Editorial

The Third International Conference on L_1-Norm and Related Methods was held in Neuchâtel, Switzerland, from August 11-15, 1997, as a Satellite Meeting to the 51st International Statistical Institute Session in Istanbul. Thirty seven of the ninety two papers presented at the Conference appeared in Volume 31 of the *Lectures Notes - Monograph Series* published by the Institute of Mathematical Statistics, Hayward, California.

The present issue of *Student* contains a selection of another thirteen refereed papers presented during the Conference. As *Student* is used to offer a variety of statistical materials, we thought it would be instructive to bring together here a collection of papers which would demonstrate the necessity of future research in different fields of statistics using as base other measures of distance than the traditional L_2-norm.

The Editor

Nearly all the participants of the Third L_1-Norm Conference,
August 11, 1997.

Introduction [1]

Yadolah Dodge

While the method of least squares (and its generalizations) have served statisticians well for a good many years (mainly because of mathematical convenience and ease of computation), and enjoys certain well known properties within strictly Gaussian parametric models, it is recognized that outliers, which arise from heavy-tailed distributions, have an unusually large influence on the estimates obtained by these methods. Indeed, one single outlier can have an arbitrary large effect on the estimate. Outlier diagnostics have been developed to detect observations with a large influence on the least squares estimation. For excellent books related to such diagnostics the reader is referred to Cook and Weisberg (1982, 1994) and Chatterjee and Hadi (1988).

Parallel to diagnostic techniques, robust methods with varying degrees of robustness and computational complexity have been developed to modify the LS method so that the outliers have less influence on the final estimates. Among others are the bounded influence estimators, the repeated median, the least median of squares and the regression quantile methods.

In 1964, Huber published what is now considered to be a classic paper on robust estimation of location parameter and subsequently extended to that linear model. The development of selected robustness concepts since their inception in the 1960's and their current status, is given by Huber (1995).

One of the simplest robust alternatives to LS is the least absolute value method. This method, which is the subject of this volume, is a widely recognized superior method especially well-suited to longer-tailed error distributions, such as the Laplace distribution.

Depending on the field of application, the least absolute value method has been studied in several contexts under a variety of names such as minimum, or least sums of absolute errors, deviations or values; and here we

[1]First published in L_1-*Statistical Procedures and Related Topics* (1997),Ed. Y. Dodge, LNMS Vol.31, pp.xiii-xviii. Reproduced by permission of IMS, Hayward, California.

refer to it as the L_1- norm method (for minimizing the L_1- norm of the vector of deviations). The L_1 method estimates the unknown parameters in a stochastic model so as to minimize the sum of the absolute deviations of a given set of observations from the values predicted by the model.

Historically, L_1 estimation is the oldest of all robust methods. The method of least absolute deviations was introduced almost 50 years before the method of least squares, in 1757 by Roger Joseph Boscovich (1711-1787). He devised the method as a way to reconcile inconsistent measurements for estimating the shape of the earth. After Pierre Simon, Laplace adopted the method 30 years later, it saw occasional use but was soon overshadowed by the method of least squares. The popularity of least squares was at least partly due to the relative simplicity of its computations and to the supporting theory that was developed by Gauss and Laplace. Laplace, in his second memoir on the Figure of the Earth in 1789, adopted Boscovich's two criteria for a line of best fit, and gave an algebraic formulation and derivation of Boscovich's algorithm.

After nearly seventy years following the publication of Laplace's second supplement to the Théorie Analytique des Probabilités (1818), Edgeworth (1887) presented a method for linear regression using L_1 method. But since the publication of Edgeworth's work , few attempts have been made to convince the statisticians and particularly the applied users to employ this method (see Turner, 1887; Rhodes, 1930; Singleton, 1940; Karst,1958). Reasons for such a long silence may be summarized as follows :

(1) Computational difficulties in producing the numeric values of the L_1 estimates in regression. (Lack of closed form formulae similar to that of least squares).

(2) Lack of an asymptotic theory for L_1 estimation in the regression model, and more generally the nonexistence of accompanying statistical inference procedures.

(3) Insufficient evidence to show the superiority of the small sample properties of L_1 estimation compared to the LS estimators when sampling from long tailed distributions.

Following the work of Charnes, Cooper and Ferguson (1955) a renewed interest in using L_1 estimation for regression problem was created. They showed the equivalence between the L_1 problem and a linear programming problem. Wagner (1959) suggested that the L_1 problem in a linear regression of the form $y_i = \theta_0 + \theta_1 x_{1i} + \cdots + \theta_p x_{pi} + \varepsilon_i$ or in matrix form $Y = X\theta + \varepsilon$ can be solved by solving the dual of the L_1 problem. He also observed that the dual problem can be reduced to a problem with a smaller basis but where the dual variables have upper-bound restrictions. Wagner's formulation of the problem is to restate the problem of minimizing $\sum |\varepsilon_i|$ with

respect to θ where ε_i is the deviation between the observed and predicted values of the i^{th} observation Y_i, as: minimize $\sum |\varepsilon_i|$, subject to $X\theta+\varepsilon = Y$, where θ, ε unrestricted in sign.

Noting the fact that $|\varepsilon_i| = \varepsilon_{1i}+\varepsilon_{2i}$ where $\varepsilon_1 = \varepsilon$ if $\varepsilon > 0$, 0 otherwise and $\varepsilon_2 = -\varepsilon$ if $\varepsilon < 0$, 0 otherwise, that both are nonnegative and $\varepsilon_i = \varepsilon_{1i} - \varepsilon_{2i}$, we can reformulate the problem as a linear programming problem: minimize $\sum \varepsilon_{1i} + \sum \varepsilon_{2i}$, subject to $X\theta + \varepsilon_1 - \varepsilon_2 = Y$, where θ unrestricted in sign, $\varepsilon_1, \varepsilon_2 \geq 0$.

From the computational point of view, the L_1 method is now extremely simple and it requires only a routine to fit the L_1 regression. There are several computer programs available for calculation of L_1 estimates. See for example Sadovski (1974) and Farebrother (1988). For the case of multiple regresssion we can use, for example, the modified simplex algorithm of Barrodale and Roberts (1973) that exists in the IMSL library under the name RLLAV. The L_1 estimation problem with additional linear restrictions (restricted L_1 problem) is considered along the same lines in Barrodale and Roberts (1974). Arthanari and Dodge (1993) devoted a complete chapter on computational aspects of the L_1 estimation. L_1 regression estimates are also obtainable from the function l1fit in the computer language S-Plus and from the robust regression package ROBSYS (Marazzi, 1993). Detection of outlying points in both dependent and independent variables in regression model are explained in Dodge (1997).

The major difficulty for applied researchers in using L_1 estimation for many years was the lack of accompanying statistical inference procedures. Such procedures would include methods for testing general linear hypothesis, obtaining confidence intervals, analysis of variance tables and for performing multiple comparison procedures.

Bassett and Koenker (1978) developed the asymptotic theory for L_1 estimators in the regression model. Their finding is considered to be a breakthrough for the problem. Their main result is that the sampling distribution of L_1 estimators will be asymptotically normal with a specified mean and variance. Under very general assumptions they confirmed that the L_1 estimator $\widehat{\theta}$ has a normal distribution with mean θ and covariance matrix $\tau^2(X'X)^{-1}$ where τ^2/n is the asymptotic variance of the sample median from random samples of size n taken from the error distribution with a continuous and positive derivative at the median. This result is remarkably similar to that for LS. Therefore the L_1 confidence intervals for an estimable function $\lambda'\widehat{\theta}$ is

$$\lambda'\widehat{\theta} \pm z_{\alpha/2} \cdot \widehat{\tau}\sqrt{\lambda'(X'X)^{-}\lambda}$$

where $(X'X)^{-}$ is the q-inverse of $X'X$ and $\widehat{\tau}$ is an estimate of τ given in

McKean and Schrader (1987).

Koenker and Bassett (1982) investigate the asymptotic distribution of three alternative L_1 test statistics of a linear hypothesis in the standard linear model. They showed that the three test statistics, which correspond to the Wald, likelihood ratio and Lagrange multiplier tests, under mild regularity conditions on the design and error distribution, have the same limiting chi-square behavior. For a complete treatment of L_1 regression the reader is referred to Chapter 4 of Birkes and Dodge (1993).

With the availability of many computationally efficient algorithms and developed inference procedures for testing general linear hypotheses, for obtaining confidence intervals, selection of variables, analysis of variance tables and multiple comparison, it is hoped that L_1 estimation methods will be employed more often by researchers in applied sciences than before.

Certainly, there are many other areas of statistical data analysis based on the L_1-norm (such as density estimation, time series analysis, multivariate analysis and classification methods) that could have been discussed here. But, unfortunately, limitation of space and time have caused many interesting and important lines of research to be treated lightly or not at all. While it is now evident that no single robust procedure is best by any criteria, it may be appropriate (or at least reasonable) to use adaptive convex combinations of L_1 with other methods rather than a single criterion to estimate the unknown parameters. However, for the error distributions for which the median is superior to the mean as an estimator of location, L_1 estimation is certainly preferred to least squares and strongly recommended for use in these cases.

Bloomfield and Steiger (1983), Devroye and Györfi (1985), and Gonin and Money (1989) are the only books entirely devoted to L_1 topic. The authors of these three texts had the courage to pull together a rich and diverse literature in this field. I hope that the proceedings contained in this volume and its predecessors, Dodge (1987, 1992), will encourage someone to write a fourth.

References

[1] Arthanari, T.S. and Dodge, Y. (1993). *Mathematical Programming in Statistics.* New York: Wiley, Classical Edition.

[2] Barrodale, I. and Roberts, F.D.K. (1973). An improved algorithm for discrete L_1 linear approximation. *SIAM J. Numer. Anal.* **10,** 839-848.

[3] Barrodale, I. and Roberts, F.D.K. (1974). Algorithm 478: Solution of an overdetermined system of equations in the L_1 norm. *Commun.*

Assoc. Comput. Mach. **17,** 319.

[4] Bassett, G.W. and Koenker, R. (1978). Asymptotic theory of least absolute error. *J. Am. Statist. Assoc.* **73,** 618-22.

[5] Birkes, D. and Dodge, Y (1993). *Alternative Methods of Regression.* New York: Wiley.

[6] Bloomfield, P. and Steiger, W.L. (1983). *Least Absolute Deviations: Theory, Applications, and Algorithms.* Boston: Birkäuser.

[7] Boscovich, R.J. (1757). De litteraria expeditione per pontificiam synopsis amplioris operis, ac habentur plura eius ex exempla ria etiam sensorum impressa. Bononiensi Scientiarum et Artium Instituto Atque Academia Commentarii 4, 353-96.

[8] Charnes, A., Cooper, W.W. and Fergusson, R.O. (1955). Optimal estimation of executive compensation by linear programming. *Manage. Sci.* **1,** 138.

[9] Chatterjee, S. and Hadi, A.S. (1988). *Sensitivity Analysis in Linear Regression.* New York: Wiley.

[10] Cook, R.D. and Weisberg, S. (1982). *Residuals and Influence in Regression.* New York: Chapman and Hall.

[11] Cook, R.D. and Weisberg, S. (1994). *An Introduction to Regression Graphics.* New York: Wiley.

[12] Devroye, L. and Györfi, L. (1985). *Nonparametric Density Estimation, the L_1 View.* New York: Wiley.

[13] Dielman, T.E. and Rose, L.E. (1995). A bootstrap approach to hypothesis testing in least absolute value regression. *Comp. Statist. Data Anal.* **20,** 119-130.

[14] Dielman, T.E. and Rose, L.E. (1996). A note on the hypothesis testing in LAV multiple regression: A small sample comparison. *Comp. Statist. Data Anal.* **21,** 463-470.

[15] Dodge, Y. (1987). (Ed). *Statistical Data Analysis Based on the L_1-Norm and Related Methods.* Amsterdam: North-Holland.

[16] Dodge, Y. (1992). *L_1-Statistical Data Analysis and Related Methods.* Amsterdam: North Holland.

[17] Dodge, Y. (1997). LAD regression for detecting outliers in response and explanatory variables. *J. Multiv. Anal.* **61,** 144-158.

[18] Edgeworth, F.Y. (1887). On observations relating to several quantities. *Phil. Mag.* (5th Series) **24,** 222.

[19] Ellis, S.P. and Morgenthaler, S. (1992). Leverage and breakdown in L_1 regression. *J. Am. Statist. Assoc.* **87,** 143-148.

[20] Farebrother, R.W. (1988). A simple recursive procedure for the L_1 norm fitting of a straight line. *Appl. Statist.* **37,** 457-489.

[21] Gonin, R. and Money, A.H. (1989). *Nonlinear L_p-Norm Estimation.*

New York: Marcel Dekber.

[22] Huber, P.J. (1987). The place of the L_1-norm in robust estimation. In *Statistical Data Analysis - Based on L_1-Norm and Related Methods*, Ed. Y. Dodge, pp. 23-33. Amsterdam: North Holland.

[23] Huber, P.J. (1995). Robustness: Where are we now ? *Student* **1,** 75-86.

[24] Karst, O.J. (1958). Linear curve fitting using least deviations. *J. Am. Statist. Assoc.* **53,** 118-132.

[25] Koenker, R. and Bassett, G.W. (1982). Tests of hypotheses and L_1 estimation. *Econometrica* **50,** 1577-83.

[26] Koenker, R. (1987). A comparison of asymptotic testing methods for L_1-regression. In this volume.

[27] Laplace, P.S. (1789). Sur les degrés mesurés des méridiens, et sur les longueurs observées sur pendule. Histoire de l'Académie royale des inscriptions et belles lettres, avec les Memoires de littérature tirez des registres de cette académie, Année 1789. Paris. "Mémoires", 18-43.

[28] Laplace, P.S. (1793). Sur quelques points du système du monde. Mémoires de l'Académie Royale des Sciences de Paris, 1-87. Reprinted in Oeuvres Complètes de Laplace, Vol.11, pp. 477-558. Paris: Gauthier-Villars, 1895.

[29] Laplace, P.S. (1818). Deuxième supplément to Laplace (1812).

[30] Legendre, A.M. (1805). Nouvelles méthodes pour la détermination des orbites des comètes. Paris: Courcier. Appendice sur la méthode des moindres carrés, 72-80.

[31] Marazzi, A. (1993). *Algorithms, Routines, and S Functions for Robust Statistics.* Wadsworth and Brooks/Cole, California.

[32] McKean, J.W. and Schrader, R.M. (1987). Least absolute errors analysis of variance. In *Statistical Data Analysis Based on the L_1-Norm and Related Methods*, Ed. Y. Dodge, pp 297-321. Amsterdam: North Holland.

[33] Rhodes, E.C. (1930). Reducing observations by the methode of minimum deviations. *Phil. Mag.* (7th Series) **9,** 974.

[34] Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust Regression and Outlier Detection.* New York: Wiley.

[35] Sadowski, A.N. (1974). Algorithm AS74. L_1-norm fit of a straight line. *Appl. Statist.* **23,** 244-248.

[36] Singleton, R.R. (1940). A method of minimizing the sum of absolute values of deviations. *Ann. Math. Statist.* **11,** 301-310.

[37] Turner, H.H. (1887). On Mr. Edgeworth's method of reducing observations relating to several quantities. *Phil. Mag.* (5th Series) **24,** 466-70.

[38] Wagner, H.M. (1959). Linear programming techniques for regression analysis. *J. Am. Statist. Assoc.* **54,** 206.

Editorial

The Third International Conference on L_1-Norm and Related Methods was held in Neuchâtel, Switzerland, from August 11-15, 1997, as a Satellite Meeting to the 51st International Statistical Institute Session in Istanbul. Thirty seven of the ninety two papers presented at the Conference appeared in Volume 31 of the *Lectures Notes - Monograph Series* published by the Institute of Mathematical Statistics, Hayward, California.

The present issue of *Student* contains a selection of another thirteen refereed papers presented during the Conference. As *Student* is used to offer a variety of statistical materials, we thought it would be instructive to bring together here a collection of papers which would demonstrate the necessity of future research in different fields of statistics using as base other measures of distance than the traditional L_2-norm.

The Editor

Nearly all the participants of the Third L_1-Norm Conference,
August 11, 1997.

Introduction [1]

Yadolah Dodge

While the method of least squares (and its generalizations) have served statisticians well for a good many years (mainly because of mathematical convenience and ease of computation), and enjoys certain well known properties within strictly Gaussian parametric models, it is recognized that outliers, which arise from heavy-tailed distributions, have an unusually large influence on the estimates obtained by these methods. Indeed, one single outlier can have an arbitrary large effect on the estimate. Outlier diagnostics have been developed to detect observations with a large influence on the least squares estimation. For excellent books related to such diagnostics the reader is referred to Cook and Weisberg (1982, 1994) and Chatterjee and Hadi (1988).

Parallel to diagnostic techniques, robust methods with varying degrees of robustness and computational complexity have been developed to modify the LS method so that the outliers have less influence on the final estimates. Among others are the bounded influence estimators, the repeated median, the least median of squares and the regression quantile methods.

In 1964, Huber published what is now considered to be a classic paper on robust estimation of location parameter and subsequently extended to that linear model. The development of selected robustness concepts since their inception in the 1960's and their current status, is given by Huber (1995).

One of the simplest robust alternatives to LS is the least absolute value method. This method, which is the subject of this volume, is a widely recognized superior method especially well-suited to longer-tailed error distributions, such as the Laplace distribution.

Depending on the field of application, the least absolute value method has been studied in several contexts under a variety of names such as minimum, or least sums of absolute errors, deviations or values; and here we

[1]First published in L_1-*Statistical Procedures and Related Topics* (1997),Ed. Y. Dodge, LNMS Vol.31, pp.xiii-xviii. Reproduced by permission of IMS, Hayward, California.

refer to it as the L_1- norm method (for minimizing the L_1- norm of the vector of deviations). The L_1 method estimates the unknown parameters in a stochastic model so as to minimize the sum of the absolute deviations of a given set of observations from the values predicted by the model.

Historically, L_1 estimation is the oldest of all robust methods. The method of least absolute deviations was introduced almost 50 years before the method of least squares, in 1757 by Roger Joseph Boscovich (1711-1787). He devised the method as a way to reconcile inconsistent measurements for estimating the shape of the earth. After Pierre Simon, Laplace adopted the method 30 years later, it saw occasional use but was soon overshadowed by the method of least squares. The popularity of least squares was at least partly due to the relative simplicity of its computations and to the supporting theory that was developed by Gauss and Laplace. Laplace, in his second memoir on the Figure of the Earth in 1789, adopted Boscovich's two criteria for a line of best fit, and gave an algebraic formulation and derivation of Boscovich's algorithm.

After nearly seventy years following the publication of Laplace's second supplement to the Théorie Analytique des Probabilités (1818), Edgeworth (1887) presented a method for linear regression using L_1 method. But since the publication of Edgeworth's work , few attempts have been made to convince the statisticians and particularly the applied users to employ this method (see Turner, 1887; Rhodes, 1930; Singleton, 1940; Karst,1958). Reasons for such a long silence may be summarized as follows :

(1) Computational difficulties in producing the numeric values of the L_1 estimates in regression. (Lack of closed form formulae similar to that of least squares).

(2) Lack of an asymptotic theory for L_1 estimation in the regression model, and more generally the nonexistence of accompanying statistical inference procedures.

(3) Insufficient evidence to show the superiority of the small sample properties of L_1 estimation compared to the LS estimators when sampling from long tailed distributions.

Following the work of Charnes, Cooper and Ferguson (1955) a renewed interest in using L_1 estimation for regression problem was created. They showed the equivalence between the L_1 problem and a linear programming problem. Wagner (1959) suggested that the L_1 problem in a linear regression of the form $y_i = \theta_0 + \theta_1 x_{1i} + \cdots + \theta_p x_{pi} + \varepsilon_i$ or in matrix form $Y = X\theta + \varepsilon$ can be solved by solving the dual of the L_1 problem. He also observed that the dual problem can be reduced to a problem with a smaller basis but where the dual variables have upper-bound restrictions. Wagner's formulation of the problem is to restate the problem of minimizing $\sum |\varepsilon_i|$ with

respect to θ where ε_i is the deviation between the observed and predicted values of the i^{th} observation Y_i, as: minimize $\sum |\varepsilon_i|$, subject to $X\theta+\varepsilon = Y$, where θ, ε unrestricted in sign.

Noting the fact that $|\varepsilon_i| = \varepsilon_{1i}+\varepsilon_{2i}$ where $\varepsilon_1 = \varepsilon$ if $\varepsilon > 0$, 0 otherwise and $\varepsilon_2 = -\varepsilon$ if $\varepsilon < 0$, 0 otherwise, that both are nonnegative and $\varepsilon_i = \varepsilon_{1i} - \varepsilon_{2i}$, we can reformulate the problem as a linear programming problem: minimize $\sum \varepsilon_{1i} + \sum \varepsilon_{2i}$, subject to $X\theta + \varepsilon_1 - \varepsilon_2 = Y$, where θ unrestricted in sign, $\varepsilon_1, \varepsilon_2 \geq 0$.

From the computational point of view, the L_1 method is now extremely simple and it requires only a routine to fit the L_1 regression. There are several computer programs available for calculation of L_1 estimates. See for example Sadovski (1974) and Farebrother (1988). For the case of multiple regresssion we can use, for example, the modified simplex algorithm of Barrodale and Roberts (1973) that exists in the IMSL library under the name RLLAV. The L_1 estimation problem with additional linear restrictions (restricted L_1 problem) is considered along the same lines in Barrodale and Roberts (1974). Arthanari and Dodge (1993) devoted a complete chapter on computational aspects of the L_1 estimation. L_1 regression estimates are also obtainable from the function l1fit in the computer language S-Plus and from the robust regression package ROBSYS (Marazzi, 1993). Detection of outlying points in both dependent and independent variables in regression model are explained in Dodge (1997).

The major difficulty for applied researchers in using L_1 estimation for many years was the lack of accompanying statistical inference procedures. Such procedures would include methods for testing general linear hypothesis, obtaining confidence intervals, analysis of variance tables and for performing multiple comparison procedures.

Bassett and Koenker (1978) developed the asymptotic theory for L_1 estimators in the regression model. Their finding is considered to be a breakthrough for the problem. Their main result is that the sampling distribution of L_1 estimators will be asymptotically normal with a specified mean and variance. Under very general assumptions they confirmed that the L_1 estimator $\widehat{\theta}$ has a normal distribution with mean θ and covariance matrix $\tau^2(X'X)^{-1}$ where τ^2/n is the asymptotic variance of the sample median from random samples of size n taken from the error distribution with a continuous and positive derivative at the median. This result is remarkably similar to that for LS. Therefore the L_1 confidence intervals for an estimable function $\lambda'\widehat{\theta}$ is

$$\lambda'\widehat{\theta} \pm z_{\alpha/2} \cdot \widehat{\tau}\sqrt{\lambda'(X'X)^{-}\lambda}$$

where $(X'X)^{-}$ is the q-inverse of $X'X$ and $\widehat{\tau}$ is an estimate of τ given in

McKean and Schrader (1987).

Koenker and Bassett (1982) investigate the asymptotic distribution of three alternative L_1 test statistics of a linear hypothesis in the standard linear model. They showed that the three test statistics, which correspond to the Wald, likelihood ratio and Lagrange multiplier tests, under mild regularity conditions on the design and error distribution, have the same limiting chi-square behavior. For a complete treatment of L_1 regression the reader is referred to Chapter 4 of Birkes and Dodge (1993).

With the availability of many computationally efficient algorithms and developed inference procedures for testing general linear hypotheses, for obtaining confidence intervals, selection of variables, analysis of variance tables and multiple comparison, it is hoped that L_1 estimation methods will be employed more often by researchers in applied sciences than before.

Certainly, there are many other areas of statistical data analysis based on the L_1-norm (such as density estimation, time series analysis, multivariate analysis and classification methods) that could have been discussed here. But, unfortunately, limitation of space and time have caused many interesting and important lines of research to be treated lightly or not at all. While it is now evident that no single robust procedure is best by any criteria, it may be appropriate (or at least reasonable) to use adaptive convex combinations of L_1 with other methods rather than a single criterion to estimate the unknown parameters. However, for the error distributions for which the median is superior to the mean as an estimator of location, L_1 estimation is certainly preferred to least squares and strongly recommended for use in these cases.

Bloomfield and Steiger (1983), Devroye and Györfi (1985), and Gonin and Money (1989) are the only books entirely devoted to L_1 topic. The authors of these three texts had the courage to pull together a rich and diverse literature in this field. I hope that the proceedings contained in this volume and its predecessors, Dodge (1987, 1992), will encourage someone to write a fourth.

References

[1] Arthanari, T.S. and Dodge, Y. (1993). *Mathematical Programming in Statistics.* New York: Wiley, Classical Edition.

[2] Barrodale, I. and Roberts, F.D.K. (1973). An improved algorithm for discrete L_1 linear approximation. *SIAM J. Numer. Anal.* **10,** 839-848.

[3] Barrodale, I. and Roberts, F.D.K. (1974). Algorithm 478: Solution of an overdetermined system of equations in the L_1 norm. *Commun.*

Assoc. Comput. Mach. **17,** 319.

[4] Bassett, G.W. and Koenker, R. (1978). Asymptotic theory of least absolute error. *J. Am. Statist. Assoc.* **73,** 618-22.

[5] Birkes, D. and Dodge, Y (1993). *Alternative Methods of Regression.* New York: Wiley.

[6] Bloomfield, P. and Steiger, W.L. (1983). *Least Absolute Deviations: Theory, Applications, and Algorithms.* Boston: Birkäuser.

[7] Boscovich, R.J. (1757). De litteraria expeditione per pontificiam synopsis amplioris operis, ac habentur plura eius ex exempla ria etiam sensorum impressa. Bononiensi Scientiarum et Artium Instituto Atque Academia Commentarii 4, 353-96.

[8] Charnes, A., Cooper, W.W. and Fergusson, R.O. (1955). Optimal estimation of executive compensation by linear programming. *Manage. Sci.* **1,** 138.

[9] Chatterjee, S. and Hadi, A.S. (1988). *Sensitivity Analysis in Linear Regression.* New York: Wiley.

[10] Cook, R.D. and Weisberg, S. (1982). *Residuals and Influence in Regression.* New York: Chapman and Hall.

[11] Cook, R.D. and Weisberg, S. (1994). *An Introduction to Regression Graphics.* New York: Wiley.

[12] Devroye, L. and Györfi, L. (1985). *Nonparametric Density Estimation, the L_1 View.* New York: Wiley.

[13] Dielman, T.E. and Rose, L.E. (1995). A bootstrap approach to hypothesis testing in least absolute value regression. *Comp. Statist. Data Anal.* **20,** 119-130.

[14] Dielman, T.E. and Rose, L.E. (1996). A note on the hypothesis testing in LAV multiple regression: A small sample comparison. *Comp. Statist. Data Anal.* **21,** 463-470.

[15] Dodge, Y. (1987). (Ed). *Statistical Data Analysis Based on the L_1-Norm and Related Methods.* Amsterdam: North-Holland.

[16] Dodge, Y. (1992). *L_1-Statistical Data Analysis and Related Methods.* Amsterdam: North Holland.

[17] Dodge, Y. (1997). LAD regression for detecting outliers in response and explanatory variables. *J. Multiv. Anal.* **61,** 144-158.

[18] Edgeworth, F.Y. (1887). On observations relating to several quantities. *Phil. Mag.* (5th Series) **24,** 222.

[19] Ellis, S.P. and Morgenthaler, S. (1992). Leverage and breakdown in L_1 regression. *J. Am. Statist. Assoc.* **87,** 143-148.

[20] Farebrother, R.W. (1988). A simple recursive procedure for the L_1 norm fitting of a straight line. *Appl. Statist.* **37,** 457-489.

[21] Gonin, R. and Money, A.H. (1989). *Nonlinear L_p-Norm Estimation.*

New York: Marcel Dekber.

[22] Huber, P.J. (1987). The place of the L_1-norm in robust estimation. In *Statistical Data Analysis - Based on L_1-Norm and Related Methods*, Ed. Y. Dodge, pp. 23-33. Amsterdam: North Holland.

[23] Huber, P.J. (1995). Robustness: Where are we now ? *Student* **1,** 75-86.

[24] Karst, O.J. (1958). Linear curve fitting using least deviations. *J. Am. Statist. Assoc.* **53,** 118-132.

[25] Koenker, R. and Bassett, G.W. (1982). Tests of hypotheses and L_1 estimation. *Econometrica* **50,** 1577-83.

[26] Koenker, R. (1987). A comparison of asymptotic testing methods for L_1-regression. In this volume.

[27] Laplace, P.S. (1789). Sur les degrés mesurés des méridiens, et sur les longueurs observées sur pendule. Histoire de l'Académie royale des inscriptions et belles lettres, avec les Memoires de littérature tirez des registres de cette académie, Année 1789. Paris. "Mémoires", 18-43.

[28] Laplace, P.S. (1793). Sur quelques points du système du monde. Mémoires de l'Académie Royale des Sciences de Paris, 1-87. Reprinted in Oeuvres Complètes de Laplace, Vol.11, pp. 477-558. Paris: Gauthier-Villars, 1895.

[29] Laplace, P.S. (1818). Deuxième supplément to Laplace (1812).

[30] Legendre, A.M. (1805). Nouvelles méthodes pour la détermination des orbites des comètes. Paris: Courcier. Appendice sur la méthode des moindres carrés, 72-80.

[31] Marazzi, A. (1993). *Algorithms, Routines, and S Functions for Robust Statistics.* Wadsworth and Brooks/Cole, California.

[32] McKean, J.W. and Schrader, R.M. (1987). Least absolute errors analysis of variance. In *Statistical Data Analysis Based on the L_1-Norm and Related Methods*, Ed. Y. Dodge, pp 297-321. Amsterdam: North Holland.

[33] Rhodes, E.C. (1930). Reducing observations by the methode of minimum deviations. *Phil. Mag.* (7th Series) **9,** 974.

[34] Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust Regression and Outlier Detection.* New York: Wiley.

[35] Sadowski, A.N. (1974). Algorithm AS74. L_1-norm fit of a straight line. *Appl. Statist.* **23,** 244-248.

[36] Singleton, R.R. (1940). A method of minimizing the sum of absolute values of deviations. *Ann. Math. Statist.* **11,** 301-310.

[37] Turner, H.H. (1887). On Mr. Edgeworth's method of reducing observations relating to several quantities. *Phil. Mag.* (5th Series) **24,** 466-70.

[38] Wagner, H.M. (1959). Linear programming techniques for regression analysis. *J. Am. Statist. Assoc.* **54,** 206.

Student (1997), Vol.2, No.2, 77-84

L_1-Convergence of a Class of Algorithms for Global Optimization

Chang C.Y. Dorea

Universidade de Brasilia, Brazil

Abstract: Sufficient conditions for the convergence of annealing type algorithms are usually derived by analysing the existence of the equilibrium distribution of the associated Markov chain. In this paper we use the L_1-norm approach. It turns out that the desired convergence can be obtained under much simpler conditions.

Key words: L_1-norm, simulated annealing, random search.

1 Introduction

Let f be a real-valued function defined on $S \subset R^d$ where $S = \{1, 2, \ldots, N\}$ or $0 < m(S) < \infty$, m standing for the Lebesgue measure. As an alternative to deterministic procedures, random search methods can be used to estimate the global minimum (or maximum) of the objective function f on S:

$$y_* = \min_{x \in S}\{f(x)\} \text{ and } S_* = \{x : x \in S,\ f(x) = y_*\}. \tag{1}$$

When S is discrete, the best known random procedures are the annealing type methods: if $X_k = i$ is the current solution, a potential sucessor $j \in \eta(i)$ is generated with probability Q_{ij}, $\eta(i)$ being an appropriately chosen neighborhood of i ; the switch from i to j will then be implemented according to a positive acceptance probability function (a.p.f.) $a_c(j|i)$ which depends on a control parameter $c > 0$; otherwise one takes $X_{k+1} = X_k$. By choosing properly the sampling stochastic matrix $Q = (Q_{ij})$ and $a_c(\cdot|\cdot)$ one avoids being trapped in local minimum points, guarantees the independence of the

initial solution X_0 and assures the desired convergence:

$$\lim_{c\downarrow 0}\lim_{n\to\infty} P(X_n \in S_*) = 1 \text{ and } \lim_{c\downarrow 0}\lim_{n\to\infty} P(f(X_n) = y_*) = 1. \quad (2)$$

In the classical simulated annealing algorithm (Kirkpatrick et al., 1983) the Metropolis a.p.f. $a_c(j|i) = \min\{1, exp(-(f(j)-f(i))/c)\}$ is used. Many authors have investigated the convergence of this algorithm: Gidas (1985), Lundy and Mees (1986), Anily and Federgruen (1987), Hajek (1988), etc. Sufficient conditions for (2) to hold usually require some reversibility condition on the matrix $P_{ij} = P(X_{k+1} = j \mid X_k = i)$ and assumptions of the type: for each $j \in S$ there exists n_j such that $P(X_{n_j} \in S_* \mid X_0 = j) > 0$ (Tierney, 1994). If general a.p.f. is considered then $a_c(\cdot|\cdot)$ must satisfy rather complex regularity conditions (Anily and Federgruen, 1987). Most of these conditions are derived by analysing the existence of the equilibrium distribution of the Markov chain $\{X_n\}_{n\geq 0}$.

For the Metropolis a.p.f. Dekkers and Aarts (1991) studied the convergence of the algorithm in the continuous case ($0 < m(S) < \infty$). Using the theory of Markov chains with general state space (Doob, 1953) and assuming conditions that include the uniform continuity of f and the finiteness of the set of local and global minima they proved that:

$$\lim_{c\downarrow 0}\lim_{n\to\infty} P(|f(X_n) - y_*| < \epsilon \mid X_0 = x) \geq 1 - \epsilon,\ \forall \epsilon > 0,\ \forall x \in S. \quad (3)$$

Using a different approach but also assuming the uniform continuity Bélisle (1992) showed that:

$$\lim_{c\downarrow 0}\lim_{n\to\infty} P(|f(X_n) - y_*| < \epsilon \mid X_0 = x) = 1,\ \forall \epsilon > 0,\ \forall x \in S. \quad (4)$$

From the practical point of view the uniform continuity of f is a restrictive condition, since f is the unknown function to be minimized.

In this paper we analyse the asymptotic behavior of the L_1-norm $||f(X_n) - y_*||_x = E\{|f(X_n) - y_*| \mid X_0 = x\}$. It will be shown that

$$\lim_{c\downarrow 0}\lim_{n\to\infty} ||f(X_n) - y_*||_x = 0 \ ,\forall\ x \in S. \quad (5)$$

As a side result we also obtain (2), (3) and (4). It turns out that, by using the L_1-norm approach, the conditions for the convergence are much simpler than those mentioned above. Particularly, in the continuous case, the uniform continuity and the finiteness of the set of local and global minima are no longer required. In Section 2 the discrete case is studied. Though the results obtained are similar to those found in the literature, it illustrates the aproach to be used in the continuous case (Section 3). Several examples of the most used a.p.f. are shown to satisfy our conditions.

2 The discrete case

Let $S = \{1, 2, \ldots, N\}$ so that $\{X_n\}_n \geq 0$ becomes a Markov chain with transition probabilities $P_{ij}(c) = P(X_{n+1} = j \mid X_n = i)$ given by:

$$P_{ij}(c) = \begin{cases} Q_{ij} a_c(j|i) & \text{if } j \neq i \\ 1 - \sum_{l \neq i} Q_{il} a_c(l|i) & \text{if } j = i. \end{cases} \tag{6}$$

Condition 1 (a) $Q_{ij} > 0$ for all i and all j.

(b) There exist functions $b_c(\cdot) > 0$ and $d_c(\cdot) > 0$ such that $b_c(0)$ is independent of c , $\lim_{c \downarrow 0} d_c(l) = 0$ for $l > 0$ and

$$\begin{aligned} f(j) - f(i) \leq l \quad &\Rightarrow \quad a_c(j|i) \geq b_c(l) \\ f(j) - f(i) \geq l \quad &\Rightarrow \quad a_c(j|i) \leq d_c(l). \end{aligned} \tag{7}$$

Note that Condition 1(b) is easily satisfied if the a.p.f. can be written as a function of the difference $f(j) - f(i)$. The following examples illustrate this assertion:

Example 1 *(a) For the Metropolis case just take* $b_c(l) = d_c(l) = exp(-l/c)$.
(b) Let $a_c(j|i) = \min\{1, exp(-(f(j) - f(i))/cf(j)f(i))\}$ *(cf. Bohachavsky et al. (1986)) then we have (7) by taking* $b_c(l) = exp(-l/cl_1^2)$ *and* $d_c(l) = exp(-l/cl_0^2)$ *where* $l_0 = \min_{i \in S} f(i)$ *and* $l_1 = \max_{i \in S} f(i)$.
(c) For the Hasting's a.p.f. ($\rho \geq 1$):

$$a_c(j|i) = \frac{1 + 2\left[\frac{1}{2} \min\left\{\frac{Q_{ij}}{Q_{ji}} exp\left(\frac{f(j) - f(i)}{c}\right), \frac{Q_{ji}}{Q_{ij}} exp\left(\frac{f(i) - f(j)}{c}\right)\right\}\right]^\rho}{1 + \frac{Q_{ij}}{Q_{ji}} exp\left(\frac{f(j) - f(i)}{c}\right)}$$

and for $\delta_1 > 0$ *and* $\delta_2 > 0$ *satisfying* $\delta_1 < \frac{Q_{ij}}{Q_{ji}} < \delta_2$ *we can take*

$$b_c(l) = \frac{1 + \frac{1}{2^{\rho-1}} \left(\frac{\exp(-l/c)}{\delta_2}\right)^\rho}{1 + \delta_2 \exp(l/c)} \quad \text{and} \quad d_c(l) = \frac{1 + \frac{1}{2^{\rho-1}} \left(\frac{\exp(-l/c)}{\delta_1}\right)^\rho}{1 + \delta_1 \exp(l/c)}.$$

Theorem 1 *Under Condition 1 we have:*
(a) There exist $\gamma_c > 0$ *and* $\delta_c \geq 0$ *with* $\lim_{c \downarrow 0} \frac{\delta_c}{\gamma_c} = 0$ *and such that*

$$L_c(i) = E\{f(X_1) - y_* \mid X_0 = i\} \leq \delta_c + (f(i) - y_*)(1 - \gamma_c). \tag{8}$$

(b) We have (2) and there exists a function $\varphi(c)$ such that

$$\lim_{c\downarrow 0}\varphi(c)=0 \quad \textit{and} \quad \lim_{n\to\infty}||f(X_n)-y_*||_1\leq\varphi(c). \tag{9}$$

Proof: (a) First note that if f is a constant then $L_c(i)=0$. If not, from (6) we can derive

$$L_c(i)=\sum_{j\neq i}(f(j)-y_*)Q_{ij}a_c(j|i)+(f(i)-y_*)(1-\sum_{l\neq i}Q_{il}a_c(l|i)). \tag{10}$$

Define $S_i^+=\{j:f(j)>f(i)\}$, $S_i^-=\{j:f(j)<f(i)\}$, $A_c(i)=\sum_{j\in S_i^+}(f(j)-y_*)Q_{ij}a_c(j|i)$, $D_c(i)=\sum_{j\in S_i^+}Q_{ij}a_c(j|i)$ and $E_c(i)=\sum_{j\in S_i^-}\frac{f(i)-f(j)}{f(i)-y_*}Q_{ij}a_c(j|i)$. From (10) we can write

$$L_c(i)=A_c(i)+(f(i)-y_*)(1-D_c(i)-E_c(i)). \tag{11}$$

Now let $l_*=\min_{f(i)\neq f(j)}\{|f(j)-f(i)|\}$ and $l^*=\max_{f(i)\neq f(j)}\{|f(j)-f(i)|\}$. Since f is not a constant we have $l^*\geq l_*>0$. From Condition 1(b) we have

$$\begin{aligned} A_c(i) &\leq l^*d_c(l_*)\sum_{j\in S_i^+}Q_{ij}\ ,\ D_c(i)\geq b_c(l^*)\sum_{j\in S_i^+}Q_{ij} \text{ and} \\ E_c(i) &\geq \frac{l_*}{l^*}b_c(0)\sum_{j\in S_i^-}Q_{ij}. \end{aligned} \tag{12}$$

If i is such that $f(i)=y_*$ then $S_i^-=\emptyset$. In this case we have $A_c(i)\leq\delta_c$ with $\delta_c=l^*d_c(l_*)$. Take $0<\gamma_c\leq 1$ and from (11) we can write $L_c(i)\leq\delta_c$ with $\lim_{c\downarrow 0}\delta_c/\gamma_c=0$.

If $f(i)>y_*$ let $\rho=\min_{i,j}Q_{ij}>0$ (Condition 1(a)). Then since $S_i^-\neq\phi$ we have $E_c(i)\geq l^*b_c(0)\rho/l_*$. Taking $\gamma_c=l^*b_c(0)\rho/l_*$ we have $(1-D_c(i))-E_c(i)\leq(1-\gamma_c)$. Since $l_*>0$ we can take $\delta_c=l^*d_c(l_*)$ and we have $\lim_{c\downarrow 0}d_c(l_*)=\lim_{c\downarrow 0}\delta_c/\gamma_c=0$. Finally from (11) and (12) we get (8).

(b) Note that $||f(X_n)-y_*||_1=EV_n(X_0)=E\{E(f(X_n)-y_*\ |X_{n-1})|X_0\}$ where $V_n(X_0)=E(f(X_n)-y_*\ X_0)$. Since the chain is homogeneous from (8) we can write:

$$E(f(X_n)-y_*\ |X_{n-1})\leq\delta_c+(f(X_{n-1})-y_*)(1-\gamma_c).$$

Applying (8) repeatedly and using the Markov property we have

$$V_n(X_0)\leq\delta_c\sum_{k=0}^{n-1}(1-\gamma_c)^k+(f(X_0)-y_*)(1-\gamma_c)^n.$$

To prove (9) just note that $\lim_{n\to\infty} V_n(X_0) \leq \delta_c \sum_{k\geq 0}(1-\gamma_c)^k = \delta_c/\gamma_c = \varphi(c)$ and $0 < \gamma_c \leq 1$. Now (2) follows by observing that for ϵ small we have

$$P(|f(X_n) - y_*| \leq \epsilon) \geq 1 - \frac{E|f(X_n) - y_*|}{\epsilon} \geq 1 - \frac{\varphi(c)}{\epsilon}. \ \square$$

3 The continuous case

Let $0 < m(S) < \infty$ and $Q(dy|x) = q(y|x)dy$ so that $\{X_n\}_{n\geq 0}$ is a homogeneous chain with transition function given by:

$$P(X_{n+1} \in B \,|X_n = x) = \int_B q(y|x)a_c(y|x)dy + I_{(x\in B)}(1 - \int_S q(u|x)a_c(u|x)du). \tag{13}$$

In this case Condition 1 can be replaced by:

Condition 2 (a) For $\epsilon > 0$ we have $m\{y : y \in S, \ |f(y) - y_* \,| < \epsilon\} > 0$.
(b) Given $\epsilon > 0$ there exists $\delta(\epsilon) > 0$ such that

$$m(A) \leq \delta(\epsilon) \Rightarrow \int_A q(y|x)dy \leq \epsilon \ , \ \forall x \in S. \tag{14}$$

(c) There exist functions $b_c(\cdot) > 0, d_c(\cdot) > 0$ and $h(\cdot) > 0$ such that $b_c(0)$ is independent of c, $\lim_{c\downarrow 0} d_c(l) = 0$, $\lim_{c\downarrow 0} h(b_c(l))/b_c(l) = \lim_{c\downarrow 0} d_c(h(b_c(l)))/b_c(l) = 0$ and

$$\begin{aligned} f(y) - f(x) \ &\leq \ l \ \Rightarrow \ a_c(y|x) \geq b_c(l) \\ f(y) - f(x) \ &\geq \ l \ \Rightarrow \ a_c(y|x) \leq d_c(l). \end{aligned} \tag{15}$$

Remark 1 *(i) Condition 2(a) simply states that the global minima points are not isolated.*
(ii) Note that for every fixed $x_0 \in S$ if $m(A) \downarrow 0$ then $\int_A q(y|x_0)dy \downarrow 0$ thus Condition 2(b) amounts to uniform convergence with respect to x. Also it is satisfied if the sampling densities are uniformly bounded, that is, $q(y|x) \leq K < \infty$.
(iii) Most of the a.p.f. in the literature verify Condition 2(c). In particular for those considered in Example 1 we have: (a) Take $b_c(l) = d_c(l) = \exp(-l/c)$ and $h(x) = x^2$, it follows that $d_c(h(b_c(l)))/b_c(l) = \exp\{[l - \exp(-2l/c)]/c\} \to 0$ and $b_c^2(l)/b_c(l) \to 0$; (b) Take $b_c(l)$ and $d_c(l)$ as before and let $h(x) = x^2$ then we have $h(b_c(l))/b_c(l) = \exp\{-l/cl_1^2\} \to 0$ and $d_c(h(b_c(l)))/b_c(l) = \exp\{[\frac{l}{cl_1^2} - \exp(-2l/cl_1^2)]/cl_0^2\} \to 0$; (c) Take $h(x) = x^2$ and note that $d_c(h(b_c(l)))/b_c(l) \leq \exp\{-[\exp(-(\rho+1)2l/c)]/c\}/\exp\{-l/c\} \to 0$.

Theorem 2 *Let f be a bounded function on S then under Condition 2 we have:*

(a) There exist $\gamma_c > 0$ and $\delta_c \geq 0$ with $\lim_{c \downarrow 0} \frac{\delta_c}{\gamma_c} = 0$ and such that

$$L_c(x) = E\{f(X_1) - y_* \mid X_0 = x\} \leq \delta_c + (f(x) - y_*)(1 - \gamma_c). \tag{16}$$

(b) We have (4) and there exists a function $\varphi(c)$ such that

$$\lim_{c \downarrow 0} \varphi(c) = 0 \quad \textit{and} \quad \lim_{c \downarrow 0} ||f(X_n) - y_*||_1 \leq \varphi(c). \tag{17}$$

Proof: (a) The proof of (16) will be carried out in several steps:

step 1. From (13) we can write for $\alpha > 0$,

$$\begin{aligned} L_c(x) &= \int_S (f(y) - y_*) q(y|x) a_c(y|x) dy \\ &\quad + (f(x) - y_*)(1 - \int_s q(u|x) a_c(u|x) du) \\ &= A_c^\alpha(x) + D_c^\alpha(x) + (f(x) - y_*)(1 - E_c^\alpha(x) - F_c^\alpha(x)) \end{aligned} \tag{18}$$

where

$$\begin{aligned} A_c^\alpha(x) &= \int_{(f(x) \leq f(y) \leq f(x)+\alpha)} (f(y) - f(x)) q(y|x) a_c(y|x) dy \,, \\ D_c^\alpha(x) &= \int_{(f(y) > f(x)+\alpha)} (f(y) - y_*) q(y|x) a_c(y|x) dy \,, \\ E_c^\alpha(x) &= \int_{(f(y) < f(x))} \frac{f(x) - f(y)}{f(x) - y_*} q(y|x) a_c(y|x) dy \\ \text{and } F_c^\alpha(x) &= \int_{(f(y) > f(x)+\alpha)} q(y|x) a_c(y|x) dy. \end{aligned} \tag{19}$$

step 2. To estimate (18) we make a convenient partition of S. Let $0 < \epsilon < 1/3$ and $\delta(\epsilon)$ be given by (14). Define $k_0 = [3m(S)/\delta(\epsilon)]$, $C(l) = \{u : u \in s \,, f(u) \leq y_* + l\}$, $l_{k_0} = l^* = \sup_{x \in S}(f(x) - y_*)$ and $l_k = sup\{l : m(C(l)) \leq k\delta(\epsilon)/3\}$ for $k = 0, 1, \ldots, k_0 - 1$. Note that by assumption $l^* < \infty$, $m(C(0)) = 0$ and $m(C(l)) > 0$, $\forall\, l > 0$. Thus for $k = 1, \ldots, k_0 - 1$ we have $m(C(l_k)) > 0$ and $m(C(l_{k_0})) = m(S)$. For $k = 1, \ldots, k_0$ define $B_k = C(l_k) \backslash C(l_{k-1})$. If $k > k_0$ then we have $m(C(l_{k-1})) > (k-2)\delta(\epsilon)/3$, $m(C(l_k)) \leq k\delta(\epsilon)/3$ and $m(B_k) \leq 2\delta(\epsilon)/3$. Since $m(C(l_{k_0-1})) > (k_0 - 2)\delta(\epsilon)/3$ and $m(B(l_{k_0})) = m(S)$ we also have

$$m(B_{k_0}) \leq m(S) - (k_0 - 2)\frac{\delta(\epsilon)}{3} < (k_0 + 1)\frac{\delta(\epsilon)}{3} - (k_0 - 2)\frac{\delta(\epsilon)}{3} = \delta(\epsilon).$$

Using Condition 2(b) we have for $k = 1, \ldots, k_0$

$$m(B_k) \leq \delta(\epsilon) \text{ and } \int_{B_k} q(y|x)dy \leq \epsilon, \quad \forall\, x \in S. \tag{20}$$

Since $\epsilon < 1/3$ we also have

$$1 - \int_{B_{k-1} \cup B_k \cup B_{K+1}} q(y|x)dy \geq 1 - 3\epsilon = \epsilon_0 > 0\,. \tag{21}$$

In (21) we interpret $B_{k-1} = \emptyset$ if $k = 1$ and $B_{k+1} = \emptyset$ if $k = k_0$.

step 3. Let $\alpha < \rho = \min\{l_1, l_2 - l_1, \ldots, l_{k_0} - l_{k_0-1}\}$ and for given x let k such that $x \in B_k$. Then

$$B_k \cup B_{k+1} \subset \{y : f(x) \leq f(y) \leq f(x) + \alpha\}, \tag{22}$$

$$B_1 \cup \ldots \cup B_{k-1} \subset \{y : f(y) < f(x)\} \tag{23}$$

$$\text{and } B_{k+1} \cup \ldots \cup B_{k_0} \subset \{y : f(y) > f(x) + \alpha\}. \tag{24}$$

step 4. From (19), (20), (22) and the fact that $a_c(y|x) \leq 1$ we have $A_c^\alpha(x) \leq \alpha \int_{B_k \cup B_{k+1}} q(y|x)dy \leq 2\epsilon\alpha$. Since $l^* = \sup_{u \in S}(f(u) - y_*)$ it follows from (19) that

$$D_c^\alpha(x) \leq l^* \int_{(f(y) > f(x)+\alpha)} q(y|x) a_c(y|x)dy \leq l^* d_c(\alpha). \tag{25}$$

The last inequality follows from the fact that $q(\cdot|\cdot)$ is a density function and the fact that $a_c(y|x) \geq d_c(\alpha)$.

step 5. From (19) and (23) we have

$$E_c^\alpha(x) \geq \int_{B_1 \cup \ldots \cup B_{k-2}} \frac{f(x) - f(y)}{f(x) - y_*} q(y|x) a_c(y|x)dy.$$

Note that for $y \in B_1 \cup \ldots \cup B_{k-2}$, $x \in B_k$ and ρ defined in step 3 we have $f(x) - f(y) \geq \rho$. Since $f(y) \leq f(x)$ we have $a_c(y|x) \geq b_c(0)$ by Condition 2(c). Thus

$$E_c^\alpha(x) \geq \frac{\rho}{l^*} b_c(0) \int_{B_1 \cup \ldots \cup B_{k-2}} q(y|x)dy. \tag{26}$$

step 6. From (19) and (24) we have $F_c^\alpha(x) \geq \int_{B_{k+2} \cup \ldots \cup B_{k_0}} q(y|x) a_c(y|x)dy$. Since $a_c(y|x) \geq b_c(l^*)$ and $f(y) - f(x) \leq l^*$ we can write

$$F_c^\alpha(x) \geq b_c(l^*) \int_{B_{k+2} \cup \ldots \cup B_{k_0}} q(y|x)dy\,. \tag{27}$$

step 7. Note that $b_c(l) \leq d_c(l)$ and since $\lim_{c \downarrow 0} d_c(l) = 0$ we also have $\lim_{c \downarrow 0} b_c(l) = 0$ for $l > 0$. In particular, $\lim_{c \downarrow 0} b_c(l^*) = 0$. Therefore for c small we have $b_c(l^*) \leq \rho b_c(0)/l^*$. Since $q(\cdot|\cdot)$ is a density function we have from (26) and (27)

$$E_c^{\alpha}(x) + F_c^{\alpha}(x) \geq b_c(l^*)(1 - \int_{B_{k-1} \cup B_k \cup B_{k+1}} q(u|x)dy).$$

Now using (21) we get $E_c^{\alpha}(x) + F_c^{\alpha}(x) \geq b_c(l^*)\epsilon_0 = \gamma_c > 0$.

step 8. From step 3 we must choose $\alpha < \rho$ where ρ depends on ϵ. Using Condition 2(c) we take $\alpha = h(b_c(l^*))$. Since $b_c(l^*) \to 0$ and $h(b_c(l^*))/b_c(l^*) \to 0$ we have $\lim_{c \downarrow 0} h(b_c(l^*)) = 0$. Thus for c small we have $\alpha = h(b_c(l^*)) < \rho$. Now let $\delta_c = 2\epsilon h(b_c(l^*)) + l^* d_c(h(b_c(l^*))$. Using steps 1,4 and 7 we have $L_c(x) \leq \delta_c + (f(x) - y_*)(1 - \gamma_c)$. Finally from Condition 2(c) it follows that $\lim_{c \downarrow 0} \delta_c/\gamma_c = 0$.

(b) Using the same type of arguments as in Theorem 1 we obtain (17) and consequently (4). □

References

[1] Anily, S. and Federgruen, A. (1987). Simulated annealing methods with general acceptance probabilities. *J. Appl. Probab.* **24,** 657-667.

[2] Bélisle, C.J., (1992). Convergence theorems of a class of simulated algorithms on R^d. *J. Appl. Probab.* **29,** 885-895.

[3] Bohachavsky, I., Johnson, M. and Stein, M. (1986). Generalized simulated annealing for function optimization. *Technometrics* **28,** 209-217.

[4] Dekkes, A. and Aarts, E (1991). Global optimization and simulated annealing. *Mathematical Programming* **50,** 367-393.

[5] Doob, J.L. (1953). *Stochastic Processes.* Wiley, New York.

[6] Dorea, C. C. Y. (1996). On the efficiency of a continuous version of the simulated annealing algorithm. *Stat. Probab. Letters* **31,** 247-253.

[7] Gidas, B. (1985). Nonstationary Markov chains and convergence of the annealing algorithm. *J. of Statistical Physics* **39,** 73-131.

[8] Hajek, B. (1988). Cooling schedules for optimal annealing. *Mathematics of Operations Research* **13,** 311-329.

[9] Kirkpatrick, S., Gelat, Jr. C. and Vecchi, M. (1983). Optimization by simulated annealing. *Science* **220,** 671-680.

[10] Lundy, M. and Mees, A. (1986). Convergence of an anneling algorithm. *Mathematical Programming* **34,** 111-124.

[11] Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.* **22,** 1701-1762.

Student (1997), Vol.2, No.2, 85-94

On a Rank Test for Autoregressive Conditional Heteroscedasticity

Alessandro Fassò

Istituto di Statistica, Università Cattolica, Milano, Italy

Abstract: A new rank test has been recently proposed by the author for testing randomness against autoregressive conditional heteroscedastic alternatives. Monte Carlo experiments show good approximation to the normal distribution and good power properties for finite samples. In the present paper, it is shown that the rank test has good asymptotic properties as well, i.e. high asymptotic efficiency with respect to the Lee-King test, which is the normal theory reference test; in particular Pitman's efficiency is one when the parent distribution is Normal and is greater than one when the parent distribution belongs to beta, Student's t and gamma families.

Key words: Contiguity, ARCH model, robustness, asymptotic efficiency.

1 Introduction

The problem of testing randomness against *autoregressive conditional heteroscedasticity* (*ARCH*) has only recently received a satisfactory solution by the one sided test of Lee and King (1993) which is based on the *locally most mean powerful* (*LMMP*) test statistics, under normal assumptions.

In the present paper, the asymptotic power of the rank test proposed by the author in Fassò (1995) is studied; the approach is based on LeCam's contiguity theory and follows the lines of Hallin, Ingenbleek and Puri (1985) and Hallin and Puri (1988). While these authors deduce the optimal rank test for each given parent distribution in *ARMA* model testing, here, with the aim of robustness in *ARCH* model testing, the two above tests are compared for different parent distributions using the Pitman's asymptotic relative efficiency (*ARE*).

A contiguity theory similar to Section 2 below is considered by Linton (1993) in estimating a regression with *ARCH* errors. In many respects the approach of Linton is more general, in particular he does not restrict contiguity to independence, which is the point of interest in this paper, and uses a smooth condition for the error density which is weaker than used in Proposition 1. As a metter of fact here, the proofs are simpler and the assumption of symmetric error density used by Linton is avoided.

The good asymptotic properties of the rank test are reinforced by its finite sample counterparts: in Fassò (1995), Monte Carlo experiments show that -*a*- under the null, the approximation to the normal distribution of the rank statistic is noticeably superior to the Lee-King's one, -*b*- the powers of these two tests are very close and neither of them dominates the other for various finite sample sizes, -*c*- these properties carry over to testing regression residuals at least in the case of symmetric distributions.

The paper is organized as follows. In Section 2, Proposition 1 states the asymptotic normality of the log-likelihood-ratio, which is the preliminary result allowing the use of LeCam's contiguity theory.

In Section 3, the rank statistic is defined and its exact mean and variance are given under the null. Essentially the idea underlying the robust test is: use the *LMMP* statistic applied to *normalized* data instead of row data; to do this a transform asymptotically equivalent to *the* probability integral transform is used, namely the distribution-free transform given by the van der Waerden score function defined in (5) below. In Section 4, using LeCam's third Lemma, the asymptotic normality is deduced under contiguous alternatives of both the rank statistic and the Lee-King statistic. This allows power computations and the *ARE* comparisons. When the parent distribution is normal, the *ARE* is one, so (asymptotically) nothing is lost by substituting the data with their ranks; moreover figures always greater than one are obtained for various beta, Student's t and gamma distributions. The *uniform* asymptotic superiority of the rank test can then be conjectured.

2 Contiguity theory for the *ARCH* model

The linear *ARCH* model of order p is given by $\epsilon_t = \eta_t h_t$, $t = p+1, p+2, ...,$ where $h_t^2 = \sigma^2 + \sum_{j=1}^{p} \alpha_j \epsilon_{t-j}^2$, ϵ_t,are the observed data, $\sigma^2 > 0$,

$$\alpha_j \geq 0 \ \textit{and} \ \sum_{j=1}^{p} \alpha_j < 1. \tag{1}$$

Moreover η_t are independent and identically distributed as the random variable η, which is supposed absolutely continuous with cumulative dis-

tribution $F()$, density $f()$ and (except in Section 3) moments $E(\eta) = 0$, $E(\eta^2) = 1$ and $E(\eta^4) = \kappa < \infty$. Suppose further that the initial values $\epsilon_1, ..., \epsilon_p$ are independent and identically distributed as $\sigma\eta$.

Let $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)'$ be the p-dimensional column vector of the parameters of interest and consider the sequences of distributions generated by the hypotheses $H_0^n : \boldsymbol{\alpha} = 0$ and $H_1^n : \boldsymbol{\alpha}^n = \frac{1}{\sqrt{n}}\boldsymbol{\alpha}$, where α is any fixed p-vector satisfying (1). The hypotheses H_0^n and H_1^n are called contiguous if the corresponding sequences of distributions are contiguous in LeCam's sense (see e.g. Hájek and Šidák, 1967, Section VI.1).

In order to use this theory some more assumptions and symbols are needed. First suppose that standard regularity conditions are satisfied by the density $f()$ allowing interchange of differentiation and integration up to second order so that, letting $\phi(x) = -\dot{f}(x)/f(x) = -\frac{d}{dx}\ln f(x)$ and $\dot{\phi}(x) = \frac{d}{dx}\phi(x)$, we can write $E(\phi) = 0$ and $E(\phi^2) = E(\dot{\phi})$, where $\phi = \phi(\eta)$ and $\dot{\phi} = \dot{\phi}(\eta)$.

Next consider the conditional log-likelihood of $\epsilon_t|(\epsilon_1, ...\epsilon_{t-1})$, $t = p + 1, p + 2, ...$, i.e. $\ln L_{t|t-1}(\boldsymbol{\alpha}) = \ln f(\epsilon_t/h_t) - \ln h_t$ and, conditionally on the asymptotically negligeable initial values $\epsilon_1, ..., \epsilon_p$, consider the log-likelihood of $\epsilon_{p+1}, ..., \epsilon_n$, i.e. $\ln L_n(\boldsymbol{\alpha}) = \sum_{t=p+1}^{n} \ln L_{t|t-1}(\boldsymbol{\alpha})$. Let $u_t = (u_{t,1}, ..., u_{t,p})'$ be the score vector given by the partial derivatives of $\ln L_{t|t-1}(\boldsymbol{\alpha})$ computed at $\boldsymbol{\alpha}$ =0; i.e.

$$u_{t,j} = \frac{1}{2}\frac{\epsilon_{t-j}^2}{\sigma^2}(\phi(\epsilon_t/\sigma)\frac{\epsilon_t}{\sigma} - 1) = \frac{1}{2}\eta_{t-j}^2(\phi_t\eta_t - 1) \tag{2}$$

with $\phi_t = \phi(\eta_t)$.

Now let $o_p(a_n)$ such that $o_p(a_n)/a_n$ goes to zero in probability as $n \to \infty$ and let the symbol $\xrightarrow{d}$ denote convergence in distribution. Moreover let $N(\cdot, \cdot)$ denote a normal random variable.

With these symbols in the Appendix it is proven the following Proposition 1 which entails that H_0^n and H_1^n are contiguous alternatives.

Proposition 1 *Suppose that the following conditions hold: a) for* $i = 0, 1$ $\exists\ K_i : \mathbf{R} \to \mathbf{R}^+$ *so that* $E(K_i(\eta)\eta^{i+1}) < \infty$, $|\phi(ax)| \leq K_0(x)$ *and* $|\dot{\phi}(ax)| \leq K_1(x)$, $\forall 0 < a < 1$; *b)* $E(|\phi|\eta^2) < \infty$ *and c)* $\dot{\phi}(x)$ *is continuous.*

Then, as $n \to \infty$, *under* H_0^n,

$$\frac{1}{\sqrt{n}}\sum_{t=p+1}^{n} \boldsymbol{\alpha}' u_t \xrightarrow{d} N(0, \delta^2) \tag{3}$$

and

$$\ln L_n(\boldsymbol{\alpha}^n) = \ln L_n(0) + \frac{1}{\sqrt{n}}\sum_{t=p+1}^{n} \alpha' u_t - \frac{1}{2}\delta^2 + o_p(1) \tag{4}$$

where $\delta^2 = \delta(\alpha)^2 = \frac{1}{4}E(\phi^2\eta^2 - 1)(\sum_{i\neq j=1}^{p} \alpha_i\alpha_j + \kappa\sum_{i=1}^{p}\alpha_i^2)$.

Remark 1 *The regularity a) to c) above are technical conditions about integrability of the shocks* η_t *and* ϕ *and* $\dot{\phi}$. *They are usually satisfied by regular location families. For example these conditions are satisfied by the distributions used in Section 4, namely the Student's t distribution with* $\nu > 2$ *degrees of freedom, the Gamma distribution* $\Gamma(p)$ *with* $p > 0$ *and the Beta distribution* $\beta(p,q)$ *with* $p, q > 2$.

3 The rank test

Consider the van der Waerden scores

$$z_{n,t} = \psi(\frac{r_{n,t}}{n+1}), \tag{5}$$

where $\psi()$ is the inverse of the standard normal distribution function, i.e. $\psi(y)^{-1} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} exp\left(-\frac{x^2}{2}\right)dx$, and $r_n = (r_{n,1}, ..., r_{n,n})'$ is the rank vector of $\epsilon_1, ..., \epsilon_n$, i.e. $r_{n,t} = \sum_{s=1}^{n} I(\epsilon_s \leq \epsilon_t)$ where I() is the indicator function. Then the rank statistic

$$S_n = \sum_{t=p+1}^{n} z_{n,t}^2 \sum_{j=1}^{p} z_{n,t-j}^2$$

with high significant values, is proposed for testing randomness against *ARCH*.

To motivate S_n from the rank-theory point of view, consider testing H_0^n against $\alpha_j > 0$ for a fixed j. If η_t are standard normal variables, apart from negligeable terms, the score is $\frac{\partial}{\partial\alpha_j}\ln L_n(0) = \sum_{t=p+1}^{n}\eta_t^2(\eta_{t-j}^2 - 1) = S_j^0(\eta)$, say, and, using Theorem II.4.8 of Hájek and Šidák (1967), the locally most powerful rank test statistic is given by $\check{S}_j(r_n) = E(S_j^0(\eta)|r_n)$. Now, it is well known that, as $n \to \infty$, $E(\eta_t^2\eta_{t-j}^2|r_n) \cong z_{n,t}^2 z_{n,t-j}^2$; moreover the involved function $\psi()$ is available in most statistical software packages. It is then a natural choice to use the rank statistic

$$\sum_{j=1}^{p}\sum_{t=p+1}^{n} z_{n,t}^2(z_{n,t-j}^2 - 1) = \sum_{j=1}^{p}\check{S}_j(r_n) + o_p(\sqrt{n}).$$

Return to S_n now. In order to study its exact distribution under H_0^n, let

$$\mu_{i_1,...,i_k} = E(z_{n,t_1}^{i_1}...z_{n,t_k}^{i_k}) = \frac{(n-k)!}{n!}\sum_{j_1\neq...\neq j_k}\psi\left(\frac{j_1}{n+1}\right)^{i_1}...\psi\left(\frac{j_k}{n+1}\right)^{i_k},$$

where $t_h \neq t_i$, $j_h \neq j_i$, $j_h = 1, .., n$ *for* $h \neq i = 1, ..., k$. In Fassò (1995), it is shown that, under H_0^n, for any density $f()$, $E(S_n) = p(n-p)\mu_{2,2}$ and $Var(S_n) = \sigma_{0,n}^2 = A + B + C + D - p^2(n-p)^2\mu_{2,2}^2$, where

$A = p(n-p)(\mu_{4,4} + (p-1)\mu_{4,2,2})$,

$B = 2(n-p-1)((2p-1)\mu_{2,4,2} + (p-1)^2\mu_{2,2,2,2})$,

$C = 2I(n \geq p+3)I(p \geq 2)((n-p)(p_0-1)c_1$

$+(p_0(p_0+1)/2 - 1)((n-p)c_2 - c_1) - (p_0(p_0+1)(2p_0+1)/6 - 1)c_2)$,

$p_0 = \min(p, n-p-1)$, $c_1 = p((p-2)\mu_{2,2,2,2} + 2\mu_{2,4,2})$, $c_2 = \mu_{2,2,2,2} - \mu_{2,4,2}$,

$D = p^2 I(n \geq 2p+2).((n-p)(n-3p-1) + p(p+1))\mu_{2,2,2,2}$.

Note that the multidimensional sums above can be simplified into one-dimensional sums using the following relations: $\mu_{a,b} = (n\mu_a\mu_b - \mu_{a+b})/(n-1)$, $\mu_{a,b,b} = (n\mu_a\mu_{b,b} - 2\mu_{a+b,b})/(n-2)$ and $\mu_{a,b,b,b} = (n\mu_a\mu_{b,b,b} - 2\mu_{a+b,b,b})/(n-3)$.

Asymptotic normality of S_n under H_0^n is a particular case of Proposition 2. But, using (14) with $a = 0$, the regularity assumptions of Proposition 1 are no more needed. So we have the following general result.

Corollary 1 *For any density* $f()$, *as* $n \to \infty$, $\sigma_{0,n}^2/n \to 4p$ *and*

$$S_n^* = \frac{S_n - E(S_n)}{\sigma_{0,n}} \xrightarrow{d} N(0,1).$$

4 Asymptotic power and ARE

In order to compare the asymptotic power of the rank test (with respect) to the Lee-King test, the following version of the Lee-King statistic is used

$$LK_n = \frac{1}{\sqrt{n}} \sum_{t=p+1}^{n} \left(\frac{\epsilon_t^2}{\sigma^2} - 1 \right) \sum_{j=1}^{p} \left(\frac{\epsilon_{t-j}^2}{\sigma^2} - 1 \right),$$

which is asymptotically equivalent to equation (12) of Lee and King (1993) as $n \to \infty$ under (1).

The asymptotic normality of both statistics S_n and LK_n under the contiguous alternatives is then established. To do this observe that the LeCam's third Lemma (Hájek and Šidák, 1967, p. 208) extends the asymptotic normality of a statistic from the null hypothesis to contiguous alternatives if the joint distribution of the statistic and the log-likelihood ratio is asymptotically bivariate normal, $N_2(\cdot, \cdot)$, under H_0^n as stated in the following Proposition which is proven in the Appendix.

Proposition 2 *If the assumptions of Proposition 1 hold, under* H_0^n, *as* $n \to \infty$, *then*

$$\left(\ln(\frac{L_n(\alpha^n)}{L_n(0)}), \frac{S_n - E(S_n)}{\sqrt{n}}\right) \xrightarrow{d} N_2\left(\begin{pmatrix} -\delta^2/2 \\ 0 \end{pmatrix}, \begin{pmatrix} \delta^2 & \gamma_S \\ \gamma_S & 4p \end{pmatrix}\right) \quad (6)$$

where $\gamma_S = E(\phi\eta\Psi^2 - 1)E(\eta^2\Psi^2 - 1)\sum_{i=1}^p \alpha_i$ *and* $\Psi = \psi(F(\eta))$. *Moreover*

$$\left(\ln(\frac{L_n(\alpha^n)}{L_n(0)}), \frac{LK_n}{\sqrt{n}}\right) \xrightarrow{d} N_2\left(\begin{pmatrix} -\delta^2/2 \\ 0 \end{pmatrix}, \begin{pmatrix} \delta^2 & \gamma_{LK} \\ \gamma_{LK} & pVar(\eta^2)^2 \end{pmatrix}\right) \quad (7)$$

where $\gamma_{LK} = 2Var(\eta^2)\sum_{i=1}^p \alpha_i$.

The asymptotic normality of both the rank statistic and the Lee-King statistic under contiguous alternatives now follows by direct application of the above mentioned Lemma, as stated in the next corollary.

Corollary 2 *Using the notation and the assumptions of Proposition 2, then, under* H_1^n, *as* $n \to \infty$, $\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(\gamma_S, 4p)$ *and* $\frac{LK_n}{\sqrt{n}} \xrightarrow{d} N(\gamma_{LK}, pVar(\eta^2)^2)$.

As a consequence of Corollary 2 the Pitman's ARE for contiguous alternatives H_1^n is given by

$$ARE_{S,LK} = \left(\frac{\gamma_S^2}{4p}\right)\left(\frac{\gamma_{LK}^2}{pVar\,(\eta^2)^2}\right)^{-1} = \left(\frac{E(\phi\eta\Psi^2 - 1)E(\eta^2\Psi^2 - 1)}{4}\right)^2.$$

First observe that this quantity does not depend either on α, i.e. on the direction of moving away from H_0^n, or on the order of the $ARCH$ model p. Second if η is a normal random variable then $\eta = \Psi = \phi$ and $ARE = 1$, so the competing tests are, in this case, equivalent. Finally, the figures of Table 1, related to various standardized random variables, show that the ARE seems to increase by moving away from normality; in particular in the direction of skewness. The conjecture that the distribution-free *normalization* achieved by the van der Waerden scores improves the Lee-King statistic seems then a natural one.

Appendix

Proof of Proposition 1: In order to prove Proposition 1, consider the following lemma, the proof of which is immediate using integration by parts.

Lemma 1 *If the expectations involved are finite and* $E\left(|\eta|^k\right) < \infty$, *then* $E(\phi\eta^k) = kE(\eta^{k-1})$; *moreover if* $E\left(|\eta|^k\right) < \infty$, *then* $E(\dot{\phi}\eta^k) = E(\phi^2\eta^k) - kE(\phi\eta^{k-1})$.

Distribution	*ARE*	*Kurtosis*	*Skewness*
β (3,3)	1.226	2.3	
β (6,6)	1.047	2.6	
β (10,10)	1.016	2.7	
β (3,6)	1.315		0.41
β (3,10)	1.538		0.64
Student's $t(5)$	1.361	9	
Student's $t(10)$	1.069	4	
Student's $t(15)$	1.029	3.6	
Γ (3)	2.415	5	1.15
Γ (5)	1.630	4.2	0.89
Γ (10)	1.261	3.6	0.63
Γ (20)	1.120	3.3	0.45
Γ (30)	1.078	3.2	0.37

Table 1: Pitman's efficiency of rank test S_n with respect to Lee-King test and parent distribution shape indices.

Proof of (3): The asymptotic normality follows easily (e.g. from Dianananda, 1952, Theorem 2) observing that, under H_0^n, from (2) $\boldsymbol{\alpha}' u_t$ is a stationary sequence of p-dependent and incorrelated random variables with zero mean (note that the previous Lemma entails $E(\phi\eta) = 1$) and variance given by

$$Var(\alpha' u_t) = \frac{1}{4} \sum_{i,j=1}^{p} \alpha_i \alpha_j E(\eta_{t-i}^2 \eta_{t-j}^2) E(\phi_t \eta_t - 1)^2 = \delta^2. \tag{8}$$

Proof of (4): In order to study the log-likelihood ratio consider the following expansion of the conditional log-likelihood

$$\ln L_{t|t-1}(\boldsymbol{\alpha}^n) = \ln L_{t|t-1}(0) + \frac{1}{\sqrt{n}} \boldsymbol{\alpha}' u_t + \frac{1}{2n} \alpha' \tilde{Q}_{n,t}^2 \alpha \tag{9}$$

where

$$\tilde{Q}_{n,t}^2 = Q_{n,t}^2(\tilde{\theta}_{n,t} \boldsymbol{\alpha}^n) = (\tilde{q}_{n,t}(i,j))_{i,j=1,\ldots,p} \quad , \tag{10}$$

say, $Q_{n,t}^2(\alpha) = \left(\frac{\partial^2 \ln L_{t|t-1}(\alpha)}{\partial \alpha_i \partial \alpha_j} \right)_{i,j=1,\ldots,p}$ and $\tilde{\theta}_{n,t} = \theta_n(\epsilon_t, \ldots, \epsilon_{t-p}) \in [0,1]$.
The sum of the right hand side of (9) leads to

$$\ln L_n(\boldsymbol{\alpha}^n) = \ln L_n(0) + \frac{1}{\sqrt{n}} \sum_{t=p+1}^{n} \boldsymbol{\alpha}' u_t + \frac{1}{2n} \alpha' \tilde{Q}_n^2 \alpha$$

where $\tilde{Q}_n^2 = \sum_{t=p+1}^n \tilde{Q}_{n,t}^2$. Hence (4) holds if

$$\frac{1}{n}\alpha' \tilde{Q}_n^2 \alpha \to -\delta^2, \quad in \ \ probability, \quad as \quad n \to \infty. \tag{11}$$

This will be shown in the following three steps, which hold for indices $i, j = 1, ..., p$; these indices will be omitted for notational simplicity when not necessary.

1. From the definition in (10) and $0 \leq \tilde{\theta}_{n,t} \leq 1$ it holds $\tilde{q}_{n,t} = \frac{1}{2}\tilde{\eta}_{t,t-i}^2 \tilde{\eta}_{t,t-j}^2 (1 - \frac{3}{2}\phi(\tilde{\eta}_{t,t})\tilde{\eta}_{t,t} - \frac{1}{2}\dot{\phi}(\tilde{\eta}_{t,t})\tilde{\eta}_{t,t}^2)$ with $\tilde{\eta}_{t,s} = \epsilon_s \left(\sigma^2 + n^{-1/2}\tilde{\theta}_{n,t} \sum_{j=1}^p \alpha_j \epsilon_{t-j}^2\right)^{-1/2} \to \eta_s$, almost surely, as $n \to \infty$. It follows that, almost surely,

$$\tilde{q}_{n,t} \to q_t(i,j) = \frac{1}{2}\eta_{t-i}^2 \eta_{t-j}^2 (1 - \frac{3}{2}\phi_t \eta_t - \frac{1}{2}\dot{\phi}_t \eta_t^2), \quad as \quad n \to \infty. \tag{12}$$

Moreover, from condition a) of Proposition 1 it holds

$$|\tilde{q}_{n,t}| \leq \frac{1}{2}\eta_{t-i}^2 \eta_{t-j}^2 (1 + \frac{3}{2}K_0(\eta_t)|\eta_t| + \frac{1}{2}K_1(\eta_t)\eta_t^2);$$

hence, using the Dominated Convergence Theorem, the convergence in (12) holds in $L_1 - norm$ too. Now observe that $\tilde{\theta}_{n,t} = \theta_n(\epsilon_t, ..., \epsilon_{t-p})$ is measurable, as $f()$ and $Q_{n,t}^2(\boldsymbol{\alpha})$ are measurable functions of $\epsilon_t, ..., \epsilon_{t-p}$; hence, for fixed n, the sequence $\tilde{\theta}_{n,t}$, $t > p$, is a stationary sequence (see e.g. Stout (1974), Theorem 3.5.3) and the same holds for $\tilde{q}_{n,t}$. It follows that $E(|\tilde{q}_{n,t} - q_t|)$ does not depend on t, and $\frac{1}{n}\sum_{t=p+1}^n (q_{n,t} - q_t) \xrightarrow{L_1} 0$, as $n \to \infty$.

2. The sequence q_t is stationary and p-dependent, hence ergodic (see e.g. Shiryayev (1984), Theorem 2, p.381). Then almost surely, $\frac{1}{n}\sum_{t=p+1}^n q_t \to \chi$, as $n \to \infty$; where

$$\chi = \chi(i,j) = E(q_t(i,j)) = \frac{1}{4}E(\eta_{t-i}^2 \eta_{t-j}^2) E(\phi_t \eta_t - 1)^2 \tag{13}$$

and the right hand side follows from Lemma 1.

3. Steps -1- and -2- entail that $\frac{1}{n}\sum_{t=p+1}^n (\tilde{q}_{n,t} - \chi) = \frac{1}{n}\sum_{t=p+1}^n (\tilde{q}_{n,t} - q_t) + \frac{1}{n}\sum_{t=p+1}^n (q_t - \chi) \to 0$, in probability, as $n \to \infty$, finally (11) follows $\forall \alpha$ from (8) and (13).

Proof of Proposition 2: In order to prove Proposition 2, let the symbol $\cong$ denote equality apart for terms which are $o_p(1)$ as $n \to \infty$ under $H_1^n \bigcup H_o^n$ and consider the following lemma.

Lemma 2 *Let* $\hat{S}_n = \frac{1}{\sqrt{n}}\sum_{t=p+1}^n (\Psi_t^2 - 1)\sum_{j=1}^p (\Psi_{t-j}^2 - 1)$ *where* $\Psi_t = \psi(F(\eta_t))$. *Then, if* $\eta_1, ..., \eta_n$ *are independent and identically distributed continuous random variables , then* $\hat{S}_n \cong 2\sqrt{p} S_n^*$.

Proof: Let $\tilde{S}_n = \frac{1}{\sqrt{n}} \sum_{t=p+1}^{n} \Psi_t^2 \sum_{j=1}^{p} \Psi_{t-j}^2 - \sqrt{n} p \bar{\Psi}_2^2$ where $\bar{\Psi}_k = \frac{1}{n} \sum_{t=1}^{n} \Psi_t^k$. Then from the equivalence $\tilde{S}_n \cong \frac{1}{\sqrt{n}} \sum_{t=p+1}^{n} (\Psi_t^2 - \bar{\Psi}_2) \sum_{j=1}^{p} (\Psi_{t-j}^2 - \bar{\Psi}_2) \cong \hat{S}_n$, the Lemma follows using Section 4.1 of Hallin, Ingenbleek and Puri (1985) applied to $2\sqrt{p} S_n^*$ and $\tilde{S}_n$. To see this, first observe that in the present context the permutation expectation m_n of Hallin, Ingenbleek and Puri (1985) is $m_n = p\mu_{2,2} = E(S_n)/(n-p)$. Next in order to compute e_n given by Hallin, Ingenbleek and Puri (1985), Equation (4.4), observe that, from the additive structure of $\Psi_t^2 \sum_{j=1}^{p} \Psi_{t-j}^2$, only the case $p = 1$ can be considered, giving

$$\begin{aligned} e_n &= \frac{\sqrt{n-1}}{n(n-1)} \sum_{s \neq t=1}^{n} \Psi_s^2 \Psi_t^2 = \frac{1}{n\sqrt{n-1}} \left((\sum_{s=1}^{n} \Psi_s^2)^2 - \sum_{s=1}^{n} \Psi_s^4 \right) \\ &= \frac{n}{\sqrt{n-1}} \bar{\Psi}_2^2 - \frac{1}{\sqrt{n-1}} \bar{\Psi}_4. \end{aligned}$$

Hence $e_n \cong \sqrt{n} p \bar{\Psi}_2^2$ follows from $\bar{\Psi}_4 \to 3$, almost surely.

Proof of (6): In view of (4) and Lemma 2 above the asymptotic normality of x_t in the following (14) is sufficient for (6). To this end, observe that

$$a \ln \left(\frac{L_n(\alpha^n)}{L_n(0)} \right) + b\hat{S}_n \cong a \left(\frac{1}{\sqrt{n}} \sum_{t=p+1}^{n} \alpha' u_t - \frac{\delta^2}{2} \right) + b\hat{S}_n = \frac{1}{\sqrt{n}} \sum_{t=p+1}^{n} x_t - a\frac{\delta^2}{2} \tag{14}$$

say, where x_t is a stationary, finite variance and p-dependent sequence for every $a, b \in \mathbf{R}$. Hence joint asymptotic normality follows e.g. from Diananda (1952), Theorem 2, with

$$\begin{aligned} Cov \left(\ln \left(\frac{L_n(\alpha^n)}{L_n(0)} \right), \hat{S}_n \right) &\cong E\left((\phi\eta - 1)(\Psi^2 - 1) \right) E\left(\eta^2 (\Psi^2 - 1) \right) \sum_{i=1}^{p} \alpha_i \\ &= \gamma_S \end{aligned} \tag{15}$$

and marginal moments as given by Proposition 1 and Corollary 1.

Proof of (7): Normality follows in the same way as for (6) by substituting η_t for Ψ_t in (14) and (15); in particular γ_{LK} follows similarly to (15) observing that, from Lemma 1, $E(\phi\eta^2 - 1) = 2$.

References

[1] Diananda, P.H. (1952). Some probability limit theorems with statistical applications. *Proc. Cambr. Philos. Soc.* **49,** 239-246.

[2] Fassò, A. (1995). On some tests for detecting *ARCH* components (in Italian). Working paper. n.72, Università Cattolica, Milano. To appear in *Statistica* (1997) **LVII,** n.1.

[3] Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[4] Hallin, M., Ingenbleek, J. and Puri, M.L. (1985). Linear serial rank tests for randomness against ARMA alternatives. *Ann. Statist.*, **13,** 1156-1181.

[5] Hallin, M. and Puri, M.L. (1988). Optimal rank-based procedures for time series analysis: testing an ARMA model against other ARMA models. *Ann. Statist.* **16,** 402-432.

[6] Lee, J.H.H. and King, M. (1993). A locally most mean powerful based score test for *ARCH* and *GARCH* regression disturbances. *J. Business & Econ. Statist.* **11,** 17-27.

[7] Linton O. (1993) . Adaptive Estimation in *ARCH* Models. *Econometric Theory* **9,** 539-569.

[8] Shiryayev, A.N. (1984). *Probability.* Springer Verlag, New York.

[9] Stout, W.F. (1974). *Almost Sure Convergence.* Academic Press, New York.

Student (1997), Vol.2, No.2, 95-108

Dimension Choice for Sliced Inverse Regression Based on Ranks

Louis Ferré

Laboratoire de Statistique, Université Paul Sabatier, Toulouse, France

Abstract: Sliced Inverse Regression is a method for reducing the dimensionality in multivariate non parametric regression problems. While the selection of the dimensionality has been investigated for the original version, no solution has been proposed for Hsing and Carroll (1992) approach based on order statistics and associated concomitant variables. By using model selection approaches, we propose here two ways for selecting the dimensionality by estimating a loss function: first, a direct estimation is proposed and, then a Jack-Knifed estimate is investigated. Finally, the rank version is compared to classical SIR on a real life data set.

Key words: Nonparametric regression, dimension reduction, order and concomittent statistics, squared trace correlation.

1 Introduction

Consider a random variable (X, Y) with X in $\Re^p$ and Y in $\Re$ and let (X_i, Y_i) be n independent variables with same distribution as (X, Y). To simplify, consider the usual regression problem. Let r be the regression function of Y in X. Then, we have $Y = r(X) + \varepsilon$. To overcome the problem of the curse of the dimensionality, Sliced Inverse Regression has been introduced by Li (1991). Based on a more general model, the main idea of Sliced Inverse Regression is to assume that the relationship between Y and X only depends on K linear combinations of X:

$$Y = f(\beta_1' X, ..., \beta_K' X, \varepsilon). \tag{1}$$

Then, after the β's are estimated, the function f may be estimated by means of classical methods such as Kernel regression, Spline regression,

wavelet regression, etc... By assuming linear independency, the vectors $(\beta_1, ..., \beta_K)$ span a K dimensional subspace called the Effective Dimension Reduction (EDR) space. Now, the model indicates that the projection of X onto this EDR space is sufficient to fully explain the relationship between Y and X and then, it is enough to estimate the EDR space.

We assume that the expectation of X given its projection on the EDR space is linear in this projection:

$$\forall \beta \in \Re^p, \quad E\left(\beta' X | \beta_1' X, ..., \beta_K' X\right) \text{ is linear in } \left(\beta_1' X, ..., \beta_K' X\right) \tag{2}$$

This condition, condition 3.1 of Li (1991), is satisfied if X has an elliptically symmetric distribution. Under this condition, Li (1991) shows that the EDR space contains the space spanned by the $\Sigma-$orthogonal eigenvectors associated to the K non zero eigenvalues of the matrix $\Sigma^{-1}V$ where $\Sigma = var(X)$ and $V = var(E(X/Y))$. Now, an estimate of the EDR space is obtained from suitable estimates of the above matrices, say $\widehat{V}$ and $\widehat{\Sigma}$, and let $(\widehat{\beta}_1, ..., \widehat{\beta}_p)$ be the eigenvectors of $\widehat{\Sigma}^{-1}\widehat{V}$. In classical SIR, the matrix V is estimated (or rather approximated) by slicing the data and the asymptotic study has lead to an extended literature, see for instance Li (1991), Kotter (1996), Saracco (1996).

For the sake of simplicity, we will assume that the EDR space coincide with the eigensubspace generated by the non zero eigenvalues of $\Sigma^{-1}V$. However, it is well-known that SIR can miss relevant directions and alternative methods have been proposed such as SIR II (Li, 1991) or principal Hessian direction method (Li, 1992). Let us stress that in Ferré (1998), a model selection approach, based on the same loss function as the one below, is investigated for the principal Hessian direction method.

The approach of Hsing and Carroll (1992) is somewhat original since they avoid slicing by considering the eigenvalue decomposition of the matrix $I_p - \Sigma^{-1}\Lambda$ where $\Lambda = E(var(X|Y))$. As estimates they propose the following matrices:

$$\widehat{\Lambda} = \sum_{i=1}^{[n/2]} (X_{[2i]} - X_{[2i-1]})(X_{[2i]} - X_{[2i-1]})'/n \tag{3}$$

and

$$\widetilde{\Lambda} = \sum_{i=2}^{n} (X_{[i]} - X_{[i-1]})(X_{[i-1]} - X_{[i-1]})'/2n, \tag{4}$$

where $X_{[i]}$ denotes the concomitant variable in X of the order statistic in Y. Actually, this approach consists in dealing with $[n/2]$ slices with cardinal 2 when $\widehat{\Lambda}$ is used.

For the original approach (Li, 1991), and, for SIR II (Scott, 1994), the crucial choice of the dimension K is investigated by means of statistical tests. However, the retained solution is related to the choice of the level at each step of the procedure. Furthermore, for such nested tests, the knowledge of the level at each step do not implies the knowledge of the overall level that remains unknown. For those reasons, Ferré (1998) prefers a model selection approach similar in mind to the Akaike criterion or Mallows Cp used for instance in variable selection for linear regression. While a criterion based on the prediction of Y would be optimal, it main defect is that it requires the estimation of f which depends itself on the number of combinations selected. Then, it seems more relevant to only focus on the projections to choose the dimension of the subspace for projecting, without including the estimation of f.

A natural way to proceed is to study the closeness between the K "ideal" variables $(\beta_1'X, ..., \beta_K'X)$ and the "observed" $(\widehat{\beta}_1'X, ..., \widehat{\beta}_K'X)$. As mentionned by Li (1991), this can be done by means of the squared trace correlation between these variables. This quantity is actually the squared trace correlation between $\text{Span}((\Sigma^{1/2}\beta_1, ..., \Sigma^{1/2}\beta_K))$ and $\text{Span}((\Sigma^{1/2}\widehat{\beta}_1, ..., \Sigma^{1/2}\widehat{\beta}_K))$, see Ferré (1998). In practice $\text{Span}((\Sigma^{1/2}\widehat{\beta}_1, ..., \Sigma^{1/2}\widehat{\beta}_K))$ is not observable and it is necessary to include the estimation of Σ to consider what is really observed. This point leads Ferré (1998) to propose, the quantity $E(tr(\hat{\Pi}_q\Pi_q)/q)$, where Π_q, respectively $\hat{\Pi}_q$, the Σ (respectively $\widehat{\Sigma}$), orthogonal projector onto the subspace spanned by the q largest eigenvalues of $\Sigma^{-1}V$ (assumed to be simple to simplify), resp. $\widehat{\Sigma}^{-1}V$, for each potential solution $q = 1$ to p (recall that K is actually unknown), to evaluate the closeness between the "ideal" subspace and the observed one. Since the squared trace correlation is in [0,1], the loss function that we consider is:

$$R(q) = 1 - E(tr(\hat{\Pi}_q\Pi_q)/q). \tag{5}$$

Now, a value of $R(q)$ close to zero indicates that the set of the q "observable" linear combinations of X is close to the "ideal" set so that q is a feasible solution. At the opposite, a value of $R(q)$ significantly different from zero means that this set is slightly different from the ideal one so that the solution is lower that q. Then, the solution could be obtained by computing the values of $R(q)$ for $q = 1$ to p and observing how much it departs from 0. Note that the definition of the loss function is independant of the estimates of V and Σ used.

However, in practice, $R(q)$ is still unknown and different estimates can be proposed corresponding to different estimates of V. Those estimates are obtained from asymptotics expansions of the loss function $R(q)$. For the

classical version of SIR, such an estimate has been proposed and studied in Ferré (1998). For the Hsing and Carroll approach, similar results are given in Ferré (1996) and recalled in Section 2. While in the first case, the estimate is easily computable, the second case requires the computation of several non parametric regressions which might be boring. In order avoid complex computing, we suggest here to use a Jack-knife method to estimate $R(q)$. The use of such a resampling method is appealing here because the choice of the number of observations to removed at each step is guided by the retained estimate. Indeed, while a classical Jack-knife have to be used when $\widetilde{\Lambda}$ is performed, it seems relevant to remove two observations when $\widehat{\Lambda}$ is preferred. In Section 3, we present with more details the loss function and we give an asymptotic expansion. On the one hand, this expansion shows how to get an estimate of $R(q)$, and, on the other hand, how it may be used to compute the estimation without performing the whole Jack-knife procedure.

Section 4 is devoted to applications where the analysis is applied to the "Boston housing" data set, studied in Kotter (1996) and initially collected by Harrison and Rubinfield (1978). Particularly, a comparison on this data set of Li and Hsing and Carroll approaches, gives some insight on the real efficiency of the latter.

2 Direct estimation of R(q)

The computation of this version of SIR reveals a defect. Indeed, some eigenvalues of $\hat{\Sigma}^{-1}\widehat{\Lambda}$ are larger than one showing that $\hat{\Sigma} - \widehat{\Lambda}$ is not positive definite. This trouble does not actually affect the analysis since negative eigenvalues will correspond to zero or at least to the smallest eigen values of $I - \Sigma^{-1}\Lambda$. However, this could be damaging for evaluating the dimensionality. To overcome this problem, we will deal with $\hat{V}$, the projection of $\hat{\Sigma} - \widehat{\Lambda}$ onto the cone of positive definite matrices. From a theoretical point of view and due to contracting properties of projections, we easily have that $\hat{V}$ converges in probability to V since $\hat{\Sigma} - \widehat{\Lambda}$ converges to V. In practice, this leads to replace the negative eigenvalues by zero. Now let K^* be the number of positive eigenvalues of $\hat{\Sigma}^{-1}\hat{V}$ and let $\widehat{\lambda}_k$, for $k = 1$ to K^* be these eigenvalues ranged in descending order.

Now let us denote

$$\begin{aligned} m(Y) &= E(X|Y),\ \epsilon = X - m(Y), \\ V(y) &= E(\epsilon\epsilon'|Y = y),\ M(y) = m(y)m(y)'. \end{aligned}$$

We also denote by $\epsilon_{[i]}$ the concomitant variable in ϵ to the ith order

statistic in Y and let $(\lambda_i)_{i=1,...,K+1}$ be the eigenvalues of $I - \Sigma^{-1}\Lambda$ ranged in decreasing order ($\lambda_{K+1} = 0$) and we denote by P_i, for $i = 1, ...K+1$, the corresponding eigenprojectors. Note that we assume here, and to simplify, that the eigenvalues of $\Sigma^{-1}V$ are simple (see however the last paragraph of the section). We also let $U_1 = \sqrt{n}(\widehat{\Lambda} - \Lambda)$ and $U_2 = \sqrt{n}(\widehat{\Sigma} - \Sigma)$. Finally, let us precise that the norm for matrices used in the sequel is the Hilbert-Schmitt norm and we recall that, for any random centred matrices A and B, the covariance between A and B is defined by $Cov(A, B) = E(vec(A)vec(B)^T)$.

To establish our results, we need some assumptions. The first assumptions are required to apply the perturbation theory for linear operator (see e.g. Kato, 1966, even if the perturbations used here are not classical since the metric of the space is also perturbed):

$$\sqrt{\frac{1}{n}} < \inf_{i<j} \lambda_i - \lambda_j/2C, \tag{6}$$

where C is a real constant such that

$$\left\|\sqrt{n}(\widehat{\Sigma}^{-1}\widehat{\Lambda} - \Sigma^{-1}\Lambda)\right\| < C \text{ in probability.}$$

Then, technical assumptions are needed to insure the consistency of $\widehat{\Lambda}$

$$\forall\, B > 0, \lim_{n\to\infty} \sup_{-B\leq y_{(1)}\leq\cdots\leq y_{(n)}\leq B} n^{-1/4} \sum_{i=2}^{n} \left\|m(y_{(i)}) - m(y_{(i-1)})\right\| = 0, \tag{7}$$

$\exists\, B_0 > 0$, such that:

$$\exists\, \widetilde{m}, \text{ a real function, non decreasing over } [B_0, +\infty[\,,$$

such that:

$$\lim_{y\longrightarrow+\infty} \widetilde{m}(y)P(|Y| > y) = 0 \tag{8}$$

$$\text{and } \|m(x) - m(y)\| \leq |\widetilde{m}(|x|) - \widetilde{m}(|y|)|\,, \text{ if } (|x|\,, |y|) \in [B_0, +\infty[^2\,,$$

$$X \text{ is normally distributed.} \tag{9}$$

Note that this latter assumption implies (2) and that it leads to the expression of $Var(U_2)$. Finally, to simplify calculations, we will assume that

$$\text{given } Y,\ \epsilon \text{ has a symmetric distribution.} \tag{10}$$

Now, we get the following theorem proved in Ferré (1996):

Theorem 1 *Under the conditions above, we have:*

$$R(q) = 1 - \frac{1}{nq}[\sum_{i=1}^{q}\sum_{j=q+1}^{K+1}\frac{1}{(\lambda_j-\lambda_i)^2}tr((\Sigma^{-1}P_j\otimes\Sigma^{-1}P_i)(Var(U_1)$$

$$+((1-\lambda_j)(1-\lambda_i))Var(U_2)-(2-(\lambda_i+\lambda_j))Cov(U_1,U_2)))]+O(n^{-3/2}),$$

where

$$Var(U_1) = Var(\epsilon\epsilon') + (I_{p^2}+K_{p^2})E(V(Y)\otimes V(Y)),$$

$$Var(U_2) = (I_{p^2}+K_{p^2})(\Sigma\otimes\Sigma),$$

$$Cov(U_1,U_2) = Var(\epsilon\epsilon') + 2(I_{p^2}+K_{p^2})E(V(Y)\otimes V(Y)) + 2Cov(V(Y),M(Y)),$$

and K_{p^2} is the commutation matrix, i.e. the matrix that satisfies for any matrix A, $K_{p^2}vec(A^T) = vec(A)$.

Now for practical use, it is enough to replace Σ by $\widehat{\Sigma}$, the eigen elements of $\Sigma^{-1}\Lambda$ by those of $\widehat{\Sigma}^{-1}\widehat{\Lambda}$ and $Var(U_1)$ and $Cov(U_1,U_2)$ by consistent estimates. However, the latter estimations require the computations of several (univariate) non parametric regressions what can be a bit cumbersome. Furthermore, since the expansion of $R(q)$ still depends on K, we suggest to replace it in practice by K^*.

The expansion of the criterion reveals the crucial role played by the difference between successive eigenvalues. Indeed, both its expression and assumption (6) show that the criterion will be affected by introducing an eigenvalue too close to the previous one relatively to the value of the sample size n. Then, for the same "true" model, the criterion will lead to retain less linear combinations if the sample size is small than if it is large: if the sample size is too small, some directions are too poorly estimated and they must not be included in the following computations. Another consequence is that the criterion will detect multiple eigenvalues. This will be appealing since the structure of the error in the model is spherical. Then the criterion will perform well if we have only one multiple eigenvalue. At the opposite, it might be irrelevant if more than one multiple eigenvalue is present. A way to deal with is to identify those possible multiple eigenvalues according to the values of the criterion and to apply the adequate criterion calculated with non zero multiple eigenvalues. Recall that the indentification of multiple eigenvalues is not straightforward since the observed eigenvalues are always

simple. Now, if K' denotes the number of distinct eigenvalues, the criterion can be computed for $q = 1$ to K' and its expression becomes:

$$\begin{aligned} R(q) &= 1 - \frac{1}{n\sum_{i=1}^{q} rank(P_i)}[\sum_{i=1}^{q}\sum_{j=q+1}^{K'+1} \frac{1}{(\lambda_j - \lambda_i)^2} tr((\Sigma^{-1}P_j \otimes \Sigma^{-1}P_i) \\ &\quad (Var(U_1) + ((1-\lambda_j)(1-\lambda_i))Var(U_2) \\ &\quad -(2-(\lambda_i+\lambda_j))Cov(U_1,U_2)))] + O(n^{-3/2}), \end{aligned}$$

where P_i, for $i = 1$ to K' is the eigenprojector associated with the eigenvalue λ_i. This result is obtained under Condition (6) applied to the distincts eigenvalues and the proof is similar to the one of Theorem 1. An estimate will be obtained by replacing each multiple eigenvalue λ_i, by the average of the eigenvalues of $\widehat{\Sigma}^{-1}\widehat{V}$ which are observations of λ_i.

3 The Jack-knife approach

A natural alternative for estimating $R(q)$ is given by using resampling. Let us consider now the use of the Jack-Knife method. While in a slicing procedure, it would be necessary to perform a delete d Jack-knife that requires the selection of the value of d, a natural choice appears here since, by taking $d = 1$ when $\widetilde{\Lambda}$ is used or $d = 2$ when $\widehat{\Lambda}$ is preferred, a whole slice is deleted if the deletion is performed on the concomitant variables. Then, in the former case, classical Jack-knife is required while in the latter a delete 2 procedure is needed.

Now let us denote by Λ^{-i} and Σ^{-i} respectively, the estimates of Σ and Λ, when $X_{[i]}$ or $X_{[2i]}$ is removed. Let Π_q^{-i} be the Σ^{-i}-orthogonal projector associated with the q largest eigenvalues of $\Sigma^{-i-1}\Lambda^{-i}$. The Jack-knife estimate of $R(q)$ is:

$$JK(q) = 1 - \frac{1}{n}\sum_{i=1}^{m} tr(\Pi_q^{-i}\hat{\Pi}_q)/q = 1 - tr(\Pi_q^{*}\hat{\Pi}_q)/q, \tag{11}$$

where $\Pi_q^* = \frac{1}{n}\sum_{i=1}^{m} tr(\Pi_q^{-i})$ and $m = n$ or $[n/2]$. Then, by performing m analysis, the criterion $\hat{R}(q)$ can be easily evaluated and, by plotting it versus q, the stability of the estimated subspaces may be evidenced. In the sequel and for the sake of simplicity, we will only deal with the estimate $\widehat{\Lambda}$ while similar results may be easily stated for $\widetilde{\Lambda}$. In that particular case, we denote Λ^{-i} by $\widehat{\Lambda}^{-2i}$ and Σ^{-i} by $\hat{\Sigma}^{-2i}$. Recall that $\widehat{V}$ is actually the projection of $\widehat{\Sigma} - \widehat{\Lambda}$ onto the cone of positive definite matrices so that $\widehat{\Sigma}^{-1}\widehat{V}$ has rank K^*.

We also denote $A_{2i} = (X_{[2i]} - X_{[2i-1]})(X_{[2i]} - X_{[2i-1]})^T$, $B_{2i} = (X_{[2i]} - \bar{X})(X_{[2i]} - \bar{X})^T + (X_{[2i-1]} - \bar{X})(X_{[2i-1]} - \bar{X})^T$, and $(\hat{\lambda}_j)_{j=1,\ldots,K^*+1}$ the eigen-

values of $\widehat{\Sigma}^{-1}\widehat{\Lambda}$ ranged in descending order, with associated eigenprojectors $(\widehat{P}_j)$, $j = 1, ..., K^*$ and with $\widehat{\lambda}_{K^*} + 1 = 0$. Then, an asymptotic expansion of $\hat{R}(q)$, for $q = 1$ to K^*, is given in the theorem below.

Theorem 2 *Let C be a constant such that*

$$\sqrt{n-2}\|\widehat{\Sigma}^{-1}\widehat{V} - \widehat{\Sigma}^{(-2i)^{-1}}\widehat{V}^{-2i}\| < C. \tag{12}$$

If n satisfies, for $i = 1, ..., K^* + 1$,

$$\frac{1}{\sqrt{n-2}} < \inf_{i<j}(\hat{\lambda}_i - \hat{\lambda}_j)/C, \tag{13}$$

then we have, under assumption (2):

$$J\hat{K}(q) = \frac{1}{q(n-2)^2}[\sum_{j=1}^{q}\sum_{l=q+1}^{K^*+1}\frac{1}{(\hat{\lambda}_j - \hat{\lambda}_l)^2}tr[(\hat{\Sigma}^{-1}P_j \otimes \hat{\Sigma}^{-1}P_l)]$$

$$[\sum_{i=1}^{[n/2]}\frac{1}{n}vec(A_{2i} - (1-\hat{\lambda}_j)B_{2i})vec(A_{2i} - (1-\hat{\lambda}_j)B_{2i})^T$$

$$+(\widehat{\lambda}_j - \widehat{\lambda}_l)(vec(B_{2i})vec(A_{2i} - (1-\widehat{\lambda}_l)B_{2i})^T)]$$

$$\text{with } JK(q) = J\hat{K}(q) + O((n-2)^{-3}).$$

and, if $\widetilde{m}$, defined in (7), satisfies $\widetilde{m}^8 P(|Y| > y) \to 0$ as $n \to \infty$, then, $(n-2)J\hat{K}_q$ converges in probability to $R(q)$.

The proof of Theorem 2 is given in the Appendix. This results has two interests. First, it shows how to simply get an estimate of $R(q)$ and, secondly, it allows us to obtain the estimation without performing $[n/2]$ analysis. Furthermore, it is possible to verify the conditions of validity of the approximation for each couple of eigenvalue.

4 Applications

For the application, we will consider the Boston Housing Data set (Harrison and Rubinfeld, 1978). This data set has been used to illustrated SIR in Kotter (1996). It consists in 506 observations of 14 environmental variables for the Boston metropolitan area. The response variable is the "median value of owner-occupied homes". A nonparametric regression on the 13 variables is not realistic and a reduction of the dimensionality is necessary.

First, Hsing and Carroll version of SIR (HC analysis) has been performed. Table 1 gives the eigenvalues of $I - \widehat{\Sigma}^{-1}\widehat{\Lambda}$. First point is that only nine eigenvalues are positive so that we take $K^* = 9$. For $q = 1, ..., 9$, the computations of $JK(q)$ and $\hat{JK}(q)$ are plotted in Figure 1.

These curves show that, for this data set, something occurs for dimension two. Indeed we clearly see a peak for $q = 2$ while the first EDR direction seems correctly estimated. As previously mentioned, $\hat{JK}(q)$ penalizes close eigenvalues. Now the values in Table 1 indicate that observed eigenvalues two and three are indeed closed together. Then not surprising is the behavior of the criterion. But now the question is: have we $\lambda_2 = \lambda_3$? If this was true, others estimates should reveal this feature.

q	1	2	3	4	5	6	7
$\hat{\lambda}_q$	0.830	0.456	0.419	0.233	0.147	0.094	0.077
q	8	9	10	11	12	13	
$\hat{\lambda}_q$	0.032	0.007	−0.115	−0.170	−0.176	−0.228	

Table 1: Overall set of eigenvalues for HC analysis for the Boston housing data set.

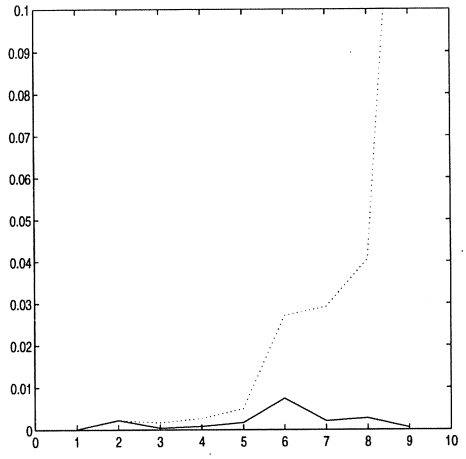


Figure 1: $JK(q)$ and $\hat{JK}(q)$ versus q for the Boston housing data set.

Then we suggest to compare those results to the one obtained by applying the classical SIR, this efficiency of this method being well-known. The number of slices has been set equal to 17 which corresponds to about 30 observations by slice. Table 2 gives the obtained eigenvalues. Interesting is

to notice that while the first and second eigenvalues are essentially the same as above, the third and successive ones appear to be quite different from a method to another. Particularly, the third eigenvalue is now significantly different from the second, rejecting the equality of λ_2 and λ_3.

Now, the estimate of the squared trace correlation for Li's version of SIR indicates that a three dimensional model seems relevant here while the Khi squared test recommended by Li(1991) leads to retain five dimensions ($\alpha = .05$).

q	1	2	3	4	5	6	7
$\hat{\lambda}_q$	0.802	0.451	0.199	0.088	0.059	0.049	0.042
q	8	9	10	11	12	13	
$\hat{\lambda}_q$	0.024	0.019	0.000	0.000	0.000	0.000	

Table 2: Set of eigenvalues for SIR for the Boston housing data set.

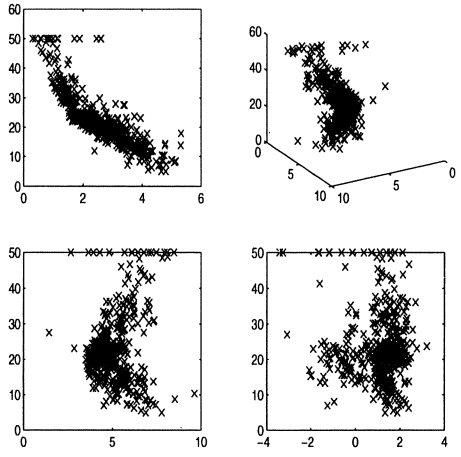


Figure 2: Scatter plots of Y versus $\hat{\beta}_1'X$, $\hat{\beta}_2'X$ and $\hat{\beta}_3'X$ for HC analysis for the Boston housing data set.

However, the comparison of the scatter plots of Y versus the three first indexes obtained by the two methods reveals interesting features. Figure 2 gives those scatter plots for Hsing and Carroll version and Figure 3 gives them for classical SIR. While for the first two indexes the plots are essentially the same, the scatter plots of Y against $\widehat{\beta}_3'X$ are different from an analysis to the other. Then it seems that, in this example, HC analysis

in unable to capture some relevant information. One explanation may be that this method correspond to dealing with $[n/2]$ slices with two observations in them and something like undersmoothing effect probably appears. Then, estimations are very perturbed by random noise and only one or two directions can be detected. Note the efficiency of the criterion JK which eliminate the directions incorrectly estimated.

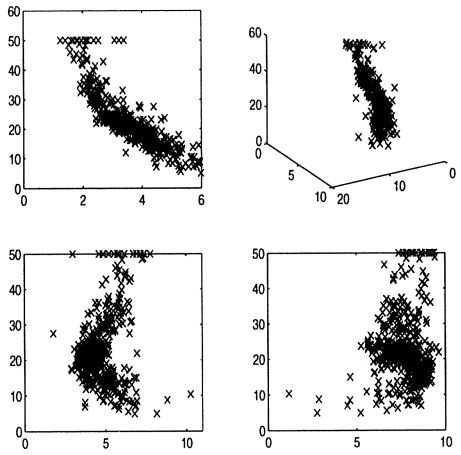


Figure 3: Scatter plots of Y versus indexes $\hat{\beta}_1'X$, $\hat{\beta}_2'X$ et $\hat{\beta}_3'X$ for SIR for the Boston housing data set.

Appendix

The proof of Theorem 2 requires the following lemma presented in Ferré (1998):

Lemma 1 *Let A_ε, B_ε, A and B be matrices $(p\times p)$ such that $A_\varepsilon = A+\varepsilon\, U_A$ and $B_\varepsilon = B + \varepsilon U_B$, A_ε et B_ε being random. Assume that B_ε and B are regular and that non-zero eigenvalues of $B_\varepsilon^{-1}A_\varepsilon$ and $B^{-1}A$ are simple. Let ν_i and P_i be the ith eigenvalue and associated B-orthogonal eigenprojector of $B^{-1}A$, for $i = 1$ to $K+1$, where $K = rank(B^{-1}A)$. The ν_i are ranged in descending order. Let $U_i = U_1 - \nu_i U_2$. Finally let $\widehat{\Pi}_q$ (resp. Π_q) the projector B_ε(resp. B)-orthogonal onto the subspace spanned by the q largest eigenvalues of $B_\varepsilon^{-1}A_\varepsilon$(resp. $B^{-1}A$). Now, if there exists a constant C such that $\frac{1}{\varepsilon}||B^{-1}A - B_\varepsilon^{-1}A_\varepsilon|| < C$ in probability and $\varepsilon < \inf_{i<j}(|\nu_i - \nu_j|)/2C$, we have:*

$$1 - tr(\widehat{\Pi}_q \Pi_q)/q = \frac{1}{q}\varepsilon^2 \sum_{j=1}^{q} \sum_{l=q+1}^{K+1} \frac{1}{(\nu_l - \nu_j)^2} tr(U_j P_j B^{-1} U_j P_l B^{-1} + (\nu_l - \nu_j) U_2 P_j B^{-1} U_j P_l B^{-1}) + O(\varepsilon^3).$$

Proof of Theorem 2:

The basis of this approximation, relies on the fact that the $\widetilde{\Lambda}^{(-i)}$ and $\widehat{\Lambda}^{(-2i)}$ can be decomposed as:

$$\widetilde{\Lambda}^{(-i)} = \widetilde{\Lambda} + \frac{1}{n-1}(\widetilde{\Lambda} + \tilde{A}_i) \tag{14}$$

and

$$\widehat{\Lambda}^{(-2i)} = \widehat{\Lambda} + \frac{1}{n-2}(2\widehat{\Lambda} + \widehat{A}_{2i}) \tag{15}$$

with

$$\begin{aligned} \tilde{A}_i &= \frac{1}{2}((X_{[i+1]} - X_{[i]})(X_{[i]} - X_{[i-1]})^T \\ &\quad + (X_{[i]} - X_{[i-1]})(X_{[i+1]} - X_{[i]})^T, \end{aligned}$$

if $i \neq 1$ and $i \neq n$;

$$\begin{aligned} \tilde{A}_1 &= (X_{[2]} - X_{[1]})(X_{[2]} - X_{[1]})^T/2; \\ \tilde{A}_n &= (X_{[n]} - X_{[n-1]})(X_{[n]} - X_{[n-1]})^T/2 \end{aligned}$$

and

$$\widehat{A}_{2i} = (X_{[2i]} - X_{[2i-1]})(X_{[2i]} - X_{[2i-1]})^T.$$

Similars decompositions exist for $\hat{\Sigma}$. If we denote respectively by $\tilde{\Sigma}^{(-i)}$ and $\widehat{\Sigma}^{(-2i)}$ the estimates of Σ at the ith step of the Jack-knife procedure, we get :

$$\tilde{\Sigma}^{(-i)} = \Sigma + \frac{1}{n-1}(\Sigma - Z_i Z_i^T) - \frac{1}{(n-1)^2} Z_i Z_i^T \tag{16}$$

and

$$\begin{aligned} \widehat{\Sigma}^{(-2i)} &= \Sigma + \frac{1}{n-2}(2\Sigma - Z_{[i]} Z_{[i]}^T - Z_{[i+1]} Z_{[i+1]}^T) \\ &\quad - \frac{1}{(n-2)^2}(Z_{[i]} + Z_{[i+1]})(Z_{[i]} + Z_{[i+1]})^T, \end{aligned} \tag{17}$$

where $Z_{[i]}$ is the ith centered version of $X_{[i]}$.

Since the proof is essentially the same whatever the method used, we only proof results for $\widehat{\Lambda}$.

Now, by letting $A = \widehat{\Sigma} - \widehat{\Lambda}$, $B = \widehat{\Sigma}$, $A_\varepsilon = \widehat{\Sigma}^{(-2i)} - \widehat{\Lambda}^{(-2i)}$ and $B_\varepsilon = \widehat{\Sigma}^{(-2i)}$,we get, $\nu_i = \lambda_i$, $U_A = 2(\widehat{\Sigma} - \widehat{\Lambda}) + (A_{2i} - B_{2i})$, $U_B = 2\widehat{\Sigma} - B_{2i}$, $U_j = 2((1-\lambda_j)\widehat{\Sigma} - \widehat{\Lambda}) + (A_i - (1-\lambda_j)B_i)$ and $K = K^*$. The Lemma above applies and we get

$$1 - tr(\widehat{\Pi}_q^{(2i)}\widehat{\Pi}_q)/q = \frac{1}{q(n-2)^2}\sum_{j=1}^{q}\sum_{l=q+1}^{K^*+1}\frac{1}{(\lambda_l-\lambda_j)^2}\, tr(((2(1-\lambda_j)\widehat{\Sigma} - \widehat{\Lambda})$$

$$+(A_i - (1-\lambda_j)B_i))P_j\widehat{\Sigma}^{-1}\ ((2(1-\lambda_j)\widehat{\Sigma} - \widehat{\Lambda}) + (A_i - (1-\lambda_j)B_i))P_l\widehat{\Sigma}^{-1}$$

$$+(\lambda_l-\lambda_j)(\left(2\widehat{\Sigma} - B_{2i}\right)P_j\widehat{\Sigma}^{-1}((2(1-\lambda_j)\widehat{\Sigma} - \widehat{\Lambda}) + (A_i - (1-\lambda_j)B_i))P_l\widehat{\Sigma}^{-1}))$$

$$+O(1/(n-2)^3).$$

Now, since P_j and P_l are $\widehat{\Sigma}$-orthogonal eigenprojectors, we have $P_j\widehat{\Sigma}^{-1}(\widehat{\Sigma} - \widehat{\Lambda})P_l = P_l\widehat{\Sigma}^{-1}(\widehat{\Sigma} - \widehat{\Lambda})P_j = 0$. By using again the relation $tr(M_1M_2M_3M_4) = tr((M_2 \otimes M_4)vec(M_3)vec(M_1)')$ for any matrices M_1, M_2, M_3 and M_4, the result is obtained by averaging over the $[n/2]$ replications and arranging.

Now, by using Lemmas A1, A2 and A3 of Hsing and Carroll(1992) with condition $\tilde{m}^8(y)P(|Y| > y) \to 0$ as $n \to \infty$, and by noting $\xi_{[2i]} = (\varepsilon_{[2i]} - \varepsilon_{[2i-1]})(\varepsilon_{[2i]} - \varepsilon_{[2i-1]})^T$, we have on the one hand:

$$\frac{1}{n-2}\sum_{i=1}^{[2i]} vec(A_{[2i]})vec(A_{[2i]})^T =$$

$$\frac{1}{n-2}\sum_{i=1}^{[2i]} vec(\xi_{[2i]})vec(\xi_{[2i]})^T - E(vec(\varepsilon\varepsilon^T)vec(\varepsilon\varepsilon^T)^T) + o_p(1).$$

On the other hand, by using again Lemma A3 of Hsing and Carroll, we get:

$$Var(\hat{\Lambda}) = \frac{1}{n}\sum_{i=1}^{[2i]} vec(\xi_{[2i]})vec(\xi_{[2i]})^T - E(vec(\varepsilon\varepsilon^T)vec(\varepsilon\varepsilon^T)^T) + o_p(1/n)$$

and then $\frac{1}{n-2}\sum_{i=1}^{[2i]} vec(A_{[2i]})vec(A_{[2i]})^T$ converge in probability to $Var(\hat{\Lambda})$. Similar arguments lead to show that $\frac{1}{n-2}\sum_{i=1}^{[2i]} vec(B_{[2i]})vec(B_{[2i]})^T$ converges in probability to $Var(\hat{\Sigma})$.

The proof is complete after replacement of $\frac{1}{n-2}\sum_{i=1}^{[2i]} vec(A_{[2i]})vec(A_{[2i]})^T$ and $\frac{1}{n-2}\sum_{i=1}^{[2i]} vec(B_{[2i]})vec(B_{[2i]})^T$ respectively by $Var(\hat{\Lambda})$ and $Var(\hat{\Sigma})$ and arranging.

References

[1] Daudin J.J., Duby C., Trecourt P. (1989). PCA stability studied by the bootstrap and the infinitesimal jackknife method. *Statistics* **20,** 255-270.

[2] Ferré L. (1996). Choix de la dimension en Regression Inverse par Tranches. *Compte-Rendus de l'Académie des Sciences*, tome 323, Série I, 403-406.

[3] Ferré L. (1998). Determining the dimension in Sliced Inverse Regression and related methods. To appear in *J. Amer. Statist Assoc.*

[4] Harrisson D. and Rubinfield D.C. (1978). Hedonic prices and the demand for clean air. *J. of Environmental Economics and Management* **8,** 276-290.

[5] Hsing T. and Carrol R.J. (1992). An asymptotic theory for sliced inverse regression. *Ann. Statist.* **20,** 1040-1061.

[6] Kato T. (1996). *Perturbation Theory for Linear Opaerators.* Springer Verlag, New York.

[7] Kotter T. (1996). An asymptotic result for sliced inverse regression. *Comput. Statist.* **11,** 113-136.

[8] Li K.C. (1992). On principal Hessian directions for data visualization and dimension reduction: Another application of Stein's Lemma. *J. Amer. Statist. Assoc.* **87**, 1025-1039.

[9] Li K.C. (1991). Slice inverse regression for dimension reduction. *J. Amer. Statist. Assoc.* **86,** 319-342.

[10] Magnus J.R. and Neudecker (1979). The commutation matrix: Some properties and applications. *Ann. Statist.* **7,** 381-394.

[11] Saracco J. (1996). An asymptotic theory for sliced inverse regression. Pre-print.

[12] Schott J.R. (1994). Determining the dimensionality in sliced inverse regression. *J. Amer. Statist. Assoc.* **89,** 141-148.

Student (1997), Vol.2, No.2, 109-117

Nonparametric Conditional Quantile Estimation in the Presence of Long Memory

Sucharita Ghosh

Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf, Switzerland

Jan Beran

University of Konstanz, Germany

John Innes

Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf, Switzerland

Abstract: In this paper we consider nonparametric quantile estimation for time series that are time dependent transformations of a stationary Gaussian process with long-range dependence. Nonstationary and nongaussian processes are included in this framework. Time dependent quantiles are estimated by kernel smoothing. Asymptotic results on mean squared error, optimal bandwidth and limiting distribution are obtained. Two data examples are presented.

Key words: Time series, Gaussian process, stationarity.

1 Introduction

We consider the problem of estimating conditional quantiles of a stochastic process that is generated by a time dependent transformation of a zero mean stationary Gaussian process Z_i $(i \in N)$ with long memory. Here, Z_i is said to have long memory, if the asymptotic decay of the autocovariances

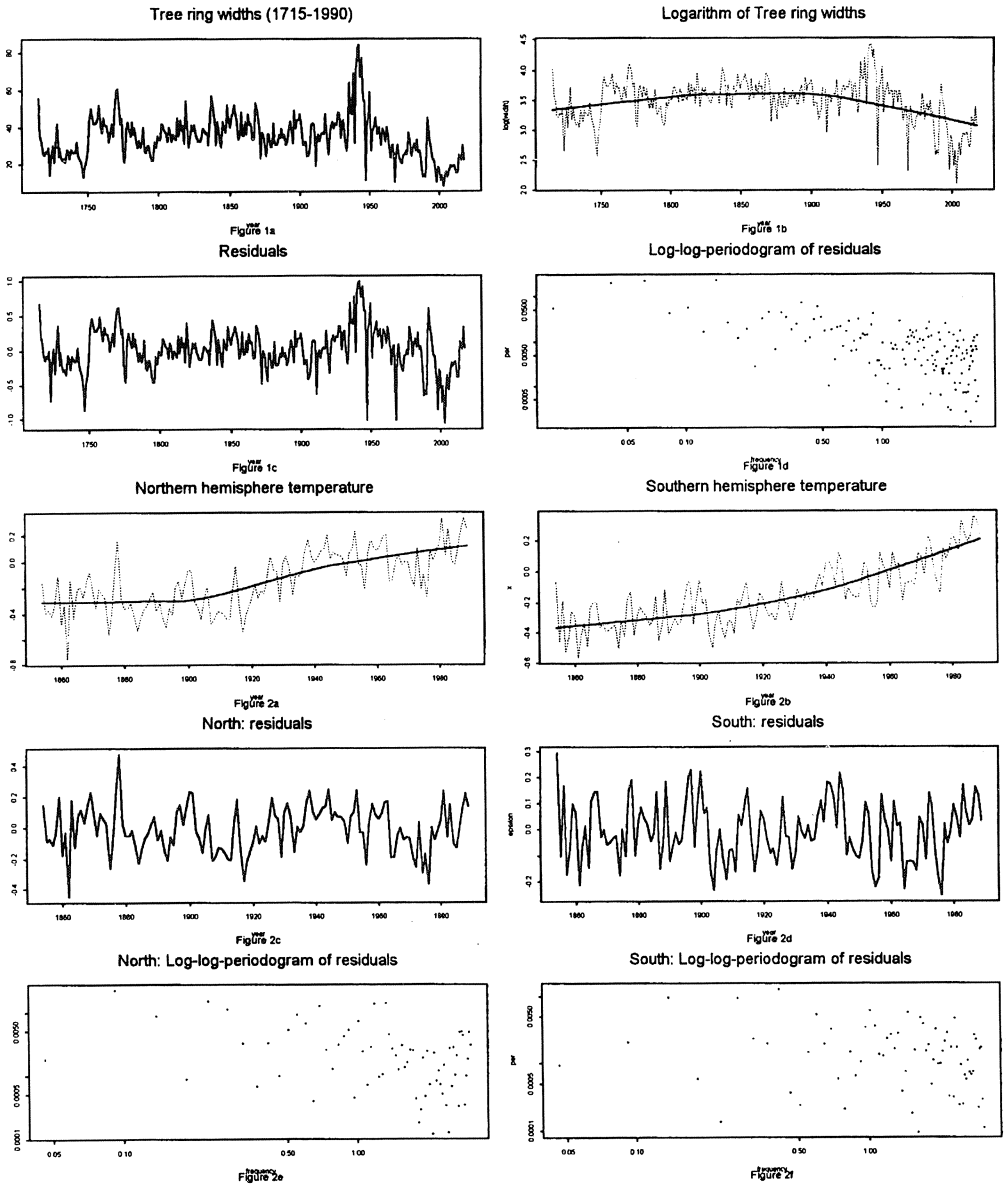
$\gamma(k) = \text{cov}(Z_i, Z_{i+k})$ is given by $\gamma(k) \sim C_Z \ |k|^{2H-2}$, as $|k| \to \infty$ with $H \in (\frac{1}{2}, 1)$. The symbol "$\sim$" means that the ratio of the two sides converges to one. Note that, since $H > 1/2$, the covariances are not summable. The slow decay of the covariances implies a pole of the spectral density $h(\lambda) = (2\pi)^{-1} \sum \gamma(k) exp(ik\lambda)$ at the origin. For an introduction to long-range dependence and references see e.g. Beran (1994). We start with two examples that motivated this research.

Example 1 *Figure 1a displays the tree ring widths (in 1/100 mm) of an individual Pilgerodendron uniferum tree sampled from Lago Presidente Rios in Chile. The time period covers the years 1715-1990. The logarithm of the series is displayed in figure 1b. The series exhibits an apparent trend. The trend in the logarithmic plot is estimated by the Splus-function lowess. The resulting residuals (figure 1c) appear to be randomly distributed (without any systematic trend), the standard deviation however may not be constant in time. Moreover, the log-log-plot of the periodogram (figure 1d) indicates long-range dependence. (This remains to be true even after heteroscedasticity is taken into account.)*

Example 2 *Figures 2a and 2b display the yearly averages of differences between observed monthly temperatures and the average monthly temperatures based on the period 1950-1979, for the northern and southern hemisphere respectively. The time period is 1854-1989. The residuals obtained from removing the trends (using the Splus-function lowess), are displayed in figures 2c and d. The log-log-periodograms (figures 2e and f) indicate long memory. Quantile estimation for these series is especially important as it is of interest to investigate the long-term behaviour of low and high temperatures (compare e.g. Katz and Brown 1992)*

These examples motivate to consider quantile estimation for processes that are time-dependent linear or nonlinear transformations of a Gaussian long-memory process. Most papers in the literature that deal with parametric or nonparametric quantile estimation and quantile regression assume independent observations (see e.g. Koenker and Bassett 1978, Bassett and Koenker 1982, Koenker et al. 1994, Stute 1986, Horváth and Yandell 1988, Samanta 1989, Bhattacharya and Gangopahyay 1990, Chaudhuri 1991, for further references see e.g. Dodge 1992). Quantile regression for short-range dependent processes is considered in Portnoy (1991). Nonparametric quantile estimation for short-memory proceses is considered in Abberger (1996, 1997). Nonparametric estimation of a mean curve in the long-memory context is considered for instance in Hall and Hart (1990) and Csörgö and

Mielniczuk (1995b). Nonparametric density estimation for stationary long-memory processes is considered in Csörgö and Mielniczuk (1995a). In the present paper, these results are used to extend nonparametric quantile regression to the long-memory case, including nonstationary and nongaussian processes.



2 Basic definitions

Let $\{Z_i, i = 1, 2, ...\}$ be a stationary Gaussian process with $E(Z_i) = 0$, $var(Z_i) = 1$ and $\gamma_Z(k) = cov(Z_i, Z_{i+k}) \sim C_Z|k|^{2H-2}$ (as $|k| \to \infty$) with $C_Z > 0$ and $1 - \frac{1}{2m} < H < 1$ for some $m \in \{1, 2, 3, ...\}$. Denote by $H_k(.)$ $(k = 1, 2, ...)$ the Hermite polynomial of order k. Let $G : R \times [0, 1] \to R$ be a Lebesgue measurable function such that for fixed $x \in [0, 1]$ and $y \in R$, and a standard normal random variable Z, $E[1\{G(Z, x) \leq y\}H_k(Z)] = 0$ $(1 \leq k < m)$ and $E[1\{G(Z, x) \leq y\}H_m(Z)] \neq 0$, i.e. $1\{G(Z, x) \leq y\} - E[1\{G(Z, x) \leq y\}]$ has Hermite rank m. Furthermore, for each $x \in [0, 1]$, define $Y = Y(x) = G(Z_{[nx]}, x)$. Note in particular that for $x = i/n$ $(i \in \{1, 2, ..., n\})$ we have $Y(i/n) = G(Z_i, i/n)$. Denote by $F_x(y)$ the conditional distribution $F_x(y) = P(Y(x) \leq y|x)$. We assume that $\frac{\partial}{\partial y}F_x(y) = f_x(y)$, and $\frac{\partial^2}{\partial x^2}F_x(y)$ exist. Also, for given y, let $c_k(x, y)$ be continuous functions of x and y where $1\{G(Z, x) \leq y\} - F_x(y) = \sum_{k=m}^{\infty}(c_k(x, y)/k!)H_k(Z)$.

3 Estimation of $F_x(y)$

$F_x(y)$ is estimated by kernel smoothing (see e.g. Härdle 1991, Gasser et al. 1985) using a kernel K that is a symmetric Lipschitz continuous density function with support $[-1, 1]$. Define a sequence of bandwidths b_n such that $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$. Suppose that we observe $(x_i, Y(x_i)), i = 1, 2, ..., n$ where $x_i = i/n$. Thus $Y(x_i) = G(Z_i, i/n)$. Then

$$\hat{F}_x(y) = \frac{1}{nb_n}\sum_{i=1}^{n} K(\frac{x_i - x}{b_n})1\{Y_i \leq y\}.$$

Note that results in Dehling and Taqqu (1989) imply that, as a function of y, the normalized empirical process $(\hat{F}_x(y) - F_x(y))/\sqrt{\text{var}(\hat{F}_x(y))}$ converges to the same random variable, instead of a process (also see Giraitis and Surgailis 1994). This is in contrast to the short memory case. The following result is proved in the appendix.

Theorem 1 *As $n \to \infty$, we have*

a) Bias:

$$E[\hat{F}_x(y)|x] - F_x(y) = \frac{{b_n}^2}{2}\frac{\partial^2}{\partial x^2}F_x(y)\int u^2K(u)du + o({b_n}^2) = A(x, y)\,{b_n}^2 + o({b_n}^2).$$

b) Variance:

$$\text{var}[\hat{F}_x(y)|x] = (nb_n)^{m(2H-2)}C_Z^m\frac{c_m^2(x, y)}{m!}\int\int K(u)K(v)|u - v|^{m(2H-2)}dudv$$

$$+o((nb_n)^{m(2H-2)}) = B(x,y)(nb_n)^{m(2H-2)} + o((nb_n)^{m(2H-2)}).$$

c) Mean square error: If $b_n^4(nb_n)^{min[1,(m+1)(2-2H)]} \to \infty$, *then*

$$MSE(\hat{F}_x(y)|x) = E[(\hat{F}_x(y) - F_x(y))^2|x]$$
$$= A^2\ b_n^4 + B\ (nb_n)^{m(2H-2)} + o(max[b_n^4, (nb_n)^{m(2H-2)}]).$$

d) Optimal bandwidth: Under the additional assumption in c), the asymptotic mean square error is minimized by the bandwidth

$$b_n = C \cdot n^{m(2H-2)/(4+m(2-2H))}$$

with

$$C = (\frac{m(2-2H)B}{4A^2})^{1/(4+m(2-2H))}.$$

4 Estimation of quantiles

For $0 < \alpha < 1$, let $\theta_\alpha(x)$ be the $\alpha-$quantile given x, i.e. $F_x(\theta_\alpha(x)) = \alpha$. Assume that $\theta_\alpha(x)$ is unique. Define $\hat{\theta}(x) = \inf\{y : \hat{F}_x(y) \geq \alpha\}$. Then we have

Theorem 2 *Under the assumptions of Theorem 1,*

a) $E[\hat{\theta}_\alpha(x) - \theta_\alpha(x)|x] = -b_n^2\ A(x,\theta_\alpha(x))/f_x(\theta_\alpha(x)) + o(b_n^2)$;

b)

$$\text{var}[\hat{\theta}_\alpha(x)|x] = \frac{B(x,\theta_\alpha(x))}{f_x^2(\theta_\alpha(x))}(nb_n)^{m(2H-2)} + o((nb_n)^{m(2H-2)})$$

c) If $b_n^4(nb_n)^{min[1,(m+1)(2-2H)]} \to \infty$, *then*

$$MSE(\hat{\theta}_\alpha(x)|x) = E[(\hat{\theta}_\alpha(x) - \theta_\alpha(x))^2|x]$$
$$= \frac{A^2}{f_x^2(\theta_\alpha(x))}{b_n}^4 + \frac{B}{f_x^2(\theta_\alpha(x))}(nb_n)^{m(2H-2)} + o(max[{b_n}^4, (nb_n)^{m(2H-2)}]).$$

d) The asymptotically optimal bandwidth b_n *is as in Theorem 1(d).*

Figures 3 and 4 show the application of the nonparametric quantile estimation to the data sets introduced above. It should be noted that optimal data driven bandwidth selection for quantile estimation of long-memory processes is an open problem. As an approximation, m was assumed to be one and b_n was set equal to $n^{-(2-2\hat{H})/(6-2\hat{H})}$ with $\hat{H}$ estimated from the residuals obtained in the analysis in the first part of this paper.

Tree rings: median, lower/upper quartiles

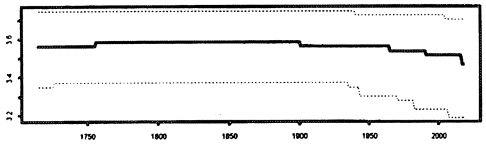


Figure 3a

Tree rings: lower/upper quartile minus median

Figure 3b

Tree rings: IQR

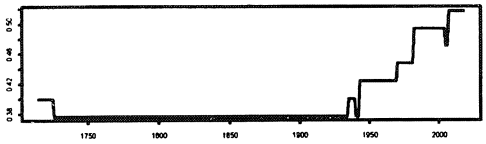


Figure 3c

North: Median, lower/upper quartiles

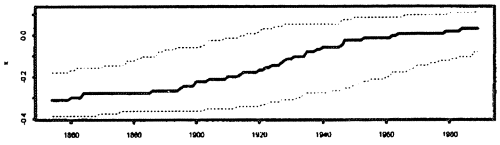


Figure 4a

South: Median, lower/upper quartiles

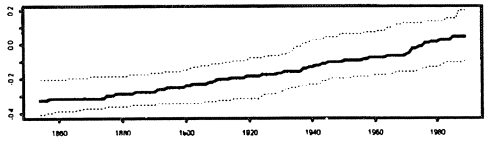


Figure 4b

North: Lower/upper quartile minus median

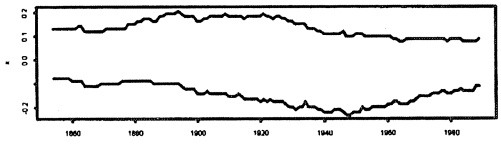


Figure 4c

South: Lower/upper quartile minus median

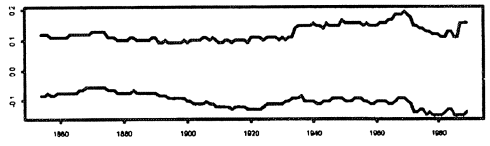


Figure 4d

North: IQR

Figure 4e

South: IQR

Figure 4f

Upper q. minus median for N & S

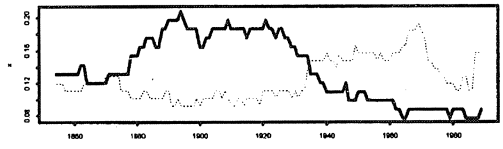


Figure 4g

Lower q. minus median for N & S

Figure 4h

Example 1 *Figure 3a displays the estimated lower and upper quartiles and the median. Figure 3b shows the difference between the upper quartile and the median as well as between the lower quartile and the median. Finally the estimated interquartile range is displayed in figure 3c. The figures support the conjecture that the standard deviation is not constant.*

Example 2 *Figure 4a and b displays the estimated lower and upper quartiles and the medians for the two hemispheres. Figure 4c and d show the difference between the upper quartile and the median as well between the lower quartile and the median. The lower and upper quartiles for the northern hemisphere show opposite behaviour, with a time shift. For the southern hemisphere the quartiles seem to follow a different pattern. The estimated interquartile ranges are displayed in figures 4e and f. Figures 4g and 4h allow a comparison (of lower and upper quartiles respectively) between the northern and southern hemisphere.*

For $m = 1$, the asymptotic distribution of the estimates is given by

Theorem 3 *Let $m = 1$. Then the following holds:*

a) For fixed $y \in R$, and $(x_1, x_2, ..., x_k) \in [0,1]^k$, where $k \geq 1$, define $U_n = (U_n(1), ..., U_n(k))^t$ with $U_n(i) = \{\hat{F}_{x_i}(y) - E[\hat{F}_{x_i}(y)|x_i]\}/\sqrt{B(x_i, y)}$. Then $(nb_n)^{(2-2H)} U_n \rightarrow_d (Z_1, ..., Z_k)$ where Z_i are iid standard normal.

b) For fixed $(x_1, x_2, ..., x_k) \in [0,1]^k$, and $\alpha \in (0,1)$, define $V_n = (V_n(1), ..., V_n(k))^t$ with $V_n(i) = \{\hat{\theta}_\alpha(x_i) - \theta_\alpha(x_i)\} f_{x_i}(\theta_\alpha(x_i))/\sqrt{B(x_i, \theta_\alpha(x_i))}$. Then $(nb_n)^{(2-2H)} V_n \rightarrow_d (Z_1, ..., Z_k)$ where Z_i are iid standard normal.

Appendix

Proof of Theorem 1:

a) $E[\hat{F}_x(y)|x]$ does not depend on the correlation structure of the process $Z(\cdot)$. The result then follows by standard arguments.

b) From Dehling and Taqqu (1989) it follows that, as a function of y, $[\hat{F}_x(y) - F_x(y)]/\sqrt{\text{var}(\hat{F}_x(y))}$ converges to the same random variable. (This is in contrast to short-memory processes where a limiting *process* is obtained.) Moreover, using Hermite expansion we have

$$\text{var}[\hat{F}_x(y)|x] = \sum_{k=m}^{\infty} \frac{1}{(nb_n)^2 k!} \sum_{i=1}^{n} \sum_{j=1}^{n} K(\frac{x - x_i}{b_n}) K(\frac{x - x_j}{b_n}) c_k(x_i, y) c_k(x_j, y) (\gamma_Z(i-j))^k.$$

Using $(\gamma_Z(s))^k \sim C_Z^k s^{k(2H-2)}$ as $s \to \infty$,

$$\text{var}[\hat{F}_x(y)|x] = \frac{C_Z^m c_m^2(x,y)}{m!}(nb_n)^{(2H-2)m}\int_{-1}^{1}\int_{-1}^{1} K(u)K(v)|u-v|^{(2H-2)m}dudv$$

$$+O((nb_n)^{-min[1,(m+1)(2-2H)]})$$

Proof of Theorem 2:

First note that $\hat{\theta}_\alpha(x) \to_p \theta_\alpha(x)$. As in Pollard (1984, p.98), we have $(nb_n)^{m(2-2H)}[\hat{\theta}_\alpha(x) - \theta_\alpha(x)] = H((nb_n)^{m(2-2H)}[\hat{F}_x(.) - F_x(.)], \hat{\theta}_\alpha, \xi_n, \psi_n)$ where $\xi_n = o_p(1)$, $\psi_n = o_p(1)$ and $H(\eta, a, b, c) = [-\eta(a) - b]/[f_x(\theta_\alpha(x)) + c]$. The result then follows from the Continuous Mapping Theorem (Pollard 1984, p. 70).

Proof of Theorem 3:

a) From the proof of Theorem 2 b) and $m = 1$ we have $|\hat{F}_x(y) - (nb_n)^{-1}\sum_{i=1}^n K((i-l)/(nb_n))c_1(i/n, y)Z_i| \to_p 0$. The result then follows from Theorem 1(i) in Csörgö and Mielniczuk (1995).

b) The result follows from a) and the arguments in the proof of Theorem 2.

Acknowledgments: We thank Mr. Gustav Schneiter (WSL, Birmensdorf) for handling the tree ring data. We also thank Professor Richard Smith (University of North Carolina, Chapel Hill, USA) for providing us with the temperature data from the data base of the Climate Research Unit of the University of East Anglia, Norwich, England.

References

[1] Abberger, K. (1996). Nichtparametrische Schätzung bedingter Quantile in Zeitreihen. Konstanzer Dissertationen. Band 530, Hartung-Gorre, Konstanz.

[2] Abberger, K. (1997). Quantile smoothing in financial time series. *Statistical Papers* **28**, 125-148.

[3] Bassett, G.W. and Koenker, R.W. (1982). An empirical quantile function for linear models with iid errors. *J. Amer. Statist. Assoc.* **77**, 407-415.

[4] Beran, J. (1994). *Statistics for Long-Memory Processes.* Chapman and Hall, New York.

[5] Bhattacharya, P.K. and Gangopadhyay, A.K. (1990). Kernel and nearest neighbour estimation of a conditional quantile. *Ann. Statist.* **18**, 1400-1412.

[6] Chaudhuri (1991). Nonparametric estimates of regression quantiles and their Bahadur representations. *Ann. Statist.* **19,** 760-777.

[7] Csörgö, S. and Mielnickuk, J. (1995a). Density estimation under long-range dependence. *Ann. Statist.* **23,** 990-999.

[8] Csörgö, S. and Mielnickuk, J. (1995b). Nonparametric regression under long-range dependent normal errors. *Ann. Statist.* **23,** 1000-1014.

[9] Dehling, H. and Taqqu, M.S. (1989). The empirical process of some long-range dependent sequences with an application to U-statistics. *Ann. Statist.* **17,** 1767-1783.

[10] Dodge, Y. (Editor) (1992). *L_1-Statistical Analysis and Related Methods.* Elsevier Science.

[11] Gasser T., Müller, H.G., Mammitsch, V. (1985). Kernels for nonparametric curve estimation. *J. Roy. Statist. Soc. B* **47,** 238-252.

[12] Giraitis, L. and Surgailis, D. (1994). A central limit theorem for the empirical process of a long-memory linear sequence. Beiträge zur Statistik, No. 24, Institut für Angewandte Mathematik, Universität Heidelberg.

[13] Hall, P. and Hart, J.D. (1990). Nonparametric regression with long-range dependence. *Stoch. Proc. Appl.* **36,** 339-351.

[14] Härdle W. (1991). *Smoothing Techniques.* Springer, New York.

[15] Horváth, L. and Yandell, B.S. (1988). Asymptotics of conditional empirical processes. *J. Multiv. Anal.* **26,** 184-206.

[16] Katz, R.W. and Brown, B.G. (1992). Extreme events in a changing climate: Variability is more important than averages. *Climatic Change* **21,** 289-302.

[17] Koenker, R.W. and Bassett, G.W. (1978). Regression quantiles. *Econometrica* **46,** 33-50.

[18] Koenker, R.W., Pin, N., Portnoy, S. (1994). Quantile smoothing splines. *Biometrika* **81,** 673-680.

[19] Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer, New York.

[20] Portnoy, S. (1991). Asymptotic behaviour of regression quantiles in nonstationary, dependent cases. *J. Multiv. Anal.* **38,** 100-113, 1991.

[21] Samanta, M. (1989). Non-parametric estimation of conditional quantiles. *Statist. Prob. Letter* **7,** 407-412.

[22] Stute, W. (1986). Conditional empirical processes. *Ann. Statist.* **14,** 638-647.

Student (1997), Vol.2, No.2, 119-130

L_p-Norm Estimation for Nonlinear Regression Models

Massimiliano Giacalone

University of Naples "Federico II", Italy

Abstract: We examine the use of L_p-norm estimators in the framework of nonlinear regression models, assuming an Exponential Power Function as error distribution. In this work, we suggest a new criterion to jointly estimate the L_p-norm exponent p and the regression parameters. This approach is motivated by theoretical error distribution considerations. These distributions are elements of a class of density functions (E.P.F.), which are related to L_p-norm estimators. Finally, we present a simulation study that leads us to conclude that L_p-norm estimators are a suitable tool for studying nonlinear regression problems in the case of nonnormal symmetric error distributions.

Key words: Exponential Power Function, adaptive procedures, kurtosis indexes.

1 Introduction

This study deals with the construction of an adaptive estimation procedure for nonlinear regression models. In this context, the L_p-norm estimation methods are investigated for different values of p over a range of error distributions with varying kurtosis. A previous work showed the same procedures applied to the location model (Giacalone, 1996).

As it is well known, in order to obtain the L_p-norm estimator of the unknown regression parameter vector $\underline{\theta}$, we minimize the sum of the p-th power of the absolute deviations of the observed points from the regression function:

$$S_p(\underline{\theta}) = \sum_{i=1}^{n} |y_i - g(\underline{x}_i, \underline{\theta})|^p \qquad 1 < p < \infty \tag{1}$$

A rule for selecting the most appropriate value of p for any given error distribution is proposed, based on the Geary length of tails index and on the Pearson kurtosis of the error distribution. The relative efficiencies of the different examined estimators are discussed. In general, the true error distribution, and hence the kurtosis, is not known. In order to overcome this problem, the performances of the relationships between p and the two Kurtosis sample indexes as well as the related algorithm are examined. The new adaptive scheme suggested is then compared with that of several commonly used proposed alternatives. The E.P.F., here considered as underlying error distribution, is a family of density functions proposed by Subbotin (1923) and studied by Vianelli (1963), Lunetta (1963) and Mineo (1989). A brief introduction is given at this point. The density function is:

$$f_p(z) = \frac{1}{2p^{1/p}\sigma_p \Gamma\left(1+\frac{1}{p}\right)} \exp\left[-\frac{1}{p}\left|\frac{z-M_p}{\sigma_p}\right|^p\right] \qquad (2)$$

with $\sigma_p > 0, \quad p > 0, \quad -\infty < z < +\infty,$
where $M_p = E(z)$ is the location parameter, $\sigma_p = (E[|z - M_p|^p])^{1/p}$ is the scale parameter, and p is the shape parameter. In fact, as p varies from 0 to ∞ the (2) assumes several shapes with different length of tails and kurtosis. Considering the Pearson kurtosis index β_2 we distinguish:

1. $0 < p < 1$: double exponential distributions, cuspidate, very long tailed and $\beta_2 > 6$;
2. $p = 1$: the Laplace distribution, cuspidate, long tailed with $\beta_2 = 6$;
3. $1 < p < 2$: leptokurtic distributions with long tails and $3 < \beta_2 < 6$;
4. $p = 2$: the Gaussian normal distribution with $\beta_2 = 3$;
5. $p > 2$: platikurtic distributions with short tails and $1.8 < \beta_2 < 3$;
6. $p \to \infty$: the uniform distribution with $\beta_2 \to 1.8$.

As previously stated, we assume that the error distribution of the regression model is a member of this class of symmetric functions. In this case, the optimal exponent p for the L_p-norm estimators of the regression parameters is the shape parameter p of the E.P.F. Our estimate of p is based on the two indexes, length of tails and kurtosis, strictly related to the shaped parameter and to the sample residuals.

2 L_p-norm nonlinear regression

Let us consider a sample of n observed data $(\underline{x}_i, y_i)$, where y_i is the dependent variable and $\underline{x}_i$ the independent nonrandom predictors. The general nonlinear regression model is:

$$y_i = g\left(\underline{x}_i, \underline{\theta}\right) + e_i \tag{3}$$

where g is a derivable function, $\underline{\theta} = (\theta_0, \theta_1, ..., \theta_k)$ is the unknown real parameter vector to be estimated, and the random errors e_i are independent and identically distributed variables according to E.P.F. scheme with mean zero and σ_p constant.

Under the supposed assumptions the loglikelihood related to the sample is given by:

$$\begin{aligned} L\left(\underline{\theta}, \sigma_p, p\right) &= -n \log\left[2p^{1/p}\sigma_p \Gamma\left(1+\frac{1}{p}\right)\right] + \\ &\quad - \left[\left(p\sigma_p\right)^{-1} \sum \left|y_i - g\left(\underline{x}_i, \underline{\theta}\right)\right|^p\right] \end{aligned} \tag{4}$$

where we consider $z = y_i$ and $M_p = g(\underline{x}_i, \underline{\theta})$

When p is known it is easy to calculate the first partial derivatives with respect to θ to get the system of n nonlinear equations with n equations and $k+1$ variables:

$$\frac{\delta L}{\delta \theta_j} = \sum_{i=1}^{n} \left|y_i - g\left(\underline{x}_i, \underline{\theta}\right)\right|^{p-1} \text{sign } \left(y_i - g\left(\underline{x}_i, \underline{\theta}\right)\right) \frac{\delta g}{\delta \theta_j} = 0$$

The solution of this system gives us the maximum likelihood estimators of the regression parameters. The same equations are obtained by minimizing the sum of the p-th power of the absolute deviations of the observed points from the regression function, by applying the L_p-norm estimators:

$$\sum \left|y_i - g\left(\underline{x}_i, \underline{\theta}\right)\right|^p = \min \qquad \text{with } p \geq 1 \tag{5}$$

This result shows that the optimal exponent p is equal to the shape parameter of the E.P.F. assumed as underlying error distribution and is very useful in connecting the L_p-norm estimators to the E.P.F. theoretical scheme. When the value of p is unknown, we consider two related problems: the estimate of a suitable exponent p based on the sample data and the choice of the minimization algorithm used to obtain the regression parameters estimation. Even if the least squares, the least absolute deviations and the minimax estimators are particular cases of L_p-norm estimators (respectively with $p = 1$, $p = 2$, $p = \infty$), there is no theoretical reason for values of p other than 1, 2, ∞, not to be considered. We do not consider the case $p < 1$, that is the quasi-norm problem (Ekblom et al., 1969).

It should be noted that $p = 1$, $p = 2$, $p = \infty$ provide exact solutions, while the other values of p give rise to a nonlinear programming problem,

whose solution can only be found in a given level of convergence. Different values of p have been proposed in the literature.

Forsythe (1972) suggested that $p = 1.5$ might be a good compromise value, as it provides estimates which are substantially better than the least squares when the error distribution has long tails, and is not bad when the errors have a Normal distribution. Harter (1977) proposed p with the following rule: if $\hat{\beta}_2 > 3.8$ use $p = 1$ (the regression of the least absolute values), if $2.2 < \hat{\beta}_2 < 3.8$ use $p = 2$ (the regression of the least squares) if $\hat{\beta}_2 < 2.2$ use $p = \infty$ (minimax or Chebychev regression), where $\hat{\beta}_2$ is the sample kurtosis.

Money et al. (1982) and Sposito et al. (1983) respectively obtained the following criteria for the choice of p by means of an extended simulation study:

$$\hat{p} = 9/\hat{\beta}_2^2 + 1 \qquad \text{for} \qquad 1 \leq p < \infty \tag{6}$$

$$\hat{p} = 6/\hat{\beta}_2 \qquad \text{for} \qquad 1 \leq p < 2 \tag{7}$$

Recently, the choice of a suitable value of p, for the linear regression case, was based on the relationships between some particular indexes of the E.P.F. and the sample residuals. This approach (Mineo, 1989) introduced the "generalized kurtosis" index to estimate p assuming the E.P.F as residual distribution. The likelihood estimation of p was considered by Agrò (1995).

3 The shape parameter, the Pearson kurtosis, and the length of tails.

For the density (2), we observe that the theoretical moment of order k is a function of the shape parameter p:

$$E\left|z - M_p\right|^{k} = \left(p\sigma_p^p\right)^{-k/p} \frac{\Gamma\left(\frac{k+1}{p}\right)}{\Gamma\left(\frac{1}{p}\right)} = \mu_k \tag{8}$$

This important relation shows that the ratios of the moments of order $2k$ and k^2 depend only on the shape parameter p.

From (8) we obtain the theoretic relation called "Generalized Kurtosis":

$$\beta_k = \frac{\mu_{2k}}{\mu_k^2} = \frac{\Gamma\left(\frac{1}{p}\right)\Gamma\left(\frac{2k+1}{p}\right)}{\left[\Gamma\left(\frac{k+1}{p}\right)\right]^2}$$

if $k = 2$ we can write the Pearson Kurtosis index:

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\Gamma\left(\frac{1}{p}\right)\Gamma\left(\frac{5}{p}\right)}{\left[\Gamma\left(\frac{3}{p}\right)\right]^2} \tag{9}$$

If $k = 1$, considering the square root of the reciprocal, we have the Geary length of tails index:

$$I = \frac{\mu_1}{\sqrt{\mu_2}} = \frac{\Gamma\left(\frac{2}{p}\right)}{\sqrt{\Gamma\left(\frac{1}{p}\right)\Gamma\left(\frac{3}{p}\right)}} \tag{10}$$

The indexes I and β_2 show a different behavior according to the variation of p. For example if p increases from 1 to 4 the index β_2 decreases from 6 to 2.1884, while I increase from 0.7071 to 0.8409. (see Table 1).

Let us now consider the following sample kurtosis of the resulting residuals obtained by the ratios of empirical moments. An estimation of β_2 is given by:

$$\beta_2^* = \frac{n\sum_i (\varepsilon_i - \bar{\varepsilon})^4}{\left[\sum_i (\varepsilon_i - \bar{\varepsilon})^2\right]^2}$$

This estimation is strongly influenced in samples with many outliers because of the fourth moment, while the one based on I,

$$I^* = \frac{\sum_i |\varepsilon_i - \bar{\varepsilon}|}{\sqrt{n}\sqrt{\sum_i |\varepsilon_i - \bar{\varepsilon}|}}$$

is particularly disturbed in samples with many values centered around the location parameter value. In Table 1 are shown the values of I and β_2 for p varying from 0.5 to 10 derived from equations (9) and (10). In the opposite way, calculating the sample values of I and β_2, it is possible to obtain, by inverse interpolation, two different estimations of p.

p	β_2	I	p	β_2	I	p	β_2	I
0.5	25.20000	.54772	3.7	2.24068	.83771	6.9	1.95463	.85587
0.6	15.57876	.59685	3.8	2.22208	.83879	7.0	1.95099	.85611
0.7	11.06208	.63441	3.9	2.20470	.83990	7.1	1.94748	.85634
0.8	8.56514	.66392	4.0	2.18844	.84090	7.2	1.94409	.85657
0.9	7.02556	.68766	4.1	2.17320	.84184	7.3	1.94082	.85678
1.0	6.00000	.70711	4.2	2.15889	.84273	7.4	1.93767	.85699
1.1	5.27660	.72330	4.3	2.14543	.84357	7.5	1.93462	.85719
1.2	4.74348	.73696	4.4	2.13276	.84436	7.6	1.93167	.85739
1.3	4.33681	.74861	4.5	2.12081	.84511	7.7	1.92883	.85758
1.4	4.01786	.75891	4.6	2.10952	.84583	7.8	1.92607	.85776
1.5	3.76195	.76738	4.7	2.09885	.84651	7.9	1.92341	.85794
1.6	3.55270	.77503	4.8	2.08875	.84715	8.0	1.92083	.85811
1.7	3.37882	.78178	4.9	2.07918	.84776	8.1	1.91833	.85827
1.8	3.23236	.78776	5.0	2.07010	.84834	8.2	1.91592	.85843
1.9	3.10751	.79310	5.1	2.05327	.84942	8.3	1.91357	.85859
2.0	3.00000	.79788	5.2	2.04547	.84993	8.4	1.91130	.85874
2.1	2.90460	.80206	5.3	2.03803	.85041	8.5	1.90910	.85889
2.2	2.82473	.80609	5.4	2.03095	.85087	8.6	1.90696	.85903
2.3	2.75252	.80963	5.5	2.02418	.85131	8.7	1.90489	.85917
2.4	2.68841	.81285	5.6	2.01773	.85173	8.8	1.90288	.85930
2.5	2.63116	.81580	5.7	2.01155	.85213	8.9	1.90093	.85943
2.6	2.57977	.81849	5.8	2.00565	.85252	9.0	1.89903	.85956
2.7	2.53342	.82097	5.9	2.00000	.85289	9.1	1.89719	.85968
2.8	2.49143	.82326	6.0	1.99459	.85324	9.2	1.89539	.85970
2.9	2.45325	.82537	6.1	1.98930	.85358	9.3	1.89452	.85981
3.0	2.41840	.82732	6.2	1.98442	.85391	9.4	1.89365	.85991
3.1	2.38648	.82914	6.3	1.97965	.85422	9.5	1.89196	.86002
3.2	2.35716	.83082	6.4	1.97506	.85452	9.6	1.89031	.86013
3.3	2.33015	.83239	6.5	1.97065	.85481	9.7	1.88871	.86024
3.4	2.30827	.83993	6.6	1.96641	.85500	9.8	1.88715	.86034
3.5	2.28286	.83522	6.7	1.96234	.85536	9.9	1.88564	.86044
3.6	2.26064	.83651	6.8	1.95841	.85562	10.0	1.88416	.86054

Table 1: Theoretical values of β_2 and I, evaluated for p varying from 0.5 to 10.

Gonin and Money (1987) considered the unbiased estimates of the second and fourth order sample moments with correction factors depending on the sample size n:

$$\hat{\mu}_2 = \frac{1}{n-1}\sum_i (\varepsilon_i - \bar{\varepsilon})^2$$

$$\hat{\mu}_4 = \frac{(n^2-2n+3)}{(n-1)(n-2)(n-3)}\sum_i (\varepsilon_i - \bar{\varepsilon})^4 - \frac{3(n-1)(2n-3)}{n(n-2)(n-3)}\hat{\mu}_2^2$$

The same result is available using the r-th K-statistics and their relations with the sample moments (Lunetta, 1966, Kendall-Stuart, 1966).

The ratio of $\hat{\mu}_4$ and $\hat{\mu}_2^2$ gives an unbiased estimator of β_2:

$$\hat{\beta}_2 = \frac{\hat{\mu}_4}{\hat{\mu}_2^2} \tag{11}$$

For the I empirical index we can only consider the correction factor of the second sample moment, consequently we use:

$$\hat{I} = \frac{\sum_i |\varepsilon_i - \bar{\varepsilon}|}{\sqrt{\sum_i |\varepsilon_i - \bar{\varepsilon}|^2}} \frac{\sqrt{n-1}}{n} \tag{12}$$

4 The proposed algorithm

The proposed algorithm is based on a two steps alternating procedure that firstly estimates the $\underline{\theta}$ parameter vector by means of the classical conjugated gradient algorithm (Fletcher-Reeves, 1964) and secondly estimates p using a joint inverse function of I and β_2 obtained comparing empirical and theoretical moments. The algorithm stops when the variation of p is not significant. In order to obtain our estimate, we minimize the difference between empirical and theoretical indexes (13) to avoid some convergence problems encountered when we investigate a different algorithm considering this difference equal to zero (Giacalone, 1994).

The function used to estimate p is therefore the following:

$$\left[\left(I - \hat{I}\right) : 0.86054\right]^2 + \left[\left(\beta_2 - \hat{\beta}_2\right) : 25.2\right]^2 = \min \tag{13}$$

where I, $\hat{I}$, β_2, $\hat{\beta}_2$ are respectively given by (10), (12), (9), (11).

For simplicity we can express this condition as $[f(p)]^2 + [g(p)]^2 = \min$.

We observe that the indexes I and β_2 show different variability and different average order size related to the variation of p. For a joint evaluation we have to eliminate the difference in average order size, therefore setting $0 < f(p) < 1$ and $0 < g(p) < 1$.

The maximum theoretical values of $f(p)$ and $g(p)$ are the chosen standardization factors. Unfortunately the β_2 index diverges to ∞, for $p \to 0$. In order to standardize $g(p)$, we use the max β_2 value equal to 25.2 and corresponding to $p = 0.5$, which is the lower bound in our simulation study. In an analogous way, to standardize $f(p)$, we use the max I value equal to 0.86054 and related to $p = 10$, which is the upper bound in our simulation study.

The proposed algorithm is then specified in the following steps:

STEP 0: Set $i = 0$ and $p_0 = 2$;

STEP 1: Fit the model to the data using the previous step value p_i;

STEP 2: Compute the estimated residuals $\varepsilon_i = y_i - g(\underline{x}_i, \underline{\theta})$, their average $\bar{\epsilon}$, and insert these quantities in the (13) which is equal to the sum of two squared functions to minimize;

STEP 3: Minimize the function (13) to obtain p_{i+1}, new estimate of p;

STEP 4: Compare the estimated p_{i+1} with the previous p_i, and if $|p_{i+1} - p_i| > 0.01$ then set $i = i + 1$ and repeat Steps 1-4, otherwise:

STEP 5: Stop the algorithm assuming the values $\hat{\theta}_i = \theta_{ji}$, as L_p-norm estimators for the parameters θ_i' and the value $p = p_i$ as joint estimation of p.

In Step 1 a nonlinear L_p-norm estimation is considered. The problem could be resolved using the optimality conditions encountered in unconstrained optimization (McCormick, 1983). The minimization algorithm (Fletcher-Reeves, 1964) is used because it takes the special structure of the problem (1) directly into account. In Step 3 we use the parabolic interpolation method (Everitt, 1987) to find the minimum sum of squared functions (13). The convergence of the proposed algorithm was empirically verified by simulating samples of different sizes for fixed theoretical values of p.

5 Design of the simulation study and conclusive notes

The performance of the above method was tested using a simulation study. The unbiasedness and the asymptotic behavior of the new estimation procedure for nonlinear regression model parameters and for the shape parameter p was verified by a Montecarlo experiment (see Tables 2 and 4). We considered 500 samples of sizes $n = 50$, 100, 200, generated from E.P.F., for 6 different values of p, ranging from 1.0 to 3.5 with step 0.5. The algorithm used to generate the pseudo-random standardized deviates ε_i (for $p \geq 1$) was proposed by Chiodi (1986). The samples x_i ($i = 1, 2, \ldots n$) were generated from a uniform distribution (0.5, 1.5). The values of $z_i = y_i$ are given by the simple exponential model:

$$y_i = \theta_2 e^{\theta_1 x_i} + \varepsilon_i \qquad i = 1, 2 \ldots n \tag{14}$$

We used model (14) with 500 samples estimates and $\theta_1' = 0.5$, $\theta_2' = 1.0$, $\sigma_p = 1.0$ as parameter values, to obtain the corresponding empirical frequency distributions and the relative analysed constants (mean, variance and mean squared error).

p	$M(\theta_1)$	$V(\theta_1)$	$M(\theta_2)$	$V(\theta_2)$	$M(\theta_1)$	$V(\theta_1)$	$M(\theta_2)$	$V(\theta_2)$
	$n = 50$				$n = 200$			
1.0	.4516	.1283	1.0671	.1645	.4975	.0249	1.0186	.0338
1.5	.4791	.1167	1.0426	.1534	.4974	.0226	1.0037	.0283
2.0	.4738	.0994	1.0548	.1266	.4915	.0178	1.0158	.0232
2.5	.4747	.0782	1.0418	.1013	.4918	.0145	1.0089	.0196
3.0	.4744	.0657	1.0579	.0785	.4927	.0116	1.0106	.0168
3.5	.5204	.0587	1.0409	.0692	.4998	.0093	1.0049	.0143

Table 2: Mean and variance of θ_1, θ_2, for a simple exponential model (14) ($\theta_1' = 0.5$, $\theta_2' = 1.0$, $\sigma_p = 1$) estimated with $L_{p_{\min}}$ method, on 500 samples of size $n = 50$, $n = 200$.

Here we present a comparative analysis using five different L_p-norm estimations:

Least squares estimators (L_2);

L_p-norm estimators with theoretical p of the E.P.F. (L_p);

L_p-norm estimators with p estimated as in (6) (Gonin and Money, 1989) ($L_{p_{gm}}$);

L_p-norm estimators using the maximum likelihood estimate of p (Agrò, 1995) (L_{p+}).

L_p-norm estimators with p estimated as in our proposal (13) ($L_{p_{\min}}$).

From the experimental results, partially reported in Tables 2 and 3, we can observe that for any p, the parameter estimates of θ_1 and θ_2 are biased for $n = 50$. Their variances decrease for increasing values of n. This is true for all the methods used and for all the theoretical values of p and depends on the nonlinearity of the model that yields the unbiasedness of the estimates only for middle-large samples sizes (see when $n = 200$). Considering the relative efficiency (see Table 3) we note that all the L_p-norm estimators give us better parameter estimates with respect to the least squares method especially for theoretical values of p away from 2.

We can also observe the gain in efficiency using the L_p-norm estimators that is higher for the L_p method in all cases considered except when $p = 2$ where L_2 and L_p methods are equal. The $L_{p_{gm}}$, L_{p+}, $L_{p_{\min}}$ methods give us different efficiency parameter values depending on the criterion used to estimate p and could be considered half way between L_2 and L_p methods. The most important difference between the L_p-norm estimators examined are presented in Table 4 where we consider the empirical sample distributions of p estimates.

For all methods considered, the estimate of p is generally biased. This bias is higher for $p > 2$, and for $n = 50$. In an analogous way, the variance of the p estimates increases as p rises and decreases in proportion to the

increase in sample size n. For the cases $p = 1.0$ and $p = 1.5$ the proposed method $L_{p_{\min}}$ seems to be the most efficient considering the variance and the mean squared error of the p estimate. The worst behavior of the $L_{p_{gm}}$ method depends on the bias of the p estimate that is reflected in the values of the mean squared error.

p	1.0		1.5		2.0		2.5		3.0		3.5	
	θ_1	θ_2	θ_1	θ_2	θ_1	θ_2	θ_1	θ_2	θ_1	θ_2	θ_1	θ_2
	$n = 50$											
L_p	2.25	1.88	1.33	1.25	1.00	1.00	1.11	1.24	1.45	1.29	1.82	1.39
$L_{p_{gm}}$	1.88	1.69	1.61	1.36	0.69	0.93	1.07	1.32	1.22	1.18	1.61	1.28
L_{p_+}	1.76	1.66	1.41	1.27	0.68	0.95	0.81	1.11	1.08	1.03	1.65	1.29
$L_{p_{\min}}$	2.16	1.77	1.34	1.48	0.85	0.98	0.89	0.99	1.33	1.24	1.73	1.32
	$n = 200$											
L_p	1.88	1.65	1.45	1.19	1.00	1.00	1.19	1.16	1.25	1.17	1.48	1.50
$L_{p_{gm}}$	1.27	1.22	1.23	1.08	0.96	0.95	1.03	1.01	1.07	1.08	1.08	1.01
L_{p_+}	1.85	1.54	1.09	1.07	0.98	0.95	1.02	1.05	1.13	1.11	1.28	1.26
$L_{p_{\min}}$	1.81	1.52	1.31	1.09	0.97	0.94	1.13	1.07	1.20	1.14	1.32	1.31

Table 3: Relative efficiency of L_p-norm estimators with respect to the least squares.

Case $p = 2$ proves to be the most efficient with $L_{p_{gm}}$ related to $L_{p_{\min}}$ and L_{p_+} for the given sample sizes. Gonin and Money's method achieves, for all the examined theoretical values of p, estimates around the value $p = 2$. For the cases $p = 2.5$ $p = 3.0$ and $p = 3.5$ Gonin and Money's procedure gives a lower variance and mean squared error compared to the L_{p_+} and $L_{p_{\min}}$ methods, but presents a very high bias for $n = 200$ sample size. The most important result is related to the asymptotic behavior of the p estimates. The methods L_{p_+} and $L_{p_{\min}}$ seem to show a possible asymptotic convergence to a Normal distribution while the $L_{p_{gm}}$ method shows that the bias increases even when the sample size increases (see $p = 3.5$).

Looking at the simulation results, it seems reasonable to distinguish two cases. For $p < 2$, the $L_{p_{\min}}$ method here proposed gives us the best performance. The case $p > 2$ is well dealt using either the $L_{p_{\min}}$ or the L_{p_+} methods. Agrò's method achieves the best results when the sample size is high as the maximum likelihood methods generally do.

Finally, our simulation study shows how our algorithm achieves good efficient estimates for regression and shape parameters when compared to the results obtained by the Gonin and Money's procedure. In particular, the asymptotic unbiasedness is a fundamental aspect for the proper function of every estimation procedure. In non-normal symmetric distribution, better

performance of variances of p and parameter estimates are obtained when using the proposed $L_{p_{\min}}$ method.

	Methods: $L_{p_{gm}}$, L_{p_+}, $L_{p_{\min}}$					
p	1.0	1.5	2.0	2.5	3.0	3.5
			$n = 50$			
M(p_{gm})	1.3349	1.8927	2.2161	2.5012	2.6938	2.8734
V(p_{gm})	0.1097	0.1546	0.2236	0.2479	0.2566	0.2637
MSE(p_{gm})	0.2219	0.3088	0.2703	0.2479	0.3504	0.6563
M(p_+)	1.4957	1.9985	2.5561	3.0392	3.5245	4.0822
V(p_+)	0.3515	0.4846	1.8152	2.1054	2.4296	3.1064
MSE(p_+)	0.5972	0.7331	2.1244	2.3961	2.7047	3.4453
M($p_{\min}$)	1.2374	1.7485	2.2656	2.7222	3.2114	3.4099
V($p_{\min}$)	0.1221	0.3577	0.6451	1.0566	1.4437	1.6687
MSE($p_{\min}$)	0.1784	0.4135	0.7156	1.1060	1.4884	1.6768
			$n = 200$			
M(p_{gm})	1.3307	1.7143	2.0596	2.3458	2.5733	2.7622
V(p_{gm})	0.0207	0.0446	0.0559	0.0589	0.0514	0.0585
MSE(p_{gm})	0.1301	0.0905	0.0594	0.0827	0.2335	0.6028
M(p_+)	1.2157	1.7386	2.0861	2.5182	2.9188	3.3094
V(p_+)	0.0495	0.0221	0.0624	0.1608	0.2747	0.3564
MSE(p_+)	0.0960	0.0790	0.0698	0.1611	0.2813	0.3927
M($p_{\min}$)	1.0784	1.5553	2.0641	2.5311	3.0423	3.5289
V($p_{\min}$)	0.0348	0.0476	0.0975	0.2741	0.4996	0.6135
MSE($p_{\min}$)	0.0411	0.0507	0.1016	0.2751	0.5014	0.6143

Table 4: Mean, variance and mean squares error of p estimated on samples of sizes $n = 50$, $n = 200$ generated on E.P.F., considering different methods to estimate p.

References

[1] Agrò, G. (1995). Maximum likelihood estimation for the exponential power function parameters. *Comm. Statist. Simul. Comput.* **24**, 523-536.

[2] Chiodi, M. (1986). Procedures for generating pseudo-random numbers from a normal distribution of order p. *Riv. Stat. Applic.* **19**, 7-26.

[3] Ekblom and Henriksson, S. (1969). L_p criteria for the estimation of location parameters. *SIAM J. Appl. Math.* **17**, 1130-1141.

[4] Everitt, B.S. (1987). *Introduction to Optimization Methods and their Application in Statistics.* Chapman and Hall, London.

[5] Fletcher, R. and Reeves, C.M. (1964). Function minimization by conjugate gradients. *Comput. J.* **7,** 149-154.

[6] Forsythe, A.B. (1972). Robust estimation of straight line regression coefficients by minimizing p-th power deviations. *Technometrics* **14,** 159-166.

[7] Giacalone, M. (1994). Stimatori di norma L_p. Applicazioni alla domanda di beni alimentari in Italia. In *Convegno: La Statistica per la Gestione Aziendale e l'Analisi di Mercato*, 141-157. Bressanone.

[8] Giacalone, M. (1996). Parameter evaluation of an exponential power function by simulation study. In *Shorts Communications and Posters, Compstat 96.* Barcelona.

[9] Gonin, R. and Money, A.H. (1987). A review of computational methods for solving the nonlinear L_1 norm estimation problem. In *Statistical Data Analysis Based on the L_1 Norm and Related Methods*, Ed. Y. Dodge. North Holland.

[10] Gonin, R. and Money A.H. (1989). *Nonlinear L_p-Norm Estimation.* Marcel Dekker, New York.

[11] Harter, H.L. (1977). The nonuniqueness of absolute values regression. *Commun. Statist. Simul. Comput.* **6,** 829-838.

[12] Kendall, M.C. and Stuart, A. (1966). *The Advanced Theory of Statistics.* Griffin, London.

[13] Lunetta, G. (1963). Di una Generalizzazione dello schema della curva normale. *Ann. Facoltà Econ. Comm.* **17,** 237-244. Palermo.

[14] Lunetta, G. (1966). Di alcune distribuzioni deducibili da una generalizzazione dello schema della curva normale. *Ann. Facoltà Econ. Comm.* **20,** 117-143. Palermo

[15] Mineo, A. (1989). The norm-p estimation of location, scale and simple linear regression parameters. In *Lectures Notes in Statistics-Statistical Modelling,* 222-223. Trento Springer-Verlag.

[16] Money, A.H. et al. (1982). The linear regression model: L_p-norm estimation and the choice of p. *Commun. Statist. Simul. Comput.* **11,** 89-109.

[17] Sposito, V.A. (1982). On unbiased L_p regression. *J. Amer. Statist. Assoc.* **77,** 652-653.

[18] Subbotin, M. (1923). On the law of frequency of error. *Mathematicheskii Sbornik.*

[19] Vianelli, S. (1963). La misura della variabilità condizionata in uno schema generale delle curve normali di frequenza. *Statistica* **23.**

Student (1997), Vol.2, No.2, 131-138

Least Orthogonal Absolute Deviations Problem for Exponential Function

Tomislav Marošević and Dragan Jukić

Department of Mathematics, University of Osijek, Croatia

Abstract: We consider the existence problem of the optimal parameters for the exponential function, in the sense of the least orthogonal absolute deviations, and prove the existence of such optimal parameters for monotonic data.

Key words: Exponentiel model, L_1-estimation, optimal parameters.

1 Introduction

Observations in applied sciences are often regarded as satisfying a mathematical model of a certain type. In this paper we are going to consider the exponential model

$$f(t; b, c) = be^{ct}, \tag{1}$$

which is often used in applied research (see [5], [10], [13]). The unknown parameters b, c have to be determined on the basis of the data (p_i, t_i, f_i), $i = 1, 2, \ldots, m$, $(m > 2)$, where $p_i > 0$ are the data weights, t_i are the measured values of the independent variable, and f_i are the respective measured values of the dependent variable.

The estimation of the parameters b, c is usually done in the sense of the ordinary least squares (i.e. the l_2 norm), by minimizing the functional (see [1], [8])

$$F(b, c) = \sum_{i=1}^{m} p_i (be^{ct_i} - f_i)^2 .$$

In this case one usually assumes that only the measured values $f_1, \ldots, f_m$ of the dependent variable contain additive errors, and that the errors are normally distributed with mean 0 and variance $\sigma^2 I$.

Some theorems about the existence of the best approximation of the unknown parameters b, c in the sense of the ordinary least squares can be found in [5], [8], [11]. The existence of optimal parameters for the exponential function in some other l_p norms (see [7]) is considered in [4], [5], [11]. When the errors follow the Laplace distribution, the use of the least absolute deviations (i.e. the l_1 norm) is recommended (see [2], [6]).

In the more general situation, the measurements of the independent variables t_i are also assumed to contain errors (see [3], [16]). If one assumes that f_i has unknown additive error ε_i and t_i has unknown additive error δ_i, then $f_i = be^{c(t_i+\delta_i)} + \varepsilon_i$, $i = 1, \ldots, m$. In this case the parameters can be estimated in the sense of the total least squares (TLS approach) (see [3], [13]). This approach leads to minimization of the weighted sum of squares of the distances from the data points (t_i, f_i) to the curve $f(t; b, c) = be^{ct}$, i.e. minimization of the functional

$$G(b, c, \delta_1, \delta_2, \ldots, \delta_m) = \sum_{i=1}^{m} p_i \left((be^{c(t_i+\delta_i)} - f_i)^2 + \delta_i^2 \right) .$$

The existence problem of the optimal parameters for the exponential function (1) in the sense of the total least squares is considered in [13].

In this paper we consider the existence problem of the optimal parameters for the exponential function in the sense of the least orthogonal absolute deviations (LOAD) approach (see [9], [14], [15]). In this case, the problem of estimating the parameters b, c is reduced to the problem of minimizing the functional

$$S(b, c, \boldsymbol{\delta}) = \sum_{i=1}^{m} p_i \sqrt{(be^{c(t_i+\delta_i)} - f_i)^2 + \delta_i^2}\,, \tag{2}$$

where $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_m)^T \in \mathbf{R}^m$. Note that $S(b, c, \boldsymbol{\delta})$ is the weighted sum of the distances from the points (t_i, f_i) to the curve $f(t; b, c) = be^{ct}$ (see Figure 1). The LOAD approach has potentially significant applications to statistical estimation and to curve fitting.

We are going to show that for monotonic data, there exists a point $(b^*, c^*, \boldsymbol{\delta}^*)$ which minimizes the functional S. The idea of the proof is based on [8] and [13].

2 The existence problem

Suppose we are given the data (p_i, t_i, f_i), $i = 1, \ldots, m$, $m > 2$, where $p_i > 0$ are some positive weights and $t_1 < \ldots < t_m$. Since the numbers f_i usually denote a quantity of something, we assume in the remainder of this paper

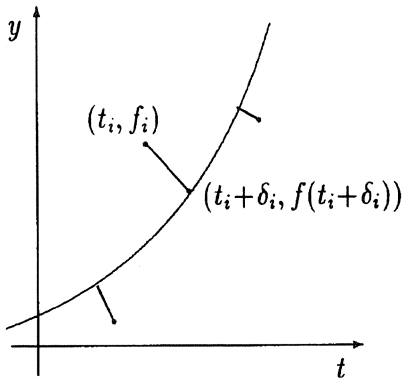


Figure 1: Least orthogonal absolute deviations problem.

that the given data are such that $f_i > 0$, $i = 1, \ldots, m$. Furthermore, if $0 < f_1 \leq f_2 \leq \ldots \leq f_m$, then it is appropriate to describe such data by an increasing function. Similarly, if $f_1 \geq f_2 \geq \ldots \geq f_m > 0$, then it is appropriate to use a decreasing function.

We have the following theorem.

Theorem 1 *Let the given data* (p_i, t_i, f_i), $i = 1, \ldots, m$, $m > 2$, *be such that* $t_1 < \ldots < t_m$ *and* $f_i > 0$, $i = 1, \ldots, m$.

(i) If $f_1 \leq f_2 \leq \ldots \leq f_m$, *then there exists a point* $(b^*, c^*, \boldsymbol{\delta}^*) \in U \times \mathbf{R}^m$,

$$U = \left\{(b, c) \in \mathbf{R}^2 \ : \ b > 0, c \geq 0\right\},$$

at which the functional S *defined by (2) attains its infimum on the set* $U \times \mathbf{R}^m$.

(ii) If $f_1 \geq f_2 \geq \ldots \geq f_m$, *then there exists a point* $(b^*, c^*, \boldsymbol{\delta}^*) \in V \times \mathbf{R}^m$,

$$V = \left\{(b, c) \in \mathbf{R}^2 \ : \ b > 0, c \leq 0\right\},$$

at which the functional S *defined by (2) attains its infimum on the set* $V \times \mathbf{R}^m$.

Proof:

Case (i): Since $S \geq 0$, there exists $S^* := \inf_{(b,c,\boldsymbol{\delta}) \in U \times \mathbf{R}^m} S(b, c, \boldsymbol{\delta})$. Let $(b_n, c_n, \boldsymbol{\delta}^n)$ be a sequence in $U \times \mathbf{R}^m$ such that

$$S^* = \lim_{n \to \infty} S(b_n, c_n, \boldsymbol{\delta}^n) = \lim_{n \to \infty} \sum_{i=1}^{m} p_i \sqrt{(b_n e^{c_n(t_i + \delta_i^n)} - f_i)^2 + (\delta_i^n)^2}\,. \quad (3)$$

Note that the sequences $(b_n e^{c_n(t_i+\delta_i^n)})$, (δ_i^n), $i = 1, \ldots, m$, are bounded. If they were not, we would have $\overline{\lim_n} S(b_n, c_n, \boldsymbol{\delta}^n) = \infty$. Without loss of generality, by the Bolzano-Weierstrass theorem, we may assume that these sequences converge (otherwise, we take convergent subsequences). Let

$$\lim_{n\to\infty} b_n e^{c_n(t_i+\delta_i^n)} = r_i\,, \qquad \lim_{n\to\infty} \delta_i^n = \delta_i^*, \qquad i = 1, \ldots, m.$$

First, we show that the sequence (c_n) is bounded. We prove this by contradiction. Suppose the sequence (c_n) is unbounded. Without loss of generality, by taking appropriate subsequence if necessary, we may assume that $c_n \to \infty$. In order to show that the infimum of the functional S cannot be attained in such a way, we are going to find a point in $U \times \mathbf{R}^m$ at which the functional S attains a value which is smaller than $\lim_n S(b_n, c_n, \boldsymbol{\delta}^n)$. For that purpose, let us denote:

$$\begin{aligned} M &:= \{1, \ldots, m\}, \quad \tau_0 := \max\{t_1 + \delta_1^*, \ldots, t_m + \delta_m^*\}, \\ M_0 &:= \{i \in M : t_i + \delta_i^* = \tau_0\}. \end{aligned}$$

Note that $r_i = 0$, $\forall i \in M \setminus M_0$. Namely, if $i_0 \in M_0$, then $\forall i \in M \setminus M_0$ the following holds:

$$\begin{aligned} r_i &= \lim_{n\to\infty} b_n e^{c_n(t_i+\delta_i^n)} = \lim_{n\to\infty} b_n e^{c_n(t_{i_0}+\delta_{i_0}^n)} \cdot \lim_{n\to\infty} e^{c_n[(t_i+\delta_i^n)-(t_{i_0}+\delta_{i_0}^n)]} \\ &= r_{i_0} \cdot 0 = 0\,. \end{aligned}$$

Furthermore, we have

$$\lim_{n\to\infty} \sqrt{(b_n e^{c_n(t_i+\delta_i^n)} - f_i)^2 + (\delta_i^n)^2} \geq \begin{cases} \sqrt{f_i^2 + \delta_i^{*2}}\,, & \text{for } i \in M \setminus M_0\,, \\ |\delta_i^*| = |\tau_0 - t_i|\,, & \text{for } i \in M_0\,. \end{cases} \tag{4}$$

Thus from (4), for the limit (3) we obtain

$$S^* \geq \sum_{i\in M\setminus M_0} p_i \sqrt{f_i^2 + \delta_i^{*2}} + \sum_{i\in M_0} p_i |\tau_0 - t_i| =: S_0\,. \tag{5}$$

Let us show that there exists a point at which the functional S attains a value smaller than S_0. For that purpose, we define the functions $\hat{b}$, $\hat{\delta}_i$: $\langle 0, \infty\rangle \to \mathbf{R}$, $i = 1, \ldots, m$, by

$$\hat{b}(c) = b_0 \cdot e^{-c\tau_0}\,,$$

$$\hat{\delta}_i(c) = \begin{cases} \delta_i^*, & \text{if } i \in M \setminus M_0 \\ \tau_0 - t_i + \frac{1}{c} \ln \frac{f_i}{b_0}, & \text{if } i \in M_0\,, \end{cases}$$

where b_0 is taken to be f_m if $\{i \in M_0 : t_i \geq \tau_0\} = \emptyset$, and if $\{i \in M_0 : t_i \geq \tau_0\} \neq \emptyset$ then we take $b_0 = \min\{f_i : i \in M_0 \ \& \ t_i \geq \tau_0\}$.

Now we get

$$S(\hat{b}(c), c, \hat{\boldsymbol{\delta}}(c)) = \sum_{i \in M \setminus M_0} p_i \sqrt{(b_0 e^{c(t_i + \delta_i^* - \tau_0)} - f_i)^2 + \delta_i^{*2}} + \sum_{i \in M_0} p_i |\tau_0 - t_i + \frac{1}{c} \ln \frac{f_i}{b_0}| . \tag{6}$$

From the definition of b_0, it is not difficult to see that for sufficiently large $c \in \mathbf{R}$

$$\sqrt{f_i^2 + \delta_i^{*2}} > \sqrt{(b_0 e^{c(t_i + \delta_i^* - \tau_0)} - f_i)^2 + \delta_i^{*2}}, \qquad \forall i \in M \setminus M_0$$

holds, and

$$|\tau_0 - t_i| \geq |\tau_0 - t_i + \frac{1}{c} \ln \frac{f_i}{b_0}|, \qquad \forall i \in M_0 .$$

Therefore, from (5) and (6) we conclude that $S^* \geq S_0 > S(\hat{b}(c), c, \hat{\boldsymbol{\delta}}(c))$. This means that the infimum of the functional S cannot be obtained in such a way ($c_n \to \infty$).

Therefore, the sequence (c_n) is bounded. By the Bolzano-Weierstrass theorem, we may assume that the sequence (c_n) is convergent. Let $c_n \to c^*$. Then, it follows from (3) that the sequence (b_n) must be bounded. Analogously, we can assume that $b_n \to b^* \geq 0$.

By continuity of the functional S, we get

$$\inf_{(b,c,\boldsymbol{\delta}) \in U \times \mathbf{R}^m} S(b, c, \boldsymbol{\delta}) = \lim_{n \to \infty} S(b_n, c_n, \boldsymbol{\delta}^n) = S(b^*, c^*, \boldsymbol{\delta}^*) .$$

For $b = 0$, $c \in [0, \infty\rangle$ and $\boldsymbol{\delta} \in \mathbf{R}^m$ we have $S(0, c, \boldsymbol{\delta}) > S(f_1, 0, \mathbf{0})$ and hence we conclude that $(b^*, c^*) \in U$.

Case (ii): Since for $c < 0$

$$\sum_{i=1}^m p_i \sqrt{(be^{c(t_i + \delta_i)} - f_i)^2 + \delta_i^2} = \sum_{i=1}^m p_i \sqrt{(be^{-c(-t_i + (-\delta_i))} - f_i)^2 + (-\delta_i)^2},$$

and the data $(-t_i, f_i)$, $i = 1, \ldots, m$ have the increasing property, this case reduces to *Case (i)*. □

Since

$$\sum_{i=1}^m p_i \sqrt{(be^{c(t_i + \delta_i)} - f_i)^2 + \delta_i^2} = \sum_{i=1}^m p_i \sqrt{(-be^{c(t_i + \delta_i)} - (-f_i))^2 + \delta_i^2},$$

from Theorem 1 we get the following

Corollary 1 *Let the given data* (p_i, t_i, f_i), $i = 1, \ldots, m$, $m > 2$, *be such that* $t_1 < \ldots < t_m$ *and* $f_i < 0$, $i = 1, \ldots, m$.

(i) If $f_1 \geq f_2 \geq \ldots \geq f_m$, *then there exists a point* $(b^*, c^*, \delta^*) \in X \times \mathbf{R}^m$,

$$X = \left\{(b, c) \in \mathbf{R}^2 \ : \ b < 0, c \geq 0\right\},$$

at which the functional S *defined by (2) attains its infimum on the set* $X \times \mathbf{R}^m$.

(ii) If $f_1 \leq f_2 \leq \ldots \leq f_m$, *then there exists a point* $(b^*, c^*, \delta^*) \in Y \times \mathbf{R}^m$,

$$Y = \left\{(b, c) \in \mathbf{R}^2 \ : \ b < 0, c \leq 0\right\},$$

at which the functional S *defined by (2) attains its infimum on the set* $Y \times \mathbf{R}^m$.

Remark 1 *In the approximation problem for the exponential function one can consider the so-called total least absolute deviations (see [12], [17]). Then, instead of minimizing the functional* S, *one is faced with the minimization problem for the functional*

$$T(b, c, \delta) = \sum_{i=1}^{m} p_i \left(|be^{c(t_i+\delta_i)} - f_i| + |\delta_i|\right) .$$

Analogously, one can also show that Theorem 1 is applicable for the functional T, *i.e. Theorem 1 is applicable in the case of the approximation in the total* l_1 *norm.*

t_i	3	4	5	6	7	8
f_i	3.064	4.865	4.898	6.133	8.290	8.911
with outlier f_1	0.064	4.865	4.898	6.133	8.290	8.911

opt. parameters	LOAD	TLS
without outliers	$b^* = 1.977,\ c^* = 0.188$	$b^*_{tls} = 1.894,\ c^*_{tls} = 0.2002$
with outlier	$b^* = 1.976,\ c^* = 0.189$	$b^*_{tls} = 0.693,\ c^*_{tls} = 0.356$

Table 1: Optimal parameters for some artificial data.

Example 1 *In order to illustrate the approximation problem we are dealing with, we mention two ad hoc chosen examples, which were done using Mathematica (see [18]).*

a) Least orthogonal absolute deviations and total least squares. In Table 1 we give some artificial data (generated by means of the function $f(t) = 2e^{0.2t}$, *with* $p_i = 1, i = 1, \ldots, m$), *and the optimal parameters* b^*, c^* *obtained in the sense of the least orthogonal absolute deviations, and* b^*_{tls}, c^*_{tls} *obtained in the sense of the total least squares. The data with outlier, and the respective curves* $f(t) = b^* e^{c^* t}$ *and* $f_{tls}(t) = b^*_{tls} e^{c^*_{tls} t}$ *are shown in Figure 2 a. We can see that the least orthogonal absolute deviations criterion was not as sensitive to the outlier as the total least squares criterion.*

b) Least orthogonal absolute deviations and the l_1 *norm. In Table 2 we give another set of artificial data (generated by means of the function* $f(t) = e^t$, $p_i = 1, i = 1, \ldots, m$), *and the optimal parameters* b^*, c^* *obtained in the sense of the least orthogonal absolute deviations, and* $b^*_{l_1}$, $c^*_{l_1}$ *obtained in the sense of the least absolute deviations. The data and the respective curves* $f(t) = b^* e^{c^* t}$ *and* $f_{l_1}(t) = b^*_{l_1} e^{c^*_{l_1} t}$ *are shown in Figure 2 b.*

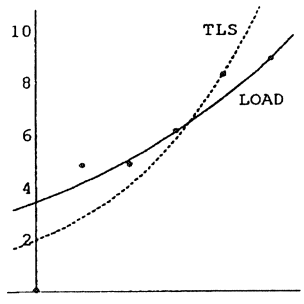


Figure 2 a. LOAD and TLS

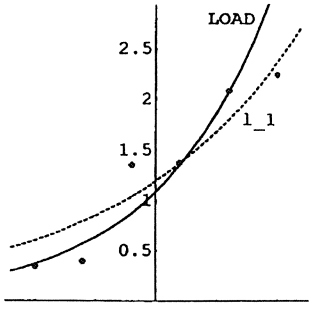


Figure 2 b. LOAD and l_1

t_i	-1	-0.6	-0.2	0.2	0.6	1
f_i	0.340	0.391	1.351	1.371	2.084	2.243
	LOAD			l_1		
opt. parameters	$b^* = 1.088,\ c^* = 1.084$			$b^*_{l_1} = 1.196,\ c^*_{l_1} = 0.686$		

Table 2: Optimal parameters for another set of artificial data.

References

[1] Björck, Å. (1996). *Numerical Methods for Least Squares Problems.* SIAM, Philadelphia.

[2] Bloomfield, P. and Steiger, W.L. (1983). *Least Absolute Deviations – Theory, Applications and Algorithms.* Birkhäuser, Boston.

[3] Boggs, P.T., Byrd, R.H. and Schnabel, R. B. (1987). A stable and efficient algorithm for nonlinear orthogonal distance regression. *SIAM J.Sci.Stat.Comput.* **8**, 1052–1078.

[4] Braess, D. (1986). *Nonlinear Approximation Theory.* Springer-Verlag, Berlin.

[5] Demidenko, E.Z. (1989). *Optimization and Regression* (in Russian). Nauka, Moscow.

[6] Dodge, Y. (Ed.) (1987). *Statistical Data Analysis Based on the L_1-norm and Related Methods.* Elsevier, Amsterdam.

[7] Gonin, R. and Money, A.H. (1989). *Nonlinear l_p-Norm Estimation.* Marcel Dekker, New York.

[8] Jukić, D. and Scitovski, R. (1997). Existence of optimal solution for exponential model by least squares. *J. Comput. Appl. Math.* **78**, 317–328.

[9] Nyquist, H. (1988). Least orthogonal absolute deviations. *Comput. Statist. Data Anal.* **6**, 361–367.

[10] Osborne, M.R. and Smyth, G.K. (1995). A modified Prony algorithm for exponential function fitting. *SIAM J. Sci. Comput.* **16**, 119–138.

[11] Rice, J.R. (1969). *The Approximation of Functions, Vol. II – Nonlinear and Multivariate Theory.* Addison-Wesley, Reading, MA.

[12] Rosen, J.B., Park, H. and Glick, J. (1996). Total least norm formulation and solution for structured problems. *SIAM J. Matrix Anal. Appl.* **17**, 110–126.

[13] Scitovski, R. and Jukić, D. (1996). Total least squares problem for exponential function. *Inverse Problems* **12**, 341–349.

[14] Späth, H. (1982). On discrete linear orthogonal L_p approximation. *ZAMM* **62**, T354–T355.

[15] Späth, H. and Watson, G.A. (1987). On orthogonal linear l_1 approximation. *Numerische Mathematik* **51**, 531–543.

[16] Steil, G.M., Volund, A., Kahn, S.E. and Bergman, R.N. (1993). Reduced sample number for calculation of insulin sensitivity and glucose effectiveness from the Minimal model. *Diabetes* **42**, 250–256.

[17] Watson, G.A. (1987). The total approximation problem. In *Approximation Theory IV,* Eds. Chui, C.K., Schumaker, L.L. and Ward, J.D., 723-728. Academic Press.

[18] Wolfram, S. (1991). *Mathematica – A System for Doing Mathematics by Computer.* Addison-Wesley.

Student (1997), Vol.2, No.2, 195-201

The Median Function on Structured Metric Spaces

F. R. McMorris

Department of Mathematics, University of Louisville, USA

Abstract: When (X, d) is a finite metric space and $\pi = (x_1, \ldots, x_k) \in X^k$, a median for π is a element x of X for which $\sum_{i=1}^{k} d(x, x_i)$ is minimum. The function that returns the set of all medians for any tuple π is called the median function on X. A brief survey is given of some of the results concerning the median function, starting with an arbitrary metric space and finishing with the case where X is a set of hypergraphs and d is the metric based on the L_1-norm. A simplistic maximum likelihood interpretation for the median function is also given.

Key words: Consensus, hypergraph, median, metric space.

1 Introduction

Let (X, d) be a finite metric space and $X^* = \bigcup_{k>1} X^k$. Elements of X^* are called *profiles* and denoted $\pi = (x_1, \ldots, x_k)$, $\pi' = (y_1, \ldots, y_m)$, etc. A *median* for the profile $\pi = (x_1, \ldots, x_k)$ is an element x of X for which $\sum_{i=1}^{k} d(x, x_i)$ is minimum. Thinking of $\sum_{i=1}^{k} d(x, x_i)$ as a measure of "remoteness" of x to the profile π, then the medians for π can be considered as the "closest" points of X to π. Thus it is not surprising that finding medians is important in many studies involving consensus and location. Among the many references are [6], [7], [12], [25], and [27]. The function that produces medians is the median function, also called the median procedure. Specifically, the *median function* on X is the map $M : X^* \to 2^X \backslash \{\emptyset\}$ defined by $M(\pi) = \{x : x \text{ is a median for } \pi\}$ for all $\pi \in X^*$. (2^X denotes the set of all subsets of X.)

In the following sections some results concerning the median function are briefly reviewed, starting with M defined on an arbitrary metric space and

then considering two L_1-metric spaces where complete characterizations of M are obtained.

2 The median function on finite metric spaces.

We first note a few things that can be said about the median function on an arbitrary finite metric space (X, d). In order to do this we need to introduce some properties for a *consensus function* on X, which is simply a function $F : X^* \to 2^X \backslash \{\emptyset\}$. Of course any "consensus" aspects of such a function will be the result of additional consensus-like conditions that F enjoys.

Consider the following properties for a consensus function F on X:

Anonymity (A): For every profile $\pi = (x_1, \ldots x_k) \in X^*$ and permutation σ of $\{1, \ldots k\}$, $F(\pi) = F(\pi^\sigma)$, where $\pi^\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(k)})$.

Betweenness (B): $F((x, y)) = \{z : d(x, y) = d(x, z) + d(z, y)\}$.

Consistency (C): If $F(\pi) \cap F(\pi') \neq \varnothing$ for profiles π and π', then $F(\pi\pi') = F(\pi) \cap F(\pi')$, where $\pi\pi'$ is the concatenation of π and π'.

Faithfulness (F): $F((x)) = \{x\}$ for all $x \in X$.

It is easy to see that if F satisfies (B) and (C), then it satisfies (F): $F((x)) = F((x))$ so by (C) we have $F((x, x)) = F((x))$, and $F((x, x)) = \{x\}$ by (B) because $d(x, y) = 0$ if and only if $x = y$.

It is also not hard to show that M satisfies (A), (B) and (C) (and hence also (F)) on any metric space ([3], [20]), but these properties do not in general characterize M among consensus functions. Now structure must be added to the metric space. Indeed it is an interesting open problem to find those conditions that need to be imposed on X in order that the above three very simple properties characterize M. Some success has been attained in characterizing M on (X, d) when X is among a restricted class of graphs or ordered sets and d is the geodesic metric on the diagram of X. The titles of the following selected references should give the reader an idea of this line of research. ([4], [15], [16], [21], [26]).

3 Hierarchies

We now move to a lower, more detailed, level of analysis where the points in the metric space are themselves discrete structures. In particular, they will be classifications of a finite set of entities S. Almost every classification scheme of S involves the notion of "cluster", where the clusters are usually constructed so that objects of the same cluster are more similar to one another than to objects of another cluster. Since a *hypergraph* on S is

simply a set of nonempty subsets of S, a classification scheme is foremost just a hypergraph. A classification of S will usually include all the singletons and the whole set so we require, for any hypergraph H on S, $\{x\} \in H$ for all $x \in S$ and also $S \in H$. We refer to elements $A \in H$ as *clusters* of the hypergraph H. Let $\mathcal{H}$ denote the set of all hypergraphs on S and $n = |S|$ denote the number of elements of the set S. As Gordon [13] notes, a hypergraph $H \in \mathcal{H}$ can be considered as a $\{0,1\}$-vector $(\alpha_1, \ldots, \alpha_m)$ where $m = 2^n - n - 2$ and $\alpha_i = 1$ if and only if the set indexed by i is a cluster of H. The L_1-norm of H is then $\sum_{i=1}^{m} |\alpha_i|$, the number of nontrivial clusters of H. Letting d be the metric based on this L_1-norm we have for $H_1, H_2 \in \mathcal{H}$

$$d(H_1, H_2) = |H_1| + |H_2| - 2|H_1 \cap H_2|$$

where $|H_i|$ is the number of clusters in H_i. This, of course, is the well known *symmetric difference* metric. From now on the metric spaces considered will be sets of hypergraphs, equipped with the symmetric difference metric.

A classification of S will usually be structured into something tree-like called a hierarchy. In our context a *hierarchy* (also called an *n-tree*) on S is a hypergraph T such that $A \cap B \in \{A, B, \varnothing\}$ for every cluster $A, B \in T$. Let $\mathcal{T}$ denote the set of all hierarchies on S. Because the median function on $(\mathcal{T}, d)$ is defined in terms of d, one way of "knowing" M is to characterize this symmetric difference metric. This was done in [17] generalizing work found in [2], [9], [10], and [14]. Also see [12] for several examples of this approach.

In order to present a characterization of M on $(\mathcal{T}, d)$ one more property needs to be defined. It is named after the Marquis de Condorcet who, according to Young [28], actually was suggesting the median function in 1785 [11]. If $\pi = (T_1, \ldots, T_k)$ is a profile of hierarchies and $A \subseteq S$, let $\gamma(A, \pi) = \frac{|\{i: A \in T_i\}|}{k}$. A consensus function F on $(\mathcal{T}, d)$ is $\frac{1}{2}$-*condorcet* if for any $A \subseteq S$ and profile $\pi = (T_1, \ldots, T_k) \in \mathcal{T}^*$ such that $\gamma(A, \pi) = \frac{1}{2}$, the following holds:

$T \in F(\pi)$ if and only if $T \cup \{A\} \in F(\pi)$ provided $T \cup \{A\}$ is a hierarchy.

The following result appears in [23] in an abstract version, and improved earlier results in [3] and [5].

Theorem 1 *Let F be a consensus function on $(\mathcal{T}, d)$. Then F is the median function if and only if F satisfies properties (C) and (F) and is $\frac{1}{2}$-condorcet.*

4 Weak hierarchies

A hypergraph W on S is a *weak hierarchy* if and only if $A \cap B \cap C \in \{A \cap B, A \cap C, B \cap C\}$ for all clusters $A, B, C \in W$. Weak hierarchies were introduced in [1] and [8] to generalize hierarchies in a way that allows partial overlap of clusters. Let $\mathcal{W}$ denote the set of all weak hierarchies on S.

In contrast to the situation for hierarchies, there exist consensus functions on $(\mathcal{W}, d)$ that satisfy the three conditions of Theorem 1 yet are not the median function [24]. Unfortunately the major additional condition is fairly complicated. For the profile $\pi = (W_1, \ldots W_k) \in \mathcal{W}^*$ and set $A \subseteq S$ define

$$w(A) = k(2\gamma(A, \pi) - 1)$$

and

$$J(\pi) = \left\{A : \gamma(A, \pi) \geq \frac{1}{2}\right\}.$$

Note that $w(A) \geq 0$ if and only if $A \in J(\pi)$. For each $W \in \mathcal{W}$ set

$$w_\pi(W) = \sum w(A)$$

where the sum is taken over all $A \in J(\pi)$ for which $A \notin W$. When $J(\pi) \subseteq W$, set $w_\pi(W) = 0$. A consensus function F on $\mathcal{W}$ is $\frac{1}{2}$-*weighted* if, for any profile $\pi \in \mathcal{W}^*$, $F(\pi) \subseteq \{W \in \mathcal{W} : w_\pi(W)$ is minimized $\}$.

In [24] the following is proved:

Theorem 2 *The median function on $(\mathcal{W}, d)$ is the maximum element of the set of all consensus functions on $\mathcal{W}$ which satisfy properties (C) and (F) and are $\frac{1}{2}$-condorcet and $\frac{1}{2}$-weighted.*

Clearly it would be nice to find a more simple list of properties that characterize M on $(\mathcal{W}, d)$.

5 Arbitrary hypergraphs and maximum likelihood

For $\pi = (H_1, \ldots, H_k) \in \mathcal{H}^*$, let $Maj(\pi) = \{A : \gamma(A, \pi) > \frac{1}{2}\}$. The resulting consensus function $Maj : \mathcal{H}^* \to 2^{\mathcal{H}} - \{\varnothing\}$ is the *majority rule*. When $\mathcal{H} = \mathcal{T}$, it is easy to see that $Maj(\pi) \in \mathcal{T}$ for all $\pi \in \mathcal{T}^*$ [17], but care must be taken for other classes of hypergraphs. For example, there exist profiles π of weak hierarchies for which $Maj(\pi) \notin \mathcal{W}$ [21]. Using Maj, finding medians in $(\mathcal{H}, d)$ is easy. Let $E(\pi) = \{A : \gamma(A, \pi) = \frac{1}{2}\}$. An argument given in [5] can be used to show that $H \in \mathcal{H}$ is a median for π if and only if $H = Maj(\pi) \cup K$ where $K \subseteq E(\pi)$.

Continuing in the most general hypergraph framework consider the situation where there are k "procedures" (voters, algorithms, . . .) that assess the objects in S to form profiles in $\mathcal{H}^*$. Let $\mathcal{C}$ be a restricted class of hypergraphs such as $\mathcal{W}$ or $\mathcal{T}$. When dealing with voter preferences, Condorcet made assumptions for voter rankings analogous to the following ones for hypergraphs [28].

The Condorcet assumptions:

1. There is a "true" hypergraph H' in $\mathcal{C}$.

2. Each of the k procedures makes a judgment about each subset of S as to whether it is a cluster in H' or not.

3. Each procedure makes a correct judgment about a given set with probability p, where $\frac{1}{2} < p < 1$. p is the same for all procedures, and all sets are equally difficult to judge.

4. Judgments about different sets are independent and so are the judgments about the same set by different procedures.

Following [19], the goal is to find those hypergraphs in $\mathcal{C}$ most likely to be true, given the results of all the k procedures. Since each procedure is making $m = 2^n - n - 2$ yes/no decisions, these data can be put into one vector $\mathbf{q} = (q_1, \ldots, q_m)$ where q_i is the number of procedures that chose the set labeled by i. Thus the likelihood of observing $\mathbf{q}$ given that the hypergraph H is true is

$$L(\mathbf{q}; H) = p^{c(H)}(1-p)^{i(H)}$$

where $c(H)$ is the total number of correct choices based on H being true, and $i(H)$ is the total number of incorrect choices based on H being true.

We want to find those $H \in \mathcal{C}$ that maximize $L(\mathbf{q}; H)$, given the observed data $\mathbf{q}$. Since $i(H) = km - c(H)$ and p is constant with $\frac{1}{2} < p < 1$, we need only maximize $c(H) \log\left(\frac{p}{1-p}\right)$. But maximizing $c(H)$ means minimizing $\sum_{i=1}^{k} d(H, H_i)$ and so we have a maximum likelihood interpretation, albeit a simplistic one, of the median function.

Acknowledgments: This research was supported from the United States Office of Naval Research grant N00014-95-1-0109.

References

[1] H.-J. Bandelt and A.W.M. Dress (1989). Weak hierarchies associated with similarity measures– an additive clustering technique. *Bull. Math. Biol.* **43,** 133-166.

[2] J. P. Barthélemy (1979). Caractérisations axiomatiques de la distance de la différence symetrique entre des relations binaires. *Math. Sci. hum.* **67,** 85-113.

[3] J. P. Barthélemy and M. F. Janowitz (1991). A formal theory of consensus. *SIAM J. Discrete Math.* **4,** 305-322.

[4] J. P. Barthélemy, B. Leclerc and B. Monjardet (1986). On the use of ordered sets in problems of comparison and consensus of classifications. *J. Classification* **3,** 187-224.

[5] J. P. Barthélemy and F. R. McMorris (1986). The median procedure for n-trees. *J. Classification* **3,** 329-334.

[6] J. P. Barthélemy and B. Monjardet (1981). The median procedure in cluster analysis and social choice theory. *Math. Social. Sci.* **1,** 235-268.

[7] J. P. Barthélemy and B. Monjardet (1988). The median procedure in data analysis: New results and open problems. In *Classification and Related Methods of Data Analysis*, Ed. H. H. Bock, 309-316. Elsevier Science, Amsterdam.

[8] A. Batbedat (1988). Les isomorphismes hts et hte. *Metron* **46,** 47-59.

[9] K. P. Bogart (1973). Preference structures I: Distances between transitive preference relations. *J. Math. Sociology* **3,** 49-67.

[10] K. P. Bogart (1975). Preference structures II: Distances between intransitive preference relations. *SIAM J. Applied Math.* **29,** 254-262.

[11] M. Condorcet (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris.

[12] W. D. Cook and M. Kress (1992). *Ordinal Information and Preference Structures: Decision Models and Applications.* Prentice Hall, Englewood Cliffs.

[13] A. D. Gordon (1977). L_1-norm and L_2-norm methodology in cluster analysis. *In this volume.*

[14] J. G. Kerneny and J. L. Snell (1962). *Mathematical Models in the Social Sciences.* Ginn and Co., NY.

[15] B. Leclerc (1993). Lattice valuations, medians and majorities. *Discrete Math.* **111,** 345-356.

[16] B. Leclerc (1994). Median for weight metrics in the covering graphs of semilattices. *Discrete Appl. Math.* **49,** 281-297.

[17] T. Margush (1982). Distances between trees. *Discrete Appl. Math.* **4,** 281-290.

[18] T. Margush and F. R. McMorris (1981). Consensus n-trees. *Bull. Math. Biol.* **43,** 239-244.

[19] F. R. McMorris (1990). The median procedure for n-trees as a maximum likelihood method. *J. Classification* **7,** 77-80.

[20] F. R. McMorris, H. M. Mulder and R. C. Powers (1997). The median

function on median graphs and semilattices. Submitted.

[21] F. R. McMorris, H. M. Mulder and F. S. Roberts (1997). The median procedure on median graphs. Submitted.

[22] F. R. McMorris and R. C. Powers (1991). Consensus weak hierarchies. *Bull. Math. Biol.* **53,** 679-684.

[23] F. R. McMorris and R. C. Powers (1995). The median procedure in a formal model of consensus. *SIAM J. Discrete Math.* **14,** 507-516.

[24] F. R. McMorris and R. C. Powers (1997). The median function on weak hierarchies. To appear in *Mathematical Hierarchies and Biology*, Eds. B. Mirkin et al.. DIMACS Series, AMS, Providence, R.I.

[25] P. B. Mirchandani and R. L. Francis Eds. (1990). *Discrete Location Theory.* Wiley, New York.

[26] B. Monjardet (1980). Théorie et application de la médiane dans les treilles distributifs finis. *Ann. Discrete Math.* **9,** 87-91.

[27] R. Vohra (1996). An axiomatic characterization of some locations in trees. *European J. Operational Res.* **90,** 78-84.

[28] H. P. Young (1988). Condorcet's theory of voting. *Am. Pol. Sci. Rev.* **82,** 1231-1244.

Student (1997), Vol.2, No.2, 203-213

Fitting L_2 Norm Classification Models to Complex Data Sets

Maurizio Vichi

University "G.D'Annunzio" di Chieti, Pescara, Italy

Abstract: In this paper methodologies for fitting classification models (dendrograms and partitions) to two and three-way arrays of dissimilarities minimizing a L_2 norm loss function are examined. A new algorithm for fitting several hierarchical classifications to quite large three-way arrays is also discussed.

Key words: Partition, dendrogram, L_2-norm, consensus classification, ultrametrics.

1 Fitting L_2 norm classification models to a dissimilarity matrix.

In this paper we provide an efficient solution to the algorithms for solving the L_2 norm problems of fitting classification models (dendrograms and partitions) to two and three-way arrays of dissimilarity data. A new algorithm for fitting several dendrograms to quite large three-way arrays is also introduced. For the three-way array of dissimilarities ${}^3\mathbf{D} \equiv \{\mathbf{D}_1, \mathbf{D}_2, ..., \mathbf{D}_r\}$ two alternative working hypotheses are postulated: i) A classification model applies with equal validity to dissimilarity matrices $\mathbf{D}_h$ $(h = 1, ..., r)$ forming the three-way array, since the $\mathbf{D}_h$ are the same except for errors and the classifications associated to each $\mathbf{D}_h$ (via a clustering algorithm) are replicates of each other, there being no systematic differences. In this situation a sole classification fits the three-way array of dissimilarities. ii) Some $\mathbf{D}_h$ differ in a systematic way, while some others are equal to the previous except for errors, so that more classification models of the same type are needed for fitting the three-way array of dissimilarities. The material described is based on a mathematical programming approach.

1.1 Fitting Hierarchical Partitions

The L_2-norm problem of fitting a dendrogram to dissimilarity data can be formally stated:

$$\begin{cases} \min \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (d_{ij} - u_{ij})^2 = \min \frac{1}{2} \|\mathbf{D} - \mathbf{U}\|_2^2 \\ \text{subject to} \\ u_{ij} \leq \max\{u_{ik}, u_{jk}\} \\ u_{ik} \leq \max\{u_{ij}, u_{jk}\} \; i = 1, ..., n-2; j = i+1, ..., n-1; k = j+1, ..., n \\ u_{jk} \leq \max\{u_{ik}, u_{ij}\} \end{cases} \tag{P1}$$

where $\mathbf{D} = [d_{ij}]$ and $\mathbf{U} = [u_{ij}]$ are the dissimilarity and ultrametric matrices and $\|\cdot\|_2$ is the L_2-norm. The optimal solution of (P1) represents the best L_2-norm ultrametric approximation of $\mathbf{D}$ and therefore the closest dendrogram to $\mathbf{D}$. (P1) is known to be NP-hard problem in the class of the NP-complete problems (Krivánek and Morávek 1986).

The constraints in (P1), define a finite ultrametric space with triplets of points forming acute isosceles triangles. The ultrametric conditions in (P1) can be rewritten permuting indices (i, j, k):

$$u_{ik} - u_{jk} = 0, \text{ for } u_{ij} \leq \min(u_{ik}, u_{jk}) \tag{2}$$
$$i = 1, ..., n-2; j = i+1, ..., n-1; k = j+1, ..., n.$$

These $O(n^3)$ constraints can be synthesised into a single one:

$$\sum_{\substack{i=1 \\ u_{ij} \leq \min(u_{ik}, u_{jk})}}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} (u_{ik} - u_{jk})^2 = 0 \tag{3}$$

that can substitute those in (P1). The objective function $(1/2)\ ||\mathbf{D} - \mathbf{U}||_2^2$ plus (3) has been reformulated (Carroll and Pruzansky 1980) as an unconstrained problem using a penalty function for the constraints and solving a sequence of parametrized unconstrained optimizations, using originally a gradient-based procedure, then the conjugate gradient (DeSoete 1984) and later the truncated-Newton or the quasi-Newton (Vichi, 1993), converging to the solution of (P1). The global solution cannot always be guaranteed since (P1) has a non convex feasible region.

Based on the analysis of several simulated and observed data, we found it more efficient to solve (P1), for small data sets ($n < 50$) directly using a sequential quadratic programming algorithm (SQP) (Powell 1983). The solution was always found with a superlinear rate of convergence. When

quadratic constraints are used (as in the formulation of Carroll and Pruzansky) the computational space complexity is reduced but the computational time complexity is usually greatly increased. For n larger than 50, the space complexity becomes quickly unfeasible. In this case instead of solving the $(1/2)\ ||\mathbf{D}-\mathbf{U}||_2^2$ plus (3) we found computationally more efficient to group the constraints (2) into more than one quadratic, and to solve the relative constraint problem with SQP. This last approach allows bounding correctly the feasible region of (P1) so that the SQP can take well informed decisions regarding directions of search and step length. According to our results on a battery of experiments we found it optimal, in terms of time complexity, to solve problems with a number of quadratic constraints between $O(n)$ and $O(n \log n)$. When a large number of quadratic constraints ($\geq O(n^2)$) is used the time increases drastically because the instability of the SQP algorithm increases.

Very recently we have been experimenting the possibility to solve (P1) by recognizing matrix $\mathbf{U}$ as ultrametric with $O(n^2 \log n)$ time and space instead of $O(n^3)$. This is accomplished using the recognition procedure suggested by Bandelt (1990), for sequential algorithms but not used for the solution of (P1) and by Dahlhaus (1993) for parallel algorithms. It consists in sorting u_{ij} in $O(n^2 \log n)$ with heap-sort; taking the minimum u_{ij} of the ordered list; verifying if $u_{ik} - u_{jk} = 0$, for all $k \in I - \{i, j\}$; deleting object i; and proceeding in the ordered list with the smallest u_{jk} $j, k \neq i$, until only two objects are left. Already for small data sets this procedure reduces significantly the time complexity as was observed in our experiments.

As an initial guess for (P1) the average linkage algorithm (UPGMA) solution can be used. An alternative is to compute : (i) the subdominant ultrametric (lagest lower bound) of the observed dissimilarity matrix (via the single linkage); (ii) the least upper bound ultrametric (via complete linkage); (iii) compute the mean matrix of these two ultrametrics; (iv) repeat (i), (ii), (iii) on the current mean matrix until convergence is achieved. Simeone and Vichi (1996), present an algorithm that performs a sequence of elementary tree operations so as to obtain at each step, in a greedy fashion, a dendrogram that is as close as possible, in a L_2-norm sense, to a single dissimilarity matrix or to a set of dissimilarity matrices (or dendrograms, this last case for consensus purposes).

1.2 Fitting partitions

Let us first state the bijection between partitions and ranked trees with two ranks, i.e., the bijection between partitions and ultrametric matrices with off-diagonal elements in $\{a, b\}$, with, say, $a = 1$ and $b = 2$ without loss of

generality. The proof is omitted for reasons of space. Using this bijection the problem of fitting a partition to $\mathbf{D}$ can be seen as the problem of fitting a ranked tree with only two ranks to $\mathbf{D}$. The mathematical formulation of this last problem is (P1) plus the further constraints

$$u_{ij} \in \{1,2\}\ i = 1,...,n-1; j = i+1,...,n \tag{4}$$

The optimal solution of (P1) plus (4) detects the ultrametric matrix with entries integers 1 or 2 closest to $\mathbf{D}$, i.e. the ranked tree with two ranks and therefore the partition of I closest to $\mathbf{D}$.

An initial feasible solution for (P1) plus (4) is given carrying out one of the procedures reviewed in the previous section or applying the UPGMA on the dissimilarity matrix $\mathbf{D}$ rescaled to interval [0,2] and taking the partition at level 1.5 of the dendrogram.

An integer linear programming co-occurrence formulation of (P1) plus (4) can be defined fixing $x_{ij} = 2 - u_{ij}$ for $i, j \in I$ and recalling that $d_{ij} \in [1,2]$. Thus, $x_{ij} = 1$ or 0, if object i and j respectively belong or do not belong to the same class. Hence, the objective function of (P1) can be rewritten :

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(d_{ij}-2)^2 + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} x_{ij}b_{ij}$$

where $b_{ij} = 2d_{ij} - 3$. Therefore, minimizing $(1/2)\ ||\mathbf{D} - \mathbf{U}||^2$ is equivalent to minimizing:

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} x_{ij}b_{ij}$$

The ultrametric constraints in (P1) can be rewritten as follows

$$\begin{aligned} u_{ij} &\leq \max\{u_{ik}, u_{jk}\} \leq u_{ik} + u_{jk} - 1; \\ u_{ik} &\leq \max\{u_{ij}, u_{jk}\} \leq u_{ij} + u_{jk} - 1; \\ u_{jk} &\leq \max\{u_{ik}, u_{ij}\} \leq u_{ik} + u_{ij} - 1. \end{aligned}$$

Passing to variables x_{ij} the objective function and constraints in (P1) become:

$$\begin{cases} \min \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} b_{ij}x_{ij} \\ \text{subject to} \\ x_{ij} \in \{0,1\}\ i = 1,...n-1; j = i+1,...,n \\ x_{ik} + x_{jk} - x_{ij} \leq 1 \\ x_{ij} + x_{jk} - x_{ik} \leq 1\ i = 1,...,n-2; j = i+1,...,n-1; k = j+1,...,n \\ x_{ik} + x_{ij} - x_{jk} \leq 1 \end{cases} \tag{P2}$$

(P2) is an integer linear programming problem also known as a clique partitioning problem; it has been mathematically formulated by Régnier (1965). An efficient algorithm for (P2), based on a cutting-plane algorithm, is described in Grötshel and Wakabayahi (1989). They reported that duality gap was not frequent and they could solve problems with a number of objects up to 158. For an initial feasible solution of (P2) the same initial guesses of problem (P1) plus (4) can be used. In particular Gordon and Vichi (1996) in their experiments, to compute partitions of a given set of partitions, found that UPGMA often provided an optimal solution to the clique partitioning problem.

2 Fitting L_2 norm classification models to three-way data sets of dissimilarities

Often the data set to analyse with classification techniques is a three-way array of dissimilarities ${}^3\mathbf{D}$ defined: (a) by evaluating in r different occasions the (dis)similarity between couples of n objects; or (b) by computing a prefixed measure of (dis)similarity between objects of a three-way data array. This last is formed by the observed values of a set of n objects, described by a set of k variables, in a set of r occasions. The situation (a) is frequent in psychometric studies, while (b) in economics and in social sciences. The array ${}^3\mathbf{D} \equiv \{\mathbf{D}_1, \mathbf{D}_2, ..., \mathbf{D}_r\}$ is here seen as a set of r dissimilarity matrices $\mathbf{D}_h$, each of dimension $(n \times n)$. We are interested in investigating the possible presence of classification models (partitions or dendrograms) among objects in different occasions (one-mode classification of ${}^3\mathbf{D}$). In this section we will consider only the classification of objects, supposing these are the same in different occasions.

Before proceeding to analyse the ${}^3\mathbf{D}$ data, we will consider two hypotheses, one alternative to the other:

i) A classification model applies with equal validity to all dissimilarity matrices of ${}^3\mathbf{D}$ because $\mathbf{D}_h$ $h = 1, ..., r$ are the same, except for errors (in measurements, sampling errors, etceteras). In other words the classifications models (partitions and dendrograms), expressed by ultrametric matrices $\mathbf{U}_h$, defined applying a clustering algorithm to each $\mathbf{D}_h$, are practically the same; they are replicates of each other, there being no systematic differences and a sole consensus classification $\mathbf{U}$ is sufficient to represent all the $\mathbf{U}_h$.

ii) More classification models of the same type apply to the set $\{\mathbf{D}_1, \mathbf{D}_2, ..., \mathbf{D}_r\}$. This is a relaxation of the first working hypothesis. Some $\mathbf{D}_h$ could be assumed to differ in a systematic way, while others to be equal to the previous except for errors. Thus, the implication is that different classes of

dissimilarity matrices are found in ${}^3\mathbf{D}$, with practically equal classifications within classes. Therefore, each g-th class is represented by a sole consensus classification $\widehat{\mathbf{U}}_g$ different for each class.

According to the two working hypotheses it is clear that we can fit a classification model *directly* on the ${}^3\mathbf{D} \equiv \{\mathbf{D}_1, \mathbf{D}_2, ..., \mathbf{D}_r\}$ dissimilarity data or *indirectly* on ${}^3\mathbf{U} \equiv \{\mathbf{U}_1, \mathbf{U}_2, ..., \mathbf{U}_r\}$, i.e. the ultrametric matrices obtained by a clustering algorithm applied on each $\mathbf{D}_h$. The methodology here described can be generally applied with equal validity to both fitting cases. Here, we will often consider the indirect fitting approach. It should be mentioned briefly that a possible classification approach for ${}^3\mathbf{D}$ data is to repeatedly apply a classification methodology to each dissimilarity matrix forming ${}^3\mathbf{D}$ and then to compare the r repeated analyses. This approach is very nonparsimonious because of the high number of parameters $(n \times n \times r)$. It implies that objects in different occasions have no shared (dis)similarity structure and the classification models $\mathbf{U}_h$ associated to $\mathbf{D}_h$ are presumed to be totally unrelated to each other in any fashion. Furthermore, the results of the classifications are difficult to interpret since researchers are faced with the task of having to compare r separate analyses.

2.1 Fitting a sole classification model to a three-way array of dissimilarities

A methodology to treat the first working hypothesis has to define an average classification representative of $\{\mathbf{U}_1, \mathbf{U}_2, ..., \mathbf{U}_r\}$. Consensus methods have this characteristic; the goal is to find replicated components, i.e. classes belonging to more than one classification. The *average dendrogram* (Lapointe and Cocumel, 1991, Vichi, 1993) is the best least squares (LS) approximation of the ultrametric matrix, obtained solving (P1) with $\mathbf{D} = \overline{\mathbf{U}} = (1/r)(\mathbf{U}_1 + \mathbf{U}_2 + ... + \mathbf{U}_r)$ and the algorithms discussed in Section 1.1. A different LS consensus is the best L_2-norm ultrametric convex combination $\mathbf{U} = [u_{ij}] = \sum_{j=1}^{r} a_j \mathbf{U}_j$, i.e. the weighted average dendrogram called *virtual dendrogram* (VD) (Vichi, 1995):

$$\begin{cases} \min \sum_{h=1}^{r} \|\mathbf{U}_h - \mathbf{U}\|^2 \\ \text{under the constraints} \\ \sum_{j=1}^{n} a_j = 1 \\ a_j \geq 0 \text{ for } j = 1, ..., r \\ \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{\substack{l=j+1 \\ u_{ij} \leq \min(u_{jl}, u_{il})}}^{n} (u_{il} - u_{lj})^2 = 0 \end{cases} \tag{P3}$$

A VD allows to evaluate the relative contribution of each $\mathbf{U}_j$ in determining

the weighted average dendrogram. The solution of (P3) is obtained with a SQP algorithm.

Carroll, Clark and DeSarbo (1984), propose *INdividual Differences in TREE Structures*, (INDTREES) a hierarchical clustering procedure, in which a unique type of trees for the r occasions is found, allowing to have different lengths of the branches between internal nodes of the trees.

Carroll and Arabie (1983) develop a non-hierarchical overlapping clustering method INDCLUS (*INdividual Differences CLUStering*) a clustering counterpart to INDSCAL. Just as INDSCAL assumes a common set of dimensions, for the r data sets, INDCLUS assumes a common set of clusters, possibly overlapping, whose numerical weights vary as a joint function of the cluster and source of data being considered (Arabie, Carroll, and DeSarbo 1987). The similarities are modelled as follows:

$$s_{ijk} \cong \sum_{l=1}^{c-1} w_{kl} m_{il} m_{jl} + a_k,$$

where $m_{il} = 1$ (resp., 0) if the i-th object belongs to the l-th class ($i = 1, ..., n$; $l = 1, ..., c-1$), w_{kl} is the weight which the k-th observer assigns to the l-th class ($k = 1, ..., m$; $l = 1, ..., c-1$), and a_k is an additive constant for the k-th observer (or the weight that observer assigns to a c-th class comprising the complete set of objects).

2.2 Fitting several classification models to the three-way array of dissimilarities

A methodology to treat the second working hypothesis can be developed following two approaches. The first is defined as *factorial classification of the three-way array of dissimilarities.* The aim is to identify a relatively small number of factorial classifications, i.e. not-directly-observable classification models explaining common information of $\{\mathbf{U}_1, \mathbf{U}_2, ..., \mathbf{U}_r\}$.

Principal Classification Analysis (Vichi, 1997) (PRINCLA), identifies c Principal Classifications (PCs), i.e., hierarchical classifications that are convex combinations of $(\mathbf{U}_1, \mathbf{U}_2, ..., \mathbf{U}_r)$. The o.f. to minimize is:

$$\sum_{g=1}^{c} \sum_{h=1}^{r} \left\| \mathbf{U}_h - \widehat{\mathbf{U}}_g \right\|_2^2 \tag{5}$$

where $\widehat{\mathbf{U}}_g$ is a virtual dendrogram, i.e., an ultrametric matrix as defined in (P3). The first PC is the convex combination that is the closest, in a L_2-norm sense, to $(\mathbf{U}_1, \mathbf{U}_2, ..., \mathbf{U}_r)$. The second PC is the next closest convex combination and it is "uncorrelated" with the first in the sense that

these PCs are defined by different original classifications only. Successive PCs are less close to $(\mathbf{U}_1, \mathbf{U}_2, ..., \mathbf{U}_r)$ and are uncorrelated with each other (Figure 1a). The constraints necessary to define uncorrelated PCs are:

$$\sum a_{hg} a_{hg'} = 0 \text{ for } g = 1, ..., m, \; g' = 1, ..., m \; (g \neq g'), \tag{6}$$

where a_{hg} and $a_{hg'}$, $h = 1, ..., r$ are the coefficients of the g-th and g'-th PCs. Therefore, for each g-th PC the relative contribution a_{hg} of each $\mathbf{U}_h$ in determining the PC can be evaluated. The methodology for computing the PCs is a SQP algorithm that was experimented successfully. In the SQP it was found useful to use the $O(n^2 \log n)$ procedure for recognising ultrametricity of PCs, since it reduced significantly the time complexity. In our experience each PC is generally defined by utrametric matrices having valued trees of the same form (equal n-tree) but different length of the branches. In this case any combination of ultrametrics is ultrametric and the problem of finding PCs reduces to detect the best coefficients a_{hg}. PRINCLA is appropriate when the working hypothesis ii) is really verified, so as in economic and social phenomena where the trend of the objects, in a short period of time, have often a smoothed shape. Hence, for each set of dissimilarity matrices relative to close time points, there is a PC averaging the classifications associated to the dissimilarities.

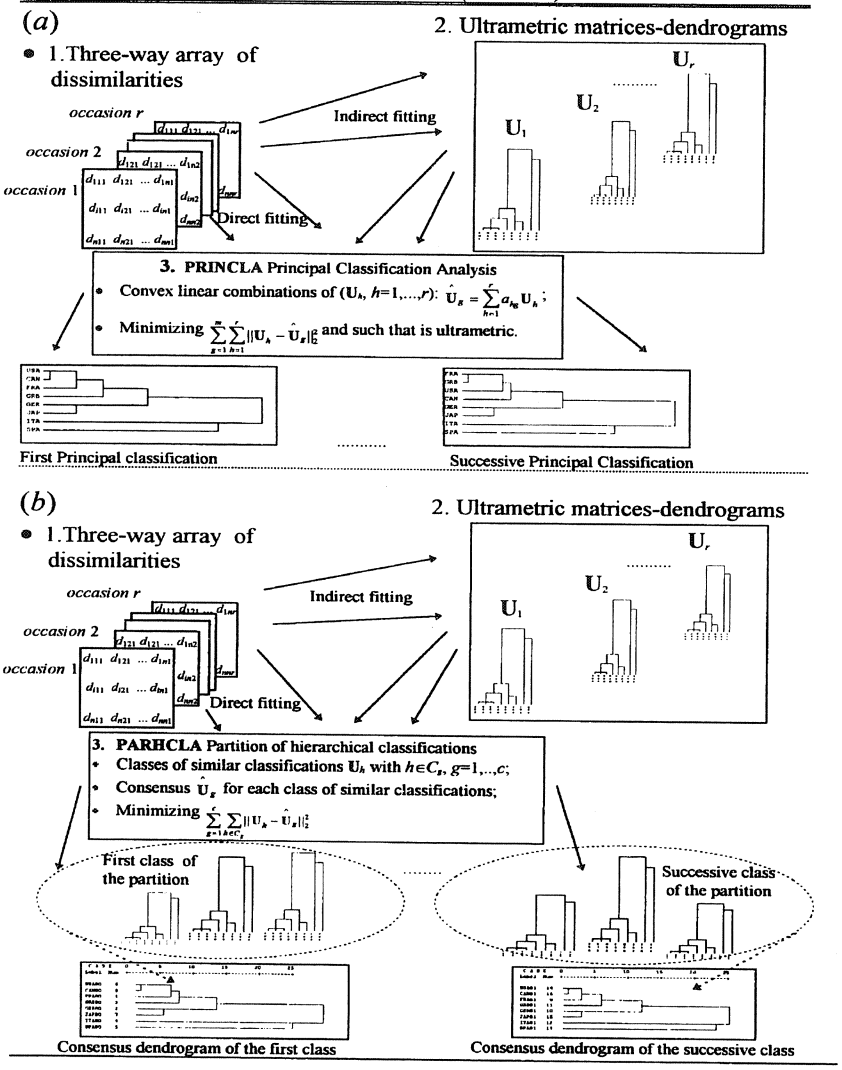
A second approach to treat the working hypothesis ii) can be defined as *partitions of classification models*. The goal is to detect a number c of classes of classification models (partitions, dendrograms) that form a partition of $\{\mathbf{U}_1, \mathbf{U}_2, ..., \mathbf{U}_r\}$, differently from the factorial approach where this last condition is not required. Therefore, we have a problem of classification, and in particular a partitioning problem, where the "objects" to classify are partitions or dendrograms, in general classification models of the same type, and for each class a consensus classification has to be detected. When the "objects" are dendrograms, the aim is to find a partition of the given set of dendrograms in c classes, with the property that the classifications belonging to a class are the closest (according to the L_2-norm loss function) to a consensus hierarchical classification that also is identified. The o. f. to minimize is:

$$\sum_{g=1}^{c} \sum_{h \in C_g} \left\| \mathbf{U}_h - \widehat{\mathbf{U}}_g \right\|_2^2 \tag{7}$$

where $\widehat{\mathbf{U}}_g$ is now an ultrametric matrix, different from a PC, representing the consensus hierarchical classification of the g-th class and $P_c \equiv \{C_1, C_2, ..., C_c\}$ is a partition of $\{1, 2, .., r\}$ identifying classes of similar dendrograms. Thus, a partition of similar hierarchical classifications and

a consensus classification for each class of the partition are simultaneously defined (Figure 1b). The algorithm used to minimize (7) is a specialised version of the SQP algorithm used also for detecting the PCs.

Figure 1: Analysis of a three-way array of dissimilarities: (*a*) Principal Classification Analysis (PRINCLA) ; (b) Partition of hierarchical classifications (PARHCLA).



A new heuristic algorithm for determining a partition of hierarchical classifications of quite a large three-way array is given by a recursive application of the two main stages of the algorithm. The first stage is based

on the application of UPGMA on the mean matrix of the $\mathbf{U}_h$ associated to each class C_g; and the second stage is an iterative relocation procedure based on the k-means algorithm, used for assigning each $\mathbf{U}_h$ to the closest $\widehat{\mathbf{U}}_g$ in a L_2-norm sense. When "objects" to classify are partitions the aim is to identify a (secondary) partition of a given set of (primary) partitions (Gordon and Vichi, 1996) with the property that primary partitions in the same class are perceived as similar to one another. The classes are obtained using either constrained or unconstrained clustering algorithms.

References

[1] Arabie, P., Carroll, J,D. and DeSarbo W. (1987). *Three-way Scaling and Clustering.* Sage Publications.

[2] Bandelt, H.J. (1990). Recognition of the tree metrics. *SIAM J. of Discrete Math.* **3,** 1-6.

[3] Carroll, J. D. and Arabie P. (1983). INDCLUS: An individual differences generalization of the ADCLUS model and MAPCLUS algorithm. *Psychometrika* **48,** 157-169.

[4] Carroll, J. D., Clark, L A., DeSarbo, W.S. (1984). The representation of three-way proximity data by single and multiple tree structure models. *J. of Classification* **1,** 24-74.

[5] Carroll, J. D. and Pruzansky, S. (1980). Discrete and hybrid scaling models. In *Similarity and Choice*, Eds. E. D. Lantermann and Feger. 108-139. Huber, Bern.

[6] Dahlhaus, E. (1993). Fast parallel recognition of ultrametrics and tree metrics. *SIAM J. of Discrete Math.* **6,** 523-532.

[7] DeSoete, G. (1984). A least squares algorithm for fitting an ultrametric tree to a dissimilarity matrix. *Pattern Recognition Letters* **2,** 133-137.

[8] Gordon, A. D. and Vichi, M. (1996). Partitions of partitions. Submitted for publication.

[9] Grötshel, M. and Wakabayahi, Y. (1989). A cutting plane algorithm for a clustering problem. *Mathematical Programming* **45,** 59-96.

[10] Kriványek, M. and Morávek, J. (1986). NP-Hard problems in hierarchical-tree clustering. *Acta Informatica* **23,** 311-323.

[11] Lapointe F. J. and Cucumel G. (1991). The average consensus. Abstract of the presentation at the IFCS-91.

[12] Powell, M.J.D. (1983). Variable metric methods for constrained optimization. In *Mathematical Programming: The State of the Art*, Eds. A. Bachem et al, 288-311. Springer Verlag.

[13] Régnier S. (1965). Sur quelques aspects mathématiques des problèmes

de classification automatique. I.C.C. Bulletin, Rome.

[14] Simeone, B. and Vichi, M. (1996). Consensus of hierarchical classifications. Paper of 5th IFCS Conference, JAP.

[15] Vichi M. (1993). Un algoritmo dei minimi quadrati per interpolare un insieme di classificazioni gerarchiche con una classificazione consenso. *Metron* **51,** 139-163.

[16] Vichi M. (1997). Principal classification analysis: a method for generating consensus dendrograms and its application to three-way data. To appear in *Comp. Statist. Data Anal.*

Student (1997), Vol.2, No.2, 215-216

From the Editor's Notebook

The Missing Diagonal

The center of a univariate data set $\{x_1, ..., x_n\}$ can be defined as a point μ that minimizes a norm of the vector of distances $\mathbf{y}' = (|x_1 - \mu|, ..., |x_n - \mu|)$. As the median and the mean are the minimizers of respectively the L_1- and the L_2-norm of $\mathbf{y}$, they are two alternatives to describe the center of a univariate data set.

The center $\boldsymbol{\mu} = (\mu_1, ..., \mu_p)$ of a multivariate data set $\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ with $\mathbf{x}_i = (x_{i1}, ..., x_{ip})$ for $i = 1, ..., n$ can also be defined as a minimizer of a norm of a vector of distances. In multivariate situations however, there are several kinds of distances. Let us consider the vector of L_1-distances $\mathbf{y}_1' = (||\mathbf{x}_1 - \boldsymbol{\mu}||_1, ..., ||\mathbf{x}_n - \boldsymbol{\mu}||_1)$ and the vector of L_2-distances $\mathbf{y}_2' = (||\mathbf{x}_1 - \boldsymbol{\mu}||_2, ..., ||\mathbf{x}_n - \boldsymbol{\mu}||_2)$. We define the L_1-median and the L_1-mean as the minimizers of respectively the L_1- and the L_2-norm of $\mathbf{y}_1$; and then the L_2-median and the L_2-mean as the minimizers of respectively the L_1- and the L_2-norm of $\mathbf{y}_2$. We have hence here two alternatives to define the median and two alternatives to define the mean of a multivariate data set.

Therefore the L_1-median is defined as the point $\boldsymbol{\mu}$ that minimizes

$$\|\mathbf{y}_1\|_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{p} |x_{ij} - \mu_j| \right| = \sum_{i=1}^{n} \sum_{j=1}^{p} |x_{ij} - \mu_j| , \tag{1}$$

the L_1-mean as the point $\boldsymbol{\mu}$ that minimizes

$$\|\mathbf{y}_1\|_2^2 = \sum_{i=1}^{n} \left(\sum_{j=1}^{p} |x_{ij} - \mu_j| \right)^2 , \tag{2}$$

the L_2-median as the point $\boldsymbol{\mu}$ that minimizes

$$\|\mathbf{y}_2\|_1 = \sum_{i=1}^{n} \left| \sqrt{\sum_{j=1}^{p} (x_{ij} - \mu_j)^2} \right| = \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{p} (x_{ij} - \mu_j)^2}, \tag{3}$$

and the L_2-mean as the point $\boldsymbol{\mu}$ that minimizes

$$\|\mathbf{y}_2\|_2^2 = \sum_{i=1}^{n} \sum_{j=1}^{p} (x_{ij} - \mu_j)^2 . \tag{4}$$

Whereas the L_1-median is the vector of coordinate medians, the L_2-mean is the vector of coordinate means and the L_2-median is the *spatial median* introduced by Gini and Galvani (1929), the L_1-mean appears to be a new concept (see Dodge and Rousson, 1997).

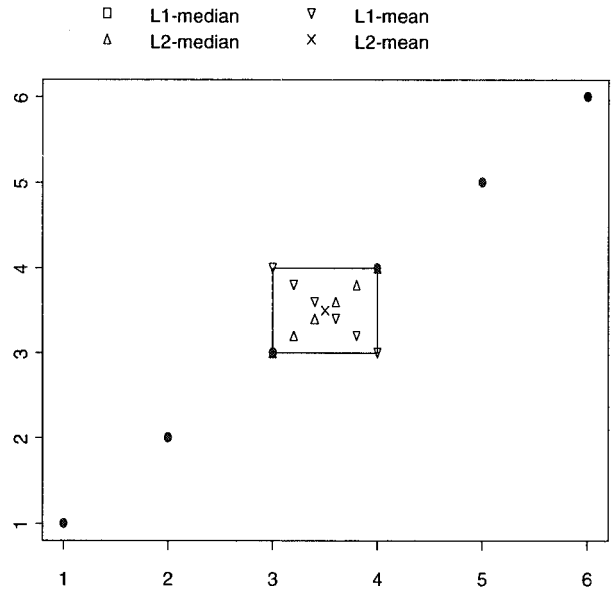


Figure 1: Four alternatives to describe the center of a bivariate data set.

If we consider a bivariate data set of six points (or of any even number of points) that are equaly spaced on a line with slope 1 (see Figure 1), and if we draw a square with the third point as bottom left corner and with the fourth point as top right corner, we have the following scheme: all points contained in the square are the L_1-medians of the data set; the center of the square is the L_2-mean; points that are lying on the first diagonal are the L_2-medians; and points that are lying on the second diagonal are the L_1-means (see Figure 1). Thus the introduction of the L_1-mean does not only fill a gap in the field of multivariate descriptive statistics, it also brings the missing diagonal in the square described here above.

References

[1] Dodge, Y. and Rousson, V. (1997). On medians and means of multivariate data sets. *IMS Bull.* **26,** No.2.

[2] Gini, C. and Galvani, L. (1929). Di talune estensioni dei concetti di media ai caratteri qualitativi. *Metron* **8**, 3-209.

Data & Statistics in *Student (1997), Vol.2, No.2, 219-227*

Aprotinin in Children Undergoing Cardiac Surgery for Cyanotic and Acyanotic Heart Disease

Ali Dodge M.D.

Cardiovascular Surgery Division, CHUV, Lausanne, Switzerland

Ekkehardt S. Altpeter M.D.

Department of Social- and Preventive Medicine, University of Bern, Switzerland

Abstract: In the present paper the authors investigate the therapeutic effect of aprotinin in surgery of cyanotic and acyanotic congenital heart diseases. The apparent lack of benefit, as well as the high cost and the potential for complications raises the question of its use for acyanotic patients. In cyanotic disease however, the authors believe the bleeding tendency along with longer cross-clamp, bypass and total intervention times justify the use of aprotinin in order to diminish loss of blood and the need for blood products.

Key words: Aprotinin, cyanotic heart disease, blood loss, optimal dose.

1 Introduction

Intra and postoperative blood loss during cardiac surgery with cardiopulmonary bypass is a major issue (see [5], [6], [7]). Of further concern are the particular problems in congenital heart surgery with multiple alterations in the coagulation mechanisms and specific disorders related to cyanotic heart diseases. Aprotinin, a serum proteinase inhibitor, has effectively been employed in adult cardiac surgery at empirically high doses in order to reduce postoperative blood loss and lessen transfusional need with encouraging results (see [1], [4], [9], [11]). However, after initial widespread enthusiasm,

the potential for systemic complications, such as vascular occlusion, renal insufficiency and allergic phenomena, as well as the elevated cost for high doses have motivated further investigation and trials (see [8], [10], [12]).

The purpose of the present paper is to study the therapeutic effect of aprotinin in surgery of cyanotic and acyanotic congenital heart disease on a variety of clinical and paraclinical variables including blood losses and need for blood transfusions.

2 Data and methods

Between 1991 and 1994, 139 children underwent surgery for congenital heart diseases with cardiopulmonary bypass (CPB) at the Division of Cardiovascular Surgery of the Centre Hospitalier Universitaire Vaudois in Lausanne, Switzerland.

Cyanotic heart disease was present in 45 cases. The remaining 94 cases were acyanotic. Previous palliative procedures needing redo surgery to obtain complete correction with residual left-to-right shunting were counted as acyanotic disease.

Since June 1991, aprotinin was routinely administered to 67 patients (35 cyanotic and 32 acyanotic). The thirteen cyanotic children who did not receive aprotinin were operated before the middle of 1991.

Surgery was performed with continuous CPB, using heparin (300 UI/kg) and a membrane oxygenator with flows of 2.4 l min^{-1} m^2. At discontinuation of bypass, heparin was reversed by protamine sulfate in a 1 to 0.8 ratio. When aprotinin was used, the mean dose of 19'750 Kallikrein Inactivator Units (KIU) per kg was added once to the priming solution before initiating bypass.

Blood samples taken one day preoperatively included hemoglobin, hematocrit, prothrombin time, thromboplastin time, platelet count, arterial oxygen saturation and fibrinogen. A second sample was taken upon arrival at the Surgical Intensive Care Unit (SICU) and included hemoglobin, hematocrit, prothrombin time, thromboplastin and thrombin times and platelet count. The final sample measured the same variables at 24 hours after surgery.

Chest drainage was measured hourly during the first 24 hours, beginning upon arrival at the SICU. No lower limit for hemoglobin concentration was defined to indicate the need for transfusion and this was evaluated according to each patient's clinical status. Expectedly, cyanotic children with chronically elevated hemoglobin values were transfused at higher levels than acyanotic children.

The patients were classified into four groups: cyanotic children with aprotinin (C/A, $n = 32$) or without aprotinin (C/NA, $n = 13$), and acyan-

otic children with aprotinin (NC/A, $n = 35$) or without aprotinin (NC/NA, $n = 59$).

Comparison between groups was done by calculating 95% confidence intervals for the average of the 12 following variables: age (years), preoperative blood sample (hemoglobin (mg/dl), % hematocrit), cross-clamp time (min), bypass time (min), total surgery time (min), heparin dose (mg/kg), protamine dose (mg/kg), blood loss at 6 and 24 hours (ml/kg) and finally the total blood transfusion (ml/kg) and the total fresh frozen plasma transfused (ml/kg)cumulated up to 24 hours after surgery (see data of Table 1).

The variable "total blood transfusion" is the sum of blood priming, blood transfused during bypass and blood transfused during the first 24 hours postoperatively. Likewise, the variable "total fresh frozen plasma (FFP) transfusion" is the sum of FFP during bypass and during the first 24 hours postoperatively.

The 95% confidence intervals were calculated by means of bootstrapping (see [3]), also called sampling-resampling, and adjustment of the level of significance $\alpha = 5\%$ by Bonferroni (α/k) for multiple comparison, where $k = 4$ is the number of groups.

3 Results

Mean age was 6.1 years (range: 0 to 16.3). The 139 cases included 6 redo operations in the cyanotic group and 13 in the acyanotic group (total redo operations, $n = 19$). Hospital mortality (under 30 days) was 9.3% ($n = 13$). In the cyanotic group, three cases needed reexploration for bleeding and surgical hemostasis, two of which were in the non-treated group (C/NA). In the acyanotic group, reexploration for bleeding was necessary in only one case in the non-treated group (NC/NA).

The effect of aprotinin upon intraoperative bleeding proved to be quite different between cyanotic and acyanotic groups (see Tables 2 and 3). In the former, blood loss at 6 and 24 hours postoperatively were 47.3% and 57.2% lesser in the treated group, clearly suggesting the influence of aprotinin in reducing blood loss. Correspondingly, total blood transfusion was 26.6% higher in the non-treated group. The difference was however not statistically significant. Total fresh frozen plasma transfusion was unexpectedly 17.8% higher in the treated group, despite shorter bypass time, but this was also not significant.

In the acyanotic group, aprotinin had no substantial therapeutic effect, as its expected benefit was not noticed with regards to blood loss. Correspondingly, no difference in the use of blood or FFP was observed. These results may be related to the low doses used in this study, the variance heterogeneity or the missing randomization.

Table 1: Data on 139 operated children. Variables measured are: 1=age (years), 2=Hb preop. (mg/dl), 3=Ht preop (%), 4=cross-clamp time (min), 5=bypass time (min), 6=total surgery time (min), 7=total heparin (mg/kg), 8=protamine (mg/kg), 9=blood loss 6h (ml/kg), 10=blood loss 24h (ml/kg), 11=total blood transfusion (ml/kg), 12=total FFP transfused (ml/kg).

group	1	2	3	4	5	6	7	8	9	10	11	12
C/A	4.1	169	56	72	119	240	5.30	3.03	10.83	19.27	53.03	0.00
C/A	4.0	170	66		53	155	4.37	3.88	3.40	10.72	33.98	29.13
C/A	5.2	173	60	27	50	140	4.44	4.03	17.74	20.71	46.77	28.23
C/A	9.5	176	62	75	113	235	5.63	3.75	26.13	34.70	43.75	12.50
C/A	4.5	168	57	32	67	165	4.21	3.97	13.97	30.48	27.78	27.78
C/A	6.9	190	60	29	133	280	5.00	3.75	4.65	7.68	0.00	0.00
C/A	7.1	135	43	104	126	340	6.15	3.85	28.67	31.02	42.56	26.67
C/A	1.9	184	58	73	118	225	6.50	6.00	10.00	22.80	30.00	0.00
C/A	0.0	190	60	110	166	270	14.00	8.00	18.00	53.76	282.00	26.00
C/A	0.4	194	61		103	200	5.42	0.00	8.67	12.40	21.67	8.33
C/A	8.5	136	40	58	89	220	4.32	3.95	5.58	11.37	13.16	15.79
C/A	11.8	163	51	65	101	250	5.17	3.45	7.79	16.97	10.69	0.00
C/A	13.3	157	45	86	159	300	4.67	3.92	6.14	15.25	9.04	0.00
C/A	6.9	175	53	76	117	240	4.57	4.31	4.57	13.40	10.15	17.26
C/A	11.1	207	61		64	240	4.39	3.75	2.64	0.00	0.00	0.00
C/A	16.2	160	47	87	120	230	6.05	5.23	6.33	29.80	22.21	6.98
C/A	4.4	187	57	71	140	350	6.50	4.78	7.52	21.40	59.87	22.93
C/A	2.8	192	64	73	101	200	6.00	4.00	8.50	17.28	43.00	2.80
C/A	5.8	179	54	80	125	225	4.86	4.29	3.94	10.70	0.00	0.00
C/A	8.6	149	45	52	161	265	11.94	5.56	8.22	17.47	30.00	47.22
C/A	2.1	141	41	91	130	246	6.09	6.09	6.09	12.10	46.96	6.96
C/A	2.0	163	49	64	115	220	5.08	5.08	10.74	29.11	36.07	28.69
C/A	1.4	180	56	76	147	230	3.88	3.40	7.18	15.84	48.54	4.85
C/A	0.6	182	57	84	138	245	5.00	5.00	16.50	31.20	56.25	22.50
C/A	0.0	150	44	41	144	290	9.09	7.58	4.55	20.36	124.24	18.18
C/A	0.9	156	46	60	120	195	4.69	4.32	10.99	17.19	43.21	0.00
C/A	6.4	175	52	39	75	200	4.44	4.44	12.33	19.87	41.67	13.89
C/A	2.3	194	61	130	250	460	9.30	8.72	22.33	31.81	53.02	75.12
C/A	13.6	217	72	60	118	240	4.93	12.84	12.36	27.51	25.07	8.66
C/A	0.5	161	52	67	140	235	5.56	4.76	18.73	49.90	47.62	19.05
C/A	0.3	180	59	25	61	215	7.45	5.32	11.06	24.51	68.09	0.00
C/A	5.1	199	62	63	100	215	5.00	4.84	20.56	43.35	49.19	22.58
C/NA	1.7	140	42	82	150	270	11.64	4.11	26.99	87.12	194.52	0.00
C/NA	4.7	197	60	86	120	245	6.54	6.54	17.66	33.64	33.64	2.80
C/NA	7.0	143	69	70	110	220	6.45	3.23	5.16	16.41	29.03	0.00
C/NA	8.7	203	61	91	120	250	4.14	3.86	63.27	72.00	90.91	27.27
C/NA	7.1	214	64	93	159	280	4.32	2.73	6.95	20.51	13.64	0.00
C/NA	1.9	156	44	54	81	170	4.74	3.02	7.67	17.79	59.66	0.00
C/NA	7.4	142	42	70	100	220	4.80	4.37	3.62	9.64	0.00	0.00
C/NA	4.1	159	48	64	120	210	5.77	4.55	8.39	19.30	101.75	25.17
C/NA	7.4	180	54	158	191	300	4.63	4.81	85.48	98.58	143.33	31.48
C/NA	0.6	203	65	96	127	230	5.10	3.06	12.65	38.69	0.00	0.00
C/NA	6.4	203	63	90	175	280	4.57	2.44	68.60	84.29	72.56	53.66
C/NA	1.4	212	67	57	107	180	4.31	3.02	7.59	19.24	10.34	4.31
C/NA	2.9	145	42	71	114	185	4.82	2.52	3.31	11.05	43.17	3.60
NC/A	1.8	105	30	45	75	180	5.96	5.13	14.49	23.38	71.79	0.00
NC/A	11.7	120	36	178	264	405	10.77	3.85	21.58	27.69	36.54	11.54
NC/A	5.0	147	44	74	127	240	6.09	6.09	68.09	76.80	85.22	0.00

group	1	2	3	4	5	6	7	8	9	10	11	12
NC/A	11.4	140	39	64	110	255	4.36	2.84	4.82	7.91	28.41	0.00
NC/A	8.4	139	42	74	110	250	5.00	5.00	6.55	11.76	15.00	0.00
NC/A	0.6	152	45	55	87	200	9.40	8.55	6.32	9.03	44.44	0.00
NC/A	3.1	132	39	77	100	240	4.29	2.86	4.57	11.66	27.14	0.00
NC/A	0.5	102	31	82	114	265	6.06	6.06	12.42	27.64	75.76	20.76
NC/A	3.2	138	39		143	240	4.69	3.45	5.52	10.26	57.93	17.24
NC/A	0.3	126	38	106	195	310	6.77	7.29	15.83	69.00	93.75	0.00
NC/A	2.2	118	37	13	29	130	5.50	5.00	9.30	19.68	50.00	0.00
NC/A	5.7	124	38	18	38	180	4.32	3.41	6.95	11.67	4.55	0.00
NC/A	9.1	118	36	28	41	120	5.56	4.44	3.24	8.64	13.33	0.00
NC/A	7.6	125	38	47	67	180	5.00	5.00	20.75	25.20	27.50	0.00
NC/A	0.9	134	41	57	85	190	5.67	5.67	14.95	23.75	90.52	0.00
NC/A	2.1	106	33	34	53	155	5.43	4.35	3.26	16.43	44.57	0.00
NC/A	5.9	132	39	29	45	190	4.71	4.71	2.41	11.29	17.65	0.00
NC/A	4.9	130	40	40	65	180	5.41	5.07	6.69	23.03	0.00	0.00
NC/A	14.3	127	39	45	119	220	6.34	5.22	2.54	3.58	0.00	0.00
NC/A	3.0	123	38	37	63	250	7.20	10.59	4.41	18.92	46.61	21.19
NC/A	3.0	123	38	37	63	250	7.20	10.59	4.41	18.92	46.61	0.00
NC/A	10.6	139	41	52	73	175	5.29	4.29	1.71	10.70	1.57	0.00
NC/A	9.7	109	34	35	60	160	4.00	3.00	2.84	6.72	16.00	0.00
NC/A	8.7	138	40	29	121	230	4.73	4.73	10.45	13.36	21.21	11.36
NC/A	3.8	119	37	20	53	155	5.07	4.35	9.35	15.48	7.25	14.49
NC/A	10.8	116	35	45	78	190	8.28	6.90	8.93	11.59	17.24	5.17
NC/A	3.3	133	40	77	112		7.78	6.94	3.56	9.47	25.00	0.00
NC/A	2.5	121	37	21	39	150	5.00	5.00	2.50	10.80	50.00	0.00
NC/A	2.5	129	40	41	59	180	5.61	5.61	10.92	30.86	51.02	0.00
NC/A	5.2	141	41	57	79	180	5.83	4.17	4.25	2.00	17.50	0.00
NC/A	4.8	126	38	42	75	205	7.78	7.41	6.15	15.64	33.33	0.00
NC/A	6.8	129	39	93	157	320	8.95	3.95	4.21	9.73	28.95	0.00
NC/A	7.1	164	50	46	103	185	5.79	3.16	7.37	14.27	53.68	0.00
NC/A	1.0	126	41	84	90	250	4.92	4.92	11.54	26.95	69.23	9.23
NC/A	3.6	135	40	55	92	190	4.64	3.57	2.50	20.57	35.71	0.00
NC/NA	16.3	134	40	30	62	155	4.75	4.39	4.22	9.49	0.00	0.00
NC/NA	1.2	107	32	34	82	205	5.67	4.12	3.61	12.62	28.87	0.00
NC/NA	12.9	99	31	60	100	180	5.50	4.17	8.23	19.84	25.00	8.33
NC/NA	6.1	118	36	53	82	195	4.86	4.32	9.03	15.70	10.81	0.00
NC/NA	12.9	130	39	48	99	220	5.42	4.22	2.94	9.43	2.41	0.00
NC/NA	2.7	106	33	49	71	185	6.00	6.00	7.10	15.36	60.00	0.00
NC/NA	14.3	102	30	20	45	160	4.35	3.24	4.32	9.41	10.81	0.00
NC/NA	13.2	131	39	80	110	215	6.55	6.03	6.24	12.25	0.00	0.00
NC/NA	5.4	135	41	81	116	215	9.58	8.33	9.25	28.00	41.67	30.83
NC/NA	1.0	100	30	25	46	180	5.08	5.08	13.05	25.22	61.69	13.56
NC/NA	15.5	119	40	63	75	235	6.35	4.05	6.62	10.05	14.86	0.00
NC/NA	15.8	126	38	44	80	165	4.27	3.66	1.88	5.74	9.76	0.00
NC/NA	1.7	136	41	25	74	160	4.27	2.44	8.17	16.98	68.29	0.00
NC/NA	3.8	132	39	17	33	165	4.13	3.97	5.95	15.43	0.00	0.00
NC/NA	5.0	125	38	40	60	150	4.67	3.33	3.93	9.76	30.00	0.00
NC/NA	6.8	133	39	35	82	140	4.08	2.65	2.45	0.00	6.12	0.00
NC/NA	9.3	126	37	39	64	150	4.81	3.70	2.56	4.80	0.00	0.00
NC/NA	2.9	120	36	19	75	165	5.05	4.04	6.67	15.27	64.65	0.00
NC/NA	5.1	136	40	15	28	105	4.21	2.63	0.74	6.32	21.05	0.00
NC/NA	5.3	90	27	15	27	135	4.30	4.30	4.46	11.35	32.26	0.00
NC/NA	8.6	121	36	76	125	280	5.25	6.00	7.00	16.08	11.00	12.00
NC/NA	12.1	144	41	43	60	155	4.69	3.47	2.88	9.83	0.00	0.00
NC/NA	11.1	142	41	42	69	190	5.97	3.29	2.10	8.49	6.17	0.00
NC/NA	3.2	130	38	20	47	175	4.74	4.31	6.29	12.62	51.72	0.00
NC/NA	6.0	120	38	65	93	215	6.03	2.84	7.45	16.85	60.28	8.51

group	1	2	3	4	5	6	7	8	9	10	11	12
NC/NA	6.6	131	40	25	42	130	5.18	5.88	3.41	7.62	0.00	0.00
NC/NA	5.9	141	41	34	62	165	4.62	4.10	3.03	6.77	10.26	0.00
NC/NA	0.4	101	29	23	52	155	4.55	9.09	3.45	15.71	94.55	0.00
NC/NA	5.4	149	44	40	70	190	5.06	4.21	2.42	10.79	0.00	0.00
NC/NA	5.6	128	37	18	54	140	4.29	4.00	3.77	12.62	17.14	0.00
NC/NA	2.9	116	36	41	77	200	4.42	2.65	6.11	25.70	53.10	0.00
NC/NA	7.2	123	38	28	49	150	4.44	4.67	4.77	16.82	14.02	0.00
NC/NA	14.5	127	40	48	88	180	5.70	4.64	10.21	15.02	23.87	0.00
NC/NA	13.2	128	38	24	57	135	4.38	3.13	3.25	10.58	7.81	0.00
NC/NA	16.0	159	47	38	70	165	4.32	3.07	7.37	17.56	6.67	8.77
NC/NA	13.0	133	40	39	49	120	5.42	5.42	3.25	10.46	5.92	0.00
NC/NA	10.4	121	36	40	83	260	6.20	5.00	4.04	15.17	21.40	0.00
NC/NA	1.0	106	33	55	80	190	8.59	5.63	15.92	34.14	80.28	0.00
NC/NA	0.5	138	42	70	128	300	8.65	6.73	6.54	38.77	96.15	24.04
NC/NA	7.3	134	39	71	116	235	5.00	5.00	16.12	28.80	20.80	13.00
NC/NA	7.7	127	39	29	57	130	4.16	3.60	4.94	14.23	63.95	0.00
NC/NA	5.6	134	41	31	60	140	4.33	4.00	8.60	18.40	45.33	0.00
NC/NA	5.7	159	48	57	109	215	4.54	2.31	19.77	36.18	76.92	0.00
NC/NA	10.3	119	36	38	77	200	4.15	2.93	5.95	12.53	21.95	0.00
NC/NA	2.5	138	41	46	84	170	7.08	6.25	4.58	4.00	46.67	0.00
NC/NA	12.4	131	42	74	135	240	5.00	4.14	6.03	20.52	31.03	6.90
NC/NA	3.1	134	40	37	69	150	4.90	3.85	4.33	17.54	14.42	0.00
NC/NA	9.0	141	44	147	245	360	5.50	4.69	3.56	8.48	7.81	9.38
NC/NA	6.9	99	29	55	90	180	6.15	5.38	24.54	33.60	46.15	26.92
NC/NA	4.2	99	30	37	65	195	5.00	4.50	5.10	19.92	30.00	0.00
NC/NA	5.9	110	35	47	52	180	6.98	4.19	11.06	19.31	53.07	0.00
NC/NA	5.3	130	38	33	69	170	4.39	4.17	6.67	12.53	42.22	18.89
NC/NA	1.0	124	36	40	60	180	7.34	11.72	18.91	35.63	103.91	0.00
NC/NA	8.8	138	40	19	35	145	4.29	3.57	0.86	6.43	0.00	0.00
NC/NA	4.4	122	36	67	120	240	5.94	4.55	15.80	38.43	51.05	6.99
NC/NA	5.3	145	45	55	110	325	5.59	3.53	7.94	7.48	29.41	0.00
NC/NA	11.4	130	39	55	85	320	4.36	3.85	9.36	18.77	0.00	7.69
NC/NA	3.7	132	38	20	43	150	4.19	4.19	4.84	14.55	19.35	0.00
NC/NA	12.5	149	44	78	116	230	5.00	5.00	14.92	44.60	29.75	41.67

Variable	C/A ($n = 32$)	C/NA ($n = 13$)
age (years)	4.8 (2.9 to 6.8)	4.2 (2.3 to 6.4)
Hb preop. (mg/dl)	174 (165 to 183)	176 (156 to 194)
Ht preop. (%)	55 (51 to 58)	55 (49 to 62)
cross-clamp time (min)	63.5 (47.6 to 76.7)	83.2 (68.5 to 102.8)
bypass time (min)	117 (101 to 134)	128 (109 to 150)
total surgery time (min)	243 (219 to 268)	233 (206 to 259)
total heparin (mg/kg)	6.2 (5.3 to 7.5)	5.7 (4.8 to 7.4)
protamine (mg/kg)	5.2 (4.4 to 6.3)	3.9 (3.2 to 4.9)
blood loss 6h (ml/kg)	11.5 (8.8 to 14.7)	24.3 (8.0 to 44.8)
blood loss 24h (ml/kg)	23.9 (18.6 to 30.4)	41.8 (22.8 to 66.3)
total blood transfusion (ml/kg)	46.2 (28.7 to 80.6)	62.9 (28.2 to 106.5)
total FFP transfused (ml/kg)	15.8 (9.1 to 25.6)	11.4 (1.2 to 24.1)

Table 2: Averages and 95% confidence intervals in the cyanotic group.

Variable	NC/A ($n = 35$)	NC/NA ($n = 59$)
age (years)	4.8 (3.4 to 6.4)	7.0 (5.5 to 8.3)
Hb preop. (mg/dl)	128 (123 to 134)	126 (122 to 131)
Ht preop. (%)	39 (37 to 40)	38 (37 to 39)
cross-clamp time (min)	52.4 (39.2 to 67.2)	44.0 (26.8 to 52.3)
bypass time (min)	90.4 (73.3 to 114)	77.1 (67.7 to 89.8)
total surgery time (min)	212 (193 to 237)	188 (173 to 206)
total heparin (mg/kg)	6.2 (5.5 to 7.0)	5.4 (5.0 to 5.8)
protamine (mg/kg)	5.5 (4.7 to 6.1)	4.6 (4.0 to 5.2)
blood loss 6h (ml/kg)	9.4 (6.0 to 16.3)	7.0 (5.5 to 8.8)
blood loss 24h (ml/kg)	20.0 (14.7 to 27.4)	17.1 (14.3 to 20.6)
total blood transfusion (ml/kg)	39.1 (28.8 to 51.3)	31.1 (21.3 to 41.0)
total FFP transfused (ml/kg)	3.3 (0.8 to 6.1)	4.0 (1.6 to 7.4)

Table 3: Averages and 95% confidence intervals in the acyanotic group.

4 Discussion

The reduction of postoperative blood loss by aprotinin and its optimal dose regimen remain controversial subjects in pediatric cardiac surgery (see [5], [2]).

In general practice, the use of aprotinin in congenital heart disease has so far been done by using the same doses per kilogram as in adult cardiac surgery. At our institution, the decision for low dose aprotinin (19'750 KIU/kg) was also derived from our adult dose regimen, being one single dose injected in the prime volume, ranging from 1 to 2 million KIU per patient. This is approximately one fifth that of the Royston regimen (see [11]). No reaction or illness with these low doses was noted during our study.

The authors believe that the low doses of aprotinin used in this study may have hidden its full beneficial effect. One double-blind randomized study at much higher doses has shown a relative effect of aprotinin in congenital heart disease, although these results were not significant and no distinction was made between cyanotic and acyanotic disease (see [8]).

Although careful surgical hemostasis remains the undeniable standard for achieving a dry blood field, aprotinin seems to be an appreciable asset in this regard. In our experience, children with cyanotic disease benefitted the most from aprotinin treatment, with substantial reduction in postoperative blood loss and the need for whole blood transfusion. This benefit has yet to be demonstrated in acyanotic disease, more specifically with the doses given at our institution. Considering the cost and the potential for systemic complications as well as the lack of evidence showing a noteworthy effect of aprotinin in acyanotic disease, the authors cannot recommend its use for

these patients.

A prospective-randomized study of cyanotic patients using 40'000 KIU/kg of aprotinin (twice the actual dose) could help to elucidate one of the more intriguing questions in pediatric cardiac surgery today, namely what is the optimal minimal dose of aprotinin to be used?

Address for correspondence: Ali Dodge M.D., RUSH-Presbyterian St.Luke's Medical Center, Department of Cardiovascular Thoracic Surgery, 1653 W. Congress Parkway, Chicago IL 60612, USA. Phone: ++1 312 942-6370. Fax: ++1 312 666-0242.

References

[1] Bidstrup B.P., Royston D. and Taylor K.M. (1989). Reduction in blood loss and blood use after cardiopulmonary bypass with high dose aprotinin (Trasylol). *J. Thorac. Cardiovasc. Surg.* **97,** 364-372.

[2] Boldt J., Zickmann, Schindler E., Welters A., Drapper G. and Hempelmann G. (1994). Influence of aprotinin on the thrombomodulin/protein C system in pediatric cardiac operations. *J. Thorac. Cardiovasc. Surg.* **107,** 1215-1221.

[3] Efron B. and Tibshirani R.J. (1993). *An Introduction to the Bootstrap.* Chapman & Hall, New York.

[4] Elliot M.J. and Allen A. (1990). Aprotinin in pediatric cardiac surgery. *Perfusion* **5,** 73-76.

[5] Evans D.A., Hodler R.L., Brawn W.J. and Sethia B. (1994). Postoperative blood loss following cardiopulmonary bypass in children. *Eur. J. Cardiothorac. Surg.* **8,** 25-29.

[6] Garcia H.V., Bufollo E., Nader H.B. and Dietrich C.P. (1994). ATP reduces blood loss produced by heparin in caridopulmonary bypass operations. *Ann. Thorac. Surg.* **57,** 956-959.

[7] Hashimoto K.M., Yamagashi M., Sasaki T., Nakano M. and Kurasawa H. (1994). Heparin and antithrombin III levels during cardiopulmonary bypass. Correlation with subclinical plasma coagulation. *Ann. Thorac. Surg.* **58,** 799–805.

[8] Heryknopf F., Luchese F., Pereira E., Kalil R., Prates P. and Nesralla I.A. (1994). Aprotinin in children undergoing correction of congenital heart defects. *J. Thorac. Cardiovasc. Surg.* **108,** 517-521.

[9] Kawasuji M.K., Ueyama K., Sakakibara N., Tedoriya T., Matsunaga Y., Misaki T. and Watanabe Y. (1993). Effects of low-dose aprotinin on coagulation and fibrinolysis in cardiopulmonary bypass. *Ann. Thorac. Surg.* **55,** 1205-1209.

[10] O'Regan D.J., Giannopoulos N., Mediratta N., Kendall S.W.H., Forni A., Pillai R. and Westaby S. (1994). Topical aprotinin in cardiac operations. *Ann. Thorac. Surg.* **58,** 778-781.

[11] Royston D. (1992). High-dose aprotinin therapy. A review of the first five years experience. *J. Cardiothorac. Vasc. Anesthesia* **6,** 76-100.

[12] Westaby S., Forni A., Dunning J., Giannopoulos N., O'Regan D., Drossos G. and Pillai R. (1994). Aprotinin and bleeding in profoundly hypothermic perfusion. *Eur. J. Cardiothorac. Surg.* **8,** 82-86.