MULTIPLE LINEAR REGRESSION WITH CONSTRAINED COEFFICIENTS: APPLICATION OF THE LAGRANGE MULTIPLIER

C. R. Kikawa [1]
Department of Mathematics and Statistics, Tshwane University of Technology
e-mail: *richard.kikawa@gmail.com*

P. H. Kloppers
Department of Mathematics and Statistics, Tshwane University of Technology

Key words: Design matrix, Goal function, Lagrange multiplier, Linear systems, Regression coefficients, Schur complement, System blocking.

Abstract: In this paper, we present two unfamiliar novel estimation techniques (UNET) for the constrained regression coefficients in the frame-work of a standard multiple linear regression model. Estimation of a linear regression problem with constraints on the regression coefficients are firstly derived by minimising a formulated goal function that minimises the total sum of the squared errors, plus the sum of the linear constraints multiplied by a Lagrangian. We also show that the solution to the system of equations can be obtained without differentiating the goal function, rather expressed in terms of the known matrices. This is achieved by employing properties of a blocked linear system. The UNET is justified by a numerical simulated system of linear equations in 3-dimensions. The UNET yields estimates that are comparable to those generated by the Schur complement principle.

1. Introduction

Regression analysis is a statistical tool used to investigate relationships between variables (Legendre, 1805). Multiple linear regression (Draper and Smith, 1998) is an approach used to determine the linear relationship that exists between a set of regression variables (x_i) and a given response variable (y):

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_N X_N + \varepsilon.$$

The regression coefficients β_i's provide an indication of the effect of the regressors on the outcome variable. Parameter estimation methods for regression models were first published by Legendre (1805) in form of least squares. Currently, a number of techniques for estimating the regression coefficients of a linear problem are well known and for computer enhanced software to do this are readily available (Draper and Smith, 1998). As a result multiple regression is currently being applied

[1] Corresponding author.

AMS: 62J05, 62P99, 65C20

to a wider domain of challenges faced in both theoretical and practical arrays (Pazzani and Bay, 2001). The ultimate goal of this application is to obtain information from data, which information is to guide in decision making and planning purposes.

In practice, multiple regression can produce models that are unreliable to experts due to a confusion on the sign of the regression coefficients (Mullet, 1976). However, this problem is usually addressed by introducing a constrained form of regression that can produce more acceptable results (Pazzani and Bay, 2001). However, in the presence of constrained coefficients the estimation procedures are not straight forward as in the unconstrained model.

In this study, the estimation approaches of a linear regression model with constrained coefficients are firstly derived by considering minimisation of an objective function that minimises the total sum of the squared errors plus the sum of the linear constants multiplied by a Lagrange multiplier (Heath, 2005; Lasdon, 2002; Sokolnikoff and Redheffer, 1966). Secondly, using the properties of blocked linear systems (Chen, Anderson, Deistler and Filler, 2012; Alanelli and Hadjidimos, 2004; Axelsson, 1985) a second method estimation approach is derived. In this, the design or influence matrix (Cardini, 2013; Searle, 1971) of this system is a natural blocked matrix and this property is used to solve the system. In this solution it is not necessary to differentiate the formulated objective function, but the solution can be expressed from the constrained part using the known matrices.

We briefly discuss the concepts of blocked linear systems and the solutions there of presented in the first section of the paper. We also present a revised solution of a multiple linear regression system.

2. Blocked system of Linear equation

Structures of linear equations with circulant coefficient matrices can be found in a number of applications. For example, in finite difference approximations to elliptic functions subject to boundary initial value conditions (Chan and Chan, 1992; Wood, 1971) and in estimations of periodic systems by employing the splines method (Ahlberg, Nilson and Walsh, 1967; Zavyalov, Kvasov and Miroshnichenko, 1980). For problems involving multidimensional arrays, the coefficients of the design matrices have a block circulant arrangement (Mingkui, 1987).

Definition 1 We define

$$Y_\delta = \begin{bmatrix} Y_\delta \\ \vdots \\ Y_{\delta+T-1} \end{bmatrix},$$

$$Z_\delta = \begin{bmatrix} Z_\delta \\ \vdots \\ Z_{\delta+T-1} \end{bmatrix}$$

$\delta = 0, T, 2T, \ldots,$
where $Y_\delta \in \Re^p$ are the explained variables and $Z_\delta \in \Re^m$ are the explanatory variables.

Then, a blocked system of linear equations can be defined by

$$x_{\delta+T} = A_b x_\delta + B_b z_\delta$$

$$Y_{\delta+T} = C_b x_\delta + D_b z_\delta$$

where $x_\delta \in \Re^n$ is the state variable, $A_b = A^T$, $B_b = [A^{T-1}BA^{T-2}B\ldots B]$,

$$C_b = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^T \\ CA^{T-1} \end{bmatrix},$$

$$D_b = \begin{bmatrix} C & 0 & \ldots & 0 \\ CB & D & \ldots & 0 \\ \vdots & \vdots & \vdots & \\ CA^{T-1}B & CA^{T-2}B & \ldots & D \end{bmatrix}.$$

We define an operator P such that it satisfies $Px_\delta = x_{\delta+T}, PY_\delta = Y_{\delta+T}, PU_\delta = U_{\delta+T}$. In this case the transfer function of the blocked linear system of equations, see Definition 1 (Chen et al., 2012; Callier and Desoer, 1994) is given by

$$W(P) = D_b + C_b(PI - A_b)^{-1}B_b.$$

Assumption 1: The explained vector has a dimension (P) greater or equal to that of the explanatory vector (m), that is, $(p \geq m)$ and the normal rank $W(P)$ is m.

We investigate some properties of the blocked linear system and use them to solve a linear system of equations.

3. Existing results

In this section, the equivalents of some already available results for the linear systems are re-visited.

Consider the system of linear equations:

$$\mathbf{Ax} = \mathbf{y}$$

with matrix $\mathbf{A}_{n\times n}$ and vectors $\mathbf{x}_{n\times 1}$ and $\mathbf{y}_{n\times 1}$. Suppose that the system can be blocked into

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \tag{1}$$

Assume that the dimensions of A_1 are $r \times r$ and for vectors $\mathbf{x}_1$ and $\mathbf{y}_1$the dimensions are $r \times 1$ and taking the dimensions of all the other components to be in line with this. If $\mathbf{A}$ is non-singular (i.e. invertible) and

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}^{-1} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix},$$

the solution to system (1) is given by

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

$$= \begin{bmatrix} B_1 y_1 + B_2 y_2 \\ B_3 y_1 + B_4 y_2 \end{bmatrix}.$$

Furthermore, it can be shown that

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}^{-1} = \begin{bmatrix} A_1^{-1} + A_1^{-1} A_2 S^{-1} A_3 A_1^{-1} & -A_1^{-1} A_2 S^{-1} \\ -S^{-1} A_3 A_1^{-1} & S^{-1} \end{bmatrix},$$

with $S = A_4 - A_3 A_1^{-1} A_2$, known as the Schur complement (Zhang, 2005; Zong, 2009; Boyd and Vandenberghe, 2004) of A_1. The solution of the system (1) now becomes

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} A_1^{-1}\mathbf{y}_1 + A_1^{-1} A_2 S^{-1} A_3 A_1^{-1}\mathbf{y}_1 - A_1^{-1} A_2 S^{-1}\mathbf{y}_2 \\ S^{-1}\mathbf{y}_2 - S^{-1} A_3 A_1^{-1}\mathbf{y}_1 \end{bmatrix}. \tag{2}$$

4. Multiple Linear Regression

The solution of a multivariate multi-response linear regression system of equations in which the design matrices are non-singular (Searle, 1971) is now revisited. For simplicity we only consider a system with three explanatory variables.

Consider a dataset

$$\left[\begin{array}{c|ccc} u_1 & x_1 & y_1 & z_1 \\ u_2 & x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots & \vdots \\ u_n & x_n & y_n & z_n \end{array}\right] = \left[\mathbf{u} \mid \mathbf{W}\right],$$

where $\mathbf{u}$ is a vector of the explained variables and $\mathbf{W}$ the set of explanatory variables. It is known that for the model

$$u_j = \beta_1 x_j + \beta_2 y_j + \beta_3 z_j + \varepsilon_j. \tag{3}$$

The least squares estimation for $\boldsymbol{\upsilon}^\top = (\beta_1, \beta_2, \beta_3)$ that minimizes the objective function

$$\sum_{j=1}^{n} \varepsilon_j^2 = \sum_{j=1}^{n} \left(u_j - \beta_1 x_j - \beta_2 y_j - \beta_3 z_j\right)^2,$$

is

$$\widehat{\boldsymbol{\upsilon}} = \left(\mathbf{W}^\top \mathbf{W}\right)^{-1} \tilde{\mathbf{u}}, \tag{4}$$

where

$$\left(\mathbf{W}^\top \mathbf{W}\right) = \begin{bmatrix} \sum x_j^2 & \sum x_j y_j & \sum x_j z_j \\ \sum x_j y_j & \sum y_j^2 & \sum y_j z_j \\ \sum x_j z_j & \sum y_j z_j & \sum z_j^2 \end{bmatrix}, \tag{5}$$

and

$$\tilde{\mathbf{u}} = \left(\mathbf{W}^{\top}\mathbf{u}\right) = \begin{bmatrix} \sum x_j u_j \\ \sum y_j u_j \\ \sum z_j u_j \end{bmatrix}.$$

Hence the solution $\widehat{\boldsymbol{\upsilon}}$ will exist provided that $\mathbf{W}$ is of *full column rank*, that is, the columns of the explanatory variables are linearly independent, (Searle, 1971; Wardlaw, 2005; Basilevsky, 1983). This is a typical multiple regression problem and any statistical software package could be used to solve the system.

5. New results

5.1. Multiple Linear Regression with constraints

In this section, we consider the derivation of a solution of a multivariate model with constrained explanatory variables. We take the classical linear model,

$$\mathbf{u} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $\mathbf{u} = [u_1 \dots u_n]^{\top}$ is the vector, and $\mathbf{X}$ is an $n \times (n > p)$ influence or design matrix (Rodriguez-Yam, Davis and Scharf, 2004) assumed to be of *full column rank*, the elements in the vector of errors $\boldsymbol{\varepsilon}$ are assumed to be independent and normally distributed $N(0, \sigma^2)$, and $\boldsymbol{\beta}$ is a vector of coefficients of the explanatory variables.

Now due to physical conditions on the system, the components of $\boldsymbol{\beta}$, are required to be linearly constrained, that is $\mathbf{C}\boldsymbol{\beta} = \mathbf{0}$, where $\mathbf{C}$ is the so called constraint matrix. Now the proposed estimation of such constrained models are derived using two techniques.

5.1.1. The Differentiation Technique

Again consider a dataset

$$u_j = \beta_1 x_j + \beta_2 y_j + \beta_3 z_j + \varepsilon_j.$$

Let us assume that, due to physical conditions, two constraints $\alpha_1\beta_1 + \alpha_2\beta_2 = 0$ and $\theta_1\beta_2 + \theta_2\beta_3 = 0$ with $\alpha, \theta \in R$, are to be imposed on the model. The objective function to estimate υ now is

$$\Im = \frac{1}{2}\sum(u_j - \beta_1 x_j - \beta_2 y_j - \beta_3 z_j)^2 + \phi_1(\alpha_1\beta_1 + \alpha_2\beta_2) + \phi_2(\theta_1\beta_2 + \theta_2\beta_3) \tag{6}$$

with ϕ_1 and ϕ_2 introduced as the Lagrange multiplier (Carpenter, 2005; Sokolnikoff and Redheffer, 1966) . Differentiating (6) with respect to $\beta_1, \beta_2, \beta_3, \phi_1$ and ϕ_2 and equating each resultant equation to zero, we obtain

$$
\begin{aligned}
\frac{\partial \Im}{\partial \beta_1} &= \sum (u_j - \beta_1 x_j - \beta_2 y_j - \beta_3 z_j)(-x_j) + \phi_1 \alpha_1 = 0, \\
\frac{\partial \Im}{\partial \beta_2} &= \sum (u_j - \beta_1 x_j - \beta_2 y_j - \beta_3 z_j)(-y_j) + \phi_1 \alpha_1 + \phi_2 \theta_1 = 0, \\
\frac{\partial \Im}{\partial \beta_3} &= \sum (u_j - \beta_1 x_j - \beta_2 y_j - \beta_3 z_j)(-z_j) + \phi_2 \theta_2 = 0, \\
\frac{\partial \Im}{\partial \phi_1} &= \alpha_1 \beta_1 + \alpha_2 \beta_2 = 0, \\
\frac{\partial \Im}{\partial \phi_2} &= \theta_1 \beta_2 + \theta_2 \beta_3 = 0.
\end{aligned}
$$

These equations of the objective function can be expressed in matrix notation as

$$
\begin{bmatrix}
\sum x_j^2 & \sum x_j y_j & \sum x_j z_j & \alpha_1 & 0 \\
\sum x_j y_j & \sum y_j^2 & \sum y_j z_j & \alpha_2 & \theta_1 \\
\sum x_j z_j & \sum y_j z_j & \sum z_j^2 & 0 & \theta_2 \\
\alpha_1 & \alpha_2 & 0 & 0 & 0 \\
0 & \theta_1 & \theta_2 & 0 & 0
\end{bmatrix}
\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \phi_1 \\ \phi_2 \end{bmatrix}
=
\begin{bmatrix} \sum x_i u_j \\ \sum y_j u_j \\ \sum z_j u_j \\ 0 \\ 0 \end{bmatrix}. \tag{7}
$$

We note that the 3 by 3 matrix appearing in the upper left side of the first matrix in (7) is the same as (5) for the unconstrained system.

The constrained part of the objective function (6), $\phi_1(\alpha_1\beta_1 + \alpha_2\beta_2) + \phi_2(\theta_1\beta_2 + \theta_2\beta_3$, can be written as

$$
\begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \end{bmatrix}
\begin{bmatrix} \alpha_1 & 0 \\ \alpha_2 & \theta_1 \\ 0 & \theta_2 \end{bmatrix}
\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}
\equiv \boldsymbol{\upsilon}^\top \mathbf{C} \boldsymbol{\phi},
$$

with $\mathbf{C}$ the constrained matrix.

From this, it is observed that system (7) is determined only by $\mathbf{W}^\top\mathbf{W}$ of (5) and $\mathbf{C}$, and hence can be obtained without differentiating the objective function (6).

The solution for the unknown parameters is

$$
\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\phi}_1 \\ \hat{\phi}_2 \end{bmatrix}
=
\begin{bmatrix}
\sum x_j^2 & \sum x_j y_j & \sum x_j z_j & \alpha_1 & 0 \\
\sum x_j y_j & \sum y_j^2 & \sum y_j z_j & \alpha_2 & \theta_1 \\
\sum x_j z_j & \sum y_j z_j & \sum z_j^2 & 0 & \theta_2 \\
\alpha_1 & \alpha_2 & 0 & 0 & 0 \\
0 & \theta_1 & \theta_2 & 0 & 0
\end{bmatrix}^{-1}
\begin{bmatrix} \sum x_i u_j \\ \sum y_j u_j \\ \sum z_j u_j \\ 0 \\ 0 \end{bmatrix}. \tag{8}
$$

In this case a 5×5 matrix must be inverted to obtain the solution.

Although a system with only three explanatory variables was investigated, the above results can easily be generalised to higher order systems with more than two constraints.

5.1.2. The Block System Technique

Considerable work has been done on the estimation of constrained parameters of a linear system using the blocking technique (Chen et al., 2012; Zong, 2009). However, in this section we present a modified block system technique involving the Lagrange multiplier (Sawyer, 2002).

From (7) it is easy to see that the system can be blocked into

$$\begin{bmatrix} \mathbf{W}^\top\mathbf{W} & \mathbf{C} \\ \mathbf{C}^\top & \underline{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\upsilon} \\ \boldsymbol{\phi} \end{bmatrix} = \begin{bmatrix} \mathbf{u} \\ \mathbf{0} \end{bmatrix},$$

with $\underline{0}$ the 2×2 zero matrix and $\mathbf{0}$ the 2×1 zero vector. By (2) it then follows that, with $S = -\mathbf{C}^\top(\mathbf{W}^\top\mathbf{W})^{-1}\mathbf{C}$ the estimations for $\boldsymbol{\upsilon}$ and $\boldsymbol{\phi}$ are:

$$\begin{bmatrix} \widehat{\widehat{\boldsymbol{\upsilon}}} \\ \widehat{\boldsymbol{\phi}} \end{bmatrix} = \begin{bmatrix} (\mathbf{W}^\top\mathbf{W})^{-1}\tilde{\mathbf{u}} + (\mathbf{W}^\top\mathbf{W})^{-1}\mathbf{C}S^{-1}\mathbf{C}^\top(\mathbf{W}^\top\mathbf{W})^{-1}\tilde{\mathbf{u}} \\ -S^{-1}\mathbf{C}^\top(\mathbf{W}^\top\mathbf{W})^{-1}\tilde{\mathbf{u}} \end{bmatrix}.$$

In this case, the estimation for $\boldsymbol{\upsilon}$ is denoted by $\widehat{\widehat{\boldsymbol{\upsilon}}}$ to distinguish it from $\widehat{\boldsymbol{\upsilon}}$ in (4). This solution can be simplified to

$$\widehat{\boldsymbol{\phi}} = -S^{-1}\mathbf{C}^\top(\mathbf{W}^\top\mathbf{W})^{-1}\tilde{\mathbf{u}}, \tag{9}$$

$$\begin{aligned} \widehat{\widehat{\boldsymbol{\upsilon}}} &= (\mathbf{W}^\top\mathbf{W})^{-1}\tilde{\mathbf{u}} + (\mathbf{W}^\top\mathbf{W})^{-1}\mathbf{C}S^{-1}\mathbf{C}^\top(\mathbf{W}^\top\mathbf{W})^{-1}\tilde{\mathbf{u}} \\ &= (\mathbf{W}^\top\mathbf{W})^{-1}\tilde{\mathbf{u}} - (\mathbf{W}^\top\mathbf{W})^{-1}\mathbf{C}\widehat{\boldsymbol{\phi}}, \\ &= \widehat{\boldsymbol{\upsilon}} - (\mathbf{W}^\top\mathbf{W})^{-1}\mathbf{C}\widehat{\boldsymbol{\phi}}, \end{aligned} \tag{10}$$

with $\widehat{\boldsymbol{\upsilon}}$ the solution to the unconstrained system given in (4).

6. Numerical application

In this section, we present a numerical study to validate the performance of the proposed derivations.

Consider a dataset,

$$\begin{bmatrix} \mathbf{u} \,\Big|\, \mathbf{x} & \mathbf{y} & \mathbf{z} \end{bmatrix} = \left[\begin{array}{c|ccc} 20.6 & 1 & 11 & 5 \\ 37.5 & 5 & 15 & 2 \\ 51.8 & 8 & 18 & 5 \\ 35.8 & 1 & 23 & 7 \\ 98.9 & 15 & 39 & 10 \\ 115.2 & 19 & 42 & 12 \\ 131.6 & 21 & 51 & 13 \\ 157.1 & 25 & 61 & 16 \\ 190.5 & 31 & 75 & 20 \\ 214.5 & 35 & 85 & 23 \end{array}\right]$$

Firstly suppose model (3) must be fitted to the data. In this case

$$\mathbf{W}^\top\mathbf{W} = \begin{bmatrix} 3929 & 9532 & 2538 \\ 9532 & 23656 & 6324 \\ 2538 & 6324 & 1701 \end{bmatrix}, \tag{11}$$

$$\tilde{\mathbf{u}} = \mathbf{W}^\top\mathbf{u} = \begin{bmatrix} 24434.7 \\ 60055.1 \\ 16026.9 \end{bmatrix},$$

$$\widehat{\boldsymbol{\upsilon}} = \begin{bmatrix} 2.67150 \\ 1.47429 \\ -0.04514 \end{bmatrix},$$

and the fitted model is

$$u_j = 2.67150x_j + 1.17429y_j - 0.04514z_j. \tag{12}$$

Secondly suppose that two constraints, $\beta_1 = 3\beta_2$ and $\beta_2 = \frac{1}{2}\beta_3$ must be enforced on the regression coefficients in (12). In this case the system of equations (7) is,

$$\begin{bmatrix} 3929 & 9532 & 2538 & 1 & 0 \\ 9532 & 23656 & 6324 & -3 & 1 \\ 2538 & 6324 & 1701 & 0 & -\frac{1}{2} \\ 1 & 3 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \phi_1 \\ \phi_2 \end{bmatrix} = \begin{bmatrix} 24434.7 \\ 60055.1 \\ 16026.9 \\ 0 \\ 0 \end{bmatrix}. \tag{13}$$

It should be noted that (13) is not a typical regression problem, and by (8) the estimated parameters are

$$\begin{bmatrix} 3929 & 9532 & 2538 & 1 & 0 \\ 9532 & 23656 & 6324 & -3 & 1 \\ 2538 & 6324 & 1701 & 0 & -\frac{1}{2} \\ 1 & 3 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 24434.7 \\ 60055.1 \\ 16026.9 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2.77593 \\ 0.9251 \\ 1.85062 \\ 11.1488 \\ 35.9426 \end{bmatrix},$$

and the fitted model is

$$u_j = 2.77593x_j + 0.92531y_j + 1.85062z_j + 11.1488\phi_1 + 35.9426\phi_2. \tag{14}$$

We notice that the enforced constraints are met.

Thirdly, note that the constrained part of the objective function (6) now is

$$\phi_1(\beta_1 - 3\beta_2) + \phi_2(\beta_2 - \frac{1}{2}\beta_3) = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -3 & 1 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix},$$

so that the constrained matrix is

$$\mathbf{C} = \begin{bmatrix} 1 & 0 \\ -3 & 1 \\ 0 & -\frac{1}{2} \end{bmatrix},$$

and the Schur complement is

$$S = -\mathbf{C}^\top(\mathbf{W}^\top\mathbf{W})^{-1}\mathbf{C} = \begin{bmatrix} \frac{-75637}{426920} & \frac{132791}{12808760} \\ \frac{132791}{12808760} & \frac{283583}{3842280} \end{bmatrix},$$

with $\mathbf{W}^\top\mathbf{W}$ as in (11). By (9) and (10)

$$\widehat{\boldsymbol{\phi}} = -S^{-1}\mathbf{C}^\top(\mathbf{W}^\top\mathbf{W})^{-1}\tilde{\mathbf{u}} = \begin{bmatrix} 11.1488 \\ 35.9462 \end{bmatrix}$$

and

$$\widehat{\boldsymbol{\upsilon}} = \tilde{\mathbf{u}} - (\mathbf{W}^\top\mathbf{W})^{-1}\mathbf{C}^\top\widehat{\boldsymbol{\phi}} = \begin{bmatrix} 2.77593 \\ 0.92531 \\ 1.85062 \end{bmatrix}.$$

So that the model is

$$u_j = 2.77593x_j + 0.92531y_j + 1.85062z_j + 11.1488\phi_1 + 35.9426\phi_2,$$

which is the same as in (14).

7. Conclusions

In this paper, two new methods (the differentiation and blocking techniques) have been derived for estimating the constrained parameters in a multiple linear regression model. A numerical study has been presented to validate the methods and results obtained are appealing. Although the derivations in this work have been demonstrated on a system with 3-explanatory variables and two constraints, these can be generalised to a system with k-explanatory variables and C-constraints, provided that both the explanatory variables and the constraints are linearly independent.

References

AHLBERG, J. H., NILSON, E. N., AND WALSH, J. L. (1967). *The Theory of Splines and Their Applications*. Academic Press: New York.

ALANELLI, M. AND HADJIDIMOS, A. (2004). Block gauss elimination followed by a classical iterative method for the solution of linear systems. *Journal of Computational and Applied Mathematics*, **163**, 381 – 400.

AXELSSON, O. (1985). Incomplete block matrix factorization preconditioning methods. the ultimate answer? *Journal of Computational and Applied Mathematics*, **12 & 13**, 3 – 18.

BASILEVSKY, A. (1983). *Applied Matrix Algebra in the Statistical Sciences*. Elsevier Science Publishing Co. (North-Holland), Inc.: New York.

BOYD, S. AND VANDENBERGHE, L. (2004). *Convex Optimization*. Cambridge University Press: New York.

CALLIER, F. M. AND DESOER, C. A. (1994). *Linear System Theory*. Springer: New York.

CARDINI, C. (2013). Data assimilation: Observation influence diagnostic of a data assimilation system. Series: ECMWF Lecture notes.
URL: `http://www.ecmwf.int/publications`

CARPENTER, K. H. (2005). Lagrange multiplier for quadratic forms with linear constraints.
URL: `https://www.ece.k-state.edu/people/faculty/carpenter/document/lagrange.pdf`

CHAN, R. AND CHAN, T. (1992). Circulant preconditioners for elliptic problems. *Journal of Numerical Linear Algebra with Applications*, **1** (1), 77 – 101.

CHEN, W., ANDERSON, B. O. D., DEISTLER, M., AND FILLER, M. (2012). Properties of blocked linear systems. *Automatica*, **48**, 2520 – 2525.

DRAPER, N. AND SMITH, H. (1998). *Applied Regression Analysis*. third edition. John Wiley & Sons: Hoboken, NJ.

HEATH, M. T. (2005). *Scientific Computing: An Introductory Survey*. McGraw-Hill: New York.

LASDON, L. S. (2002). *Optimization Theory for Large Systems*. Dover Publications: Mineola, NY.

LEGENDRE, A. M. (1805). Nouvelles méthodes pour la détermination des orbites des comètes. *Firmin Didot*.

MINGKUI, C. (1987). On the solution of circulant linear systems. *SIAM Journal on Numerical Analysis*, **24** (3), 668 – 683.

MULLET, G. (1976). Why regression coefficients have the wrong sign. *Journal of Quality Technology*, **8** (3), 121 – 126.

PAZZANI, M. AND BAY, S. (2001). Gaining insight from multiple linear regression. *In Proceeding of Annual Meeting of the Cognitive Science Society*.

RODRIGUEZ-YAM, G., DAVIS, R. A., AND SCHARF, L. L. (2004). Efficient Gibbs sampling of truncated multivariate normal with application to constrained linear regression.
URL: `http://www.stat.columbia.edu/~rdavis/papers/CLR.pdf`

SAWYER, S. (2002). The method of Lagrange multiplier. Accessed on 27/04/2015.
URL: `http://hedibert.org/wp-content/uploads/2014/04/LM-Sawyer-2002.pdf`

SEARLE, S. R. (1971). *Linear Models*. John Wiley & Sons: New York.

SOKOLNIKOFF, I. S. AND REDHEFFER, R. M. (1966). *Mathematics of Physics and Modern Engineering*. McGraw-Hill: New York.

WARDLAW, W. P. (2005). Row rank equals column rank. *Mathematics Magazine*, **78** (4), 316 – 318.

WOOD, W. L. (1971). Periodicity effects on the iterative solution of elliptic difference equations. *SIAM Journal on Numerical Analysis*, **8**, 439 – 464.

ZAVYALOV, U., KVASOV, B., AND MIROSHNICHENKO, V. (1980). *Spline Functions Methods*. Nauka: Moscow.

ZHANG, F. (2005). *The Schur Complement and its Applications*. Springer Science and Business Media: New York.

ZONG, Z. Y. (2009). Schur complements and determinant inequalities. *Journal of Mathematical Inequalities*, **3** (2), 161 – 167.

Manuscript received, 2015-05-09, revised, 2016-01-05, accepted, 2016-04-13.