\boldsymbol{z}_n$ lies on the sphere. It is easy to see that we have

$$z_1 < \cdots < z_n.$$

Note that $\sigma \leq 1$ by necessity since $\mathcal{S}_p^{n-1}$ is the boundary of a convex set. This follows directly from the absolute homogeneity and the triangle inequality associated with the p-norm:

$$\left\| \frac{\boldsymbol{x}_n + \boldsymbol{x}'_n}{2} \right\|_p \leq \left\| \frac{\boldsymbol{x}_n}{2} \right\|_p + \left\| \frac{\boldsymbol{x}'_n}{2} \right\|_p = 1.$$

We have established that $\boldsymbol{z}_n$ lies on $\mathcal{S}_p^{n-1}$. For each $1 \leq i < j \leq n$ we now have

$$\begin{aligned} |z_j - z_i| &= \frac{\left| x_j + x'_j - x_i - x'_i \right|}{2\sigma} = \frac{|x_j - x_i| + \left| x'_j - x'_i \right|}{2\sigma} \\ &\geq |x_j - x_i|^{\frac{1}{2}} \left| x'_j - x'_i \right|^{\frac{1}{2}}, \end{aligned}$$

where the last step follows from $\sigma < 1$ and the general relation

$$\left(\frac{a+b}{2} \right)^2 = \left(\frac{a-b}{2} \right)^2 + ab \geq ab.$$

It follows that

$$|v_n(\boldsymbol{z}_n)| \geq |v_n(\boldsymbol{x}_n)|^{\frac{1}{2}} \, |v_n(\boldsymbol{x}'_n)|^{\frac{1}{2}} = v_n^{max},$$

and to not have a contradiction, we must have that the equality holds, that is $\boldsymbol{x}_n = \boldsymbol{x}'_n$, and so the maximum is unique.

Now, consider

$$v_n(-\boldsymbol{x}_n) = (-1)^{\frac{n(n-1)}{2}} v_n(\boldsymbol{x}_n),$$

which follows easily from the degree of the expansion of v_n, where every term is of degree $\frac{n(n-1)}{2}$. If follows that if $\boldsymbol{x}_n$ is a maximum of $|v_n|$ on the sphere then $-\boldsymbol{x}_n$ is also a maximum. Now, since the maximum with ordered coordinates $x_1 < \cdots < x_n$ is unique we must have that $-x_n = x_1, \cdots, -x_1 = x_n$, that is $x_i = -x_{n-i+1}$ for $1 \leq i \leq n$ and so the maxima are symmetric.

From $x_i = -x_{n-i+1}$ and the pairwise distinctness of the coordinates (the determinant is non-zero at the maximum) we have that the $n!$ permutations $(x_{\sigma_1}, \cdots, x_{\sigma_n})$ are the distinct maxima. □

Remark 1. The condition $x_i = -x_{n-i+1}$ for the ordered maximum $x_1 < \cdots < x_n$ implies that the maxima all lie in the hyperplane $x_1 + \cdots + x_n = 0$. These facts can be used to visualize v_n on $\mathcal{S}_p^{n-1}$ for $n = 4, 5, 6, 7$, while we for $n \geq 8$ have more than three degrees of freedom.

As an interesting example we provide a plot of v_4 over all points on the Euclidean hypersphere that is constructed by selecting an orthonormal basis in $\mathbb{R}^4$ that is in turn orthogonal to $(1, 1, 1, 1)$. We do not miss any extrema by doing this since all extrema, considered as points in $\mathbb{R}^4$, must be orthogonal to this vector. We provide both a 3D plot and a plot in spherical coordinates by the following transformations.

$$\boldsymbol{x}_4 = \begin{bmatrix} -1 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{4} & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix} \boldsymbol{t}. \tag{3}$$

$$\boldsymbol{t}(\theta, \phi) = \begin{bmatrix} \cos(\phi)\sin(\theta) \\ \sin(\phi) \\ \cos(\phi)\cos(\theta) \end{bmatrix}. \tag{4}$$

In Figure 2 the placement of the $4! = 24$ extreme points of v_n over $\mathcal{S}_2^3$ can be seen.

Motivated by the fact that there is only one set of coordinates and that the maxima are constructed from different orderings of these, it makes sense to instead consider the polynomial constructed from these coordinates. Our task now is to find the polynomials that define the optimizing coordinates for different $n \geq 2$ and $p \geq 1$,

$$P_p^n(x) = \prod_{i=1}^{n} (x - x_i),$$

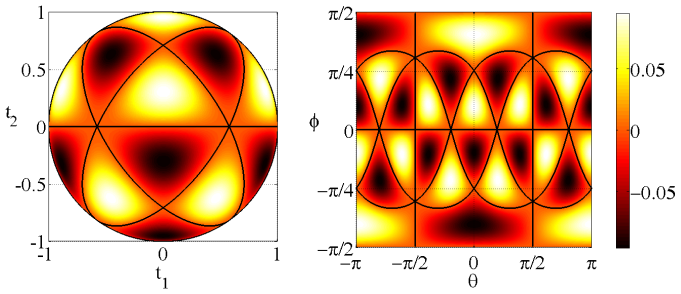


Fig. 2. The values of the Vandermonde determinant over $\mathcal{S}_2^3$.

where x_i are the distinct maximizing coordinates that depend on n and p. We have

$$\begin{aligned} P_p^{n\prime}(x_k) &= \sum_{j=1}^{n} \prod_{\substack{i=1 \\ i\neq j}}^{n} (x - x_i)\bigg|_{x=x_k} = \prod_{\substack{i=1 \\ i\neq k}}^{n} (x_k - x_i), \\ P_p^{n\prime\prime}(x_k) &= \sum_{l=1}^{n} \sum_{\substack{j=1 \\ j\neq l}}^{n} \prod_{\substack{i=1 \\ i\neq j \\ i\neq l}}^{n} (x - x_i)\bigg|_{x=x_k} = \sum_{\substack{j=1 \\ j\neq k}}^{n} \prod_{\substack{i=1 \\ i\neq j \\ i\neq k}}^{n} (x_k - x_i) + \sum_{\substack{l=1 \\ l\neq k}}^{n} \prod_{\substack{i=1 \\ i\neq l \\ i\neq k}}^{n} (x_k - x_i) \\ &= 2 \sum_{\substack{j=1 \\ j\neq k}}^{n} \prod_{\substack{i=1 \\ i\neq j \\ i\neq k}}^{n} (x_k - x_i), \end{aligned}$$

and it is easy to show that

$$\frac{P_p^{n\prime\prime}(x_k)}{P_p^{n\prime}(x_k)} = 2 \sum_{\substack{i=1 \\ i\neq k}}^{n} \frac{1}{x_k - x_i}. \tag{5}$$

Define

$$s_p^n(\boldsymbol{x}_n) = \left(\sum_{i=1}^{n} |x_i|^p \right) - 1,$$

so that s_p^n vanishes exactly on the unit sphere under the p-norm, that is $\mathcal{S}_p^{n-1} = \left\{ \boldsymbol{x}_n \in \mathbb{R}^n \ : \ s_p^n(\boldsymbol{x}_n) = 0 \right\}$. We have the partial derivatives

$$\frac{\partial s_p^n(\boldsymbol{x}_n)}{\partial x_k} = p\, |x_k|^{p-1} \operatorname{sgn}(x_k).$$

To maximize $|v_n(\boldsymbol{x}_n)|$ over the surface $s_p^n(\boldsymbol{x}_n) = 0$ we transform the objective function by applying a (strictly increasing) logarithm:

$$w_n(\boldsymbol{x}_n) = \log(|v_n(\boldsymbol{x}_n)|) = \sum_{1 \leq i \leq j} \log(|x_j - x_i|),$$

with partial derivatives

$$\frac{\partial w_n(\boldsymbol{x}_n)}{\partial x_k} = \sum_{\substack{i=1 \\ i \neq k}}^{n} \frac{1}{x_k - x_i} = \frac{1}{2} \frac{P_p^{n\prime\prime}(x_k)}{P_p^{n\prime}(x_k)}.$$

By the method of Lagrange multipliers we now have that the maxima of $|v_n(\boldsymbol{x}_n)|$ on $\mathcal{S}_p^{n-1}$ must be stationary points to the Lagrangian

$$\Lambda(\boldsymbol{x}_n, \lambda) = w_n(\boldsymbol{x}_n) - \lambda s_p^n(\boldsymbol{x}_n), \tag{6}$$

which explicitly means

$$\frac{\partial w_n(\boldsymbol{x}_n)}{\partial x_k} = \lambda \frac{\partial s_p^n(\boldsymbol{x}_n)}{\partial x_k},$$

for some multiplier $\lambda \in \mathbb{R}$. We then get

$$\frac{1}{2} \frac{P_p^{n\prime\prime}(x_k)}{P_p^{n\prime}(x_k)} = \lambda p \left| x_k \right|^{p-1} \operatorname{sgn}(x_k).$$

Letting $\rho = -2\lambda p$ we then get

$$P_p^{n\prime\prime}(x_k) + \rho \left| x_k \right|^{p-1} \operatorname{sgn}(x_k) P_p^{n\prime}(x_k) = 0. \tag{7}$$

This leads us to our first set of polynomials defining the solution to our maximization problem.

Theorem 3. *The polynomial P_2^n, of degree $n > 2$ and with a leading coefficient of 1, whose roots form the coordinates in the points $\boldsymbol{x}_n \in \mathbb{R}^n$ that maximize $|v_n(\boldsymbol{x}_n)|$ over the Euclidean hypersphere $\mathcal{S}^{n-1}$ satisfy the differential equation*

$$P_2^{n\prime\prime}(x) + n(1-n)x P_2^{n\prime}(x) + n^2(n-1) P_2^n(x) = 0. \tag{8}$$

Furthermore, the coefficients for the three terms of highest degree are

$$c_n = 1, c_{n-1} = 0, c_{n-2} = -\frac{1}{2}, \tag{9}$$

and the subsequent coefficients are defined recursively by

$$c_k = -\frac{(k+1)(k+2)}{n(n-1)(n-k)} c_{k+2}. \tag{10}$$

Proof. By Equation (7) we have

$$P_2^{n\prime\prime}(x) + \rho x P_2^{n\prime}(x) \Big|_{x=x_k} = 0, \quad 1 \leq k \leq n.$$

The left part of this equation represents n evaluations of a polynomial of degree n that vanishes on $x_1, \cdots, x_n$ and must thus be a constant multiple of P_2^n, that we defined as the polynomial that vanishes on $x_1, \cdots, x_n$, and so

$$P_2^{n\prime\prime}(x) + \rho x P_2^{n\prime}(x) + \sigma P_2^n(x) = 0.$$

Two find the coefficients σ and ρ we need to adapt this polynomial to the sphere. We require $\sum_{i=1}^{n} x_i^2 = 1$ and by the symmetry from Remark 1 we have $\sum_{i=1}^{n} x_i = 0$. The condition $c_n = 1$ is by choice. The condition $c_{n-1} = 0$ follows from the expansion of the coefficients in P_2^n.

$$P_n^p(x) = \prod_{i=1}^{n}(x - x_i) = \sum_{k=0}^{n}(-1)^{n-k} e_{n-k}(x_1, \cdots, x_n) x^k,$$

where e_k is the elementary symmetric polynomial defined by

$$e_k(x_1, \cdots, x_n) = \sum_{i_1 < \cdots < i_k} x_{i_1} x_{i_2} \cdots x_{i_k}.$$

We have $c_{n-1} = -e_1(x_1, \cdots, x_n) = -(x_1 + \cdots + x_n) = 0$. The condition $c_{n-2} = -\frac{1}{2}$ places us on the unit sphere

$$e_1(x_1, \cdots, x_n)^2 - 2e_2(x_1, \cdots, x_n) = x_1^2 + \cdots + x_n^2 = 1,$$

$$-2e_2(x_1, \cdots, x_n) = 1,$$

$$c_{n-2} = e_2(x_1, \cdots, x_n) = -\frac{1}{2}.$$

This establishes Equation (9) in the theorem.

Now, the coefficients $c_n, \cdots, c_0$ for any polynomial solution $p(x)$ of degree n to a differential equation on the form

$$p^{n\prime\prime}(x) + \rho x p^{n\prime}(x) + \sigma p^n(x) = 0,$$

must satisfy

$$\begin{aligned} \rho n c_n + \sigma c_n &= 0 \\ \rho(n-1)c_{n-1} + \sigma c_{n-1} &= 0 \\ n(n-1)c_n + \rho(n-2)c_{n-2} + \sigma c_{n-2} &= 0. \end{aligned}$$

The second of these equations is trivial since $c_{n-1} = 0$. The first and third equation simplifies to

$$\begin{aligned} \rho n + \sigma &= 0, \\ n(n-1) - \frac{1}{2}\rho(n-2) - \frac{1}{2}\sigma &= 0, \end{aligned}$$

and so

$$\begin{aligned} \rho = \frac{-2n(n-1)}{n-(n-2)} &= -n(n-1) \\ \sigma &= n^2(n-1), \end{aligned}$$

which establishes Equation (8) in the theorem.

For the recursive formula for the coefficients $c_{n-3}, \cdots, c_0$, Equation (10), we again look at the slightly more general case and retain ρ, σ. The coefficients for the polynomial p satisfying

$$p^{n\prime\prime}(x) + \rho x p^{n\prime}(x) + \sigma p^n(x) = 0,$$

must satisfy

$$c_{k+2}(k+1)(k+2) + \rho k c_k + \sigma c_k = 0,$$

that is

$$c_k = \frac{-(k+1)(k+2)}{\rho k + \sigma} c_{k+2}.$$

For P_2^n we then have

$$c_k = \frac{-(k+1)(k+2)}{-n(n-1)k + n^2(n-1)} c_{k+2}, \tag{11}$$

and Equation (10) follows. □

The case $p = \infty$ follows in a similar manner.

Theorem 4. *The polynomial P_∞^n, of degree $n > 2$ and with a leading coefficient of 1, whose roots form the coordinates in the points $\boldsymbol{x}_n \in \mathbb{R}^n$ that maximize $|v_n(\boldsymbol{x}_n)|$ over the cube $\mathcal{S}_\infty^{n-1}$ satisfy*

$$P_\infty^n(x) = (x-1)(x+1)p_\infty^n(x). \tag{12}$$

where p_∞^n is defined by the differential equation:

$$(1-x^2){p_\infty^n}''(x) - 4x{p_\infty^n}'(x) + m(m+3)p_\infty^n(x) = 0,$$

where $m = n-2$ is the degree of the polynomial p_∞^n. Furthermore, the first two coefficients for p_∞^n are $c_m = 1, c_{m-1} = 0$ and the subsequent coefficients satisfy

$$c_k = \frac{(k+1)(k+2)}{k(k+3) - m(m+3)} c_{k+2}. \tag{13}$$

Proof. It is easy to show that the coordinates -1 and $+1$ must be present in the maxima points, if they were not then we could rescale the point so that the value of $|v_n(\boldsymbol{x}_n)|$ is increased, which is not allowed. We may thus assume the ordered sequence of coordinates

$$-1 = x_1 < \cdots < x_n = +1.$$

The absolute value of the Vandermonde determinant then becomes

$$|v_n(\boldsymbol{x}_n)| = 2 \prod_{i=2}^{n-1} \left(|1 + x_i|\,|1 - x_i|\right) \prod_{1<i<j<n} |x_j - x_i|\,.$$

We now take the logarithm of this, differentiate and equate the partial derivatives to zero to find the stationary points (actually maxima), and arrive at

$$\frac{1}{x_k+1}+\frac{1}{x_k-1}+\sum_{\substack{i=2\\ i\neq k}}^{n-1}\frac{1}{x_k-x_i}=0, \quad 1<k<n,$$

which by the essence of Equation (5) can be written

$$\frac{1}{x_k+1}+\frac{1}{x_k-1}+\frac{1}{2}\frac{{p_\infty^n}''(x_k)}{{p_\infty^n}'(x_k)}=0, \quad 1<k<n, \tag{14}$$

for some polynomial p_∞^n constructed from the roots $x_2,\cdots,x_{n-1}$.

The left part of Equation (14) now identifies a differential expression on p_∞^n which we by the same method as for $p=2$ identify by a multiple of p_∞^n, that is

$$(1-x^2){p_\infty^n}''(x)-4x{p_\infty^n}'(x)+\sigma p_\infty^n(x)=0. \tag{15}$$

The constant σ is found by considering the coefficient for x^m:

$$-m(m-1)-4m+\sigma=0 \quad \Leftrightarrow \quad \sigma=m(m+3).$$

Finally

$$(1-x^2){p_\infty^n}''(x)-4x{p_\infty^n}'(x)+m(m+3)p_\infty^n(x)=0.$$

The polynomial that provide all coordinates for the maxima is then

$$P_\infty^n(x)=(x-1)(x+1)p_\infty^n. \tag{16}$$

Equation (13) follows by the same methods as for $p=2$, that is, by identifying all coefficients in the differential equation to be identically zero. □

We have provided the means to describe the coordinates of the extreme points of the Vandermonde determinant over the unit spheres under the Euclidean norm and under the infinity norm. The resulting polynomials can be identified by rescaled Hermite polynomials, $H_n(x\sqrt{n(n-1)/2})$, and Gegenbauer polynomials, $C_n^{(3/2)}(x)$, respectively. For norms other than $p=2$ and $p=\infty$ further analysis is warranted.

Acknowledgements This research was supported in part by the Swedish Research Council (621-2007-6338), Swedish Foundation for International Cooperation in Research and Higher Education (STINT), Royal Swedish Academy of Sciences, Royal Physiographic Society in Lund and Crafoord Foundation.

References

1. El-Mikkawy, M., "Vandermonde interpolation using special associated matrices", *Applied Mathematics and Computation* 141:589–595 (2003).
2. Goldwurm, M. and Lonati, V., "Pattern statistics and Vandermonde matrices", *Theoretical Computer Science* 356:153–169 (2006).

3. Klein, A. and Spreij, P., "Some results on Vandermonde matrices with an application to time series analysis", *Siam J. Matrix Anal. Appl.* 25:213–223 (2003).
4. Lundengård, K., Ogutu, C., Silvestrov, S. and Weke, P., "Asian Options, Jump-Diffusion Processes on a Lattice, and Vandermonde Matrices", *Modern Problems in Insurance Mathematics, Silvestrov, Dmitrii, Martin-Löf, Anders, Eds.* Springer-Verlag, Berlin, 337–364 (2014).
5. Lundengård, K., Österberg, J., Silvestrov, S., "Extreme points of the Vandermonde determinant on the sphere and some limits involving the generalized Vandermonde determinant", arXiv:1312.6193 (2013).
6. Serre, D., *Matrices: Theory and Applications*, Springer (2002).
7. Szegő, G., *Orthogonal polynomials*, American mathematics society (1939).

Sensitivity Analysis of the GI/M/1 Queue with Negative Customers

Sofiane Ouazine[1] and Karim Abbas[2]

[1] Department of Mathematics
University of Bejaia, Campus of Targua Ouzemour, Algeria
(e-mail: wazinesofi@gmail.com)
[2] LAMOS Laboratory,
University of Bejaia, Campus of Targua Ouzemour, Algeria
(e-mail: kabbas.dz@gmail.com)

Abstract. In this paper we discuss the applicability of the Taylor series approach to the numerical analysis of the GI/M/1 queue with negative customers. In other words, we use the Taylor series expansions to examine the robustness of the GI/M/1 (FIFO, ∞) queueing model having RCH (Removal of Customer at the Head) to perturbations in the negative customers process (the occurrence rate of RCH). We analyze numerically the sensitivity of the entries of the stationary distribution vector of the GI/M/1 queue with negative customers to those perturbations, where we exhibit these entries as polynomial functions of the occurrence rate of RCH parameter of the considered model. Numerical examples are sketched out to illustrate the accuracy of our approach.
Keywords: Taylor series expansion, Sensitivity analysis, GI/M/1 queue with negative customers, Numerical methods, Performance measures.

1 Introduction

Recently there has been a rapid increase in the literature on queueing systems with negative arrivals. Queues with negative arrivals, called G-queues, were first introduced by Gelenbe [5]. When a negative customer arrives, it immediately removes an ordinary (positive) customer if present. Negative arrivals have been interpreted as inhibiter and synchronization signals in neural and high speed communication network. For example, we can use negative arrivals to describe the signals, which are caused by the client, cancel some proceeding.

There is a lot of research on queueing system with negative arrivals. But most of these contributions considered continuous-time queueing model: Boucherie and Boxma [6], Jain and Sigman [8], Bayer and Boxma [2], Harrison and Pitel [9] all of them investigated the same M/G/1 model but with the different killing strategies for negative customers; Harrison, Patel and Pitel [10] considered the M/M/1 G-queues with breakdowns and repair; Yang [11] considered GI/M/1 model by using embedded Makov chain method.

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

In this paper we investigate the GI/M/1/N with Poisson negative arrivals to test the robustness of the model to perturbation in the negative customers process (the occurrence rate of RCH). In deed, we use the Taylor series expansions to examine the robustness of the GI/M/1/N queue to perturbations in the arrival process. Specifically, we analyse numerically the sensitivity of the entries of the stationary distribution vector of the GI/M/1/N queue to that perturbations, where we exhibit these entries as polynomial functions of the occurrence rate of RCH.

The remainder of this paper is organized as follows. In Section 2, we introduce the necessary notations for analyzing of the considered queueing model, and present closed-form expressions for the sensitivity of the stationary distribution to model parameter as a function of the deviation matrix. In Section 3, we outline the numerical framework to compute the relative absolute error in computing the stationary distribution. Concluding remarks are provided in Section 4.

2 Queueing Model Analysis

We investigate the GI/M/1/N queue with negative customers, where N is the capacity of the system including the one who is in service. Assume that customer arrivals occur at discrete-time instants τ_k, where $\tau_0 = 0$, customers arrive at the system according to a renewal process with interarrival time distribution $G(t)$ and mean $1/\lambda$. The service time T_s of each server is assumed to be distributed exponentially with service rate μ. Its density function is given by

$$s(t) = \mu e^{-\mu t}, \quad t \geq 0.$$

Additionally, we assume that there is another kind of customers, namely RCH, arriving in the system according to an independent Poisson process of parameter h. Let L_k denote the number of customers left in the system immediately after the kth departing customer. A sequence of random variables $L_k; k = 1, 2, \ldots, N$ constitutes a Markov chain. Its transition probabilities matrix is given by:

$$P = \begin{pmatrix} b_0 & a_0 & 0 & 0 & 0 & \ldots & 0 \\ b_1 & a_1 & a_0 & 0 & 0 & \ldots & 0 \\ b_2 & a_2 & a_1 & a_0 & 0 & \ldots & 0 \\ b_3 & a_3 & a_2 & a_1 & a_0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{N-1} & a_{N-1} & a_{N-2} & a_{N-3} & a_{N-4} & \ldots & a_0 \\ b_{N-1} & a_{N-1} & a_{N-2} & a_{N-3} & a_{N-4} & \ldots & a_0 \end{pmatrix}_{(N+1)\times(N+1)}$$

where $a_j = \frac{(\mu+h)^j}{(-1)^j j!} \frac{\partial^j F^*}{\partial \alpha^j}(\alpha)$, $b_j = 1 - \sum_{i=0}^{j} a_i$, $\alpha = \mu + h$ and F^* is the Laplace transformation corresponding to pdfs f(i.e $dG(t) = f(t)dt$) of the interarrival process.

Let π denote the stationary distribution of the Markov chain L_k. We define the traffic intensity ρ =(arrival rate)/(service rate) = λ/μ. It can be shown

that the Markov chain L_k is positive recurrent for all ρ. In this paper, we consider the stationary distribution π as a mapping of some real-valued parameter θ, in notation π_θ. For example, θ may denote the occurrence rate of RCH parameter of the model. We are interested in obtaining higher-order sensitivity of stationary distribution with respect to parameter θ. In the sequel we derive formulas for the higher order derivatives of π_θ with respect to θ. Then, by using these formulas we obtain a Taylor series expansions in θ for $\pi_{\theta+\Delta}$, where its coefficients are expressed in closed form as functions of the *deviation matrix* (denoted by D_θ) associated to the Markov chain L_k. It is well know that if P_θ is irreductible, then $(I - P_\theta + \Pi_\theta)$ is invertible, where $\Pi_\theta = e\,\pi_\theta$. Then, the matrix $D_\theta = (I - P_\theta + \Pi_\theta)^{-1} - \Pi_\theta$ exists and it is called the deviation matrix. The deviation matrix can be obtained in explicit form by:

$$\begin{aligned} D_\theta &= \sum_{n=0}^{\infty} (P_\theta^n - \Pi_\theta), \\ &= \sum_{n=0}^{\infty} (P_\theta - \Pi_\theta)^n - \Pi_\theta, \\ &= (I - P_\theta + \Pi_\theta)^{-1} - \Pi_\theta. \end{aligned}$$

In the following theorem we give the higher-order derivatives of the stationary distribution π_θ with respect to θ in terms of the deviation matrix D_θ, which is a key result used in the framework proposed subsequently.

Theorem 1. *[7]* Let $\theta \in \Theta$ and let $\Theta_0 \subset \Theta$, with $\Theta \subset \mathbf{R}$ be a closed interval with θ an interior point such that the Markov chain is ergodic on Θ_0. Provided that the entries of the transition matrix P_θ are n-times differentiable with respect to θ, let

$$K_\theta(n) = \sum_{\substack{1 \le m \le n; \\ 1 \le l_k \le n; \\ l_1 + \cdots + l_m = n}} \frac{n!}{l_1! \cdots l_m!} \prod_{k=1}^{m} \left(P_\theta^{(l_k)}\, D_\theta \right) ..$$

Then, it holds that

$$\pi_\theta^{(n)} \;=\; \pi_\theta\, K_\theta(n)\,, \tag{1}$$

where $P_\theta^{(k)}$ (respectively $\pi_\theta^{(k)}$) is the matrix (respectively vector) of the elementwise kth order derivative of P_θ (respectively π_θ) with respect to parameter θ.

In the following, we propose a numerical approach to compute the stationary distribution π_θ for some parameter value θ, and we demonstrate how this stationary distribution can be evaluated for the case where the control parameter θ is changed in some interval. In other words, we will compute the function $\pi(\theta+\Delta)$ on some Δ-interval. More specifically, we will approximately compute $\pi(\theta + \Delta)$ by an polynomial in Δ. To achieve this we will use the Taylor series

expansion approach established in [7]. Under some mild conditions it holds that $\pi_{\theta+\Delta}$ can be developed into a Taylor series of the following form:

$$\pi_{\theta+\Delta} = \sum_{n=0}^{k} \frac{\Delta^n}{n!} \pi_\theta^{(n)}, \tag{2}$$

where $\pi_\theta^{(n)}$ denotes the n-th order derivative of π_θ with respect to θ (see formula (1)).

We call

$$H_\theta(k, \Delta) = \sum_{n=0}^{k} \frac{\Delta^n}{n!} \pi_\theta^{(n)} \tag{3}$$

the k-th order Taylor approximation of $\pi_{\theta+\Delta}$ at θ.

Under the conditions put forward in Theorem 1 it holds for $k < n$ that:

$$\pi_\theta^{(k+1)} = \sum_{m=0}^{k} \binom{k+1}{m} \pi_\theta^{(m)} P_\theta^{(k+1-m)} D_\theta \,. \tag{4}$$

An explicit representation of the lower derivatives of π_θ is given by [1]:

$$\pi_\theta^{(1)} = \pi_\theta P_\theta^{(1)} D_\theta \tag{5}$$

and

$$\pi_\theta^{(2)} = \pi_\theta P_\theta^{(2)} D_\theta + 2\pi_\theta (P_\theta^{(1)} D_\theta)^2. \tag{6}$$

Elaborating on the recursive formula for higher order derivatives (4), the second order derivative can be written as:

$$\pi_\theta^{(2)} = \pi_\theta P_\theta^{(2)} D_\theta + 2\pi_\theta^{(1)} P_\theta^{(1)} D_\theta. \tag{7}$$

In the same vein, we obtain for the third order derivative:

$$\pi_\theta^{(3)} = \pi_\theta P_\theta^{(3)} D_\theta + 3\pi_\theta^{(2)} P_\theta^{(1)} D_\theta + 3\pi_\theta^{(1)} P_\theta^{(2)} D_\theta. \tag{8}$$

3 Numerical Application

In this section, we apply the numerical approach based on the Taylor series expansions introduced above to the GI/M/1/N queue with negative customers, where we consider the model with perturbed the occurrence rate of RCH parameter. In this case, we estimate numerically the sensitivity of the stationary distribution of the queueing model with respect to the perturbation.

Let $\Theta = (a, b) \subset \mathbf{R}$, for $0 < a < b < \infty$.

(H) For $0 \leq j \leq N$ it holds that a_j is n-times differentiable with respect to h on Θ.

Under **(H)** it holds that the first n derivatives of P exists. Let $P^{(k)}$ denote the kth order derivative of P with respect to h, then it holds that

$$P^{(k)}(i,j) = \frac{d^{(k)}}{dh^{(k)}} P(i,j), 0 \leq \quad i,j \leq N, \tag{9}$$

or, more specifically,

$$P^{(k)} = \begin{pmatrix} b_0^{(k)} & a_0^{(k)} & 0 & 0 & 0 & \dots & 0 \\ b_1^{(k)} & a_1^{(k)} & a_0^{(k)} & 0 & 0 & \dots & 0 \\ b_2^{(k)} & a_2^{(k)} & a_1^{(k)} & a_0^{(k)} & 0 & \dots & 0 \\ b_3^{(k)} & a_3^{(k)} & a_2^{(k)} & a_1^{(k)} & a_0^{(k)} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{N-1}^{(k)} & a_{N-1}^{(k)} & a_{N-2}^{(k)} & a_{N-3}^{(k)} & a_{N-4}^{(k)} & \dots & a_0^{(k)} \\ b_{N-1}^{(k)} & a_{N-1}^{(k)} & a_{N-2}^{(k)} & a_{N-3}^{(k)} & a_{N-4}^{(k)} & \dots & a_0^{(k)} \end{pmatrix}_{(N+1)\times(N+1)} \tag{10}$$

where

$$a_j^{(k)} = \frac{(-1)^k(\mu+h)^j}{(-1)^{j+k}j!} \frac{\partial^{j+k} F^*}{\partial \alpha^{j+k}}(\alpha) + \sum_{n=1}^{k} \frac{C_{k-1}^n (-1)^{k-n}(\mu+h)^{j-n}}{(-1)^{k+j-n}(j-n)!} \frac{\partial^{k+j-n} F^*}{\partial \alpha^{k+j-n}}(\alpha) + \frac{(\mu+h)^{j-k}}{(-1)^j (j-k)!} \frac{\partial^j F^*}{\partial \alpha^j}(\alpha)$$

and $b_j^{(k)} = -\sum_{i=0}^{j} a_j^{(k)}$

Consider the M/M/1/N queue with service rate μ and exponential inter-arrival time with rate λ. First, we present the numerical results obtained by applying our approach to this case. Therefore, we set $\mu = 2$, $\lambda = 1$. For the implementation of our algorithm in MATLAB, we require a finite version of our queueing model. Figures 1, 2 and 3 depict the relative error on the stationary distributions $\pi_\theta^{(k)}(i)$ for $0 \leq i \leq N$ and $k = 1, 2, 3$, of the M/M/1/N queue versus the perturbation parameter $\Delta \in [-\delta, \delta]$, where $\delta = 0.1$. As expected, the relative error on the stationary distributions decreases as the perturbation parameter h decreases.

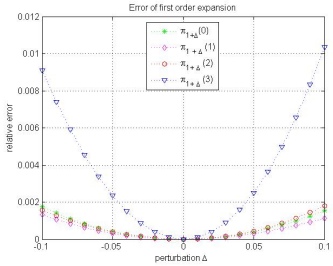


Fig. 1. The relative error in computing $\pi_{1+\Delta}$ by Taylor series of 1st order.

series coefficients are given in terms of the deviation matrix corresponding to the embedded Markov chain. We have presented some numerical examples

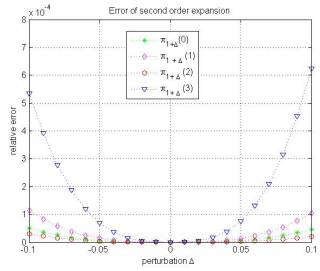


Fig. 2. The relative error in computing $\pi_{1+\Delta}$ by Taylor series of 2nd order.

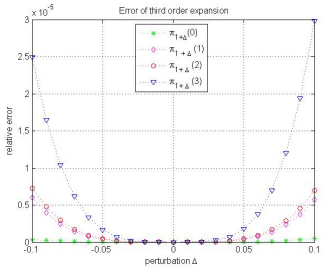


Fig. 3. The relative error in computing $\pi_{1+\Delta}$ by Taylor series of 3rd order.

that illustrate our numerical approach. In fact, the convergence rate of the Taylor series is such that already a Taylor polynomial of degree 2 or 3 yields good numerical results. As part of future work, we will further investigate the multi-server queues with vacations. We will also further provide a simplified and easily computable expression bounding the remainder of the Taylor series and, thereby provide an algorithmic way of deciding which order of the Taylor polynomial is su?cient to achieve a desired precision of the approximation Abbas, Heidergott and Aissani (2013).

4 Conclusion

This paper has developed a numerical, method to analyze the effect of the perturbation of the negative customers process in the performance measures of the queuing model considered (Stationary distribution), our numerical investigation are based on the Taylor series expansion; see [7], where the Taylor series coefficients are given in terms of the deviation matrix corresponding to the embedded Markov chain. Therefore, we have presented different examples that illustrate our numerical approach, and as illustrated by the numerical examples the convergence rate of the Taylor series is such that already a Taylor polynomial of degree 2 or 3 yields good numerical results, we will further investigate the multi-server queues with vacations. We will also further provide a simplified and easily computable expression bounding the remainder of the Taylor series and, thereby provide an algorithmic way of deciding which or-

der of the Taylor polynomial is sufficient to achieve a desired precision of the approximation [1].

References

1. Abbas, K., Heidergott, B. and Aïssani, D. A Functional Approximation for the M/G/1/N Queue. Discrete Event Dynamic Systems , pages 93–104, 2013.
2. Bayer, N., Boxma, O.J. WienerHopf analysis of an M/G/1 queue with negative customers and of a related class of random walks, Queueing Syst.23 (1996) 301316.
3. Boucherie, R.J. , Boxma, O.J. The workload in the M/G/1 queue with work removal, Probab. Eng.Inform.Sci.10 (1995) 261277.
4. Cao, X.R. Realization Probabilities: The Dynamics of Queueing Systems, Springer-Verlag, New York, 1994.
5. Gelenbe,E. Product-form queueing networks with negative and positive customers, J.Appl. Probab.28 (1991) 656663.
6. Boucherie, R.J. , Boxma, O.J. The workload in the M/G/1 queue with work removal, Probab. Eng.Inform.Sci.10 (1995) 261277.
7. Heidergott, B., Hordijk, A.: Taylor series expansions for stationary Markov chains. Advances in Applied Probability 35, 1046–1070 (2003)
8. Jain, G., Sigman, K. A pollaczek-Khintchine formula for M/G/1 queues with disasters, J.Appl. Probab.33 (1996) 11911200.
9. Harrison, P.G., Pitel, E. The M/G/1 queue with negative customers, Adv. Appl. Probab. 28 (1996) 540566.
10. Harrison, P.G., Patel, N.M. Pitel, E. Reliability modelling using G-queues, Eur. J. Oper. Res. 126 (2000) 273287.
11. Yang, W.S., Chae, K.C. A note on the GI/M/1 queue with poisson negative arrivals, J. Appl. Probab.38 (2001) 10811085.

Survival Analysis on clinical trial in Monte Carlo simulation

Megdouda Ourbih-Tari[1] and Mahdia Azzal[2]

[1] Laboratoire de mathématiques Appliquées, Faculté des Sciences Exactes, Université de Bejaia, 06000, Algeria
(E-mail: ourbihmeg@gmail.com)
[2] Laboratoire de mathématiques Appliquées, Faculté des Sciences Exactes, Université de Bejaia, 06000, Algeria
(E-mail: azzal.mahdia@gmail.com)

Abstract. This paper focuses on the statistical comparison of Refined Descriptive Sampling (RDS) method taken from the literature for Monte Carlo simulation process and the well known Simple Random Sampling (SRS) method. For this purpose, a lifetime model whose observations are right-censored of random type is used to perform a nonparametric estimation of the survival function. A real application was conducted and its parameters were computed to be used as a basis for comparison. We estimate the survival function by the product limit estimation using both sampling methods and a real population. The obtained results prove the efficiency of RDS over SRS when entries are supposed following an exponential distribution and when entries are generated from the population distribution.

Keywords: Nonparametric Estimation, Clinical Trial, Sampling, Monte Carlo, Simulation.

1 Introduction

The stochastic lifetime models have usually been used in scientific areas such as biomedical, medicine (Couchoud and Villar [8]) and engineering to measure for example, the survival duration of patients and electrical components. Survival analysis (Akbar et al. [1]) is any analysis of the occurrence of an event over time. An important and most common form of survival analysis in lifetime models is the presence of censored data. Typically, survival data are not fully observed, but rather are censored. Among the censored observations, we can find different censoring, right ,left and by intervals, each, with different types of censors whose censor of type I, type II and random type. The first nonparametric method is called actuarial method and is intended for the survival function' estimation in the presence of random right censored data. The disadvantage of this method is that it does not view the individual lifetimes. More details about the actuarial method can be found in (Ayiomamitis [4]) and (Lakhal et al. [16]).

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

Kaplan and Meier generalized the concept of empirical survival function in the presence of random right-censored data by proposing an estimator. The Kaplan-Meier (KM) or product limit estimator is the limit of the actuarial estimator when intervals are taken so small that only at most one distinct observation occurs within an interval. More details about the estimate of KM can be found in (Chen et al. [7]) and (Mauro [17]).

Several authors (Couchoud and Villar [8]) and (Gill [11]) have studied the estimator of KM and have shown that this estimator is unbiased except in the tails of lifetime distributions.

In case of large data, the actuarial method can give fairly reliable results. Otherwise it is recommended to choose the method of KM which uses all available information. The actuarial method requires less computation than the KM method, but is less accurate. In most situations encountered today in clinical research, the KM seems more appropriate and accurate. See more work concerning nonparametric estimations in (Blackstone [5]) and (Giorgi et al. [12]) and in the presence of other censored data in (Kiein and Moeschberger [14]) and (Kleinbaum and Klein[15]).

In this paper, we are interested in the efficiency of nonparametric estimation of the survival function and in its better representation through the KM estimator in lifetime models whose observations are just random right-censored using Monte Carlo (MC) simulation. We focus on this sort of data since it is the most common censoring scheme. It is important to study the efficiency of the KM estimator since generally the application area of this estimator is sensitive.

Then, we need to generate random samples using the best sampling scheme, according to the distributions of the input random variables of the lifetime model.

In simulation, a lot of statistical methods for choosing input random samples exist. The simplest and well known is Simple Random Sampling (SRS) (Fishman [9]) which is used to generate identically independent distributed (iid) random variables using some Random Number Generators in the simulation of systems containing stochastic situations (Andersson [3]) such as lifetime models. But SRS estimates obtained through simulation vary between different runs and this variation is due to its set and sequence effect, so it is an imprecise procedure. Consequently, different sampling methods have been proposed as alternatives to MC methods. These sampling methods used in simulation studies are in fact numerous. Among them, we mention Descriptive sampling (DS) (Saliby [20]) and Refined Descriptive Sampling (RDS) (Tari and Dahmani [25]). RDS was proposed to make DS safe, efficient and convenient. It was applied to simulate several statistical models (Tari and Dahmani[23]), (Tari and Dahmani [24]) and was shown to be unbiased and more accurate than SRS according to the variance criteria with respect to a class of estimators (Ourbih and Guebli [19]). Ourbih-Tari and Aloui, 2009 have shown that RDS outperforms all other sampling methods given in the literature (Ourbih and Aloui [18]). Then, we propose the use of RDS method to better estimate and represent the survival function through the Kaplan Meier estimator in lifetime models whose observations are just random right-censored.

Section 2 is devoted to describing lifetime models, survival function and KM estimator with its characteristics. By then, RDS method is given in section 3. In section 4, existing theoretical results about RDS, were verified with different simulation experiments using both sampling methods (Ourbih and Guebli [19]) and (Tari and Dahmani [25]) for the generation of random variables when those are supposed following exponential distribution. Finally, section 5 deals with a real application in a medical field concerned by a population of patients where data are obtained from the register of Algeria Pierre and Marie Curie Center. The population distribution was studied. Then, from this population, simple random samples and refined descriptive samples are generated and used to estimate the KM survival function and the variance of KM estimator. All obtained results show that RDS method better estimates and represents the survival function through the KM estimator in lifetime models. We finish this paper by a conclusion.

2 BACKGROUND

2.1 Lifetime models and survival function

The lifetime models are appropriated since the studied phenomenon is represented via positive random variables. Let us consider a variable T defined by the duration of a period separating two events with a cumulative function F and a survival function S. The latter is one of the most important characteristic which is represented by the probability that the variable T overtake a given value t given by $S(t) = P(T > t) = 1 - F(t)$. It's a decreasing continuous function satisfying the limits conditions:$S(0) = 1$ and $S(+\infty) = 0$. Survival analysis is any analysis of the occurrence of an event over time. An important form of survival analysis is the presence of censored data that are observations for which the exact value of an event is not always known. However, we still have partial information for setting a lower bound, an upper bound or two terminals called respectively right, left and interval censoring.The term of survival duration is employed generally to indicate the time elapsing till a particular event occurs. Consider $(T_1, T_2, ...T_n)$ a sample of size n of the lifetime T. The problem of lifetime without censoring data is naturally estimated by the empirical survival function $\widehat{S}_{emp}$ defined for all $t \geq 0$ by

$$\widehat{S}_{emp}(t) = \frac{1}{n} 1_{\{T_i > t\}}. \tag{1}$$

But in the presence of random right censoring, the empirical survival function of the variable T is no longer available. Indeed, it depends on the observations T_i that are not observed. In order to estimate the distribution of T in a censored case, it is necessary to propose an estimator of the survival function that have properties similar to those of the empirical survival function used in the absence of censor.

2.2 Kaplan Meier estimator

Suppose the variable of interest, T is randomly right censored by a censoring variable C which is often assumed to be independent of T. In this case,

the lifetime models will be represented by a pair of variables (Y, d) where its observations are given by:

$$Y_i = \min(T_i, C_i) \ \ and \ \ d_i = 1_{\{T_i \leq C_i\}} \tag{2}$$

Let $Y_{(1)}, Y_{(2)}, ...Y_{(n)}$ be the associated order's statistics to $Y_1, Y_2, ...Y_n$ and $d'_1, d'_2, ...d'_n$ be the corresponding ordered indicators.

The product limit estimation allows the generalization of the empirical survival function concept within the presence of censored data. It is based on the following remark:

If $t\prime < t$ the survival probability beyond the date t is equal to the following product

$$S(t) = P(T > t, \ T > t')S(t') \tag{3}$$

i.e, the computed survival at a point $t^{'}$ is preserved over the interval until the next computed point t.

If we select an other date $t'' < t'$, we will also have

$$S(t') = P(T > t', \ T > t'')S(t'') \tag{4}$$

and so on...

In the case where the time t is affected by the occurrence of an event, we estimate quantities of the form:

$$p_i = P(T > t_i / T > t_{i-1}) \tag{5}$$

where p_i is the probability of surviving during the interval of time $I_i =]t_{i-1}, t_i]$ if the i^{th} individual is living at the beginning of this interval determined by the dates of observed events.

Denote by R_i the number of living individuals that are therefore at risk of dying just before the time t_i, M_i the number of dead individuals at time t_i and $q_i = 1 - p_i$ is the probability of dying during the interval I_i knowing that one was alive at the beginning of this interval. Then the natural estimator of q_i is $\widehat{q_i} = \frac{M_i}{R_i}$.

In case the individuals are not ex-aequo and if $d'_i = 1$ i. e there is a death at time t_i, therefore $M_i = 1$, then $\widehat{p_i} = 1 - \frac{1}{R_i}$. Otherwise, if $d'_i = 0$ i. e there is a censor at time t_i, therefore $M_i = 0$, then $\widehat{p_i} = 1$.

Given that the number of possible outcomes is $R_i = n - i + 1$, the Kaplan-Meier estimator is in this case

$$\widehat{S}_{KM}(t) = \prod_{Y_{(i)} \leq t} (1 - \frac{1}{n - i + 1})^{d'_i} \tag{6}$$

In the most general case (ex-aequo are allowed) where one can observe a number of M_i outputs greater than 1 at each time t, the Kaplan-Meier survival function (Kaplan and Meier [13]) becomes:

$$\widehat{S}_{KM} = \widehat{S}_{KM}(t) = \prod_{Y_{(i)} \leq t} (1 - \frac{M_i}{R_i}) \tag{7}$$

2.3 Characteristics of Kaplan Meier estimator

A first rigorous result may be found in (Gill [11]) given by

$$E[\widehat{F}_{KM}(t) - F(t)] = Bias[\widehat{F}_{KM}(t)] \leq 0 \tag{8}$$

i.e., $\widehat{F}_{KM}(t)$ is always biased downwards.

Subsequently, it was possible to determine lower bounds of increasingly accurate through time t (Mauro [17], Zhou [26], Stute and Wang [21], Stute [22]). For an explicit expression of this bias, we cite the result of (Fleming and Harrington [10]).

For all t such that $0 \leq F(t) < 1$

$$\begin{aligned} Bias_n &= E[\widehat{F}_{KM}(t) - F(t)] \\ &= -E[1_{\left\{Y_{(n)}<t\right\}} \frac{\left[1 - \widehat{F}_{KM}(Y_{(n)})\right]\left[F(t) - F(Y_{(n)})\right]}{1 - F(Y_{(n)})}] \end{aligned} \tag{9}$$

The KM estimator is known to be unbiased in the case where $t \leq Y_{(n)}$ or if there is no censoring at all. Under censorship, when the variable of interest is at risk of being censored by a smaller one (distribution tails), estimation of functionals of F is typically negatively biased which is even greater in absolute value as the time t is large.

The variance of KM estimator for the survival function is unknown and is estimated by the following Greenwood formula which is well-known and commonly used:

$$Var_{GW}(t) = \widehat{Var(\widehat{S}_{KM}}(t)) = \widehat{S}^2_{KM}(t) \sum_{i/Y_{(i)}<t} \frac{M_i}{(R_i - M_i)\, R_i} \tag{10}$$

3 REFINED DESCRIPTIVE SAMPLING

RDS method (Tari and Dahmani [25]) was proposed as an alternative approach to Monte Carlo Simulation. It is based on a block that distributes samples of size, prime numbers randomly generated as required by the simulation. We stop the process when the simulation terminates. Let P be a randomly generated prime number using MATLAB software as given in the following procedure:

1. M=input('Give M')
2. y=0;
3. while y==0
4. P=round(M*rand(1));
5. y=isprime(P);
6. end
7. P

where M is a given large value fixed by the user that the generated prime number does not exceed.

Let us suppose that a simulation run defined by m sub-runs terminates when m prime numbers are used. Refined descriptive Samples values for the input random variable X having F^{-1} as the inverse cumulative distribution are generated using the following formulae:

$$x_i = F^{-1}(v_i) \quad for\ i = 1, 2, ..., P_q, \quad q = 1, 2, ..., m \tag{11}$$

where $v_i, i = 1, 2, ..., P$ are obtained by a non-random selection as

$$v_i = \frac{i - 0.5}{P}, \quad i = 1, 2, ..., P \tag{12}$$

These numbers are then randomized using MATLAB software as given in the following procedure and the obtained numbers are called V_i.

1. for i=1:P
2. rs=(i-0.5);
3. v=rs/P;
4. V(i)=v;
5. end
6. for i=1:P
7. p_1=round(P*rand(1));
8. while($p_1 < i$) | (p_1 >P)
9. p_1=round(P*rand(1));
10. end
11. va=V(p_1);
12. V(p_1)=V(i);
13. V(i)=va;
14. end
15. V

The refined descriptive numbers are the successive m sub-sets V^q, $q = 1, 2, ..., m$ of component $V_i, i = 1, 2, ..., P_q$.

4 SIMULATION EXPERIMENT

The distributions of both independent input variables T and C are assumed to be known as an exponential distributions of parameter λ and δ respectively, while the output variable' distribution is unknown and defined by the survival function. We designed a simulation program using MATLAB software to check the existing theoretical results about RDS. For this purpose, KM estimator and Greenwood variance are selected as performance measures to study the problem of nonparametric estimation of the survival function. A variety of simulation experiments for different N replicated runs are carried out on the lifetime model. A random samples of size n are generated in each run.

4.1 Algorithm estimating the survival function using SRS method

Initialization of the experiment.
At the beginning of each experiment, initialize the parameter N
For $j = 1toN$

Initialization of a run.
At the beginning of each run, initialize the parameters λ, δ and the sample size n

- Generate a stream of random numbers u_i from uniform distribution [0,1[using the rand function of MATLAB
- Compute observations $T_i = -\frac{1}{\lambda}\ln(1-u_i)$, we obtain the sample $(T_1, T_2, ...T_n)$
- Generate another stream of random numbers $u_i^{'}$ from uniform distribution $[0, 1[$ using the rand function of MATLAB
- Compute observations $C_i = -\frac{1}{\delta}\ln(1-u_i^{'})$, we obtain the sample $(C_1, C_2, ...C_n)$.
- Compute the pair (Y_i, d_i), $i = 1, ..., n$ such as: $Y_i = min(T_i, C_i)$ and $d_i = 1_{\{T_i \leq C_i\}}$
- Classify Y_i, $i = 1, ..., n$ in an increasing order to obtain the order statistics $Y_{(1)}, Y_{(2)}, ...Y_{(n)}$ and their corresponding ordered indicators $d'_1, d'_2, ...d'_n$.
- Compute the number M_i of death at $Y_{(i)}$.
- Compute the number R_i of individuals with risk just before $Y_{(i)}$.
- Compute $\widehat{S_{KM}}(i)$ for each observation $Y_{(i)}$ using formulae 7 for $t = i$.
- Compute Var_{GW} for each observation $Y_{(i)}$ with the Greenwood formulae 10 for $t = i$.
- Compute $Mean_n$, the average of all $\widehat{S_{KM}}(i)$ for each run using formulae 13.

$$Mean_n = Mean(\widehat{S_{KM}(i)}) = \frac{1}{n}\sum_{i=1}^{n}(\widehat{S_{KM}(i)}) \tag{13}$$

- Compute $MeanGW_n$, the average of all Var_{GW} for each run using formulae 14.

$$MeanGW_n = Mean(\widehat{Var_{GW}}) = \frac{1}{n}\sum_{i=1}^{n}(\widehat{Var_{GW}}(i)) \tag{14}$$

Compute the mean and variance of $Mean_n$ and $MeanGW_n$ for each experiment using the sample mean and sample variance.

4.2 Algorithm estimating the survival function using RDS method

The same principle and the same parameters as SRS are used to simulate the survival function by RDS.

Initialization of the experiment.
At the beginning of each experiment, initialize the parameter N
For $j = 1toN$

Initialization of a run.
At the beginning of each run, initialize the parameters λ, δ and the sample size $n = \sum_{q=1}^{m} P_q$.

(a_1) Sampling without replacement during the sub-run
(a) Give an integer M
(b) Generate randomly a prime number P as shown in section 2.
(c) Generate an array V as shown in section 2.
(d) Like in SRS, compute the T_i observations, we obtain the sample $(T_1, T_2, ...T_P)$.
(e) Generate another random sequence V' as shown in section 2.
(f) Like in SRS, compute the C_i observations, we obtain the sample $(C_1, C_2, ...C_P)$.
(g) Like in SRS, compute the different estimates
(h) Collect the sub-results from the sub-run and go to (a_1).
(i) Collect results at the end of different runs by computing the average of the different sub-results. We deduce the sample size $n = \sum_{q=1}^{m} P_q$ which will be used for SRS.
(j) collect the final results such as the mean and variance of different estimates.

4.3 Experiment and Results

In this subsection we estimate the unknown survival function S. Three experiments with $N = 10, 50$ and 100 are carried out. We chose $n = 10000$ for each run in different experiments. We summarized each experiment by computing the mean and variance of both performance measures using both MC methods.

The designed software generates the input variables with $\lambda = 1$ and $\delta = 1/2$ using RDS method. The obtained results are given in table 1.

N	estimators	RDS		SRS		Var Reduce
		Mean	Var	Mean	Var	
10	$\widehat{S_{KM}}$	0.6793	0.0032	0.5997	0.0053	39.62%
	$Var_{GW}(t)$	0.0064	0.0017	0.0241	0.0022	22.73%
50	$\widehat{S_{KM}}$	0.7177	0.0013	0.5999	0.0042	69.05%
	$Var_{GW}(t)$	0.0216	0.0056	0.0282	0.0067	16.42%
100	$\widehat{S_{KM}}$	0.6687	0.0016	0.5993	0.0032	50%
	$Var_{GW}(t)$	0.0211	0.0011	0.0291	0.0020	45%

Table 1. Simulation results using SRS and RDS for different replication runs

Table 1 demonstrates that RDS reduces both estimates $\widehat{S_{KM}}$ and $Var_{GW}(t)$ for different experiments with a significant variance reduction. Then, RDS estimates efficiently the survival function in the presence of censored data.

$\widehat{S_{KM}}$ estimates efficiently the survival function S by the Kaplan Meier estimator $\widehat{S_{KM}}$ whatever the variance used. So, we can say that RDS method is more accurate than SRS according to survival analysis.

The curve of Kaplan-Meier estimator is a step curve to which the range of the steps varies from one level to another. The illustration of the sample distribution of Kaplan-Meier estimator obtained by SRS is given in figure 1 and its sample distribution of Greenwood variance obtained by both sampling methods are given in figure 2, both can be found in the appendix.

5 A CASE STUDY IN MEDICAL FIELD

In order to define a lifetime random variable, we need to define the followings:

Date origin: which is the starting point from where the patient is observed(e.g. diagnosis, date of commencement of treatment ...).

The date of last news: This is when we got news of the individual for the last time. Thus, it may be the date of death or the date of occurrence of the event (the recovery, the first relapse...). But it can also be the date of the last visit if the individual is lost of sight.

State of last news: These states are dead, lost of sight, recovered and not recovered.

Monitoring time: This is the difference between the date of the last news and date of origin.

Date point: This is the date beyond which we will not consider the information for which we seek to know the status of each individual.

Lost of sight: About which we do not know the state at the date point

Rewind: This is the time elapsed between the date origin and the date point.

Time participation Y_i: If the latest news is prior to the date of point, Y_i is the difference between the date of last news and date of origin (called the monitoring time). If the latest news is after the date of point, Y_i is the difference between the time point and time of origin (called rewind).

5.1 Description of the survival study: A clinical trial

The collected data from the Center of Pierre and Marie Curie (CPMC) which constitutes our population are used to determine the time participation Y defined as a continuous quantitative character, the recovery daily length of patients having undergone a "Grafts Marrow" surgical operation from the 1^{st} January 2011 until 31^{st} December 2011 whose Y_i observation is the recovery length of the i^{th} patient. The mean $(\overline{Y})$, the median (Me) and the variance (Var) of the population were computed such as $\overline{Y} = 65.17$ days, $M = 35.5$ days and $Var = 0.0056$. In the frame of the analysis of a clinical trial, some data can be censored. Yet we deduce a random variable indicator d_i of the patient recovery from the collected data, valued to 1 if the patient is recovered and 0 otherwise i.e censored data. The survival function was estimated by $\widehat{S_{KM}(t)}$ and its variances as shown in table 3.

5.2 Example

In this subsection, we show how to compute Y_i and d_i when the time point, time origin, latest news, and the status are available on a small example of size 7 given in table 2.

In this subsection, we show also how to compute M_i, R_i, $\widehat{S_{KM}(t)}$ and $Var_{GW}(\widehat{S}_{KM}(t))$ when the Y_i and d_i are available on a small example of size 17 given in table 3.

i	Time origin	Last news	Status	Y_i	d_i
1	01/01/2011	22/01/2011	Recovered	21	1
2	18/01/2011	22/01/2011	Recovered	04	1
3	09/01/2011	01/03/2011	Recovered	51	1
4	16/11/2011	21/12/2011	Not Recovered	35	0
5	10/10/2011	14/12/2011	lost of sight	65	0
6	19/10/2011	20/10/2011	Not Recovered	1	0
7	03/09/2011	11/11/2011	Dead	69	0

Table 2. Example of how to compute Y_i and d_i using time point $31december2011$

Y_i	d_i	M_i	R_i	$\widehat{S_{KM}(t)}$	$Var_{GW}(\widehat{S}_{KM}(t))$
2	0	0	17	1	0
3	1	1	16	1	0
3	1	1	15	0.9375	0.0037
3	0	0	14	0.8750	0.0068
5	0	0	13	0.8750	0.0068
6	0	0	12	0.8750	0.0068
8	0	0	11	0.8750	0.0068
9	0	0	10	0.8750	0.0068
10	1	1	9	0.8750	0.0068
12	0	0	8	0.7778	0.0138
18	1	1	7	0.7778	0.0138
18	1	1	6	0.6667	0.0207
22	1	1	5	0.5556	0.0247
34	0	0	4	0.4444	0.0257
38	1	1	3	0.4444	0.0257
43	1	1	2	0.2963	0.0260
58	1	1	1	0.1481	0.0175

Table 3. Example of how to compute M_i, R_i, $\widehat{S_{KM}(t)}$ and $Var(\widehat{S}_{KM}(t))$.

Table 3 shows the Kaplan Meier estimation for some recovery length of patients expressed by day. For example, the estimation of the probability that a patient recovers, after 12 days is 77.78% and its Greenwood variance is 0.0138. Consequently, the probability that a patient dies during the first 12 days is 22.22%.

5.3 The frequency distribution of the real data

From the 318 patients population available, we define the frequency distribution of Y given in table 4.

Class	Y_i	n_i	f_i	F_i
[0; 16[	8	88	0.2767	0.2767
[16; 32[	24	56	0.1761	0.4528
[32; 48[	40	37	0.1164	0.5692
[48; 64[	56	17	0.0535	0.6227
[64; 80[	72	31	0.0975	0.7202
[80; 96[	88	20	0.0629	0.7831
[96; 112[	104	14	0.0440	0.8271
[112; 128[	120	11	0.0346	0.8617
[128; 144[	136	5	0.0157	0.8774
[144; 160[	152	5	0.0157	0.8931
[160; 176[	168	2	0.0063	0.8994
[176; 192[	184	3	0.0094	0.9088
[192; 208[	200	6	0.0189	0.9277
[208; 224[	216	2	0.0063	0.9340
[224; 240[	232	4	0.0126	0.9466
[240; 256[	248	1	0.0031	0.9497
[256; 272[	264	3	0.0094	0.9591
[272; 288[	280	3	0.0094	0.9685
[288; 304[	296	5	0.0157	0.9842
[304; 320[	312	2	0.0063	0.9905
[320; 336[	328	2	0.0063	0.9968
[336; 352[	344	1	0.0032	1.0000

Table 4. The frequency distribution of Y

5.4 Sampling scheme

In this subsection, we give a sampling scheme by the Top Hat method which will be used to sample from the frequency distribution of Y given in subsection 5.3 and from the status d distribution. The latter is a Bernoulli distribution of parameter $b = 0.5157$, where b is the probability that the patient recovered. These sampling schemes are given respectively in tables 5 and 6 where "r.n" stands for u_i obtained by the "rand" function of MATLAB when using SRS and by the number V_i when using RDS.

5.5 Experiments and Results

To estimate the survival function using a nonparametric estimation we generate artificial samples with RDS and SRS using algorithms given in subsection 4.1 and 4.2 but the exponential distribution is replaced by a real distribution. The population was studied and a sampling scheme was generated. By then, artificial samples of size n of each variable Y and d as simple random samples and refined descriptive samples were generated. Then, two samples of size n of KM estimator and Greenwood variance were generated.

In this sub-section, a comparison between both sampling methods and the studied population is carried out to show that RDS better estimate the survival function in a non parametric estimation.

Y_i	8	24	40	56	72
r:n	[0, 0.2767 [	[0.2767,0.4528[	[0.4528,0.5692[	[0.5692,0.6227[	[0.6227,0.7202[
Y_i	88	104	120	136	152
r:n	[0.7202,0.7831[	[0.7831,0.8271[	[0.8271,0.8617[	[0.8617,0.8774[	[0.8774,0.8931[
Y_i	168	184	200	216	232
r:n	[0.8931,0.8994[	[0.8994,0.9088[	[0.9088,0.9277[	[0.9277,0.9340[	[0.9340,0.9466[
Y_i	248	264	280	296	312
r:n	[0.9466,0.9497[	[0.9497,0.9591[	[0.9591,0.9685[	[0.9685,0.9842[	[0.9842,0.9905[
Y_i	328	344			
r:n	[0.9905,0.9968[	[0.9968,1[			

Table 5. The sampling scheme of the variable Y

d_i	0	1
r:n	[0,0.5157[	[0.5157,1[

Table 6. The sampling scheme of the variable d

N	estimators	RDS		SRS		Var Reduce
		Mean	Var	Mean	Var	
10	$\widehat{S_{KM}}$	0.7697	0.0017	0.6617	0.0029	41.38%
	$Var_{GW}(t)$	0.0259	0.0026	0.0214	0.0032	18.75%
50	$\widehat{S_{KM}}$	0.6726	0.0009	0.6596	0.0024	62.5%
	$Var_{GW}(t)$	0.0339	0.0077	0.0475	0.0093	17.20%
100	$\widehat{S_{KM}}$	0.6668	0.0008	0.6587	0.0014	42.86%
	$Var_{GW}(t)$	0.0298	0.0062	0.0551	0.0079	21.52%

Table 7. Simulation results of the survival analysis comparing RDS with SRS

Table 7 suggests also that RDS reduces both estimates $\widehat{S_{KM}}$ and $Var_{GW}(t)$ for different experiments with a significant variance reduction. Then, RDS estimates efficiently the survival function in the presence of censored data.

6 CONCLUSION

From the study of the population given in subsection 5.1 , we conclude that the median recovery daily length of patients having undergone a "Grafts Marrow" surgical operation is better than is its mean although the dispersion of data is low.

In the absence of a priori information on the shape of the survival function, we estimated the latter by the non-parametric method of KM. The study on the length recovery allows us to apply one of the classical methods of lifetime estimation to medical domain. By then, the estimation of the length recovery of patients was used to evaluate other characteristics of patients. On the other hand, the application of RDS method on lifetime models reduces the sample

and Greenwood variances of KM estimator and its bias. The KM method can be then adapted to medical data by using the best sampling scheme which is RDS method. We also notice that RDS and SRS KM estimators are closer to the theoretical value using the sampling error criteria. Therefore, these results strongly support the efficiency of RDS over SRS on lifetime models in the presence of censored data.

7 APPENDIX

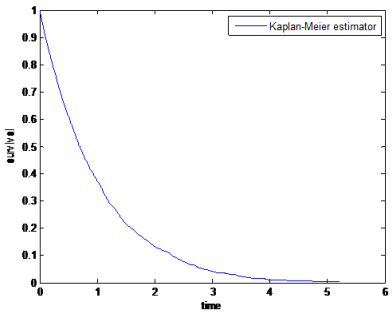


Fig. 1. The Kaplan-Meier Curve of the Survival Function when entries follow exponential distributions using SRS

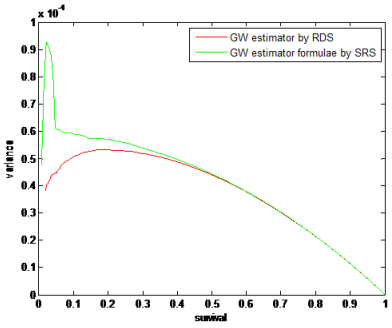


Fig. 2. The Kaplan-Meier Estimator Variance Curve by SRS and RDS when entries follow exponential distributions

References

1. A. Akbar , GR. Pasha and SFH. Naqvi. Properties of Kaplan-Meier estimator: Group comparison of survival curves. European Journal of Scientific Research, 32, 391-397, 2009.
2. A. Aloui and M. Ourbih-Tari. The use of refined descriptive sampling and applications in parallel Monte Carlo simulation. Computing and Informatics, 30, 681-700, 2011.
3. EC. Andersson. Monte Carlo methods for inference in population genetic models, Ph. D thesis, University of Washington, 2001.
4. A. Ayiomamitis. Analysis of survival data using the actuarial life table method. International Journal of Bio-Medical Computing, 19, 109-117, 1986.
5. EH. Blackstone. Actuarial and Kaplan-meier survival analysis: there is a difference. J Thorac Cardiovasc Surg, 118,5, 973-5, 1999.
6. P. Bohmer. Theorie der unabhngigen Wahrscheinlichkeiten Rapports. Mmoires et procs verbaux du septime congrs international d'actuaires, Amsterdam, 2, 327-43, 1912.
7. Y. Chen, M. Hollander and NA. Lagrberg. Small-Sample results for the Kaplan-Meier estimator. Journal of the American Statistical Association, 77,141-144, 1982.

8. C. Couchoud and E. Villar. Bias in survival analysis: The specific case of end-stage renal disease patients. Association Society of nephrologie, 7, 27-31, 2011.
9. GS. Fishman. Monte-Carlo: Concepts, algorithms and applications, Springer-Verlag, New York, 1997.
10. T. Fleming and D. Harrington. Counting processes and survival analysis, Wiley, New York, 1991.
11. RD. Gill. Censoring and stochastic integrals. Statistica Neerlandica, 34, 124, 1980.
12. R. Giorgi, A. Armanet, J. Gouvernet, P. Bonnier and M. Fieschi. Regression models for crude and relative survival: a comparative review. Rev Epidemiol Sante Publique, 53, 4, 409-17, 2005.
13. EL. Kaplan and P. Meier. Nonparametric Estimation from Incomplete Observations. American Statistical Association, 53, 457-481, 1958.
14. J P. Klein and M L. Moeschberger. Survival Analysis: Techniques for Censored and Truncated Data, 2nd ed. Statistics for Biology and Health, Springer, 2003.
15. D G. Kleinbaum and M. Klein. Survival Analysis: A Self-Learning Text, 3rd ed. Statistics for Biology and Health, Springer, 2012.
16. L.Lakhal-Chaieb, B. Abdous and Th. Duchesne. Nonparametric estimation of the conditional survival function for bivariate failure times. Canadian Journal of Statistics, 41, 3,439452, 2013.
17. D. Mauro. A combinatoric approach to the Kaplan-Meier estimator. Ann. Statist, 13, 142-149, 1985.
18. M. Ourbih-Tari and A. Aloui. Sampling methods and parallelism into Monte Carlo simulation. Journal of Statistics: Advances in Theory and Applications, 1, 169-192, 2009.
19. M. Ourbih-Tari and S. Guebli. Comparing two sampling methods in Monte Carlo simulation, In Proccedings of the 2010 European Simulation and Modelling Conference Hasselt, Belgique, 27-31, 2010.
20. E. Saliby. Descriptive Sampling: A better approach to Monte Carlo simulation. Journal of the Operational Research Society, 41, 1133-1142, 1990.
21. W. Stute and JL. Wang. The strong law under random censorship. Ann. Statist, 21, 1591-1607, 1993.
22. Stute W, The Bias of Kaplan-Meier Integrals, Scandinavian Journal of Statistics 21, 475-484 (1994)
23. M. Tari and A. Dahmani. Flowshop simulator using different sampling methods. Operational Research: An International Journal, 5, 261-272, 2005.
24. M. Tari and A. Dahmani. The three phase discrete event simulation using some sampling methods. International Journal of Applied Mathematics and Statistics, 3, 37-48, 2005.
25. M. Tari and A. Dahmani. Refined descriptive sampling: A better approach to Monte Carlo simulation. Simulation Modelling Practice and Theory, 14, 143-160, 2006.
26. M. Zhou. Two-sided bias bound of the Kaplan-Meier estimator. Probab. Theory Relat. Fields, 79, 165-173, 1988.

Youth mortality by violence in the semiarid region of Brazil

Neir Antunes Paes[1], and Everlane Suane de Araújo Silva[2]

[1] Postgraduate Programme in Decision Modelling and Health of the Department of Statistics of the Federal University of Paraíba, Cidade Universitária, João Pessoa, PB, Brazil.
(Email: antunes@de.ufpb.br)

[2] PhD Candidate in Postgraduate Programme in Decision Modelling and Health of the Department of Statistics of the Federal University of Paraíba, Brazil
(Email: everlanesuane@hotmail.com)

Abstract: The Brazilian semiarid region is the world's largest in terms of density population and extension with 22 million inhabitants in 2010. An ecological study addressing the mortality by aggression for 137 microregions of the Brazilian semiarid region to young males, in the year 2010, was performed. Two indicators were calculated for each microregion: standardized mortality rates by violence and an indicator named Reducible Gaps of Mortality. We investigated the correlation between standardized mortality rates and a set of 154 indicators that express living conditions. 18 of them were considered as significant. By means of the multivariate technique - Principal Component Analysis - the construction of a Synthetic Indicator was performed, which was categorized in four strata reflecting different living conditions and mortality rates. The results showed that microregions with high values of mortality rates by aggressions were present in all strata, thus contradicting some studies linking high rates of mortality due to aggression to low condition of life. This paradox allowed us to raise issues to identify the most vulnerable regions, and contribute to the decision making process to combat mortality by violence of the Brazilian semiarid population.
Keywords: Mortality by violence. Living conditions. Young.

1 Introduction

The semiarid region is divided into 137 microregions and, in turn, distributed in 1,135 municipalities in the geographic space of 9 (from 27) units of the federation. It is the world's largest in terms of population density and extent, with some 22.6 million people in 2010, and represent approximately 12% of the population. The indicators show that the level of development is considered as the most precarious of the country, and flattens to many African countries less developed, Brasil [1].

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

Currently, public security gains prominence among the issues that most concern the Brazilian society, and takes along with health and education, the priority for the attention of public authorities according to Waiselfisz [2]. Today the external causes in Brazil are the third most frequent cause of death, and predominate in the age group 10-24 years for males, Brasil [3]. Paradoxically, has occupied a limited place in the themes discussed by public administrators to implement policies to promote the common welfare in this region, UNICEF [4]. It is a problem to be overcome, the lack of a synthetic statistical indicator that expresses the association between living conditions and mortality from violence. In seeking to answer this question, we have as main objective the construction of a synthetic indicator for young men aged 10-24 years in the Brazilian semiarid region.

2 Methods

This is an ecological study of census basis that covers nine states. The geographical units of study are the 137 microregions that form the semiarid region in 2010. The following demographic and socioeconomic indicators were used: population, income, education and health. Data on deaths of residents were obtained from the website of the Ministry of Health through the Brazil Mortality Information System [5]. Indicators related to cause of death from aggression of the youths and indicators of living conditions were calculated: a) Standardized Mortality Rate (SMR); b) Reducible Gaps of Mortality (RGM), which means how much a region missed in preventing a particular cause of mortality compared with the SMR of a reference region (Here we considered the median of the microregions) c) 154 indicators on living conditions (health, education and income) were selected, which were obtained from Brazil Atlas of Human Development [6] produced by the United Nations Development Programme .

In order to measure the degree and direction of the linear association between the variables (living conditions and mortality rates by aggression) the Pearson correlation coefficient was calculated, and was built a correlation matrix on them, adopting a level of significance of $\alpha = 0.05$.

The multivariate technique Factor Analysis (FA) was used for the calculation of a Synthetic Indicator corresponding to each microregion (Hair et al., [7]. The number of factors extracted was obtained by the criterion of latent roots (eigenvalues) greater than one. The method of orthogonal Varimax rotation was used to facilitate the reading of the factors and simplify the columns of the matrix factors. The measures used were the adequacy of Kaiser-Meyer-Olkin (KMO) and Bartlett Sphericity Test, which indicate the degree of susceptibility or adjustment to the FA. Thus, one can evaluate that the confidence level of the data for the multivariate method of FA was successfully employed. Another measure of fitness used consisted of verifying the communality of the variables with values ≥ 0.50 as having a sufficient explanation. Thus, variables were identified and grouped into factors. A Synthetic Indicator was generated by the linear combination of the variables and coefficients generated by principal component. Therefore, the

Synthetic Indicator, which summarizes the living conditions and mortality by aggression for the semiarid microregions was given by:

$$Synthetic\ Indicator\ d_k = w_{1k}\, v_{1k} + w_{2k}\, v_{2k} + \ldots + w_{pk}\, v_{pk}$$

Where *k* represents the number of microregions (137), *p* is the number of variables (v) selected by the correlation test (18), and (*w*) the weights or coefficients generated by the main component.
After obtaining the values of the Synthetic Indicator, which represent living conditions together with mortality by aggressions, the microregions were sorted in descending order into quartiles, corresponding to four strata. As a way to show a correspondence of our indicator with a very known Synthetic Indicator (Human Development Index - HDI), to each stratum was assigned a combination that associates the degree of HDI with a degree of violence: (I) HDI medium with high degree of violence, (II) HDI medium with low degree of violence; (III) HDI low with high degree of violence (IV) HDI low with low degree of violence. This made it possible to identify the best and worst microregions in terms of living conditions associated with different degrees of deaths from violence.

3 Results and Discussion

The figures show that 53% of the municipalities in the semiarid region showed a degree of urbanization higher than 50%, much lower than the rest of the country, reaching up to about 85%. With a small urban predominance, the male mortality by aggressions exceeded significantly the female mortality in 2010. The mortality by aggressions on the total deaths for the young men was very high reaching a percentage of 62.1%. For women, this percentage reached a maximum value of 33.4%. Among men, 72 of the 37 microregions, ie, 52.2% of them had a percentage of deaths by aggressions higher than 25%. Similarly, high rates of standardized mortality were found. The rates for men ranged between 3.6 and 63.3 deaths/100.000 men, while for women the range was between 1.3 and 21.5 deaths/100.000 women.
The values of RGM suggest that the states of Ceará, Pernambuco and Alagoas stood out for having the majority of microregions in the most critical situation, with positive values, compared to the median value. This is indicative of a greater need for attention to reducing the incidence of these deaths. The States of Paraiba, Sergipe and Bahia presented the values of the microregions in two situations: below and higher in relation to the SMR of the microrregion considered as reference. The states of Rio Grande do Norte and Minas Gerais have had most of their microregions in a more favorable situation with negative values. The Piauí State, was the only one with SMR below (negative) the reference value.
The correlation matrix among the 154 variables used initially to express the conditions of life and SMR by aggression, revealed that only 18 of them were statistically significant ($p<0.05$). They correspond to the six dimensions of the

Brazil Atlas of Human Development [6]: demography, income, employment, housing, vulnerability and education, the latter being responsible for the largest number of indicators.

From the 18 selected variables, 6 factors were selected after applying the technique of factor analysis (FA). These factors explained 75.9% of the total variability in the data. They include the variables of living conditions correlated with SMR by aggression, which were used for the stratification of the microregions towards greater explicability of the multivariate analysis.

To confirm the hypothesis that the correlation matrix is an identity matrix (Table 1), Bartlett's test of sphericity produced a chi-square statistic equal to 1941.390 with degree of freedom (df) of 153, providing significant p-value = 0.000, whose decision was to reject the null hypothesis H_0: the correlation matrix is an identity matrix. Therefore, the correlation matrix is significantly different from the identity matrix. The KMO index - a measure of sampling adequacy showed a score of 0.730 indicating that the adequacy of the factor analysis method was "medium" for the treatment of the data. The communality for each variable presented a recommended value, ie, above 0.50.

The classification of variables in each factor was based on the value of the factorial loading. The six factors were selected for analysis using the latent root criterion, where the factor/component is selected if its eigenvalue is greater than one (1). The first factor explained most of the variability in the data (27.0%), and involved six indicators/variables with prevalent coefficients. This factor was formed by the dimension related to education indicators. The second factor was formed by two indicators related to the work and income. In the third factor appeared indicators of the dimension: housing, vulnerability and demography. Fourth, fifth and sixth factors included indicators related to income, work and vulnerability.

The correspondence from the Synthetic Indicator proposed in this work with the HDI indicator combined with different levels of violence is shown in Table 1, classified in four strata.

Table 1. Classification of stratum, according to the condition of life and the standardized mortality rates by aggression of the microregions of the semiarid for young males, Brazil 2010.

HDI/SMR	Stratum I	Stratum II	Stratum III	Stratum IV
	HDI medium with high degree of violence	HDI medium with low degree of violence	HDI low with high degree of violence	HDI low with low degree of violence
IDH Medium	0,611	0,601	0,586	0,578
SMR Medium	44,734	35,027	47,810	39,528

Sources of basic data: Brazil 2013 Human Development Atlas [6].
HDI - Human Development Index, SMR - Standardized Mortality Rate by aggression.

The values of the Synthetic Indicator for microregions according to the adopted classification are shown in Table 2. The microregions that comprise the stratum I, associates a standard of living expressed by an average of HDI of 0.611, classified

Table 2. Classification of microregions of the semiarid, according to the living conditions and mortality rates by aggression for males, Brazil 2010.

Stratum I-HDI medium with high degree of violence		Stratum II-HDI medium with low degree of violence		Stratum III-HDI low with high degree of violence		Stratum IV-HDI low with low degree of violence	
Microregions	*Scores*	*Microregions*	*Scores*	*Microregions*	*Scores*	*Microregions*	*Scores*
Mossoró (RN)	3667,9	Cajazeiras (PB)	487,6	Itaberaba (BA)	-42,0	Alto Médio Canindé (PI)	-511,1
Macaíba (RN)	2586,8	Santa Maria da Vitória (BA)	437,0	Serra de Santana (RN)	-86,5	Seridó Ocidental (RN)	-513,0
Cariri (CE)	2523,2	Janaúba (MG)	398,3	Campina Grande (PB)	-130,2	Traipu (AL)	-554,5
Sertão do São Francisco (AL)	2229,4	Euclides da Cunha (BA)	389,0	Cariri Oriental (PB)	-130,9	Montes Claros (MG)	-562,7
Pedra Azul (MG)	2134,8	Itapetinga (BA)	380,6	Curimataú Oriental (PB)	-165,2	Boquira (BA)	-564,0
Macau (RN)	2105,4	Guanambi (BA)	375,2	Barra (BA)	-185,0	Vale do Ipojuca (PE)	-572,7
Seridó Oriental (RN)	2086,7	Santa Quitéria (CE)	358,7	Juazeiro (BA)	-186,6	Chorozinho (CE)	-577,4
Propriá (SE)	2023,7	Médio Jaguaribe (CE)	316,7	Guarabira (PB)	-218,4	Pio IX (PI)	-613,4
Senhor do Bonfim (BA)	1995,6	Serrinha (BA)	315,9	Salinas (MG)	-220,5	Cotegipe (BA)	-619,9
Tobias Barreto (SE)	1956,5	Litoral Nordeste (RN)	315,4	Umbuzeiro (PB)	-236,2	Brejo Pernambucano (PE)	-620,2
Vale do Açu (RN)	1801,3	Pajeú (PE)	307,9	Esperança (PB)	-243,8	Sousa (PB)	-641,1
Nossa Senhora das Dores (SE)	1747,0	Almenara (MG)	298,1	Palmeira dos Índios (AL)	-245,9	Jeremoabo (BA)	-653,5
Itapipoca (CE)	1630,7	Agreste Potiguar (RN)	273,1	Batalha (AL)	-268,8	Patos (PB)	-684,0
Carira (SE)	1575,9	Chapada do Araripe (CE)	272,1	Alagoinhas (BA)	-297,8	Garanhuns (PE)	-697,8
Baixo Curu (CE)	1528,7	Sertão de Cratéus (CE)	255,3	Coreaú (CE)	-301,4	Baturité (CE)	-718,4
Chapada do Apodi (RN)	1496,0	Umarizal (RN)	250,8	Serra do Pereiro (CE)	-309,5	Itaporanga (PB)	-740,2
Bertolínia (PI)	1430,1	Seridó Ocidental Paraibano (PB)	223,8	Cariri Ocidental (PB)	-330,5	Borborema Potiguar (RN)	-743,8
Alto Médio Gurguéia (PI)	1273,2	Barro (CE)	197,4	Curimataú Ocidental (PB)	-349,8	Itabaiana (PB)	-858,6
Santo Antônio de Jesus (BA)	1248,2	Sertão de Quixeramobim (CE)	177,0	Catolé do Rocha (PB)	-354,9	Araçuaí (MG)	-900,1
Petrolina (PE)	1141,2	Valença do Piauí (PI)	171,7	Santana do Ipanema (AL)	-368,2	Arapiraca (AL)	-956,6
Pacajus (CE)	1106,2	Serra de São Miguel (RN)	169,8	Sertão de Inhamuns (CE)	-372,1	Picos (PI)	-962,5
Itaparica (PE)	1067,1	Caririaçu (CE)	145,5	Bom Jesus da Lapa (BA)	-372,6	Várzea Alegre (CE)	-982,4
Lavras da Mangabeira (CE)	1009,6	Pau dos Ferros (RN)	144,3	Seridó Oriental Paraibano (PB)	-375,1	Médio Capibaribe (PE)	-1029,3
Litoral de Aracati (CE)	928,0	Meruoca (CE)	132,4	Canindé (CE)	-383,7	Livramento do Brumado (BA)	-1110,7
Iguatu (CE)	873,6	Piancó (PB)	114,8	Brejo Santo (CE)	-414,1	Brumado (BA)	-1117,2
Baixo Jaguaribe (CE)	839,5	Jacobina (BA)	109,4	Feira de Santana (BA)	-424,0	Serrana do S. Alagoano (AL)	-1130,5
Seabra (BA)	782,9	Ipu (CE)	106,1	Serra do Teixeira (PB)	-435,2	Vale do Ipanema (PE)	-1146,5
Sertão do São Francisco (SE)	781,9	Floriano (PI)	104,8	Ribeira do Pombal (BA)	-442,4	Vitória da Conquista (BA)	-1371,7
Baixa Verde (RN)	715,7	Médio Oeste (RN)	92,2	Araripina (PE)	-449,5	Chapadas do Extremo Sul (PI)	-1452,6
Fortaleza (CE)	608,8	Capelinha (MG)	67,8	Sertão de S. Pompeu (CE)	-463,1	Brejo Paraibano (PB)	-1565,2
Salgueiro (PE)	606,4	Paulo Afonso (BA)	55,2	Alto Capibaribe (PE)	-466,9	São Raimundo Nonato (PI)	-1692,4
Sobral (CE)	570,7	Médio Curu (CE)	44,1	Campo Maior (PI)	-508,8	Sertão do Moxotó (PE)	-1742,4
Angicos (RN)	568,3	Jequié (BA)	2,1	Januária (MG)	-509,7	Agreste de Itabaiana (SE)	-1773,0
Ibiapaba (CE)	561,7	Uruburetama (CE)	-15,6	-	-	Litoral Piauiense (PI)	-2002,3
-	-	Irecê (BA)	-32,4	-	-	Grão Mogol (MG)	-2362,7

Sources of basic data: Brazil 2013 Human Development Atlas [6].
HDI: Human Development Index.

as medium degree of violence. The microregions were considered with high level of violence. In stratum II the average HDI with a score of 0.601, was also considered medium grade, and was associated with a low level of violence. Stratum III presented an average HDI of 0.586 and stratum IV with 0.578, both considered of low levels. The microregions were associated, respectively, to a high and low level of violence.

All strata included microregions with high and low values of SMR by agression, no matter the level of the HDI, medium or low. For instance, the microregion of Mossoró (RN) with the highest score (3667.9) in Stratum I was associated to a very high mortality rate. At the other extreme is the microregion of Irecê (BA) with the lowest score (-32.4) in the Stratum II, which was associated to a low mortality rate. Similarly, this fact was also observed when considering the Stratum III and IV. Strong positive association was found for microregions with a medium score of living conditions and a low mortality rate. On the other hand, microregions with low life conditions was associated with high mortality. This fact sugests a paradox for the Brazilian semiarid region.

In many cases it was possible to note the disparity between microregions belonging to the same State. The regions of almost all states were classified in all strata. It was unable to identify a unique pattern of relationship between level of living conditions and level of violence. These relationships occurred in various ways, independent of the level of development of the region.

4 Conclusions

The mortality rates by aggressions and reducible gaps of mortality revealed high magnitudes at several regions of the semiarid region. The results point to the need for development of more effective and explicit politics to prevent these deaths among youth. The link between high mortality rates by aggressions and low living condition is very usual. Nevertheless, it was found that the microregions of the Brazilian semiarid region in 2010 with high levels of mortality were present in all strata of living conditions (medium or low).

In the application of Factor Analysis, the first factor explained most of the variability in the data (27.0%), which was formed by indicators related to education dimension. Groot and Brink [8] in a study conducted in 1996 in the Netherlands, Duenhas and Gonçalves [9] in Spain, Soares [10] for a set of countries, support the idea that education is a means of crime prevention, especially violent crimes. These authors argue that education can contribute to reducing violence.

For Brazil, study carried out by Lobo and Fernandez [11] covering ten counties in the metropolitan region of Salvador for the period 1993-1999 indicated that if municipalities have access to education could contribute to reducing crime. Duenhas and Gonçalves [9] found that municipalities that spent more on education and public safety in the period 2000-2005 presented fewest number of homicides per hundred thousand inhabitants. Using longitudinal aggregate data regarding to Census 1991 and 2000, Oliveira [12] found association between crime to the size of cities. This study raises the hypothesis that cities with larger populations have

higher rates of homicide. Also argued that inequality, poverty and inefficiency of teaching contribute to increased violence in Brazil. França, Paes and Andrade [13] developed a study for Brazilian metropolitan and non-metropolitan areas for 2000. They found that the M-HDI (Municipal Human Development Index) acted in the opposite direction: a higher degree of development of these municipalities favored a lower rate of young homicide.

The conclusions drawn by the authors do not seem to express what we found for the semiarid region of Brazil. Violence is present both in very urbanized regions as in the less urbanized; in more developed regions or not; both in regions with higher educational level as the lower level. Thus, it appears that violence in the semiarid region presents multiple relationships difficult to capture. The coexistence of different living conditions associated with different levels of violence in these regions appears as a paradox, whose population coexists with an open insecurity. The existence of this paradox makes it difficult to identify any patterns of association between living conditions and violence.

Even considering limitations such as data quality for some microregions, and measurement errors in the construction of the indicators, the results suggest that one can not explain violence by a single variable, or by a short set of them. Despite the effort to capture the violence of young people, after a selection of 154 indicators/variables that resulted in 18 of them statistically significant, which represented various dimensions of living conditions, the theme does not end here.

References

1. BRAZIL. Instituto Nacional do Semiárido. Synopsis of the census for the Brazilian semiarid, http://www.insa.gov.br/censosab/publicacao/sinopse.pdf, 2013.
2. Waiselfisz, J. J., Map of Violence 2012: The New Pattern of Homicidal Violence in Brazil. São Paulo: Sangari Institut, 2011.
3. BRAZIL. Ministério da Saúde. Epidemiology of external causes in Brazil: mortality from accidents and violence in the period 2000-2009, http://portal.saude.gov.br/portal/arquivos/pdf/cap_11_saude_brasil_2010.pdf, 2010.
4. UNICEF. Fundo das Nações Unidas para a Infância - The Brazilian Semiarid, and Food Safety and Nutrition in Children and Adolescents, 2005.
5. BRAZIL. Mortality Information System . Deaths from external causes, http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sim/cnv/ext10ba.def , 2013.
6. Brazil 2013 Human Development Atlas. http://atlasbrasil.org.br/2013/pt/home/, 2014.
7. Hair, J. F. et al., Multivariate Data Analysis, 7th edition, Pearson Prentice-Hall Publishing, 2009.
8. Groot, W. and Brink, H. M. V. D., The effects of Education on Crime. "Scholar" Research Center for Education and labor Market – Department of Economics, University of Amsterdam, Amsterdam, 2002.
9. Duenhas, R. A. and Gonçalves, F. O., Education, Public Safety and Violence in Brazilian Municipalities: An Analysis of Dynamic Panel Data. Paper presented at XVII National Meeting of Population Studies, Caxambú, Brazil, 21-23 September, 2010.
10. Soares, R. R., Development, crime and punishment: accounting for the international differences in crime rates, in Journal of Development Economics, **7**3, 1, 155-184, 2004.
11. Lobo, L. F. e Fernandez, J. C., Crime in the metropolitan city of Salvador**.** Paper presented at XXXI National Meeting of Economy, Porto Seguro, Brazil, 09-12,

December, 2003.
12.Oliveira, C. A., Crime and the size of the Brazilian cities: focus of the criminal economy. Paper presented at XXXIII National Meeting of Economy, Natal, Brazil, 06 – 09 December, 2005.
13. França, M. C. and Paes, N. A. and Andrade, R. C. C., Determinants of mortality by homicide among young people in urban areas of Brazil. Paper presented at International Seminar on "Violence in Adolescence and Youth in Developing Countries". Organized by: IUSSP Scientific Panel on Young People's Life Course in Developing Countries. New Delhi, India, 09-11 October, 2012.

Discrete semi Markov patient pathways through hospital care via Markov modelling

Aleka Papadopoulou[1], Sally McClean[2] and Lalit Garg[3]
[1]Department of Mathematics,
Aristotle University of Thessaloniki,
Thessaloniki 54124, Greece
(e-mail: apapado@math.auth.gr)
[2]School of Computing and Information Engineering,
University of Ulster,
Coleraine, Northern Ireland, UK, BT52 1SA
(e-mail: si.mcclean@ulster.ac.uk)
[3]Computer Information Systems
Faculty of Information & Communication Technology
University of Malta, Malta
(e-mail: lalit.garg@um.edu.mt)

Abstract

In the present paper, we study the movement of patients through hospital care where each patient spends an amount of time in hospital, referred to as length of stay (LOS). In terms of semi-Markov modelling we can regard each patient pathway as a state of the semi-Markov model, therefore the holding time distribution of the *i*th state of the semi-Markov process is equivalent to the LOS distribution for the corresponding patient pathway. By assuming a closed system we envisage a situation where the hospital system is running at capacity, so any discharges are immediately replaced by new admissions to hospital. In the present paper a method is applied according to which we can describe first and second moments of numbers in each semi Markov patient pathway at any time via Markov modelling. Such values are useful for future capacity planning of patient demand on stretched hospital resources. The above results are illustrated numerically with healthcare data.

Keywords: Healthcare, Markov process, Semi Markov process, Population structure.

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

1 The semi Markov system via Markov modelling

Semi Markov models (Iosifescu-Manu (1972), Howard (1971), McClean (1980, 1986), Janssen (1986), Mehlman(1979), Bartholomew *et al*.(1991), Vassiliou and Papadopoulou (1992) were introduced as stochastic tools, which can provide a general framework that can accommodate a great variety of applied probability models. A semi-Markov approach provides more generality than may be required to describe the complex semantics of the models. However, the complexity of analysis in semi Markov models discourages its application to real life problems and leads to the simpler choice of a Markov model which most of the time provides inaccurate results. In the present paper a method is applied according to which we can describe the expected population structure of an open semi Markov system in discrete time with fixed size via Markov modeling.

So, let us now consider a population which is stratified into a set of states according to various characteristics and S={1,2,...,k}. The expected population structure of the system at any given time is described by the vector $\mathbf{N}(t)=[N_1(t),N_2(t),...,N_k(t)]'$ where $N_i(t)$ is the expected population in state i at time t. Also we assume that the individual transitions between the states occur according to a homogeneous semi Markov chain. In this respect let us denote by $\mathbf{P}$ the transition probability matrix of the embedded Markov chain and $\mathbf{H}(m)$ the matrix of holding time probabilities.

Let us now suppose that when an individual is in state i at time t and entered state i at time t-d (i.e. at its last transition) then the individual is in duration state (i,d) at time t. The transition probabilities between the duration states are of two types: the actual transitions (i.e. transitions from state (i,d) to (j,0) for every i and j) and the virtual transitions (i.e. transitions from state (i,d) to (i,d+1) for every i). The definition of the duration states and the calculation of the transition probabilities between them provides the tool to form an equivalent Markov model containing all the information from the semi Markov system (Papadopoulou and Vassiliou (2011)). Thus the new state space is S*={(1,0), (1,1),…(1,b_1-1),… (i,d),…,(k,0), (k,1),…(k,b_k-1)}, where b_i is the maximum possible duration in the original state i and the corresponding transition probability matrix is of the form

$$\mathbf{M}_{(b\,x\,b)} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{T}_{12} & \mathbf{T}_{13} & \dots & \dots & \mathbf{T}_{1k} \\ \mathbf{T}_{21} & \mathbf{S}_{22} & \mathbf{T}_{23} & \dots & \dots & \mathbf{T}_{2k} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{T}_{k-1,1} & \mathbf{T}_{k-1,2} & \mathbf{T}_{k-1,3} & \dots & \dots & \mathbf{T}_{k-1,k} \\ \mathbf{T}_{k1} & \mathbf{T}_{k2} & \mathbf{T}_{k3} & \dots & \dots & \mathbf{S}_{kk} \end{bmatrix}$$

where $b = \sum b_i$ while the $\mathbf{S}_{i,i}$ diagonal matrices have as super-diagonal elements the transition probabilities for the virtual transitions and as first column elements the probability of re-entry to that state and $\mathbf{T}_{i,j}$ matrices have as first column elements the transition probabilities for the actual transitions between states. We therefore define:

$$\mathbf{S}_{i,i} = \begin{bmatrix} \pi_i(1-\alpha_1^{(i)}) & \alpha_1^{(i)} & 0 & 0 & \dots & 0 \\ \pi_i(1-\alpha_2^{(i)}) & 0 & \alpha_2^{(i)} & 0 & \dots & 0 \\ \pi_i(1-\alpha_3^{(i)}) & 0 & 0 & \alpha_3^{(i)} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 \\ \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \pi_i(1-\alpha_{b_i-1}^{(i)}) & 0 & 0 & 0 & \dots & \alpha_{b_i-1}^{(i)} \\ \pi_i & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

while

$$\mathbf{T}_{i,j} = \begin{bmatrix} \pi_j(1-\alpha_1^{(i)}) & 0 & 0 & 0 & \dots & 0 \\ \pi_j(1-\alpha_2^{(i)}) & 0 & 0 & 0 & \dots & 0 \\ \pi_j(1-\alpha_3^{(i)}) & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 \\ \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \pi_j(1-\alpha_{b_i-1}^{(i)}) & 0 & 0 & 0 & \dots & 0 \\ \pi_j & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

where the first column of $\mathbf{S}_{i,\,i}$ represents probability of re-entry to state i having left one of the duration states of i, for $i=1, \dots d_i$, and the first column of $\mathbf{T}_{i,\,j}$ represents probability of re-entry to state j having left one of the duration states of i, for $j=1,\dots k$ and $i=1, \dots d_i$. The super-diagonal elements of $\mathbf{S}_{i,\,i}$ represents probability of transition from (new) state (i,d) to $(i,d+1)$ for every i. π_i is the probability of entry to (old) state i and $a_x^{(i)}$ is the probability of remaining in pathway i for at least one more unit of time given that the current holding time is x. Hence, $1\text{-}a_x^{(i)}$ is the probability of discharge from pathway i given that the current holding time is x.

All of the above matrices are defined as functions of the basic parameters of the semi Markov system. So, now we can define the expected population structure for the new model. Let

$$\mathbf{N}^*(t) = \left[E(n_{1,0}(t)), E(n_{1,1}(t)),..., E(n_{1,b_1-1}(t)),..., E(n_{i,d}(t)),..., E(n_{k,0}(t)),..., E(n_{k,b_k-1}(t))\right]$$

where $n_{i,d}(t)$ is the population in state (i,d) at time t. It is obvious that $n_i(t)=\sum_{d=0}^{b_i-1} n_{i,d}(t)$ for every i=1,2,..,k. Hence, the expected population structure of the old (semi Markov) system can be described as follows:

$$\mathbf{N}(t)=\left[\sum_{d=0}^{b_1-1}E(n_{1,d}(t)),\sum_{d=0}^{b_2-1}E(n_{2,d}(t)),...,\sum_{d=0}^{b_i-1}E(n_{i,d}(t)),...,\sum_{d=0}^{b_k-1}E(n_{k,d}(t))\right] \quad (1)$$

where

$E(n_{j,d}(t))=\sum_{i=1}^{k}\sum_{x=0}^{b_i-1} n_{i,x}(0)M_{(i,x)(j,d)}^{(t)}$ and $M_{(i,x)(j,d)}^{(t)}$ is the element of $\mathbf{M}^t$ in the position: $\left(x+1+\sum_{z=1}^{i-1}b_z(\text{row}),\ d+1+\sum_{z=1}^{j-1}b_z(\text{column})\right)$.

By applying well known properties of variances and covariances we get that

$$Cov(n_i(t),n_j(t))=Cov\left(\sum_{d=0}^{b_i-1}n_{i,d}(t),\sum_{d=0}^{b_j-1}n_{j,d}(t)\right)=\sum_{d_1=0}^{b_i-1}\sum_{d_2=0}^{b_j-1}Cov(n_{i,d_1}(t),n_{j,d_2}(t)) \quad (2)$$

It is known from (Bartholomew, 1982) that the variances and covariances of $n_z(t)$, $n_r(t)$ of a Markov system are described by

$$Cov(n_z(t),n_r(t))=\sum_{x=1}^{k}\sum_{s=1}^{k}p_{xz}p_{sr}Cov(n_x(t-1),n_s(t-1))+\sum_{x=1}^{k}(\delta_{zr}p_{xz}-p_{xz}p_{xr})E(n_x(t-1)) \quad (3)$$

with initial conditions $Cov(n_z(0),n_r(0))=0$.

Hence, from (2),(3) we get that the variances and covariances of $n_i(t)$, $n_j(t)$ of the old (semi Markov) system can be described as follows:

$$Cov(n_i(t),n_j(t))=\sum_{d_1=0}^{b_i-1}\sum_{d_2=0}^{b_j-1}Cov(n_{i,d_1}(t),n_{j,d_2}(t))=$$

$$=\sum_{d_1=0}^{b_i-1}\sum_{d_2=0}^{b_j-1}\sum_{(m,d_3)}\sum_{(q,d_4)}M_{(m,d_3)(i,d_1)}M_{(q,d_4)(j,d_2)}Cov(n_{(m,d_3)}(t-1)n_{(q,d_4)}(t-1))+$$

$$+\sum_{d_1=0}^{b_i-1}\sum_{d_2=0}^{b_j-1}\sum_{(m,d_3)}\left(\delta_{(i,d_1)(j,d_2)}M_{(m,d_3)(i,d_1)}-M_{(m,d_3)(i,d_1)}M_{(m,d_3)(j,d_2)}\right)E(n_{(m,d_3)}(t-1)), \quad (4)$$

where in the sums $\sum_{(m,d_3)} \sum_{(q,d_4)}$, $\sum_{(m,d_3)}$ the counters $(m,d_3), (q,d_4)$ take all the possible values for every m=1,2,…,k and d_3=0,1,..., $b_m - 1$, q=1,2,…,k and d_4=0,1,…, $b_q - 1$.

The initial conditions are $Cov(n_{(m,d)}(0), n_{(r,q)}(0)) = 0$.

From (4) and for i=j we get the variance of $n_i(t)$.

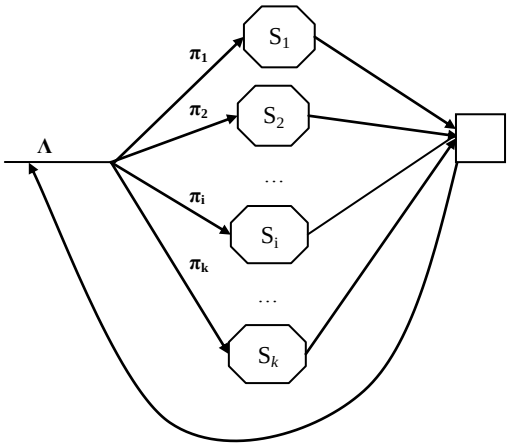


Figure 1

In Figure 1, we present the original semi Markov system while Figure 2 shows the same system transformed into a Markov system, as described above.

2. The Healthcare Application

As an example of the use of this approach in a healthcare environment, we consider the movement of patients through hospital care where each patient spends an amount of time in hospital, referred to as length of stay (LOS). In particular, we consider stroke patients where we regard stroke as a good paradigm example, affecting large numbers of patients with a resulting heavy burden on society.

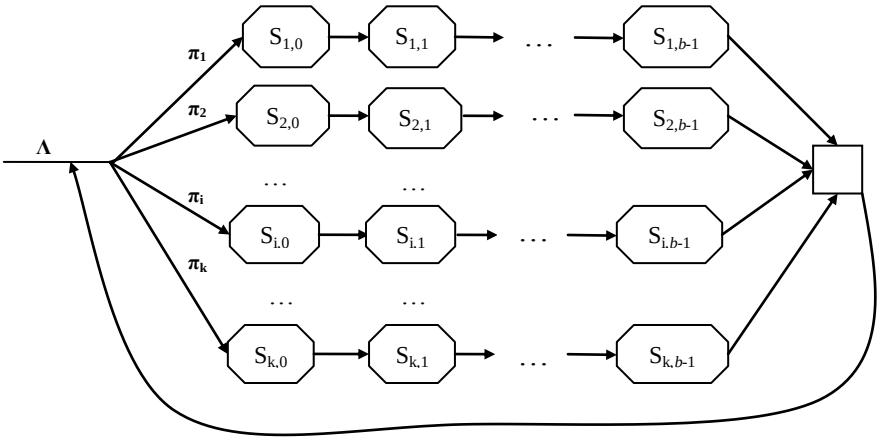


Figure 2

For example, in the UK it is estimated that Stroke disease costs over £7 billion a year, including community and social services, and costs to the labour force, as well as direct costs for hospital care.
We have previously defined stroke patient pathways (McClean et al., 2011) through hospital care, on the basis of diagnosis, gender, age and outcome, using log-rank tests to assess equality of survival distributions of LOS in hospital. Cox proportional hazards models were also employed to assess the effect of relevant covariates. On the basis of these tests we have defined 27 groups, each relating to a different patient pathway with respect to their LOS distribution. So patients in different pathways have different LOS distributions. Such pathways are characterised by the available covariates. Thus in the case of Stroke disease, examples of pathways are female, older patients, diagnosed with a haemorrhagic stroke and discharged to a private nursing home or male, younger patients with a transitory ischaemic attack, who were discharged to their own home. We note that, although we have discussed this framework specifically with respect to Stroke disease and the corresponding data available in our previous study, the concept is easily generalisable to other diseases and conditions, with different possible covariates.

In terms of the semi-Markov model discussed in the previous section, we regard each patient pathway as a state of the semi-Markov model, where pathway i is follow with probability π_i for i-1,…,k. The holding time distribution of the ith state of the semi-Markov process is therefore equivalent to the LOS distribution for the corresponding patient pathway.

By assuming a fixed size system we envisage a situation where the hospital system is running at capacity, so any discharges are immediately replaced by new admissions to the hospital.

Thus we can use the theory of the previous section to determine the distribution of population structure, in particular first and second moments of numbers in each patient pathway at any time. Such values are useful for future capacity planning of patient demand on stretched hospital resources.

Finally we discuss how to obtain the parameter estimates of our Markov representation of the original semi-Markov model. For each state of the Markov model (i, j) we define α_{ij} as the probability of transition into state $S_{i,\,j+1}$ given that the individual has already been in pathway i for duration t_j, for i=1,…,k and j=1,…b_i-1. Then β_{ij} = 1-α_{ij} is the probability of discharge from state $S_{i,j}$ given that the individual has already been in pathway i for duration t_j, for i=1,…,k and j=1,… b_i-1.

As estimators for α_{ij} and β_{ij}, we employ Kaplan-Maier type non-parametric maximum likelihood estimators (NPLMEs) (Bartholomew et al., 1991), as follows:

$\hat{\alpha}_{ij}$ = the number of pathway i patients who progress to another day in hospital divided by the number of patients that have stayed in hospital for t_j days. Also,

$\hat{\beta}_{ij} = 1-\hat{\alpha}_{ij}$, for i=1,…,k and j=1,… b-1.

References

1. Bartholomew, D. J., Forbes, A. F., and McClean, S. I., Statistical Techniques for Manpower Planning, Wiley, Chichester (1991).
2. Howard, R.A., Dynamic Probabilistic systems, Wiley, Chichester (1971).

3. Iosifescu-Manu A., Non homogeneous semi-Markov processes, Stud. Lere. Mat. **24**, 529-533 (1972).
4. Janssen, J. (Ed.), Semi-Markov Models: Theory and Applications, Plenum Press, New York. (1986).
5. McClean, S. I., A semi-Markovian model for a multigrade population, J. Appl. Probab. **17,** 846-852 (1980).
6. McClean, S. I., Semi-Markov models for manpower planning, in Semi-Markov Models: Theory and Applications (J. Janssen, Ed.), Plenum, New York. (1986).
7. McClean, S.I., Barton, M, Garg, L. and Fullerton, K., A Modeling Framework that combines Markov Models and Discrete Event Simulation for Stroke Patient Care, ACM Transactions on Modeling and Computer Simulation 21(4) (2011).
8. Mehlmann, A., Semi-Markovian manpower models in continuous time, J. Appl. Probab. **16**, 416-422 (1979)
9. Papadopoulou A.A and P-CG Vassiliou, On the variances and covariances of the duration state sizes of semi Markov systems MSMPRF2011 Conference 19-23/9, Chalkidiki, Greece (2011).
10. Vassiliou P-CG and Papadopoulou A.A., Non homogeneous semi-Markov systems and maintainability of the state sizes. J Appl Prob **27**, 756–76 (1992).

A Study of an Interval Scale for a Motivation Test

Artur Parreira[1] and Ana Lorga da Silva[2]

[1] MESTRADO CESGRANRIO, RJ, Brasil, and CPES – ECEO, ULHT, Lisboa, Portugal

(E-mail: arturmparreira@gmail.com)

[2] CPES – ECEO, ULHT, Lisboa, Portugal and CEDRIC - CNAM, France

(E-mail: ana.lorga@ulusofona.pt)

Abstract. This study deals with the development of interval scales, which combine the qualitative dimension of language with the corresponding quantitative dimension. The qualitative dimension is represented by adverbs of quantity and frequency, currently used by people, when they assign a value to things; the quantitative dimension was determined by an empirical study of the quantitative value of quantity and frequency adverbs.

This article is focused on adverbs of quantity.

A multivariate analysis is made, mainly using Classification methods (Gordon, 1999) and Principal Component Analysis (Jollife, 2002).

The interval scale was constructed to be used in a test of motivational profiles (Pinder, 2008) theoretically based in the paradigm of complex behavior processes (Parreira, 2006).

Keywords: Motivational profiles; Interval scales; Adverbs of quantity; Classification methods, PCA.

1 Introduction

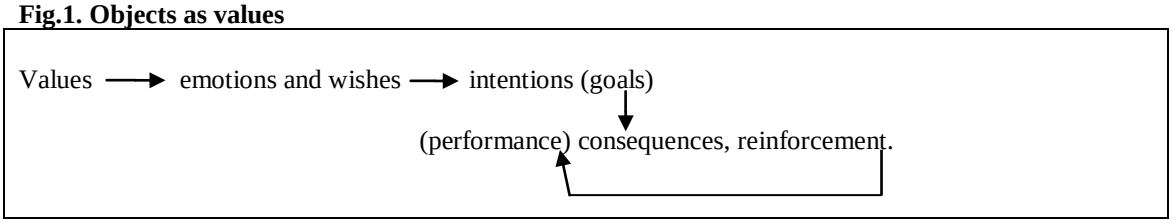
Since the interest in the phenomenon of motivation emerged in the psychosocial literature, its primary driver was a demand for effective technical solutions in organizational practice. It triggered the completion of studies in the field of theory, aiming to question and tests the established models and concepts, to propose alternatives, even to decide the interest in maintaining or abandoning the very theme of motivation (Salancik, 1983). Such attempts were, in general, ways to enlighten one or another dimension proposed by the models that resulted from the different images of the individual, delineated by the different approaches in social psychology:
- Information processor;
- Open system;
- Strategist and actor;
- Spontaneous epistemologist and builder of theories;
- Complex interpreter of himself and his context.

Perhaps the pivotal point to integrate those theories and models can be found in the open systems theory, particularly in the perspective of complex systems of action. This theoretical view stresses two dimensions, moreover focused directly or indirectly by all the models that previously tried to explain the motivational process:
- The internal dimension in the individual;
- The external dimension, the factors in the context.

This paper is about the elaboration of a motivation test, based on a motivational model supported by a systemic view of the human being. The proposed model resumes the ideas offered by both internal and

external previous theories and models, linking them to a regulatory system which includes emotions and the cognition of value objects as a part of the regulatory process.
Locke (quoted in Steers and Porter, 1983), defines value as something one strives to obtain or retain and proposes the following scheme to make clear its combination with intentions or goals:

Fig.1. Objects as values



Herzberg (1961) proposes an interpretation of the environment, based on his theory: the context contains two kinds of motivational factors that influence the individual's choices:
- Hygiene factors that influence individual behavior primarily by feelings of dissatisfaction;
- Motivating factors that influence individual behavior primarily by feelings of satisfaction.
  Both theories come from Lewin gestalt psychology, and have a similar view, even though they use a different language.

The model presented in this study seeks to integrate the two dimensions in its attempt to explain and manage organizational behavior: the internal dimension to the individual, which is the energy of emotions channeled into the production of behavior; contextual dimension, the objects in the field, that the subject views as value objects.

2. The REDA Model – emotional regulation of decisions to act

This systemic perspective of the person is just the first pillar of the REDA Model (Emotional Regulation of Decisions to Act). The model adopts a systemic perspective based on James Miller (1978).
The human individual is conceived as a system of action, alive and complex, that exists in time and in space interlaced with time, i.e., he has a history. This allows us to say that a living system is, finally, its history (Nuttin, 1980). The space-time where a living system moves is its interactional or vital field (Lewin, 1943).
Human behavior is regulated, ie, it happens according to the subject's criteria; this regulation has an energy component, emotional, it is its engine; and has a cognitive, intentional component, which is the information for decision making. The behavior is triggered as a result of the affective tone, positive or negative content of more or less intensity, that follows the emotional state of the subject, when he faces the value objects in the context. This affective tone is associated to the functioning of the 19 subsystems that Miller considers critical to the survival of living systems.
In REDA model, these 19 subsystems are aggregated into three sets: those that process energy and matter, those which process information, finally one which processes information and energy.

Table 2 - The Individual and his subsystems (REDA Model)

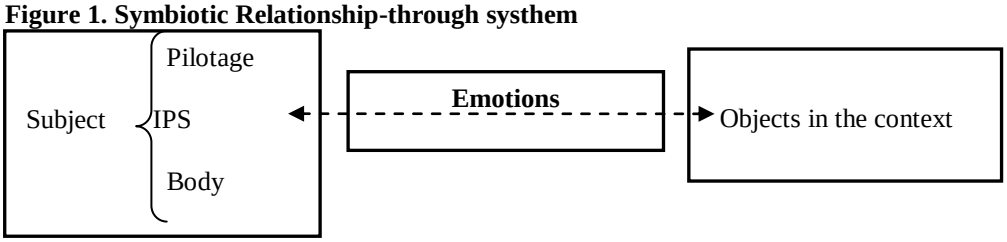
NATURE OF THE SUBSYSTEM	SUBSYSTEM
Subsystems processing matter and energy, physical support of information-processing subsystems.	**Body, physical subsystems**
Processor of instrumental information. Perception and interpretation of reality and himself, operational and technical knowledge, maps of the context, problem-solving processes.	**IPS** **Perception, intelligence, forms of reasoning.**
Regulating processor of integrative information and associated energy. It contains self-image, body schema, basic principles and assumptions, life values, criteria of action, vital goals and objectives, temporal perspective.	**Pilotage (decider subsystem)**

Source: Artur Parreira (2006)

The human subject is understood as a complex open system of action; the subsystems shown in table 2 are the fundamental structures of the human person and each one operates on its own specificity. The behavior is conceived in terms of essential connection with what is called the value objects in the context:

all objects, people, circumstances, events with which the subject interacts. It's this interactional behavior that is motivationally regulated.

Figure 1. Symbiotic Relationship-through systhem



Source: Authors elaboration

Figure 1 highlights the relationship with the environment as a part of the nature of the human system; it shows that it is processed through the handling of emotions, which are integrated into the model as the energy dimension of behavior regulation (Damasio, 2008).
In this paper, motivators and emotions are conceived as having the same nature; their difference lies only in the way we define them in relation to the subject.
When we talk about emotion as a driver for action, we use the term motivation; when we speak of emotion as an expression of the subject, we use the term emotion. The motivations are, therefore, emotions, and these states are energetic forces that trigger the action. The type of emotion and emotional intensity are perceived by the subject as indicators for the regulation of their conduct. Motivation is thus conceived as an affective or energetical process, and this energy is the engine of behavior: since each motivator is an affective system of variable intensity, this variability allows using it as regulator of behavior. Table 3 shows the motivators and regulated subsystems

To understand the process of motivation or behavior regulation, we must understand its relationship with the three sets of subsystems:
- The drivers have emotions as content, that is, as noted above, the motivations are oriented emotions to trigger the action (Izard, 2009);
- There are motivators used to regulate the activities of the body (motivators F); motivators to regulate the activities of SPI, or information processing subsystem (motivators C); and motivators to regulate the activities of pilotage (motivators S, P, R and A);
- With the motivational regulation of behavior, the subject pursues two results: maintaining and expanding his sustainability.
REDA model assumes six regulators / motivators of human behavior, denominated after Murray (1938) and Maslow (1954), since their definition of needs implicitly contains the emotions which are the core of REDA motivators. These are divided in two types of regulation, according to their goal:
- **Adjustment and maintenance**, homeostatic type, to restore the initial situation, defined as consistent with the standard of an adequate operation. It is a a negative feedback process, in which the main role is played by emotional states (negative / positive).
- **Setting expansion,** a dynamic type of feedback positive, ie, which seeks adjustment to affective representations of the future, which naturally increases the deviation from the initial state toward more complex regulatory conditions.
The first process is characteristic of the four motivators focused mainly on the internal functioning of the human system; the second is characteristic of the two motivators essentially responsive to informational objects, and their focus is the relationship with the objects of the environment.

Table 3. Negative feedback motivators (maintenance)

Motivators	**Nature of motivating**	**Regulation mode**	***Regulated behaviors***
F (maintenance)*Subsystem:* *Maintenance of well being and pleasure / avoidance of pain* ***Body***	*Positive pole: feelings of joy, satisfaction, associated with physical well-being* ↕ *Negative pole: feelings of pain (physical) fear (pain duration),*	***1. Intensity intra cognition*** *Positive pole: maintains the standard of behavior* *Negative pole: triggers behavior regulator.* ***2. Intensity ultra cognition*** *Marked loss of regulatory function; inhibition of the regulatory function of other motivators*	**Position 1** Behaviors search sensations of physical pleasure (leisure activity, eating, drinking, sleeping, cuddling, sex), stop stress, rest.; escape behaviors to pain and discomfort**.** **Position 2.** Uncontrolled behaviors, disregard of standards, imprudent risk, and the like, for excessive intensity of emotions associated.
S (maintenance) *Search for an adequate level of security protection against threats and dangers* Subsystem: ***Pilotage***	*Feeling quiet, comfortable, relaxed,* ↕ *Feelings of fear, insecurity, anxiety.*		**Position 1:** Behaviors of obedience, conformity, passivity, caution, relaxation; but also carelessness, subservience, lies, deception, concealment, when they seem more effective to avoid fear. **Position 2:** Paralysis of action, uncontrolled physical symptoms
P (maintenance) Seeking a friendly company, an escape from loneliness Subsystem: ***Pilotage***	*Feelings of acceptance, inclusion, affective company* ↕ *Feelings of loneliness, emptiness, loss of love*		**Position 1.:** Behaviors of sympathy, solidarity, support, benevolence, friendship, forgiveness, love. **Position 2:** temporize, lack of exigence, acceptance of failures, emotional dependency and lack of autonomy
R (maintenance) Seeking power, prominent position, to escape social devaluation and low status Subsystem: ***Pilotage***	*Pride; feeling valuable;important; capable; worthy of praise; powerful* ↕ *Guilt, shame; worthlessness; low esteem.*		**Position 1:** ambition behavior, self esteem, dynamism, exigence posture, goal-orientation, attraction for power**.** **Position 2:** Attitudes of arrogance, contempt of others, exhibitionism, devaluation of information

Source: Authors elaboration (2014)

Table 4. Positive feedback motivators (expansion)

Motivators	**Nature of motivating**	**Regulation mode**	***Regulated behaviors***
C *(expansion)* Desire to know, curiosity *Subsystem:* ***SPI***	*Amazement, excitement, curiosity* ↕ *Apathy, boredom, indifference to knowledge*	***1. Intensity intra cognition*** *Positive pole: triggers the action* *Negative pole: passivity, inaction* ***1. Intensity ultra cognition*** *Partial loss of regulation by inhibition of other motivating processes*	**Position 1 :** curiosity behaviors, active listening, attention to information, data exploration. **Position 2:** onlookers behavior indiscreet questions, invasion of privacy, data abuse.
A *(expansion)* Seeking autonomy and rationality in action and relationships *Subsystem: Pilotage*	*Joy in being complete and in development; to have power over things; to do well and avoid error; freedom of action and assert ho* ↕ *Absence of self liking and development; indifference for personal identity; low sensitivity to the meaning of life and personal autonomy.*		**Position 1:** autonomous decision behavior, rational analysis, assertiveness, openness to others, seeking balance in actions and affections, spiritual sensitivity. **Position 2:** attitudes of indifference and incomprehension of less rational attitudes or difficulties of others; ignorance and alienation of the field.

Source: Authors elaboration (2014)

2.1. How motivators regulate behavior?

Maintenance motivators induce the subject to maintain the standards of behavior that he is customary and that trigger positive emotions of lower or higher intensity. Thus these regulators maintain system's inertia, because the behavior is triggered by the negative pole of emotion and seeks to restore the positive emotional state.

The expansion motivators regulate the behavior of the subsystems involved in a slightly different way: like maintenance motivators they make behavior appear when the emotional energy is sufficient to trigger it; but with these motivators it is the positive emotion that fosters action, whose goa lis not to eliminate a gap in relation to a proper situation of the past, but in relation to a desirable future. It is this mode of regulating which leads to the expansion of system's capabilities and frontiers.

First postulate

Human behavior is always a response to the objects in the field, a relational response when the object is a person

The attribution of meaning and value to objects in the field is taken by the subject based on their motivational-emotional schema; consequently, human behavior isbased on the processing of affective information from the objects (Zajonc, 1980): the living space is a field of meanings associated with valences, ie, emotional intensities (Lewin, 1951). That's why systems are only partially open and rational (Morin, 1975, Simon, 1959): people only pay attention to the objects that are significant to them (ie, that are connected to one or more of our motivational regulators), ignoring much information actually existing in the context. The human subject choices heavily depends on the intensity and affective tone aroused by the objects of the vital field; this guides him to a decision to act.

Second Postulate

The choice of a behavior results from the assessment of their comparative marginal utility and is perceived as a dilemma to solve; the resolution of the dilemma is controlled by the images that the subject makes situation.

Third postulate

The regulatory action of the motivators is a function of its emotional intensity: is appropriate when the emotional intensity is integrated in the global cognition of the subject about his relationship with the environment, and is lost when i tis no longer integrated in this global cognition.

Fourth postulate

When the vital field has more and more varied resources, individuals tend to become 'motivator seekers' more than 'higiene seekers' (Myers, 1964) and regulatory function is overall enhanced.

2.2. The regulation of behavior in practice

1. When the human subject perceives standard functioning of one subsystem (type and intensity of positive emotion felt) he recognizes that he is in the emotional condition associated with standard and tends to maintain his condition or behavior (*behavioral inertia*). The alternative behavior appears to have zero marginal utility.

2. A negative affective state experienced indicates to the subject the urgency of acting to eliminate the gap indicated by the negative emotion. According to the intensity of emotion, the subject feels a pressure to produce the behavior that he perceives as adequate to bring him back to a positive emotional condition (the standard condition).

3. The energy applicable to a choosed behavior is a function of *the intensity of positive emotion associated to norm condition + the intensity of negative emotion associated with deviations from the norm*. The conversion applicable into applied emotional energy is a function of behavior instrumentality, that is, the perceived probability that the behavior will lead to the emotional condition associated to the norm, without an emotional cost superior to that of maintaining the gap. This concept of instrumentality, defined on the basis of emotions in conflict, corresponds to what Alberto Bandura (1986) called self-efficacy.

4. According to the fourth postulate, the vital field may present conditions that facilitate or difficult the return to the emotional state of the norm for several motivators. If the individual can not afford that his efforts replace that condition the intensity of the negative state remains or worsens: the subject channels his energy to substitute behaviors, eventually varying them, but if frustration is maintained, he may reach poorly regulated patterns of adjustment or even clearly deregulated. If, instead, vital field of the subject becomes richer in objects appropriate to several motivators, the frustration process becomes rare and emotional experiences are generally positive. Then individual represents reality as a source of satisfaction and value and becomes sensitive to objects and situations which Herzberg (1961), and also REDA model, designated 'motivators'. The motivational questionnaire which was elaborated in this research is a tool to measure the motivational profiles of people, based on the theoretical concepts described above.

3. The Test

The test puts value objects and situations in confrontation, in sets of three. The purpose of the test is to put the person "in front of" affectively guided choices - motivational dilemas - which are the way motivation works (first and second postulates).

Example:

Directions
The questionnaire has several sets of three statements in confrontation: evaluate which of them describes you more exactly; then do the same with the other two, checking which one applies to you more accurately; the remaining assertion is the least true about you. Now, based on the scale that accompanies the questions, assess the value each statement has for you: signal first the value you attach to the phrase you have put in the first place; then mark the value you assign to the second position; finally indicate the value of the phrase put in third position. Assigned values express the relative position they have in your life.

1.

	order	value
1a. It has been very important for me to enjoy good moments of comfort and leisure		
1b. It has been very important for me to attend spectacles of music, theater, ballet, cinema, to appreciate works of art in general		
1c. It has been a very important thing in my life dedicating myself to help people to solve their problems		

The questionnaire includes 20 sets, like this one, which cover the most important areas of daily life:
- Personal life and family (10 sets)
- Work (5 sets)
- Leisure and free time (2 sets)
- Friends and friendships (3 sets).

The test has a second part in which is recorded the frequency of emotions felt in the personal life and at work. Evaluation is made with the same type of scale, measuring the affective tone of the subject's life.

Example:

How often have you experienced in your personal life and work the situations shown in this table?		
Situations	**In personal life**	**At work**
A - Situations that generate feelings of warmth, sympathy, friendship		
R - Situations that generate feelings of anger, irritation, anger		

Test Subscales
- **F scale:** motivation to search feelings of well being and pleasure, to escape pain and displeasure
- **S scale:** motivation to seek safety, avoidance of fear
- **P scale:** motivation for searching the sympathy and love of others, a desire to belong to someone or a group, to avoid solitude
- **R scale:** motivation to assert pride, seeking consideration and status, to avoid devaluation and insignificance
- **C scale:** curiosity, desire to learn, motivation to learn and be informed; indifference to information and knowledge
- **A scale:** search for development and realization of personal values, autonomy, vital sense and rationality; refusal of the barriers to realization of the self, refusal of irrationality.

The scores of these subscales show the emotional intensity and hence the effort channelled to the behaviors associated to each motivator.

3.1. Meaning of each motivational pattern

1. Existential maturity and responsibility pattern (EMR pattern)

Is given by the expression: EMR $= \frac{\sum_{i=1}^{10} x_{i.}}{10} - \frac{\sum_{i=1}^{50} x_{i.}^{*}}{50}$, where $x_{i.}$, represents the "positive items" and $x_{i.}^{*}$ the negative one in this pattern.

The formulas are similar to the following patterns, differing only on the items and in their number.

2. Existential Teen Relaxation Pattern (ETR pattern**)**

3. Manipulative pattern , attitude of obtaining advantage (MAN pattern**)**

4. Conformist pattern, index of social conformity (CONF pattern)

Each pattern indicates an emotional strength that is in confrontation with the other patterns. This is theoretically a subtraction. The individual's behavior is guided by one pattern if its strength is superior to that os the others.

5. **Dominant Constellation** - Sum of values in the two most scored motivators

The dominant constellation explains the vocation and the overall orientation of life.

6. Rejected Constellation - Sum of the values of the two least scored Motivators

Rejected constellation explains weaknesses, phobias, automatic rejections

7. Vital energy (global motivational intensity): Averages of all Motivators + level of positive tone

8. Motivation for learning and innovative change

9. Emotional Positivity

3. Data Analysis

3.1. The Sample

In fact this study is a pré-test, the sample, has a degree of diversity not still similar to the final / ideal sample that will support the creation of norms and "rules" that will allow to interpret the test scores results, for instance the number of observed men is largely superior of the observed women.

Figure 5 – The Sample

Dimension of the sample:116

Gender

women - 64,3%
men –35,7%

Age

≤ 25	**43,3%**
]25, 35]	**38,1%**
]35 , 45]	**11,4%**
]45 , 55]	**6,2%**
≥ 55	**1,0%**

Number the years of Education

≤ 12	6,1%
]12 , 15]	92,9%
> 15	0,9%

Jobs

Professor…….....……13,3%
Human Resources Management .13,2%
Economics and management …..13,3%
Administrative jobs…… 8,2%
Students……..……..52,0%

586

3.2. The Scale

The statistical data for the test were calculated on the basis of this interval scale; as an example:

Figure 6 – a scale

2.	Extremely true	- scale value - 9.25
3.	Very true	- « « - 7.62
4.	True enough	- « « - 7.08
5.	Medially true	- « « - 4.62
6.	Little true	- « « - 2.27
7.	Not true at all	- « « - 0.54

(Parreira and Lorga da Silva, 2013)

3.2.1. Internal consistensy

Studing internal consistency according to Sijtsma (2009) for instance, the value of the realibity statistics is 0.861 for the 60 items, wich means we are in the presence of internal consistency in our motivation test. It makes sense to explore the obtained data.

3.2.2. The motivation scales – statistics and hipothesis test

Figure 7 – motivation scales

	F	A	S	C	R	P
Mean	6,687	6,6597	6,744	6,729	5,766	6,463
sd	0,863	1,087	1,059	1,029	1,044	0,967
Q_1	6,094	6,11	6,069	5,867	5,079	5,836
Q_2	6,706	6,701	6,877	6,776	5,736	6,631
Q_3	7,221	7,421	7,668	7,427	6,450	7,191

The obtained results are similar for each scale

Figure 8 – Test: Equality of means by Gender:

	t-test for Equality of Means						
	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper
F	1,432	96	,156	,26959	,18832	-,10421	,64340
	1,401	66,082	,166	,26959	,19236	-,11447	,65365
A	4,424	96	,000	,96571	,21828	,53242	1,39900
	4,344	66,668	,000	,96571	,22231	,52194	1,40948
S	1,724	96	,088	,39168	,22722	-,05934	,84270
	1,660	62,953	,102	,39168	,23596	-,07985	,86320
C	2,022	96	,046	,42600	,21068	,00780	,84419
	1,959	64,056	,055	,42600	,21749	-,00848	,86048
R	,882	96	,380	,19479	,22081	-,24352	,63311
	,926	80,716	,357	,19479	,21041	-,22388	,61347
P	3,399	96	,001	,67498	,19860	,28076	1,06920
	3,302	64,627	,002	,67498	,20441	,26671	1,08326

Figure 9 – Test: "The distribution is the same across Gender"

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of F is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,114	Retain the null hypothesis.
2	The distribution of A is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
3	The distribution of S is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,096	Retain the null hypothesis.
4	The distribution of C is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,041	Reject the null hypothesis.
5	The distribution of R is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,284	Retain the null hypothesis.
6	The distribution of P is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,005	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of F is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,114	Retain the null hypothesis.
2	The distribution of A is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
3	The distribution of S is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,096	Retain the null hypothesis.
4	The distribution of C is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,041	Retain the null hypothesis.
5	The distribution of R is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,284	Retain the null hypothesis.
6	The distribution of P is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,005	Reject the null hypothesis.

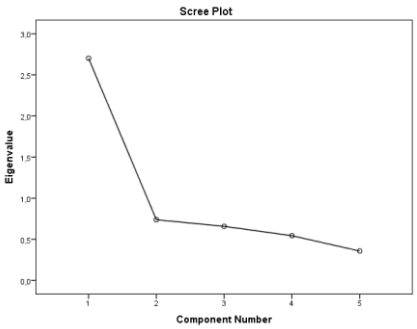
Asymptotic significances are displayed. The significance level is ,01.

Has we have made the remark there is not an equilibrium in the variable gender in this sample, the tests shown in figures **8** and **9** let us conclude that the scales **A** and **P** are those who can reflect that difference.

3.2.3 The motivation scales – Principal component Analysis and Hierachical Classificatiion Analisys.

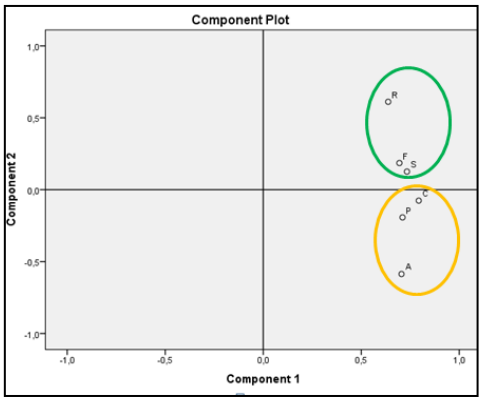
Global Analisys

Figure 10 – Número of dimensions and Groups / Partitions



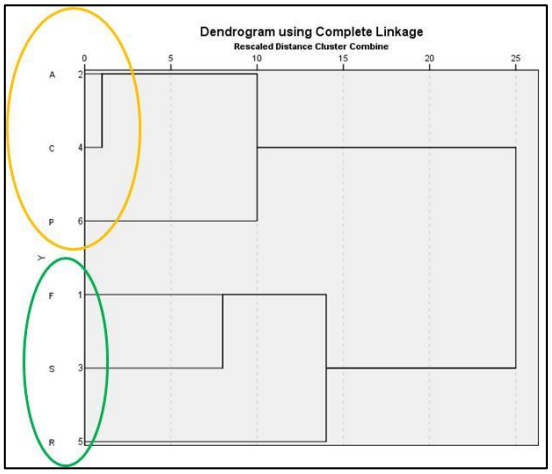
It induce us to choose two chose two dimensions and groups.

Figure 11- PCA result



All the conditions about eigenvalues and variance have been verified

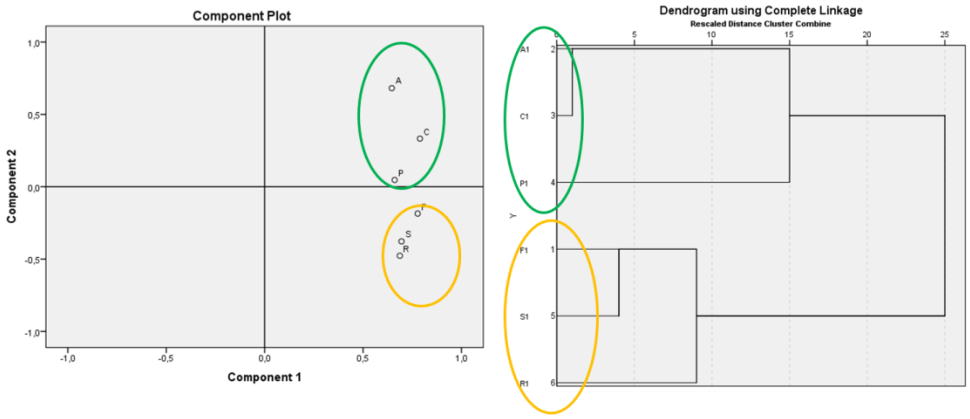
Figure 12 - Hierachical Classificatiion Analisys – Complete Linkage



This results are consistent with the theory, which states that the motivators perform two regulatory roles: maintenance of the normal condition of the human system, expansion of the existing capacity to higher levels.

By Gender

Figure 13 – PCA results and Hierachical Classificatiion Analisys – Complete Linkage



All the conditions about eigenvalues and variance have been verified.

Those results are slightly different but not incoherent, for women, the affective dimension seems to be more associated to expansion than the desire to know.

The scores of all motivators show a reasonable intensity (above 6,5). This indicates a population with positive emotional tone and vital energy, to express and act. Motivator R is situated below the average value of the other; this is congruent with the stereotype that the Portuguese have difficulties in personal statement and are prone to be passive and conformist.

A curious statement, is that women have a higher average value in A and R, in opposition to generalized opinions

The results are consistent with the theory, which states that the motivators perform two regulatory roles: maintenance of the normal condition of the human system and expansion of the existing capacity to higher levels.

3.3 Patterns – Brief analysis of EMR Pattern

Figure 14 – EMR pattern

	EMR
Mean	0,339
sd	0,841
Q_1	-0,187
Q_2	0,396
Q_3	0,868

4. Conclusion

The obtained results are coherent and consistent with the postulates of the theory. So, this work allows us to think that the test will offer indications useful to assess people adaptability to work and educational situations.

The use of interval scales combining qualitative and quantitative responses seems also an interesting way to refine our behavioral measurements. Our following work will increase the dimension and diversity of the sample, and refine the formulas about several motivational patterns. That will be done after the more extensive collect of data, in what we are engaged now.

References

Adams, J.S. (1965). Inequity in Social Exchange. In: L. Berkowitz (Ed) *Advances in Experimental Social Psychology. Vol. 2*, pp. 267-99. Nova Iorque: Academic Press.

Aldefer, C (1972). *Existence, Relatedness and Growfh*. Londres: McMillan

Borg,I. and Groenen, P.J.F. (2005), *Modern Multidimensional Scaling: Theory and Applications*: Springer Series in Statistics

Damásio, A. (2010). *O livro da Consciência*. Mem Martins: Europa América.

Freud, I, 228; II, 2041) Freud, S. (1905d / 1962). *Standard Edition*. Londres: Pelican Freud Library.

Gordon, A. D. (1999), Classification, Chapman & Hall, London.

Herzberg, F. (1966). *Work and Nature of Man.* Cleveland: World Publishing Company

Izard, C.E. (2009) Emotion Theory and Research *Annu. Rev. Psychol. 2009.60:1-25.*

Jollife, I.T. (2002), Principal Component Analysis, Springer, New York.

Lewin, K. (1936). *Principles of Topological Psychology*. N. York: McGraw-Hill.

Locke E. e Latham G. (1984) *Goal setting: A motivational technique that works.* Englewood Cliffs: Prentice-Hall

Locke, E. (2004) Goal setting theory and its applications to the world of business. *Academy of Management Executive 18(4),* pp. 124- 125.

Murray, H. A. (1938). *Exploration in personality*. New York: Oxford University Press.

Maslow, A. H. (1954). *Motivation and personality*. New York: Harper & Row.

Miller, J. (1978). *Living Systems*. Londres: McGraw-Hill

Morin, E. (1985), *O problema epistemológico da complexidade*, Europa América, Portugal

Murray, H. A. 1938. Exploration in personality. New York: Oxford University Press.

Myers, S. (1964). Who are your motivated workers?. (Reprint). *Harvard Business Review, Nº de Janeiro-Fevereiro de 1964.*

Nuttin, J. (1978). *Théorie de la Motivation Humaine*. Paris: PUF

Parreira, A. (2006), Gestão do Stress e da Qualidade de Vida, Monitor, Lisboa.

Pinder, C. (2008), Work motivation in organizational behavior (2nd ed.). Psychology Press, New York US.

Sijtsma, K (2009) On the Use, the Misure, and the Very Limited Usefulness of Cronbach's Alpha, *Psychometrika*. 74(1): 107 120.

Parreira, A. (2006) *Gestão do Stress e da Qualidade de Vida*. Lisboa: Monitor.

Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.

Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.

Is it worth using a fuzzy controller to adjust the mutation probability in a genetic algorithm when the input variable is the number of iterations?

André G. C. Pereira[1], Bernardo B. de Andrade[1], and Enrico Colossimo[2]

[1] Applied Mathematics/Statistics Program, Universidade Federal do Rio Grande do Norte, Natal/RN, Brazil
(E-mail: andre@ccet.ufrn.br, bba@ccet.ufrn.br)
[2] Statistics Department - Universidade Federal de Minas Gerais, Belo Horizonte/MG, Brasil
(E-mail: enricoc@est.ufmg.br)

Abstract. In recent years, several attempts to improve the efficiency of the Canonical Genetic Algorithm (CGA) have been presented. One of such attempts was the introduction of the elitist non-homogeneous genetic algorithm whose advantage over the CGA is that variations in the mutation probabilities are allowed. Such variations permit the algorithm to broaden/restrict its search space as the mutation probability increases/deacreases. The problem with this approach is that the way such changes will happen has to be defined before the algorithm is initiated. In general such changes happen in a decreasing way. Another way to vary the mutation probability is to use controllers. Fuzzy controllers have been widely used to adjust the parameters of a genetic algorithm. In this paper a stochastic controller is introduced along with a convergence proof of the genetic algorithm which uses such controller to adjust the mutation probability and a way to construct a stochastic controller from a fuzzy controller. Numerical comparisons between the two types of controllers and the CGA are performed. The general case with several input variables and two output variables (mutation and crossover probabilities) is work in progress.

Keywords: Stochastic controller, Fuzzy controller, Genetic Algorithms..

1 Introduction

The Canonical Genetic Algorithm (CGA), introduced in Holland[8], is a computational tool that mimics the natural genetic evolutionary process of a population that undergoes three stages: selection, crossover (mating) and mutation. In the CGA, a population of N individuals or chromossomes, $(u_1, u_2, ..., u_N)$, is considered. An evaluation function $f : E \to (0, \infty)$ assigns to each individual u_i a fitness value $0 < f(u_i) < \infty$. In the selection stage, the actual population will be resampled, individuals with higher fitness are more likely to be selected and those with low fitness tends to be eliminated (elitist selection). Following the natural evolutionary process, biological reproduction (crossover) and eventual mutation occur. In the crossover stage, individuals are independently chosen for crossover with a prescribed probability p_c. Mutation also operates independently on each individual with a prescribed probability p_m. In order to

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

be easier for implementation, each individual is represented by a binary vector of lenght l, where l depends on the desired precision. For more details as well as implementation procedures see, for example, Campos *et al.*[2], Pereira and Andrade [12], Goldberg[7] .

In optimization, CGAs are used to solve problems of the type $\max\{f(x), x \in E\}$ with the objective function satisfying $0 < f(x) < \infty$. The individuals represent the feasible solutions and the selection stage preserves with higher probability the best fitted/searched points. In the crossover stage, neighboring points are searched, allowing a refined comparison in the surrounding. In the mutation stage, random points, possibly away from the preserved ones, are visited and constitute a strategy to avoid being trapped in local optimum points.

The non-homogeneous Genetic Algorithms (NHGA) was analysed in Campos *et al.*[1] focusing on the improvement of eficiency upon the CGA, by allowing the mutation and crossover probabilities to vary under certain conditions. The elitist genetic algorithm (EGA), which was introduced in Rudolph[15], was a modification in the CGA that solved the problem of eficiency of the CGAs. A non-homogeneous version of the EGA, called elitist non-homogeneous genetic algorithm (ENHGA), was introduced in Rojas Cruz and Pereira[14] in order to improve the eficiency of the EGA. Other attempts to improve the eficiency of the CGA, without changing the mutation and crossover probabilities can be seen in [3,5], some numerical comparisons between ENHGA and EGA, can be seen in Campos *et al.*[2], and the proper way of running the ENHGA can be seen in Pereira and Andrade[12].

The advantage of the ENHGA over the EGA is that variation of the mutation probabilities (starting high and decreasing) permits the algorithm to broaden its search space at the start and restrict it later on. The way in which the mutation probabilities vary is defined before the algorithm is initiated. The ideal would be for the parameters to vary, rather than only diminish, depending on a certain measure of dispersion of the elements of the current population, as well as the number of iterations of the algorithm. To this end, controllers are introduced in the intermediate stages of the algorithm in order to adjust such changes. Various types of controllers can be used for this task, ranging from deterministic methods to those that employ fuzzy logic. Many simulation studies have used fuzzy controllers to adjust the parameters in order to improve the performance of the genetic algorithm. However, it has been only recently [13] discussed the conditions that must be met by the controller in order to ensure convergence of the genetic algorithm.

Ref. [6] describes a series of parameters of the GA and simulations were developed to illustrate that variations on those parameters interfere with the output of the algorithm. Ref. [10] reports that, based on [6], one attempt of defining the way of varying the parameters in order to improve the performance of the algorithm was unsuccessfully tried. For this reason a fuzzy controller was proposed as a tool to control the parameters substantiated only by simulations. In [14] sufficient conditions are given on the crossover and mutation probabilities in order to guarantee the convergence of the GA. In many others papers GA and fuzzy controllers are used but in a different way than the one that is

presented in this work: The GA is used to obtain a rule basis and membership functions in a dynamic way, so that the performance of the fuzzy controller is improved. An example of that use can be seen in [9] where a GA, called "improved GA", is used to adjust the rules and functions mentioned above. However, no additional piece of information is given to explain why that GA is better, but in [14] it is explained that an elitist version of this GA converges. The contribution of our work is to define a naive stochastic controller, to give sufficient conditions to the stochastic controller that is used to adjust the mutation probability, has to satisfy in order to guarantee the convergence of the GA. It also shows that this kind of controller is statistically as efficient as the fuzzy one.

In this work, a fuzzy controller and a stochastic controller were constructed where the input variable is the number of iterations N_g and the output variable is the mutation probability (p_m) . The convergence of the genetic algorithm using fuzzy controller can be seen in [13] while the convergence of the stochastic controller is presented in this work. The stochastic controller will be constructed in the simplest way possible, using a uniform distribution. This work thus illustrates that if the input variable is the number of interations, although the use of fuzzy controllers improves, in some sense, the genetic algorithm it is not worth using it because the same improvement can be obtained using a faster and lighter controller such as a stochastic controller. Many other input variables can be used in a fuzzy controller as those presented in [10] and in the references therein.

In Section 2 definitions and results concerning the non-homogeneous Markov chains that will be used in the rest of the paper are presented. In Section 3 the fuzzy controller with N_g as input variable and p_m as output variable is shortly described (which is extensively done in [4]). The stochastic controller is lighter but expresses the relevant characteristics of fuzzy controllers by means of statistical distributions. In addition convergence results are obtained. In Section 4, numerical comparisons between the elitst non-homogeneous genetic algorithm that uses a fuzzy controller and a stochastic controller for five special functions are presented and statistical properties are obtained.

2 Preliminaries

Following notation from the previous section let $f : E \to (0, \infty)$ be a function subject to a genetic algorithm in order to find

$$x^* = \text{argmax}\{f(x), x \in E\},$$

where E is a discretization of the domain of the function f. To proceed the following steps of the algorithm, such points are represented as binary vectors of length l, where l depends on the desired precision. A population of size N is considered, let $Z = \{(u_1, u_2, ..., u_N); u_i \in E, i = 1, 2, ..., N\}$ be the set of all populations of size N. Z is the state space of the Markov chain that is used to prove the convergence of the algorithm (see Campos *et al.*[1], Dorea *et al.*[5], Rojas Cruz and Pereira[14] and Rudolph[15]).

The evolution of the ENHGA is different from the evolution of the EGA just in the update of the values of the parameters p_m and p_c. Thus, the elitist algorithm can be summarized in the following sketch:

a)Choose randomly an initial population having N elements, each one being represented by a binary vector of length l, and create one more position, the $(N+1)$-th entry of the population vector, which will keep the best element from the N previous elements.

b)Repeat

1. perform selection with the first N elements
2. perform crossover with the first N elements
3. perform mutation with the first N elements
4. If the best element from this new population is better than that of the $(N+1)$-th position, change the $(N+1)$-th position by this better element, otherwise, keep the $(N+1)$-th position unchanged
5. perform p_c and p_m changes, as previously planned.

c) until some stopping criterion applies.

Denote this new state space by $\tilde{Z}$.

In Rojas Cruz and Pereira[14], it is shown that the ENHGA is a non-homogeneous Markov chain, with a finite state space $\tilde{Z}$, whose transition matrices are given by $P_n = SC_nM_n, \forall n \in \mathbb{N}$, where S, C_n, M_n are transition matrices which represent the selection, crossover and mutation stages respectively. Here the M_n is composed by the third and fourth steps described in the above sketch. In the same paper it is shown that there is a sequence $\{\alpha_n\}_{n\in\mathbb{N}}$ such that

$$\inf_{i\in\tilde{Z}, j\in\tilde{Z}^*} P_n(i,j) \geq \alpha_n,$$

where $\tilde{Z}^* \subset \tilde{Z}$, which contains all populations that have the optimum point as one of its points. The following results were obtained as corollaries to their main result.

Corollary 1: Let $\{X_n\}_{n\in\mathbb{N}}$ be the Markov chain which models the elitist non-homogeneous genetic algorithm, if the sequence above is such that $\sum_{k\geq 1}\alpha_k = \infty$ then

$$P(\lim_{n\to\infty} X_n \in \tilde{Z}^*) = 1. \quad (1)$$

A simpler condition to verify in actual implementations which guarantee the above result is

Corollary 2: Let $\{X_n\}_{n\in\mathbb{N}}$ be the Markov chain which models the elitist non-homogeneous genetic algorithm, if the mutation probabilities $\{p_m(n)\}_{n\in\mathbb{N}}$ are such that $p_m(n) > \gamma > 0$ for all $n \in \mathbb{N}$ or $\sum_n p_m(n)^l < \infty$ then (1) holds.

3 The Fuzzy Controller/The Stochastic Controller

A fuzzy controller has the ability to associate an output value with an input value. Its implementation involves four essential components: fuzzification, the inference method, the base rule and defuzzification, as shown in Figure 1 [11].

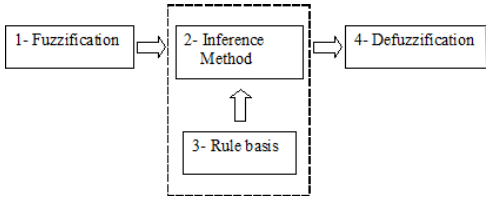


Fig. 1. Fuzzy controler scheme

Fuzzification is the first part of the process, and it consists of converting the numerical input into fuzzy sets. The second part involves definition of the inference method that provides the fuzzy output (or control) to be adopted by the controller, from each fuzzy input. The third step in the process (the rule base) involves a mathematical "translation" of the information comprising the knowledge base of the fuzzy system. Finally, defuzzification is an operation that converts a fuzzy set to a numerical value, which can be achieved using a technique such as the Center of Gravity method described by the formula (of weighted average of the membership function $\varphi_{(A)}$ of set A)

$$C(A) = \frac{\int u\varphi_{(A)}(u)du}{\int \varphi_{(A)}(u)du} \tag{2}$$

Consider the following construction of fuzzy sets for the variables N_g and p_m, using the state variables: low (L), average (A), and high (H). These sets were characterized by their membership functions, which define the extent to which a determined element does or does not belong to the set. Figures 2 and 3 show the membership functions for Ng and p_m.

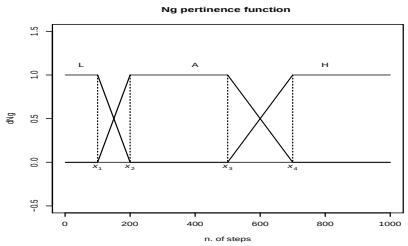


Fig. 2. Membership function for N_g

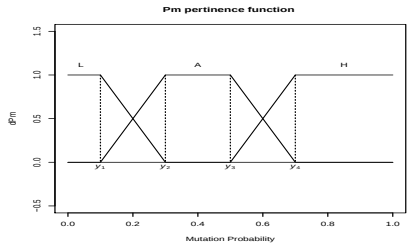


Fig. 3. Membership function for p_m

It was proved in [13] that

Theorem 1: If we use membership functions like those presented in figures 2 and 3 and the Mamdani inference method, based on max-min operators and the base rule consisting of three rules:

1. If Ng is low (L), then p_m is high (H);
2. If Ng is average (A), then p_m is average (A);
3. If Ng is high (H), then p_m is low (L).

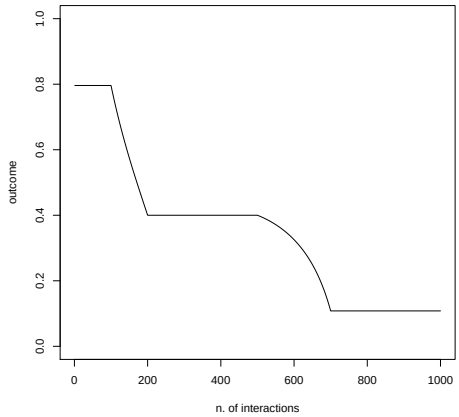


Fig. 4. Fuzzy controler outcome

Then the non-homogeneous genetic algorithm controlled by this fuzzy controller converges.

Now, let's observe more closely the possible outcomes of this controller. As time goes by, the number of iterations makes the rules be triggered, starting with $N_g = L$ what implies that $p_m = H$ is triggered, then $N_g = L$ and $N_g = A$ are trigged together and that implies $p_m = H$ and $p_m = A$ are triggered together too, and so on. Taking these rules into consideration and calculating (2) we obtain the outcome of the fuzzy controller as a function of the number of iterations illustrated in Figure 4.

Let us set up the following stochastic controller based on the outcome of the fuzzy controller

$$\text{out}(n) = \begin{cases} 0.8,\ 0 < n < 100 \\ U_1,\ 100 \leq n < 200 \\ 0.4,\ 200 \leq n < 500 \\ U_2,\ 500 \leq n < 700 \\ 0.1,\ n \geq 700 \end{cases}$$

where $U_1 \sim U[0.4, 0.8]$ and $U_2 \sim U[0.1, 0.4]$.

Remark. We note that the numbers in the definition of the stochastic controller are obtained by the membership functions. If the membership functions change, then these values are likely to change too.

Using this stochastic controller to adjust the mutation probability in the elitist non-homogeneous genetic algorithm, by corollary 2 of the former section, it converges. More generally

Theorem 2: If a stochastic controller is set up to adjust the mutation probability of an elistist non-homogeneous genetic algorithm, if its output satisfies the hypothesis of corollary 2, then (1) holds.

Remark. In Theorem 2, the input variable(s) were not estabilished, that is because we just need the output to satisfy the hypotheses of corollary 2.

4 Numerical Evaluations

We have used five test functions to compared the performance of three versions of the GA: (i) the EGA; (ii) the ENHGA using a fuzzy controller on the mutation probability p_m and (iii) the ENHGA using a stochastic controler on p_m. The functions are:

1. $f : [-2, 1] \times [-2, 1] \to \mathbb{R}$ given by

$$f(x) = 6 + x^2 - 3\cos(2\pi x) + y^2 - 3\cos(2\pi y)$$

2. $f : [-\frac{1280}{63}, \frac{1240}{63}] \times [-\frac{1280}{63}, \frac{1240}{63}] \to \mathbb{R}$ given by

$$f(x) = 0.5 - \frac{\sin(\sqrt{x^2 + y^2})^2 - 0.5}{(1 + 0.001(x^2 + y^2))^2}$$

3. $f : [-4, 2] \times [-4, 2] \to \mathbb{R}$ given by

$$f(x) = \frac{1}{0.3 + x^2 + y^2}$$

4. $f : [-2, 2] \times [-2, 2] \to \mathbb{R}$ given by

$$\begin{aligned} f(x) = [1 + (19 - 14x + 3x^2 - 14y + 6xy + 3y^2)(x + y + 1)^2] \cdot \\ [30 + (18 - 32x + 12x^2 + 48y - 36xy + 27y^2)(2x - 3y)^2] \end{aligned}$$

5. $f : [-5, 10] \times [0, 15] \to \mathbb{R}$ given by

$$f(x) = \left(y - \frac{5.1x^2}{4\pi^2} + \frac{5x}{\pi} - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos(x) + 10$$

All of the above domains were discretized so that x would lie in a grid of 2^{12} points. The maximum is known for each function within the respective grid. In the simulations we considered population sizes N of 10, 30 and 50 individuals. For each N and each function we ran 1000 trials and we recorded the number of trials in which the optimum was found within 1000 interations. These counts are displayed in the fourth column of Table 1. We also report in Table 1 the maximum number of interations needed to reach the optimum (worst case) for those instances where all 1000 trials were succesful. For example: given the first function, with $N = 10$, the EGA was succesful (before our fixed limit of 1000 interations) in 813 out of 1000 trials; still with the first function, with $N = 30$, the ENHGA with the fuzzy controller was succesful in all 1000 trials and all of them reached the maximum before the 716th interation.

Figures 5–9 show the evolution of our simulations and generally suggest that either form of controller provides an improvement over the simple EGA.

Statistics from those simulations are provided in Table 1. The results in Table 1, specifically the third column (S), show an enormous gain in performance when using a controller (either fuzzy (ii) or stochastic (iii)) over the standard algorithm, except for function 3 and perhaps 4 where all versions seem equally good. In other words, except for functions 3 and 4, the successes under algorithm (i), S_i, is much smaller than S_{ii} and S_{iii}. For function 1 with $N = 10$ the counts S seem less distant but a χ^2 test for equality of those three proportions (.813, .976, .980) reveal a significant difference (p-value < 0.001) which we attribute to S_i being smaller than S_{ii} and S_{iii}. The same test for equality of two proportions was carried out for those cases where neither S_{ii} nor S_{iii} is equal to 1000. These cases occurred in (Function, N) = (1, 10), (2, 10), (2, 30) and (5, 10). The corresponding χ^2 tests resulted in p-values bounded below by 0.18 so we can safely conclude that there is no statistical difference between the performances of algorithms (ii) and (iii) in terms of proportion of succesful trials. In addition, the worst-case statistic (Table 1, fourth column) strongly indicates that GA with the stochastic controller (iii) seems to require less interations to reach the maximum then the GA the fuzzy controller (ii): most such cases favour (iii) over (ii); the largest difference – 299 iterations – favouring (iii) was observed in case (3, 50) and the largest difference – 87 interations – favouring (ii) was observed in case (1, 30). Finally, there is an enormous gain in execution time when comparing (iii) with (ii) given the simplicity of the stochastic controller over the fuzzy controller. Since our functions were rather simple we did not experience prohibite execution times but this can be an issue in larger problems.

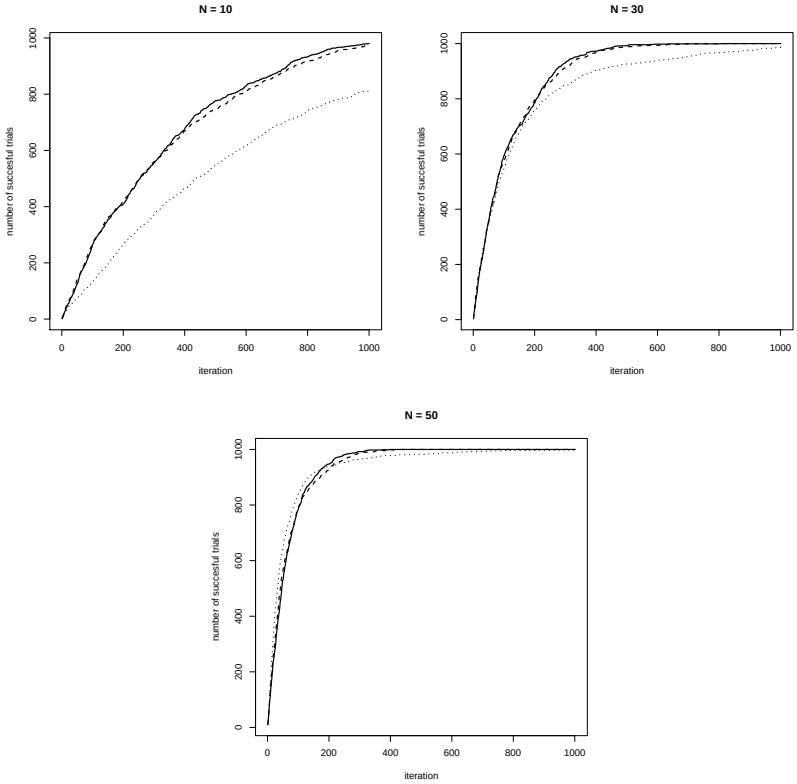


Fig. 5. Number of succesful trials across 1000 simulations each limited to 1000 iterations using function 1 and different population sizes N. Dotted line represents scheme (i)–no controller, dashed line scheme (ii)–fuzzy, and full line scheme (iii)–stochastic.

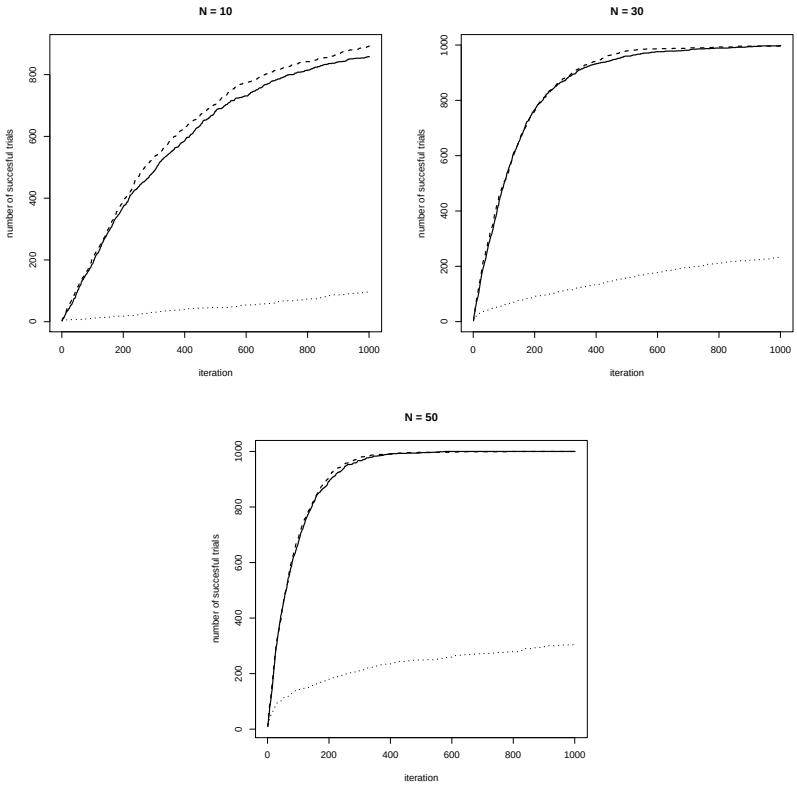


Fig. 6. Number of succesful trials across 1000 simulations each limited to 1000 iterations using function 2 and different population sizes N. Dotted line represents scheme (i)–no controller, dashed line scheme (ii)–fuzzy, and full line scheme (iii)–stochastic.

5 Conclusions

The literature has many studies in which fuzzy controllers have been used to adjust the parameters of a genetic algorithm, some of them using the number of iterations of the algorithm as input variable. This work presents a very light stochastic controller to adjust the mutation probability which does not require any expensive calculation in its evolution. The idea was to compare this stochastic controller with a fuzzy one. We have given conditions under which such controllers make the elitist non-homogeneous genetic algorithm converge. Our simulations using five test functions suggest that the elitist genetic algorithm is greatly improved by the use of the two controllers considered. Most interestingly is the fact that we observed no statistical difference between the fairly expensive fuzzy controller and the much simpler stochastic one. In fact, the performance was even better under the simple stochastic controller in many cases. Thus, when the input variable is the number of interations, the computational effort can be reduced by using the light stochastic controller to adjust the mutation rate instead of the fuzzy controller without any loss in performance.

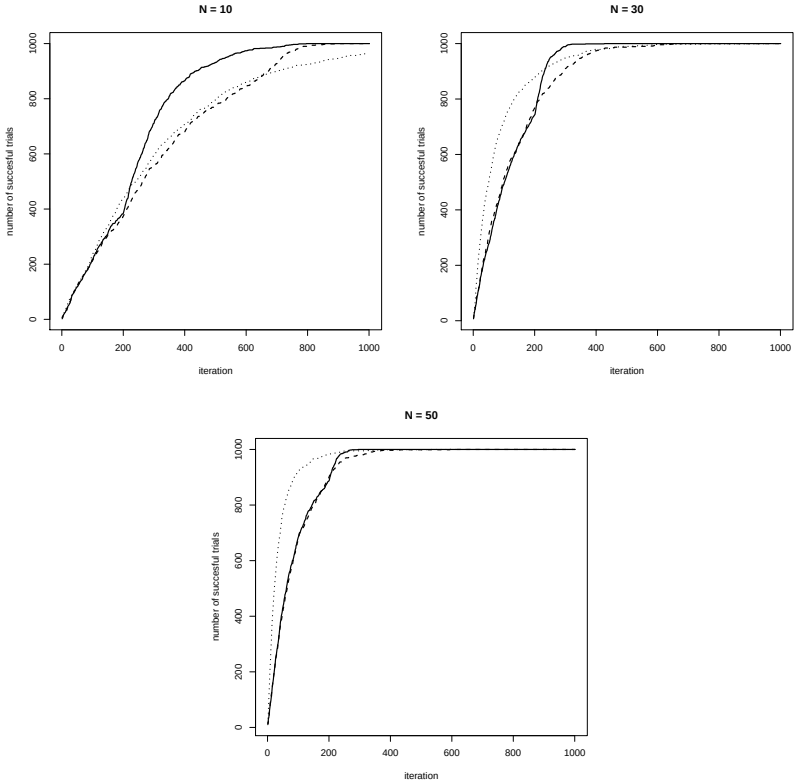


Fig. 7. Number of succesful trials across 1000 simulations each limited to 1000 iterations using function 3 and different population sizes N. Dotted line represents scheme (i)–no controller, dashed line scheme (ii)–fuzzy, and full line scheme (iii)–stochastic.

Acknowledgment

The authors are grateful to PROCAD/CAPES and CNPq for financial support.

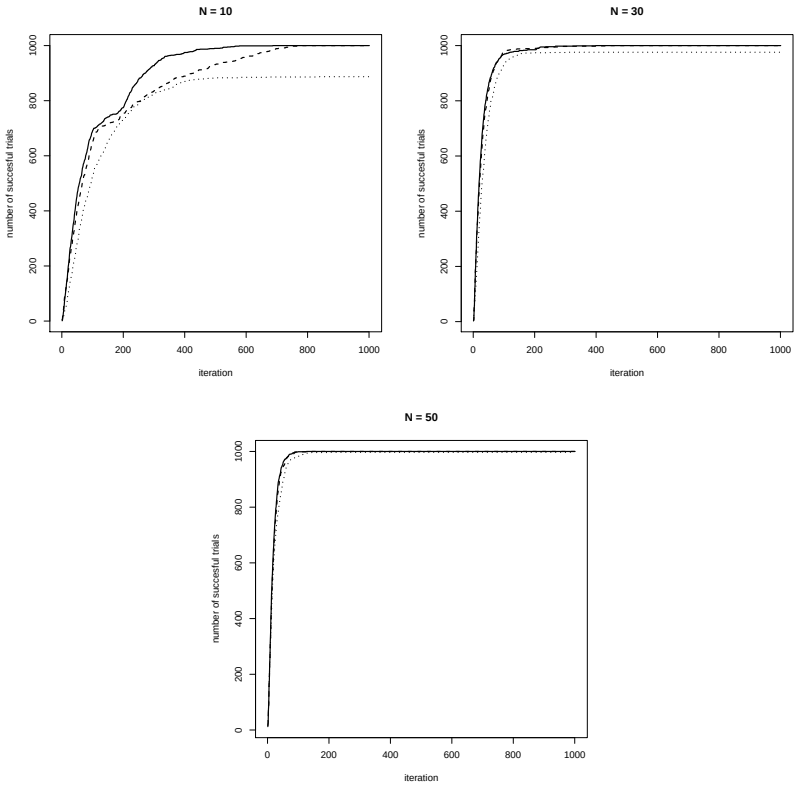


Fig. 8. Number of succesful trials across 1000 simulations each limited to 1000 iterations using function 4 and different population sizes N. Dotted line represents scheme (i)–no controller, dashed line scheme (ii)–fuzzy, and full line scheme (iii)–stochastic.

References

1.Campos, V.S.M., Pereira, A.G.C. and Rojas Cruz, J.A., *Modeling the Genetic Algorithm by a Non-Homogeneous Markov Chain: Weak and Strong Ergodicity.* Theory of Probability and its Applications, v.57,Issue 1(2012),p.185-192.

2.V.S.M. Campos, A.G.C. Pereira, L.A. Carlos and I.A.S. de Assis. Algoritmo Genético por cadena de Markov homogénea versus no-homogénea : Un estudio comparativo. *Journal of the Chilean Institute of Operations Research*, vol 2 (2012) p. 30-35.

3.V.S.M. Campos and A.G.C. Pereira, *A study on the Elitist Non-homogeneous Genetic Algorithm.* Biometrie und Medizinische Informatik. Greifswalder Seminarberichte (to appear).

4.L.A. Carlos, *Genetic Algorithms Convergence Analysis: A Markov Chain approach.* PhD Thesis presented at Electric Engineering Program of the Federal University of Rio Grande do Norte- Brazil (in portuguese)(2013).

5.Dorea, C.C.Yu, Guerra Júnior, J.A., Morgado, R., Pereira, A.G.C. *Multistage Markov Chain Modeling of the Genetic Algorithm and Convergence Results.*. Numerical Functional Analysis and Optimization, v.31, (2010) p. 164-171.

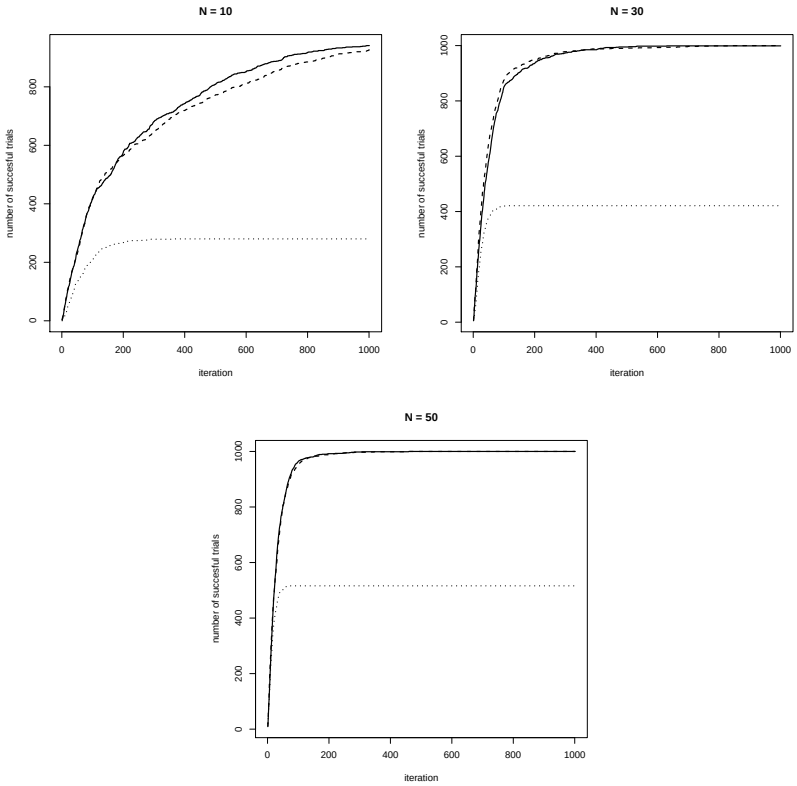


Fig. 9. Number of succesful trials across 1000 simulations each limited to 1000 iterations using function 5 and different population sizes N. Dotted line represents scheme (i)–no controller, dashed line scheme (ii)–fuzzy, and full line scheme (iii)–stochastic.

6.Grefenstette, J.J., *Optimization of control parameters for genetic algorithm*, IEEE Transactions on systems, man, and cybernetics, vol. smc-16, n.1, pp.122-128, (1986).

7.Goldberg, D.E. *Genetic Algorithms in Search, Optimizations and Machine Learning.* Assison Wesley, Teading, MA (1989).

8.Holland, J.H. *Adaptation in natural and artificial systems.* Ann Arbor: The University of Michigan Press,(1975).

9.Lam, H.K., Ling, S.H., Leung F.H.F. and Tam, P.K.S., *Function estimation using a fuzzy neural-fuzzy network and an improved genetic algorithm.* International Journal of Approximate Reasoning, Volume 36, Issue 3, pp. 243–260 (2004).

10.Lee, M.A. and Takagi,H., *Dynamic control of genetic algorithms using fuzzy logic techniques*, Procceding of 5th International conference on Genetic Algorithms (ICGA'93), Urbana-Champaign, IL, pp. 76-83, (1993).

11.Pedrycz, W., *Fuzzy Sets Engineering.* CRC Press (1995).

12.Pereira, A.G.C. and Andrade, B.B. On the Genetic Algorithm with Adaptive Mutation Rate and Statistical Applications (submitted).

13.Pereira, A.G.C.. and Roveda,J.A.F., *Convergence analysis of an elitist non-homogeneous genetic algorithm with mutation probability adjusted by a fuzzy controller*, (submitted) .

14.Rojas Cruz, J.A., Pereira, A.G.C., *The Elitist Non-homogeneous Genetic Algorithm: Almost sure convergence*. Statistics and Probability Letters, v.83 (2013), p. 2179-2185.
15.Rudolph, G. *Convergence Analysis of Canonical Genetic Algorithms*. IEEE Transactions on Neural Networks, V.5, (1994),96-101.
16.Yun, Y., Gen, M. *Performance Analysis of Adaptative Genetic Algorithms with Fuzzy Logic and Heuristics*. Fuzzy Optimization and Decision Making, V.2, (2003), 161-175.

Table 1. Simulation results considering the five functions described in the text and three versions of the GA: (i) the EGA, (ii) the ENHGA using a fuzzy controller on p_m and (iii) the ENHGA using a stochastic controler on p_m.

Function	N	GA	Successes (S)	Worst case (iterations)
1	10	(i)	813	-
		(ii)	976	-
		(iii)	980	-
	30	(i)	986	-
		(ii)	1000	716
		(iii)	1000	803
	50	(i)	998	-
		(ii)	1000	416
		(iii)	1000	398
2	10	(i)	96	-
		(ii)	894	-
		(iii)	859	-
	30	(i)	232	-
		(ii)	996	-
		(iii)	998	-
	50	(i)	304	-
		(ii)	1000	779
		(iii)	1000	574
3	10	(i)	964	-
		(ii)	1000	904
		(iii)	1000	795
	30	(i)	1000	945
		(ii)	1000	660
		(iii)	1000	426
	50	(i)	1000	517
		(ii)	1000	593
		(iii)	1000	294
4	10	(i)	887	-
		(ii)	1000	776
		(iii)	1000	686
	30	(i)	976	-
		(ii)	1000	445
		(iii)	1000	398
	50	(i)	997	-
		(ii)	1000	125
		(iii)	1000	126
5	10	(i)	280	-
		(ii)	926	-
		(iii)	941	-
	30	(i)	421	-
		(ii)	1000	833
		(iii)	999	-
	50	(i)	516	-
		(ii)	1000	453
		(iii)	1000	459

Modeling of multivariate skewness measure distribution

Margus Pihlak

Tallinn University of Technology (e-mail: margus.pihlak@ttu.ee)

Abstract. In this paper the distribution of random variables skewness measure is modeled. Firstly we present some results of matrix algebra useful in multivariate statistical analyses. Then we apply the central limit theorem on modeling of multivariate skewness measure distribution. That skewness measure is introduced in [6].
Keywords: Central limit theorem, Multivariate skewness measure, Skewness measure distribution.

1 Introduction and basic notations

In the firs section we introduce some notations used in the paper. The k-dimensional zero vector is denoted as $\mathbf{0}_k$. The transposed matrix $\mathbf{A}$ is denoted as $\mathbf{A}'$.

Let us have random vectors $\mathbf{X}_i = (\mathbf{X}_{i1}, \mathbf{X}_{i2}, \ldots, \mathbf{X}_{ik})'$ where index $i = 1, 2, ..., n$ is for observations and k denotes number of variables. These random vectors are independent and identically distributed copies (each copy for one observations) of a random k-vector $\mathbf{X}$. Let

$$\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i$$

and

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}_i - \overline{\mathbf{x}})(\mathbf{X}_i - \overline{\mathbf{x}})'$$

be the estimators of the sample mean $E(\mathbf{X}) = \mu$ and the covariance matrix $D(\mathbf{X}) = \mathbf{\Sigma}$ respectively.

Now we present matrix operations used in this paper. One of the widely used matrix operation in multivariate statistics is Kronecker product (or tensor product) $\mathbf{A} \otimes \mathbf{B}$ of matrices $\mathbf{A} : m \times n$ and $\mathbf{B} : p \times q$ which is defined as a partitioned matrix

$$\mathbf{A} \otimes \mathbf{B} = [a_{ij}\mathbf{B}], i = 1, 2, \ldots, m; j = 1, 2, \ldots, n.$$

3^{rd} *SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal*
C. H. Skiadas (Ed)

By means of Kronecker product we can present the third and the fourth order moments of vector $\mathbf{X}$:

$$m_3(\mathbf{X}) = E(\mathbf{X} \otimes \mathbf{X}' \otimes \mathbf{X})$$

and

$$m_4(\mathbf{X}) = E(\mathbf{X} \otimes \mathbf{X}' \otimes \mathbf{X} \otimes \mathbf{X}').$$

The corresponding central moments

$$\overline{m_3}(\mathbf{X}) = E\{(\mathbf{X} - \mu) \otimes (\mathbf{X} - \mu)' \otimes (\mathbf{X} - \mu)\}$$

and

$$\overline{m_4}(\mathbf{X}) = E\{(\mathbf{X} - \mu) \otimes (\mathbf{X} - \mu)' \otimes (\mathbf{X} - \mu) \otimes (\mathbf{X} - \mu)'\}.$$

The third order moment of random vector $\mathbf{X}$ is $k^2 \times k$-matrix and its fourth order moment is $k^2 \times k^2$-matrix.

The operation vec($\mathbf{X}$) denotes a mn-vector obtained from $m \times n$-matrix by stacking its columns one under another in natural order. For the properties of Kronecker product and vec-operator the interested reader is referred to [2] or [4]. In the next section skewness measure will be defined be means of the star-product of the matrices. The star-product was introduced in [7] where some basic properties of the operation were presented and proved.

Definition 1. *Let us have matrix* $\mathbf{A} : m \times n$ *and a partitioned matrix* $\mathbf{B} : \times ns$ *consisting of* $r \times s$*-blocks* $\mathbf{B}_{ij}, i = 1, 2, \ldots m; j = 1, 2, \ldots n$. *Then the star-product* $\mathbf{A} * \mathbf{B}$ *is a* $r \times s$*-matrix*

$$\mathbf{A} * \mathbf{B} = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} \mathbf{B}_{ij}.$$

The star product is inverse operation of Kronecker product in sense of increasing and decreasing of matrix dimensions. One of the star-product applications is presented in the paper [12]. Let us give an example how the star product works.

Example. Let us have matrices $A : 2 \times 2$ and partitioned matrix $B : 4 \times 4$ with 2×2-blocks $B_{11}, ..., B_{22}$. Then

$$A * B = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} * \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} =$$

$$= a_{11}B_{11} + ... + a_{22}B_{22}.$$

We also use the matrix derivative defined following H. Neudecker in [10].

Definition 2. *Let the elements of the matrix* $\mathbf{Y} : r \times s$ *be functions of matrix* $\mathbf{X} : p \times q$. *Assume that for all* $i = 1, 2, \ldots p$; $j = 1, 2, \ldots q$; $k = 1, 2, \ldots r$ *and* $l = 1, 2, \ldots s$ *partial derivatives* $\dfrac{\partial y_{kl}}{\partial x_{ij}}$ *exist and are continuous in an open set A. Then the matrix* $\dfrac{d\mathbf{Y}}{d\mathbf{X}}$ *is called matrix derivative of matrix* $\mathbf{Y} : r \times r$ *by matrix* $\mathbf{X} : p \times q$ *in a set A, if*

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d}{dvec'(\mathbf{X})} \otimes vec(\mathbf{Y})$$

where

$$\frac{d}{d\mathrm{vec}'(\mathbf{X})} = \left(\frac{\partial}{\partial x_{11}} \quad \cdots \quad \frac{\partial}{\partial x_{p1}} \quad \cdots \quad \frac{\partial}{\partial x_{1q}} \quad \cdots \quad \frac{\partial}{\partial x_{pq}} \right).$$

Matrix derivative defined by Definition 2 is called *Neudecker* matrix derivative. This matrix derivative has been in last 40 years a useful tool in multivariate statistics.

2 Multivariate measures of skewness

In this section we present multivariate skewness measure by means of matrix operation described above. A skewness measure in multivariate case was introduced in Mardia [8]. Mori et al [9] have introduced a skewness measure as a vector. B. Klar in [3] has given thorough overview of the skewness problem. In this paper is also examined asymptotic distribution of different skewness characteristics. In Kollo [6] a skewness measure vector is introduced and applied in Independent Component Analyses (ICA).

The skewness measure in multivariate case is presented through the the third order moments as

$$\mathbf{s}(\mathbf{X}) = E(\mathbf{Y} \otimes \mathbf{Y}' \otimes \mathbf{Y}) \tag{1}$$

where

$$\mathbf{Y} = \mathbf{\Sigma}^{-1/2}(\mathbf{X} - \mu).$$

In Kollo [6] a skewness measure based on (1) is introduced by means of the star product:

$$\mathbf{b}(\mathbf{X}) = \mathbf{1}_{k\times k} * \mathbf{s}(\mathbf{X}) \tag{2}$$

where

$$\mathbf{1}_{k\times k} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}.$$

In [5] the Mardia's skewness measure is presented as through the third order moment:

$$\beta = \mathrm{tr}(m_3'(\mathbf{Y} m_3(\mathbf{Y})$$

where operation tr denotes the trace of matrix. A sample estimate $\widehat{\mathbf{b}(\mathbf{X})}$ of the skewness vector (2) we can present in the form:

$$\widehat{\mathbf{b}(\mathbf{X})} = \mathbf{1}_{k\times k} * \sum_{i=1}^{n} (\mathbf{y}_i \otimes \mathbf{y}_i' \otimes \mathbf{y}_i) \tag{3}$$

where

$$\mathbf{y}_i = \mathbf{S}^{-1/2}(\mathbf{x}_i - \overline{\mathbf{x}})$$

$\overline{\mathbf{x}}$ and $\mathbf{S}$ are the sample mean vector and the sample covariance matrix of the initial sample $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$. The estimator $\widehat{\mathbf{b}(\mathbf{X})}$ is random k-vector.

3 Modeling the multivariate skewness measure distribution

In this section we model the distribution of the random variable $\widehat{\mathbf{b}(\mathbf{X})}$ defined by equality (3). From this equality concludes that $\widehat{\mathbf{b}(\mathbf{X})}$ is k-vector. Let us have a sequence of independent and identically distributed random vectors $\{\mathbf{X}_n\}_{n=1}^{\infty}$. Let $E(\mathbf{X}_n) = \mu$ and $D(\mathbf{X}_n) = \mathbf{\Sigma}$. Then according to the central limit theorem the distribution of the random vector $\sqrt{n}(\mathbf{X}_n - \mu)$ converges to the normal distribution $N(\mathbf{0}_k, \mathbf{\Sigma})$ where $\mathbf{0}_k$ denotes k-dimensional zero vector.

Let us introduce $k^2 + k$-vector

$$\mathbf{Z}_n = \begin{pmatrix} \overline{\mathbf{x}} \\ \mathrm{vec}(\mathbf{S}) \end{pmatrix}.$$

Applying the central limit this random vector we get the following convergence in distribution

$$\sqrt{n}(\mathbf{Z}_n - E(\mathbf{Z}_n)) \mapsto N(\mathbf{0}_{k^2+k}, \mathbf{\Pi})$$

where $(k^2 + k) \times (k^2 + k)$-dimensional partitioned matrix

$$\mathbf{\Pi} = \begin{pmatrix} \mathbf{\Sigma} & \overline{m_3}'(\mathbf{X}) \\ \overline{m_3}(\mathbf{X}) & \mathbf{\Pi}_4 \end{pmatrix}.$$

The $k^2 \times k^2-$block $\mathbf{\Pi}_4 = \overline{m_4}(\mathbf{X}) - \mathrm{vec}(\mathbf{\Sigma})\mathrm{vec}'(\mathbf{\Sigma})$ ([11]). This convergence can be generalized by means of the following theorem.

Theorem 1. *Let $\{\mathbf{Z}_n\}_{n=1}^{\infty}$ be a sequence of $k^2 + k$-component random vectors and ν be a fixed vector such that $\sqrt{n}(\mathbf{Z}_n - \nu)$ has the limiting distribution $N(\mathbf{0}_{k^2+k}, \mathbf{\Pi})$ as $n \to \infty$. Let the function $g : R^{k^2+k} \to R^k$ have continuous partial derivatives at $\mathbf{z}_n = \nu$. Then the distribution of random vector $\sqrt{n}(g(\mathbf{Z}_n) - g(\nu))$ converges to the normal distribution $N(\mathbf{0}_{k^2+k}, g'_{\mathbf{z}_n}\mathbf{\Pi}g_{\mathbf{z}_n})$ where $(k^2 + k) \times k$-matrix*

$$g_{\mathbf{z}_n} = \left.\frac{dg(\mathbf{z}_n)}{d\mathbf{z}_n}\right|_{\mathbf{z}_n=\nu}$$

is Neudecker matrix derivative at $\mathbf{z}_n = \nu$. The proof of Theorem 1 can be found in the book of T. W. Andreson ([1], page 132).

In our case the function

$$g(\mathbf{z}_n) = g\begin{pmatrix} \overline{\mathbf{x}} \\ \mathrm{vec}(\mathbf{S}) \end{pmatrix} = \widehat{\mathbf{b}(\mathbf{X})}.$$

Applying Theorem 1 we get the following convergence in distribution:

$$\sqrt{n}(\widehat{\mathbf{b}(\mathbf{X})} - \mathbf{b}(\mathbf{X})) \mapsto N(\mathbf{0}_k, \mathbf{\Sigma}_{\mathbf{b}}).$$

Here the $k \times k$-matrix

$$\mathbf{\Sigma}_{\mathbf{b}} = g'_{\mathbf{z}_n}\mathbf{\Pi}g_{\mathbf{z}_n}\Big|_{\mathbf{z}_n=(\,\mu \quad \mathrm{vec}'(\mathbf{\Sigma})\,)'} =$$

$$= \left(\frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\overline{\mathbf{x}}} \quad \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\mathbf{S}} \right) \begin{pmatrix} \mathbf{\Sigma} & \overline{m_3}'(\mathbf{X}) \\ \overline{m_3}(\mathbf{X}) & \mathbf{\Pi}_4 \end{pmatrix} \begin{pmatrix} \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\overline{\mathbf{x}}} \\ \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\mathbf{S}} \end{pmatrix} \Big|_{\overline{\mathbf{x}}=\mu, \mathbf{S}=\mathbf{\Sigma}} =$$

$$= \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\overline{\mathbf{x}}} \mathbf{\Sigma} \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\overline{\mathbf{x}}}' \Big|_{\overline{\mathbf{x}}=\mu, \mathbf{S}=\mathbf{\Sigma}} + \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\mathbf{S}} \overline{m_3}(\mathbf{X}) \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\overline{\mathbf{x}}}' \Big|_{\overline{\mathbf{x}}=\mu, \mathbf{S}=\mathbf{\Sigma}} +$$

$$+ \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\overline{\mathbf{x}}} \overline{m_3}(\mathbf{X})' \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\mathbf{S}}' \Big|_{\overline{\mathbf{x}}=\mu, \mathbf{S}=\mathbf{\Sigma}} + \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\mathbf{S}} \mathbf{\Pi} \frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\mathbf{S}}' \Big|_{\overline{\mathbf{x}}=\mu, \mathbf{S}=\mathbf{\Sigma}}.$$

Here $\frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\overline{\mathbf{x}}}$ and $\frac{d\widehat{\mathbf{b}(\mathbf{X})}}{d\mathbf{S}}$ are $k \times k$- and $k \times k^2$-dimensional *Neudecker* matrix derivatives respectively.

Knowing the skewness measure distribution enables to estimate asymmetry of k-dimensional data. We can find for α-confidence interval for skewness vector $\mathbf{b}(\mathbf{X})$. The problem of asymmetry is actual on environmental data for example.

Acknowledgement

This paper is supported by Estonian Ministry of Education and Science target financed theme Nr. SF0140011s09.

References

1. Anderson, T. W. (2003) *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.
2. Harville, A. (1997) *Matrix Algebra from a Statistican's Perspective*. Springer, New York.
3. Klar, B. (2002) A Treatment of Multivariate Skewness, Kurtosis, and Related Statistics. *Journal of Multivariate Analysis*, **83**, 141-165.
4. Kollo, T., von Rosen, D. (2005) *Advanced Multivariate Statistics with Matrices*. Springer, Dordrecht.
5. Kollo T., Srivastava M. S. (2004) Estimation and testing of parameters in multivariate Laplace distribution. *Comm. Statist.*, **33**, 2363-2687.
6. Kollo, T. (2008) Multivariate skewness and kurtosis measures with an application in ICA. *Journal of Multivariate Analyses*, **99**, 2328-2338.
7. MacRae, E. C. (1974) Matrix derivatives with an applications to an adaptive linear decision problem. *The Annals of Statistics*, **7**, 381-394
8. Mardia, K. V. (1970) Measures of multivariate skewness and kurtosis measures with applications. *Biometrika*, **57**, 519-530.
9. Mori, T. F., Rohatgi, V. K., Szkely. (1993) On multivariate skewness and kurtosis. *Theory Probab. Appl*, **38**, 547-551.
10. Neudecker, H. (1969) Some theorems on matrix differentiations with special reference to Kronecker matrix products. *J. Amer. Stat. Assoc.*, **64**, 953-963.
11. Parring, A-M. (1979) Estimation asymptotic characteristic function of sample (in Russian). *Acta et Commetationes Universitatis Tartuensis de Mathematica*, **492**, 86-90.
12. Pihlak, M. (2004) Matrix integral. *Linear Algebra and Its Applications*, **388**, 315-325

Diffusion Maps in the Reconstruction of Nonlinear Signals

Lúcia M. S. Pinto[1], Ricardo Fabbri[2], and Francisco D. Moura Neto[2]

[1] Escola Nacional de Ciências Estatísticas
Instituto Brasileiro de Geografia e Estatística
Rua André Cavalcanti, 106, Rio de Janeiro, RJ 20231-050, Brazil
(e-mail: lumasapin@hotmail.com)
[2] Polytechnic Institute, Rio de Janeiro State University
Rua Bonfim, 25, Nova Friburgo, RJ 28625-570, Brazil,
(e-mail: [rfabbri,fmoura]@iprj.uerj.br)

Abstract. Diffusion maps have proven to be very useful for dimensionality reduction of high dimensional data sets. This method has been introduced by Coifman *et al.* citecoifman2005geometric. Following the program set forth by Munford and Desolneux [10], which establishes a feedback architecture for data recognition and reconstruction, we construct a procedure for the regularized reconstruction of signals, based on the distance from the training data set and otimization of diffused data. The results show the robustness of the methodology.
Keywords: Diffusion maps, Dimensionality reduction, PCA, Regularization.

1 Introduction

The aim of this work is to investiate one of the most recent techniques for dimensionality reduction of data in data modeling, looking in particular at pattern recognition and the related problem of reconstruction from a few parameters.

In a general sense, given n data points in $\mathbb{R}^d$, $X^1, X^2, \ldots, X^n$, the dimensionality reduction algorithms attempt to find n points in $\mathbb{R}^k$, $Y^1, Y^2, \ldots, Y^n$, such that each Y^i represents the corresponding X^i, preferably with k much less than d (in fact, we are interested in reducing the dimensionality of the data) in such a way as to preserve, as much as possible, the inter-relation of the data points in the new set, in the same way as in the original data set. Several methods have been used with this aim since the classic PCA (Principal Components Analysis) till the Spectral MDS, Aflalo and Kimmel[1], which shows the continued importance of this topic.

In this article we explore the diffusion maps method as a powerful tool for dimensionality reduction, in special for the recognition and reconstruction of signals coming from the quatification of qualities of physical systems.

We are interested in the problem in the perspective of *pattern theory* which consists in the search for a feedback structure, where *bottom-up* and *top-down*

3rd SMTDA Conference Proceedings, 11-14 June 2014, Lisbon Portugal
C. H. Skiadas (Ed)

algorithms are re-feed and modified for a better understanding of the physical system. In general terms we can interpret a bottom-up algorithm as being represented by a function G, an algorithm, mapping the space of signals in a space of signal features,

$$\begin{array}{ccc} \text{Signals space} & \text{Analysis} & \text{Signal features space} \\ \mathbb{R}^d \supset E & \xrightarrow{G} & F \subset \mathbb{R}^k \ , \end{array}$$

that is, atributing parameters (in F) to the signals (in E).

The set of features can have high dimension. When E is a diferentiable manifold of dimension e, it would be adequate that the dimensionality reduction algorithm G would take E into $G(E) \subset F$, possibly still a manifold of dimension e in $\mathbb{R}^k$, with $e \leq k < d$, (hopefully $k \ll d$). Therefore, a signal that required d real numbers to its apecification, would be represented by just k real numbers.

In this setting, the recognition of a signal X^* would be to identify $\tilde{X} \in E$ sharing similar features with X^*, that is,

$$G(X^*) \text{ near, or in the same class as, } G(\tilde{X}) \ .$$

Having the features of a signal, $Y^* \in F$, or near to F, the aim of reconstruction would be to determine a signal, $X^* \in \mathbb{R}^d$ (not necessarily belonging to te training set E but, preferably, close), such that

$$G(X^*) \cong Y^*$$

and that X^* could be accepted as a real signal of the physical system. In ths way we can imagine a function, or an algorithm, close to an inverse of G,

$$\begin{array}{ccc} \text{Space of signals} & \text{Synthesis} & \text{Signal features space} \\ \mathbb{R}^d \supset E & \xleftarrow{H} & F \subset \mathbb{R}^k \end{array}$$

in such a way that $X^* \cong H(Y^*)$.

Appart from the *bottom-up* stage, there is the *top-down* stage, and the two algorithms interact. Given X^*, a signal with the properties close to the detected ones $G(X^*)$ (in low dimension) it is synthetized, possibly following a stochastic model sufficiently simple and compared with the input signal. In essence, one computes $\tilde{X} \cong H(G(X^*))$ and compares X^* with $\tilde{X}$, adopting a feedback architecture.

Diffusion maps have great potential in this scenario. However this method has not been conceived with the full set of tools for the construction of machine learning systems based on *recognition by synthesis* in the framework of pattern theory. One of the contributions of this work is that we exploit these ideas to shed light on the problem of recognition.

2 Diffusion maps

Diffusion maps is one of the most recent and promissing non-linear dimensionality reduction techniques. This technique allows mapping distances in a

convenient form, in the sense that a diffusion distance discussed below between the input data (training set) approaches the euclidean distance between their images by the diffusion mapping.

The initial step is to construct a graph where each element X^i of a data set $E = \{X^1, X^2 \ldots, X^n\}$ becomes a node of the graph, while the weight of edge joining X^i e X^j, w_{ij}, are recorded as the ij entry of an afinity matrix, W.

It is usual to express the afinity by means of a gaussian kernel given by $W(X^i, X^j) = \exp(-\frac{\|X^i - X^j\|^2}{\varepsilon})$, where ε depends on the problem. One can interpret $\sqrt{\varepsilon}$ as the size of a neighborhood and it is based on the knowledge of the structure and the density of the data set. This kernel defines a local geometry of the data set. Here we choose ε as a function of the diameter, r, of the data set.

Coifman and Lafon [4] present three normalizations for a family of diffusion maps: $(W^\alpha)_{ij} = \frac{w_{ij}}{d_i^\alpha d_j^\alpha}$, where $d_i^\alpha = (\sum_{k=1}^n w_{ik})^\alpha$ is the degree of the i^{th} node of the original graph to the power α and w_{ij} is computed by the gaussian kernel. They emphasize three values for α. When $\alpha = 0$ this corresponds to the classical normalized laplacian of a graph, $\alpha = 1/2$ corresponds to the Fokker-Plank operator and $\alpha = 1$ leads to the Laplace-Beltrami operator. Here we stick to $\alpha = 1$.

We normalize the weight matrix W. Let $d_i = (D)_{ii} = \sum_{j=1}^n w_{ij}$ and $p_{ij} = \frac{w_{ij}}{d_i}$. Matrix $P = D^{-1}W$, whose entries are p_{ij}, is a Markov matrix, for a Markov process where the states are the nodes of the graph and the transition probability matrix is P.

Considering increasing powers of P, $P^t = (D^{-1}W)^t$, the Markov process incorporates more and more the intrinsec geometry of the data set. Since p_{ij} is the one-step transition probability from X^i to X^j, the ij entry of P^t, p_{ij}^t, is the transition probability from X^i to X^j in t steps, that is, the probability associated with the set of all paths of lenght t leaving X^i and arriving at X^j, reconstructing the geometry of the data set from local connectivity.

3 Diffusion distance

To the Markov process described previously, there corresponds a family of diffuion distances, $D_t(X^i, X^j)$. This family measures the connectivity between points X^i and X^j by paths of lenght t in the data set. The *diffusion distance* between X^i e X^j, for each fixed t, is defined by

$$D_t(X^i, X^j) = \left(\sum_{X^r \in E} \frac{(p_{ir}^t - p_{jr}^t)^2}{\sigma_r} \right)^{1/2},$$

where $\sigma_r = \frac{d_r}{\sum_{i=1}^n d_i}$.

The difusion distance can be rewritten as

$$D_t(X^i, X^j) = (\mathrm{tr}\,(D))^{\frac{1}{2}} \left(\sum_{k=1}^{n-1} \lambda_k^{2t} (v_k(i) - v_k(j))^2 \right)^{1/2}, \qquad (1)$$

where $\boldsymbol{v_k}$, λ_k are, respectively, the eigenvectors and the eigenvalues of the Markov matrix.

Motivated by this expression, the diffusion map is defined in the following way. Let $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ be eigenvectors from the right of $P = D^{-1}W$ associated to $\lambda_0 = 1 \geq \lambda_1 \geq \ldots \geq \lambda_{n-1} \geq -1$. For each fixed t, the diffusion map is $\mathcal{D}_t$ such that

$$\mathcal{D}_t(X^i) = \begin{pmatrix} \lambda_1^t v_1(i) \\ \lambda_2^t v_2(i) \\ \vdots \\ \lambda_{n-1}^t v_{n-1}(i) \end{pmatrix}, \tag{2}$$

for each X^i in the training data set. (There is no need to use λ_0 since $\boldsymbol{v}_0$ is a constant vector.)

We can verify that $D_t(X^i, X^j) = (\mathrm{tr}\,(D))^{\frac{1}{2}} \parallel \mathcal{D}_t(X^i) - \mathcal{D}_t(X^j) \parallel$, and in this way the diffusion distance between the original data is proportional to the euclidean distance of its features.

The parameter t of the Markov process works as a type of scaling factor. The larger t is the bigger is the scale considered in the modeling of the data. By varying t one gets a kind of multi-scale analysis of the data set.

Since the absolute value of the eigenvalues are between 0 and 1, for increasing values of t in the stochastic process allow us to keep few components in the diffusion map to analyse data. In fact, for t large enough, we shall have several insignificant $(\lambda_k)^t$, and several terms in $\parallel \mathcal{D}_t(X^i) - \mathcal{D}_t(X^j) \parallel$ contributing very little for the distance between X^i and X^j, and can be neglected. Therefore, for large t, it is possible to consider few components of the diffusion map.

If, in addition, W is positive semi-definite then the eigenvalues of P are between zero and one. In this case, if we let k be the number of components, chosen as a function of t, we can rewrite (1) in an approximate way,

$$\begin{aligned} D_t(X^i, X^j) &\cong (\mathrm{tr}\,(D))^{\frac{1}{2}} \left(\sum_{s=1}^{k} \lambda_s^{2t} (v_s(i) - v_s(j))^2 \right)^{1/2} \\ &\cong (\mathrm{tr}\,(D))^{\frac{1}{2}} \parallel \mathcal{D}_t(X^i) - \mathcal{D}_t(X^j) \parallel, \end{aligned}$$

where

$$\mathcal{D}_t(X^i) \cong \begin{pmatrix} \lambda_1^t v_1(i) \\ \lambda_2^t v_2(i) \\ \vdots \\ \lambda_k^t v_k(i) \end{pmatrix}.$$

Therefore, the diffusion distance between X^i and X^j is almost the same as the euclidean distance between their images in $\mathbb{R}^k$ which, in many practical applications, has dimension $k \ll n$, Lafon and Lee [8]. We also remark that, as the scale parameter t increases, the features of the data, that is, their images by the diffusion map, tend to merge together since $\mathcal{D}_t(X^i) \to 0$, when $t \to +\infty$, for every single data point.

4 Pre-image

The pre-image problem consists in finding in the input space an element of the training set which better approximates the inverse image of an element in the reduced space. In general, the exact pre-image does not exist, or it is not unique, and we need an approximate solution Mika *et al.* [9].

We consider here the pre-image problem in the context of the reconstruction of signals by means of a cost function, which differs from previous approaches, Etyngier *et al.* [6], Arias *et al.*[2] and Arif *et al.* [3] We use Nystrom's extension of $\mathcal{D}$ to other vectors in $I\!R^d$, which do not belong to the training set, Lafon *et al.* [7]. We represent it by $\tilde{\mathcal{D}}$.

Since $\mathcal{D}$ is injective in the training set, the pre-image problem has a unique solution in the set of features of training signals. For features outside that set that question is more complicated. The problem of the pre-image of an arbitrary point of $I\!R^{n-1}$ is an ill-posed problem and, in general, the pre-image of a unique point, if it exits, will be a set of vectors in the input space Arias *et al.* [2]. In order to circunvent this difficulty and to look for adequate modifications, we can consider a regularization of the problem by means of the training set.

Assume we are given a point $b \in I\!R^{n-1}$. We look for a good approximation of a possible pre-image, x, of that point. We want that x be as close as possible to the data set, in such a way to regularize the inversion. Clearly, we also want that the image of x by the diffusion map be b or near by it. For each b we can represent these requirements by means of an objetive function $f : I\!R^d \to I\!R$, as follows,

$$f(x) = \|\tilde{\mathcal{D}}(x) - b\| + \gamma \min_k(\|x - X^k\|). \tag{3}$$

That is, given $b \in I\!R^{n-1}$, its pre-image, if it exits, will be the vector $x \in I\!R^d$ minimzing f above.The parameter γ is used so that it is possible to adjust the level of influence of the second term with respect to the first term, in the right hand side of (3). These ideas can also be used for the reconstruction of PCA.

If we wish to consider the pre-image problem for several points, $b \in I\!R^{n-1}$, we may extend the previous cost function to explicitely consider its dependence not only on x but also on b, $f : I\!R^d \times I\!R^{n-1} \to I\!R$, defined by $f(x, b) = \|\tilde{\mathcal{D}}(x) - b\| + \gamma \min_k(\|x - X^k\|)$. Therefore, we consider a function G defined by minimizing f, $G : I\!R^{n-1} \to I\!R^d$, such that

$$G(b) = \arg\min_x f(x, b). \tag{4}$$

In general $G(b)$ may be a subset of $I\!R^d$ since $f(\cdot, b)$ can have several minimum points.

The point of minimum, denoted by $\tilde{X}$, when $b = G(X^*)$, is the reconstructed sinal X^*. The residue $X^* - \tilde{X}$, has to be verified to check the quality of the reconstruction and the power of analysis and synthesis of the prposed method.

5 Experiment

We applied the discussion of extension and pre-image problem to a set of known geometric structure in $I\!R^3$, representing an helix.

We considered just 38 points in $\mathbb{R}^3$ consisting of an helix with three turns. Further, random noise was added to 189 points distributed with equal spacing between the points of the helix. The calculation of the features of these noisy points was done by means of Nystrom extension. The pre-images were computed by minimization of the cost function, equation (3), using a simulated annealing algorithm. The regularization parameter was set $\gamma = 0.09$. For the diffusion map we let $\varepsilon = 0,001r^2, t = 50$ e $\alpha = 1$.

Figure 1 presents the results of this experiment. There is a small part of it which has been amplified. In blue are represented the points of the ideal helix, the noisy points are in red, and the pre-images are in green.

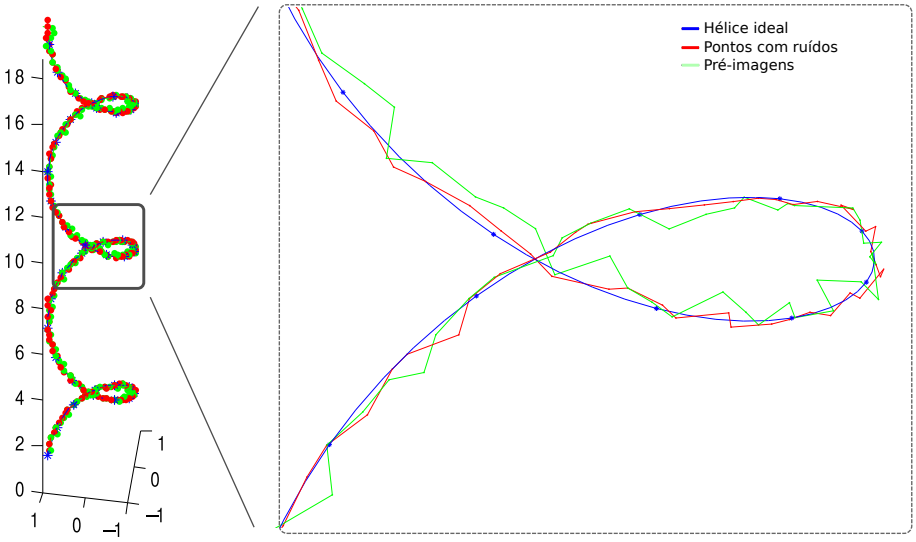


Fig. 1. Helix with a small stretch amplified where one can see its noisy points (red) and the corresponding pre-images (green) for the diffusion maps.

6 Conclusion

This article presents the diffusion map method for dimensionality reduction focussed on pattern theory in respect to the non-linear reconstruction of signals.

We also formulate and exploit a cost function to compute pre-images for the diffusion maps, which constitutes a significant contribution of this work.

References

1. Aflalo, Y. and Kimmel, R.,"Spectral multidimensional scaling", *Proceedings of the National Academy of Sciences*, 110(45):18052–18057 (2013).

2. Arias, P., Randall, G. and Sapiro, G., "Connecting the out-of-sample and pre-image problems in kernel methods", *IEEE Conference on Computer Vision and Pattern Recognition*, 1–8 (2007).
3. Arif, O., Vela, P. A. and Daley, W., *Pre-image problem in manifold learning and dimensional reduction methods*, *IEEE Ninth International Conference on Machine Learning and Applications*, 921–924 (2010).
4. Coifman, R. R. and Lafon, S., "Diffusion maps", *Applied and computational harmonic analysis*, 21(1):5–30 (2006).
5. Coifman, R.R. and Lafon, S. and Lee, A.B. and Maggioni, M. and Nadler, B. and Warner, F. and Zucker, S.W. "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps", *Proceedings of the National Academy of Sciences of the United States of America*, 102(21):7426 (2005).
6. Etyngier, P., Segonne,F. and Keriven, R. "Shape priors using manifold learning techniques", *IEEE 11th International Conference on Computer Vision*, 1–8 (2007).
7. Lafon, S., Keller, Y. and Coifman, R. R., "Data fusion and multicue data matching by diffusion maps" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1784–1797 (2006).
8. Lafon, S. and Lee, A., "Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(9):1393–1403 (2006).
9. Mika, S., Schlkopf, B., Smola, A. J., Mller, K.-R., Scholz, M. and Rtsch, G., "Kernel pca and de-noising in feature spaces", *NIPS* 11:536–542 (1998).
10. Mumford, D., and Desolneux, A., *Pattern theory: the stochastic analysis of real-world patterns*, A K Peters, Ltd., Natick (2010).