Proceedings of International Conference on Theory and Applications of Mathematics and Informatics – ICTAMI 2003, Alba Iulia

ON THE LEAST SQUARES FITTING IN A LINEAR MODEL

by
Nicoleta Breaz

Abstract. In this paper, we present a linearisable regressional model for which we obtain a full rank case theorem for uniquely fitting written in terms of initial matrix of sample data.The model considered here can be reduced to the linear one either by substitution or by written in other form.

INTRODUCTION

In [1] we also consider a linearisable regressional model.There, the explicative variables from the linearised model was $Z_k = X_k X_{k+1}, \forall k = \overline{1, p-1}$.In this paper, the linarisable regressional model considered will be reduced to a linear model with the explicative variables of the form $Z_k = X_k + X_{k+1}, \forall k = \overline{1, p-1}$.
As introduction we remember some classical notions and results from linear regression.

Definition 1
Let be Y a variable which depends on some factors exprimed by others p variables $X_1, X_2,..., X_p$. The regression is a search method for dependence of variable Y on variables $X_1, X_2,..., X_p$ and consists in determination of a functional connection f such that

$$Y = f(X_1, X_2,..., X_p) + \varepsilon \qquad (1)$$

where ε is a random term (error) which include all factors that can not be quantificated by f and which satisfies the conditions:
i) $E(\varepsilon) = 0$
ii) $Var(\varepsilon)$ has a small value

Formula (1) with conditions i) and ii) is called regressional model, variable Y is called the endogene variable and variables $X_1, X_2,..., X_p$ are called the exogen(explicative) variables.

Definition 2
The regression given by the following function is called a parametric regression

$$f(X_1, X_2,..., X_p) = f(X_1, X_2,..., X_p; \alpha_1, \alpha_2,..., \alpha_p).$$

Otherwise the regression is called a nonparametric regression.
The regression given by the following function is called a linear regression

$$f\left(X_1, X_2, ..., X_p; \alpha_1, \alpha_2, ..., \alpha_p\right) = \sum_{k=1}^{p} \alpha_k X_k .$$

Remark 3
If function f from regressional models is linear with respect to the parameters $\alpha_1, \alpha_2, ..., \alpha_p$, that is

$$f\left(X_1, X_2, ..., X_p; \alpha_1, \alpha_2, ..., \alpha_p\right) = \sum_{k=1}^{p} \alpha_k \varphi_k \left(X_1, X_2, ..., X_p\right)$$

than regression can be reduced to the linear one.

Definition 4
It is called the linear regressional model beetwen variable Y and variables $X_1, X_2, ..., X_p$, the model

$$Y = \sum_{k=1}^{p} \alpha_k X_k + \varepsilon \qquad (2)$$

Remark 5
The liniar regression problem consists in study of the variable Y behavior whit respect to the factors $X_1, X_2, ..., X_p$, the study made by " evaluation" of the regressional parameters $\alpha_1, \alpha_2, ..., \alpha_p$ and random term ε.
Let be considered a sample of n data

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad x = \left(x_1, ..., x_p\right) = \begin{pmatrix} x_{11} & x_{12} & ... & x_{1p} \\ x_{21} & x_{22} & ... & x_{2p} \\ ... & ... & ... & ... \\ x_{n1} & x_{n2} & ... & x_{np} \end{pmatrix}, n >> p$$

Than one can make the problem of evaluation for regressional parameters $\alpha^T = \left(\alpha_1, \alpha_2, ..., \alpha_p\right) \in \mathbb{R}^p$ and for error term $\varepsilon^T = \left(\varepsilon_1, \varepsilon_2, ..., \varepsilon_n\right) \in \mathbb{R}^n$, from these data.

From this point of views the fitting of the theoretical model can offering solutions. Matriceal, the model (2) can be written in form

$$y = x\alpha + \varepsilon \qquad (2')$$

and represent the linear regressional theoretical model.
By fitting this models using a condition of minimum results the fitted model

$$y = xa + e \qquad (2'')$$

where $a^T \in \mathbb{R}^p, e^T \in \mathbb{R}^n$.
It is desirable that residues $e_1, e_2, ..., e_n$ to be minimal. Then can be realised using the least squares criteria.

Definition 6
It is called the least squares fitting, the fitting which corresponds to the solutions (*a*,*e*) of the sistem $y = xa + e$, which minimise the expression

$$e^T e = \sum_{k=1}^{n} e_k^2$$

Theorem 7(full rank case)
If $rank(x) = p$ then the fitting solution by least squares criteria is uniquely given by

$$a = \left(x^T x\right)^{-1} x^T y .$$

MAIN RESULTS

In this paper we consider the folowing model

$$Y = \alpha_1(X_1 + X_2) + \alpha_2(X_2 + X_3) + ... + \alpha_{p-1}(X_{p-1} + X_p) + \varepsilon \text{ , } p \geq 2 . \qquad (3)$$

For a sample of *n* data/variables we will use the notations

$y^T = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$, $\alpha^T = (\alpha_1, \alpha_2, ..., \alpha_{p-1}) \in \mathbb{R}^{p-1}$,

$$\varepsilon^T = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n) \in \mathbb{R}^n, x = (x_1, x_2, ..., x_p) \in M_{n,p} \text{ , } x = \begin{pmatrix} x_{11} & x_{12} & ... & x_{1p} \\ x_{21} & x_{22} & ... & x_{2p} \\ ... & ... & ... & ... \\ x_{n1} & x_{n2} & ... & x_{np} \end{pmatrix}.$$

From (3) we obtain a linear model either by eliminating the brackets or by using the substitution $X_k + X_{k+1} = Z_k, \forall k = \overline{1, p-1}$.

In the first case we have a linear model with p explicative variables, $X_1, X_2, ..., X_p$ and with some constraints on the coefficients. So, the model can be written as

$$Y = \alpha_1 X_1 + (\alpha_1 + \alpha_2) X_2 + (\alpha_2 + \alpha_3) X_3 + ... + (\alpha_{p-2} + \alpha_{p-1}) X_{p-1} + \alpha_{p-1} X_p + \varepsilon .$$

The estimate $a, a = (a_1, a_2, ..., a_{p-1})^T$ of the parameter α , $\alpha = (\alpha_1, \alpha_2, ..., \alpha_{p-1})^T$ is obtained in this case, uniquely, by the least squares criteria if $rank(x) = p$. The estimate will has the form

$$a^T = (b_1, b_2 - b_1, b_3 - b_2 + b_1, ..., b_{p-1} + b_{p-2} + ... + (-1)^p b_1, b_p) \in \mathbb{R}^{p-1}$$

where $b = (b_1, b_2, ..., b_p)^T = (x^T x)^{-1} x^T y \in \mathbb{R}^p$ is the estimate of β ,

$$\beta = (\beta_1, \beta_2, ..., \beta_p)^T = (\alpha_1, \alpha_1 + \alpha_2, ..., \alpha_{p-2} + \alpha_{p-1}, \alpha_p)^T .$$

In the second case we have a linear model with p-1 explicative variables $Z_1, Z_2, ..., Z_{p-1}$ obtained with the above mentioned substitution.So, the model (3) becomes

$$Y = \alpha_1 Z_1 + \alpha_2 Z_2 + ... + \alpha_{p-1} Z_{p-1} + \varepsilon .$$

If we use the sample notations plus the new matrix, $z \in M_{n,p-1}$,

$$z = (z_1, z_2, ..., z_{p-1}) = \begin{pmatrix} x_{11} + x_{12} & x_{12} + x_{13} & ... & x_{1p-1} + x_{1p} \\ x_{21} + x_{22} & x_{22} + x_{23} & ... & x_{2p-1} + x_{2p} \\ ... & ... & ... & ... \\ x_{n1} + x_{n2} & x_{n2} + x_{n3} & ... & x_{np-1} + x_{np} \end{pmatrix},$$

we obtain the matriceal form, $y = z\alpha + \varepsilon$.

Thus, by linearising, we obtain a similar model with model treated by us in [1].For such models, according to theorem 7, the condition $rank(z) = p-1$ is required for obtaining the least squares estimate a of α .

Corollary 8
If $rank(z) = p-1$ then the least squares estimate of α is given by

$$a = (a_1, a_2, ..., a_{p-1})^T = (z^T z)^{-1} z^T y \quad .$$

In that following, we try to formulate such conditions in term of the initial matrix *x*.

Proposition 9
We have $rank\,(x) = p \Rightarrow rank\,(z) = p-1$.

Proof
We have $rank\,(x) = rank(v)$ where $v \in M_{n,p}$,

$$v = (x_1 + x_2, x_2 + x_3, ..., x_{p-1} + x_p, x_p) = \begin{pmatrix} x_{11} + x_{12} & x_{12} + x_{13} & ... & x_{1p-1} + x_{1p} & x_{1p} \\ x_{21} + x_{22} & x_{22} + x_{23} & ... & x_{2p-1} + x_{2p} & x_{2p} \\ ... & ... & ... & ... & ... \\ x_{n1} + x_{n2} & x_{n2} + x_{n3} & ... & x_{n1p-1} + x_{np} & x_{np} \end{pmatrix}.$$

Then we have $rank(v) = p$ and further, the matrix *v* has a minor of *p* order, different from zero. Without restricting the generality, let be this minor, the minor *d*, constructed with the first *p* rows. If we develop this minor in respect with the last column, we obtain $d = \sum_{k=1}^{p} (-1)^{k+p} x_{kp} d_k$, where $d_1, d_2, ..., d_p$ are the determinants of *p*-1 order that are implied by the developing.

If $d_k = 0, \forall k = \overline{1,p}$ then it results $d = 0$ that is false. Then it results that there exists a $d_k, d_k \neq 0$ and in fact, this is a minor of *p*-1 order in matrix $z \in M_{n,p-1}$. So, we have $rank\,(z) = p-1$.

Now, according to corollary 8 , the least squares estimate exists if $rank\,(x) = p$ .If we take into account, the form of the model, we can obtain the form of the estimate depending on the initial matrix *x*. For this, we state the folowing result that is obviously after some calculation.

Proposition 10
The next equality holds: $z = xA$, $A \in M_{p,p-1}$,

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & ... & 0 & 0 \\ 1 & 1 & 0 & 0 & ... & 0 & 0 \\ 0 & 1 & 1 & 0 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & ... & 1 & 1 \\ 0 & 0 & 0 & 0 & ... & 0 & 1 \end{pmatrix}.$$

The next theorem is a consequence of the corollary 8,the proposition 9 and the proposition 10.

Theorem 11
If $rank(x) = p$ then the least squares estimate is uniquely given by

$$a = \left(a_1, a_2, ..., a_{p-1}\right)^T = \left(A^T x^T xA\right)^{-1} A^T x^T y .$$

In that following, we present another condition for the existence of the least squares estimate *a*. In [2] we prove the following result:

Lemma 12
Let be $x \in M_{p,p}, x = \left(x_{ij}\right)_{1\le i, j\le p}$ and $z \in M_{p-1,p-1}, z = \left(z_{ij}\right)_{1\le i, j\le p-1}$ with $z_{ij} = x_{ij} + x_{ij+1}, \quad \forall i = \overline{1, p-1}, j = \overline{1, p-1}.$ Then $\det(z) = d_1 + d_2 + ... + d_p$ where d_i is the minor of $p-1$ order from the matrix *X*, obtained by elimination of $p-$th row and the $i-$th column, $i = \overline{1, p}$.
Now we can prove the following corollary:

Corollary 13
Let be $x \in M_{n,p}, z \in M_{n,p-1}, z_{ij} = x_{ij} + x_{ij+1}, \forall i = \overline{1,n}, j = \overline{1, p-1}$ and $x^{kj} \in M_{p-1,p-1}$ the matrices obtained from the first *p*-1 rows of x , $k = \overline{1, C_n^{p-1}}$ and by elimination of *j*-th coloumn. We denote $d_{kj} = \det\left(x_{kj}\right)$. If there exists at least one k_0 such that $\sum_{j=1}^{p} d_{k_0 j} \neq 0$ then $rank\,(z) = p-1$.

Proof

Let be k_0 such that $\sum_{j=1}^{p} d_{k_0 j} \neq 0$. We consider in the matrix *z* the minor of $p-1$ order which contains the same rows (as number of order) with the matrix $X_{k_0 j}$ and we denote this with d_z. According to lemma 12, we have $\det z = d_{k_0 1} + d_{k_0 2} + ... + d_{k_0 p} \neq 0$ and further $rank(z) = p-1$.

The next theorem will be a consequence of the corollary 8, the propositon 10 and he corollary 13.

Theorem 14

Under the hypothesis of corollary 13, the model (3) has the least squares fitted coefficients uniquely given by

$$a = (a_1, a_2, ..., a_{p-1})^T = (A^T x^T xA)^{-1} A^T x^T y .$$

Remarks 15

i) The condition given in corollary 13 implies that $rank(x) = p-1$.This last condition is not sufficient for the existence of the least squares estimate because doesn't imply $rank(z) = p-1$.

ii) If $d_{k_0 j}, j = \overline{1, p}$ defined in corollary 13 have the same signe(+ or -) then $rank(x) = p$ imples the condition from corollary 13.Then we can state that, in particular case when the sample data are such that the minors $d_{kj}, j = \overline{1, p}$ have the same signe (for each one value of *k*), the corollary 13 gives an less restrictive condition than the condition from the theorem 11.In the same time, the condition from corollary 13 is beter than the condition $rank(x) = p-1$ which is not a sufficient one. So the condition from the corollary 13 is an intermediary condition between " *p*-1 vectors from $x_k, k = \overline{1, p}$ are linearly independent ($rank(x) = p-1$)" and "the vectors $x_k, k = \overline{1, p}$ are linearly independent ($rank(x) = p$)".

REFERENCES

1. Breaz N.,Breaz D.-The full rank case for a linearisable model, Acta Universitatis Apulensis, Mathematics-Informatics,no3/2002
2. Breaz N., Mera S.-Asupra unor matrici, Prietenii matematicii,no 3/2002
3. Saporta G.-Probabilités, analyse des données et statistique, Editions Technip, Paris, 1990.
4. Stapleton J.H.-Linear Statistical Models, A Willey-Interscience Publication, Series in Probability and Statistics, New York,1995.

Author:

Nicoleta Breaz, "1 Decembrie 1918" University of Alba lulia, Romania., Computer Science and Mathematics Department, e-mail: nbreaz@uab.ro