THE FULL RANK CASE FOR A LINEARISABLE MODEL

by
Nicoleta Breaz, Daniel Breaz

Abstract. We consider a model which can be reduced to a linear one by substitution. For this model, we obtain a full rank case theorem for uniquely fitting written in terms of initial matrix of sample data.

Definition 1
Let be Y a variable which depends on influence of some factors expressed by other 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 consist in determination of a functional connection f such that

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

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

Formula (1) with conditions a) and b) is called regressional model, variable Y is called the endogene variable and variables $X_1, X_2,..., X_p$ are called the exogene variables.

Definition 2
The next regression is called a parametric regression

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

Otherwise the regression is called a nonparametric regression.
The regression bellow 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 linear one.

Definition 4
It is called the linear regressional model between 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 variable Y behavior whit respect to the factors $X_1, X_2, ..., X_p$ the study made by " evaluation" of regressional parameters $\alpha_1, \alpha_2, ..., \alpha_p$ and random term ε .

Let be considered a sample of *n* data

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

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

From this point of views the fitting of theoretic 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 minim 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 system $y = xa + e$, which minimise the expression

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

Theorem 7.(full rank case)

If $rang(x) = p$ then the fitting solution by least squares criteria is uniquely given by

$$a = (x^T x)^{-1} x^T y$$

Remark 8.

In this paper we consider the next model

$$y = \alpha_1 x_1 x_2 + \alpha_2 x_2 x_3 + ... + \alpha_{p-1} x_{p-1} x_p + \varepsilon \qquad (3)$$

where $y^T = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$, $\alpha^T = (\alpha_1, \alpha_2, ..., \alpha_{p-1}) \in \mathbb{R}^{p-1}$, $\varepsilon^T = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n) \in \mathbb{R}^n$ and x is the sample data/sample variables matrix

$$x \in M_{n.p}, \ x = (x_1, x_2, ..., x_p) = \begin{pmatrix} x_{11} & x_{12} & ... & x_{1p-1} & x_{1p} \\ x_{21} & x_{22} & ... & x_{2p-1} & x_{2p} \\ ... & ... & ... & ... & ... \\ x_{n1} & x_{n2} & ... & x_{np-1} & x_{np} \end{pmatrix}$$

With substitutions $x_i x_{i+1} = z_i, \forall i = \overline{1, p-1}$, we obtain the new matrix,

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

and the linear model $y = \alpha_1 z_1 + \alpha_2 z_2 + ... + \alpha_{p-1} z_{p-1} + \varepsilon$, which after the least squares fitting becomes $y = a_1 z_1 + a_2 z_2 + ... + a_{p-1} z_{p-1} + e$, where $a^T = (a_1, a_2, ..., a_{p-1}) \in \mathbb{R}^{p-1}$, $e^T = (e_1, e_2, ..., e_n) \in \mathbb{R}^n$. The fitting uniquely solutions results from $rang(z) = p-1$ (see theorem 7).

The purpose of our paper is to give an analogouse theorem based on initial sample variables.

Theorem 9.

If $rang(x) = p$, $p \in \{2,3\}$, and if the sample data are not nulls than uniquely exist the least squares fitting solution $a = (x_*^T x_*)^{-1} x_*^T y$ where $x_* = (x_1 * x_2, ..., x_{p-1} * x_p)$ and $x_i * x_j$ is the natural product between the vectors x_i and x_j that is the vector which can be obtained by multiplications one components of the two vectors.

Proof

Case $p = 2$:

The sample matrix is $x \in M_{n,2}$, $x = (x_1, x_2)$, $x = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ ... & ... \\ x_{n1} & x_{n2} \end{pmatrix}$ and the model

$y = \alpha_1 x_1 x_2 + \varepsilon$.

We make the substitution $x_1 x_2 = z$, which results $y = \alpha_1 z + \varepsilon$.

Because the sample data are not nulls it results $rang(z)=1$, where $z=\begin{pmatrix} x_{11}\cdot x_{12} \\ x_{21}\cdot x_{22} \\ \dots \\ x_{n1}\cdot x_{n2} \end{pmatrix}$.

One can observe that is sufficient that for a single unit of sample, data must be differed from zero. According to theorem 7 if $rang(z)=1$ then uniquely exist $a_1 = a = \left(z^T z\right)^{-1} z^T y = \left(x_*^T x_*\right)^{-1} x_*^T y$ where $x_* = (x_1 * x_2) \in M_{n,1}$.

Case $p=3$:

The sample matrix is $x \in M_{n,3}$, $x=(x_1, x_2, x_3)$, $x=\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \dots & \dots & \dots \\ x_{n1} & x_{n2} & x_{n3} \end{pmatrix}$

and the model

$y=\alpha_1 x_1 x_2 + \alpha_2 x_2 x_3 + \varepsilon$. We use the substitutions $x_1 x_2 = z_1$, $x_2 x_3 = z_2$ and we obtain $y=\alpha_1 z_1 + \alpha_2 z_2 + \varepsilon$. If $rang(\mathrm{x})=3$ then results that at least one minor of three order is differed from zero and let be this one (without restrict the generality) $d=\begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{vmatrix}$.

If $d_3 \neq 0$, developing by the second column results that at least one of the three minor from the development is not null and let be, by example, this one $d_2 = \begin{vmatrix} x_{11} & x_{13} \\ x_{21} & x_{23} \end{vmatrix} \neq 0$. On the other way we calculate a minor of two order from *z*, by example

$$d_{2_z} = \begin{vmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{vmatrix} = z_{11}z_{22} - z_{12}z_{21} = x_{11}x_{12}x_{22}x_{23} - x_{12}x_{13}x_{21}x_{22} = x_{12}x_{22}\left(x_{11}x_{23} - x_{13}x_{21}\right) =$$
$$= x_{12}x_{22}\cdot d_2$$

Because from the hypothesis, the sample data are not nulls and $rang(x)=3$, $d_2 \neq 0$ then results that $d_{2_z} \neq 0$, so $rang\,(\mathrm{z})=2$. According to theorem 7 it results that uniquely exists the solution $a=\left(z^T z\right)^{-1} z^T y = \left(x_*^T x_*\right)^{-1} x_*^T y$ where $x_* = (x_1 * x_2, x_2 * x_3)$.

Remark 10.

A weak condition, namely $rang(x) = p-1, p \in \{2,3\}$ is not sufficient because this is not implied that $rang(z) = p-1$. However, it can be given an intermediary condition between $rang\,(x) = p-1$ and $rang\,(x) = p$.

Theorem 11.

If in sample data matrix there exists a minor of second order differed from zero, at least, which not contains elements from second column, then $a = \left(x_*^T x_*\right)^{-1} x_*^T y$ with $x_* = (x_1 * x_2, x_2 * x_3)$.

Proof

By example $d_2 = \begin{vmatrix} x_{11} & x_{13} \\ x_{21} & x_{23} \end{vmatrix} = x_{11}x_{23} - x_{13}x_{21} \neq 0$ and

$$d_{2_z} = \begin{vmatrix} x_{11}x_{12} & x_{12}x_{13} \\ x_{21}x_{22} & x_{22}x_{23} \end{vmatrix} = x_{12}x_{22}\left(x_{11}x_{23} - x_{13}x_{21}\right) \neq 0$$

Remark 12.

These theorems can not be generalised for any p. A similar theorem with 11 can be given in genarall case if we define the pseudominor in follow sense.

Definition 13.

Let be $x \in M_{n,p}$. We call the pseudo-minor of $p-1$ order from matrix x, formed by the first $p-1$ rows of x, the expression

$$d^f = \sum_{\sigma \in S_{p-1}} (-1)^{sign\sigma} \left(x_{1\sigma(1)} \cdot x_{2\sigma(2)} \cdot \ldots \cdot x_{p-1\sigma(p-1)}\right)\left(x_{1\sigma(1)+1} \cdot x_{2\sigma(2)+1} \cdot \ldots \cdot x_{p-1\sigma(p-1)+1}\right) =$$

$$= \sum_{\sigma \in S_{p-1}} \left[(-1)^{sign\sigma} \cdot \prod_{i=1}^{p-1} x_{i\sigma(i)} \cdot \prod_{i=1}^{p-1} x_{i\sigma(i)+1} \right]$$

Remark 14.

In calculus of d we apply the ordinary formulla for a determinant of $p-1$ order only that the elements from products appearing in determinant are product of elements which are from minor of $p-1$ obtained by elimination of the first column and from minor of $p-1$ order obtained by elimination of last column.

$$d_3^f = \left\|\begin{matrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \end{matrix}\right\| = (x_{11}x_{22}x_{33})(x_{12}x_{23}x_{34}) + (x_{13}x_{21}x_{32})(x_{14}x_{22}x_{33}) +$$

$$+ (x_{31}x_{12}x_{23})(x_{32}x_{13}x_{24}) - (x_{13}x_{22}x_{31})(x_{14}x_{23}x_{32}) - (x_{11}x_{23}x_{32})(x_{12}x_{24}x_{33}) -$$

$$- (x_{33}x_{12}x_{21})(x_{34}x_{13}x_{22})$$

By example, d_3^f is the pseudo-minor of $p-1$ order from matrix $x \in M_{n,p}(p=4)$, formed with the first three rows of this matrix.

Theorem 15.

If in sample data matrix there exist at least one pseudo-minor of $p-1$ order different from zero then exists uniquely fitting solution $a=\left(x_*^T x_*\right)^{-1} x_*^T y$ with $x_* = \left(x_1 * x_2, x_2 * x_3, ..., x_{p-1} * x_p\right)$.

Proof

Let be the pseudo-minor of $p-1$ order, different from zero, formed with the first $p-1$ rows, without restrict the generality $\left(d_{p-1}^f \neq 0\right)$.

With substitutions $x_j x_{j+1} = z_j, \forall j = \overline{1, p-1}$, we obtain matrix $z = \left(z_{ij}\right)_{\substack{1 \le i \le n \\ 1 \le j \le p-1}}, z_{ij} = x_{ij} \cdot x_{ij+1}$

We calculate the minor of $p-1$ from z formed with the first $p-1$ rows:

$$d_z = \sum_{\sigma \in S_{p-1}} (-1)^{sign\sigma} \left(z_{1\sigma(1)} \cdot z_{2\sigma(2)} \cdot ... \cdot z_{p-1\sigma(p-1)}\right) =$$

$$= \sum_{\sigma \in S_{p-1}} (-1)^{sign\sigma} \left(x_{1\sigma(1)} \cdot x_{1\sigma(1)+1}\right)\left(x_{2\sigma(2)} \cdot x_{2\sigma(2)+1}\right) \cdot ... \cdot \left(x_{p-1\sigma(p-1)} \cdot x_{p-1\sigma(p-1)+1}\right) = d_{p-1}^f \neq 0$$

So $rang\ z = p-1$ and $a = \left(z^T z\right)^{-1} z^T y = \left(x_*^T x_*\right)^{-1} x_*^T y$.

REFERENCES

[1] Stapleton J.H.-Linear Statistical Models, A Willey-Interscience Publication, Series in Probability and Statistics, New York,1995.

[2] Saporta G.-Probabilités, analyse des données et statistique, Editions Technip, Paris, 1990.

Authors:

Nicoleta Breaz- „1 Decembrie 1918" University of Alba Iulia, str. N. Iorga, No. 13, Departament of Mathematics and Computer Science, email: nbreaz@lmm.uab.ro

Daniel Breaz- „1 Decembrie 1918" University of Alba Iulia, str. N. Iorga, No. 13, Departament of Mathematics and Computer Science, email: dbreaz@lmm.uab.ro