International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878 (UGC Approved, Sr. No. 48096)
Index Copernicus Value (2015): 62.86 | Impact Factor (2015): 3.791

A New Extension of Dynamic Programming to Evaluate Linear Model

Bonoshree Saikia

Department of Statistics, North Eastern Hill University, Shillong (India)

Abstract: ***The method Dynamic Programming (DP) can be applied on linear models which are not of full rank. The use of DP for a matrix enables to solve systems of linear equations that are unbalance and linearly dependent. This technique can be used to compute the various statistical measures such as coefficient of determination R^2, the t-statistic and F-statistic. These measures are very much important to check how well a model fits the data. DP technique can be used to compute the coefficients even when multicollinearity is present among the explanatory variables under the investigation.***

Keywords: Dynamic Programming, Linear Models, Principle of Optimality, Least Square Problem, Non-Orthogonal

1. Introduction

Richard Bellman was the first man who coined the term 'Dynamic Programming' in 1950s to describe multistage decision processes. He introduced the principle of optimality and the functional equations of DP. Bellman's statement of the 'Principle of Optimality' is that "an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" (Bellman, 1957). Dynamic programming (DP) is a technique for computing recurrence relation efficiently by sorting partial results. Data fitting (or parameter estimation) is an important technique used for modeling in many areas of disciplines. Numerical analysts, statisticians and engineers have developed techniques and nomenclature for the least squares problems of their own discipline. Bellman and Kalaba (1965) used DP to the problem of obtaining a numerical solution to an ill conditioned system of linear equations. Kalaba and Natsuyama introduced two cost functions in the literature of DP. The first cost function is the square of the length of the current discrepancy vector and the second is the square of the length of the current solution vector. The two cost function is minimized simultaneously by optimally selecting the minimum length vector solution. This algorithm introduced in DP has been tested in a number of ways by Kalaba, Natsuyama and Uneo (1999) and Itiki, Kalaba and Natsuyama (1999). The principle of optimality in DP allows converting the least square problem into a sequential decision problem. Kalaba and Natsuyama made the most of this further by showing how the least squares problem can be solved for both the case introducing two constraints into DP framework, where the columns of the matrix A are independent and the case where they are dependent vectors. Kalaba, Natsuyama and Uneo (1999) treat the inverse problem of estimating transport parameter on the basis external observation of radiant intensity. These problems are approached using associative memory neural networks whose associated least square problems is solved by using a new DP algorithm. Kalaba, Johnson and Natsuyama (2005) introduced a new algorithm, which can be used to calculate various statistical quantities needed for evaluation of the linear model. They even showed that optimal control law can be used to deal with the least square problem in the case of collinearity. This paper will show how the DP algorithm introduced by Johnson and Kalaba can be used to produce the coefficients of the least square problems.

2. Methodology

A problem of obtaining the shortest length solution of consistent set of linear algebraic equations is defined as

$$X\beta = Y \tag{1}$$

where $X = \begin{bmatrix} X_{11} & X_{21} & X_{31} & \dots & X_{p1} \\ X_{12} & X_{22} & X_{32} & \dots & X_{p2} \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ X_{1n} & X_{2n} & X_{3n} & \dots & X_{pn} \end{bmatrix}$ $Y = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix}$ $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ . \\ \beta_p \end{bmatrix}$

For the estimation of parameter vector β, $\left| X\beta - Y \right|^2$ should be minimized keeping β as small as possible. Here, DP is allowing solving the problem sequentially. Let $h_r(Y)$ be the smallest square of the length of the vector β^r with subject to restriction

$$\left| X_r\beta^r - Y \right| = \text{minimum} \qquad \text{for} \qquad r = 1, 2, \dots, p.$$

The r[th] column of the matrix X is denoted x_r. The matrix X_r is the first r columns of X and β^r be a r-dimensional vector. Here, it illustrates about the linearly dependent case, i.e., the case when x_r is linearly dependent on x_1, x_2, …,

x_{r-1}. Thus, the Bellman principle of optimality gives the following recurrence relation as

$$h_r(Y) = \min_{\beta_r} [\beta_r^2 + h_{r-1}(Y - x_r\beta_r)] \quad \text{for } r = p+1, p+2, \ldots, n. \tag{2}$$

Here, only one set of scalars are like $\alpha_1, \alpha_2, \ldots, \alpha_r$ which satisfies the following relation

$x_1\beta_1 + x_2\beta_2 + \ldots + x_p\beta_p = Y_r$

$$\text{Finally, set } Y_n = Y \tag{3}$$

$$Y_r = Y_{r+1} - x_{r+1}\beta_{r+1} \quad \text{for } r = n-1, n-2, \ldots, p. \tag{4}$$

Expressing (3) in vector form as

$$\beta^{(p)} = (X'_p X_p)^{-1} X'_p Y_p \tag{5}$$

Defining $h_p(Y)$ as

$$h_p(Y) = x_1^2 + x_2^2 + \ldots + x_p^2 = (x^{(p)})'x^{(p)}$$

$$= Y' X_p[(X'_pX_p)^{-1}]'(X'_pX_p)^{-1}X'_pY$$

$$= Y' X_p(X'_pX_p)^{-1}(X'_pX_p)^{-1}X'_pY \tag{6}$$

Subsequently,

$$h_p(Y) = Y' R_p Y \tag{7}$$

$$\text{where } R_p = X_p(X'_pX_p)^{-1}(X'_pX_p)^{-1}X'_p \tag{8}$$

and also R_p is a n x n symmetric matrix.

$$\text{Considering that, } h_{r-1}(Y) = Y' R_{r-1} Y \tag{9}$$

where R_{r-1} is n x n symmetric matrix. Now, it will be shown that $h_r(Y)$ has the form $h_r(Y) = Y' R_r Y$ in which R_r is *n* x *n* symmetric matrix with *r* = *p+1*, *p+2*, …,*n*. From (2), for *r* = *p+1, p+2, …,n.*

$$h_r(Y) = \min_{\beta_r} [\beta_r^2 + R_{r-1}(Y - x_r\beta_r)]$$

$$= \min_{\beta_r} [\beta_r^2 + (Y - x_r\beta_r)' R_{r-1}(Y - x_r\beta_r)]$$

$$= \min_{\beta_r} [\beta_r^2 + Y' R_{r-1}Y - 2Y' R_{r-1}x_r\beta_r + x'_r R_{r-1}x_r\beta_r^2]$$

$$h_r(Y) = \min_{\beta_r} [(1 + x'_r R_{r-1} x_r)\beta_r^2 + Y' R_{r-1}Y - 2Y' R_{r-1}x_r\beta_r] \tag{10}$$

Differentiating within bracket expression of RHS in (10) with respect to β_r provides

$$\frac{\partial(.)}{\partial\beta_r} = 2(1 + x\, x'_r R_{r-1}x_r)\beta_r + 2Y' R_{r-1}x_r = 0 \tag{11}$$

$$\text{Hence, } \beta_r^{OPT} = \frac{Y' R_{r-1}x_r}{1 + x'_r R_{r-1}x_r} \tag{12}$$

$$\text{or, } \beta_r^{OPT} = \frac{x'_r R_{r-1}Y}{1 + x'_r R_{r-1}x_r} \tag{13}$$

Using (12) or (13) in (10) yields

$$h_r(Y) = (1 + x'_r R_{r-1}x_r)(\beta_r^{OPT})^2 - 2x'_r R_{r-1}Y\,\beta_r^{OPT} + Y' R_{r-1}Y$$

$$= (1 + x'_r R_{r-1} x_r)\left[\frac{Y' R_{r-1}x_r x'_r R_{r-1}Y}{(1 + x'_r R_{r-1}x_r)^2}\right]$$

$$-2Y'\left[\frac{R_{r-1}x_r x'_r R_{r-1}}{1 + x'_r R_{r-1}x_r}\right]Y + Y' R_{r-1}Y = Y' R_{r-1}Y - Y'\left[\frac{R_{r-1}x_r x'_r R_{r-1}}{1 + x'_r R_{r-1}x_r}\right]Y$$

$$h_r(Y) = Y'\left[\frac{R_{r-1} - R_{r-1}x_r x'_r R_{r-1}}{1 + x'_r R_{r-1}x_r}\right]Y = Y' R_r Y \tag{14}$$

$$\text{where, } R_r = \frac{R_{r-1} - R_{r-1}x_r x'_r R_{r-1}}{1 + x'_r R_{r-1}x_r}$$

$$\text{provided } 1 + x'_r R_{r-1}x_r \neq 0. \tag{15}$$

Let

$$\gamma_r = R_{r-1}x_r \quad \text{for } r = p+1, p+2, \ldots, n. \tag{16}$$

Therefore, the relation becomes

$$R_r = \frac{R_{r-1} - \gamma_r\gamma'_r}{1 + x'_r\gamma_r} \quad \text{for } r = p+1, p+2, \ldots, n. \tag{17}$$

Hence, if x_r is linearly independent of the vectors $x_1, x_2, \ldots, x_{r-1}$ then the algorithm gives the general form by substituting (16) in (11) or (12),

$$\beta_r^{OPT} = \frac{Y'_r\gamma_r}{1 + x'_r\gamma_r} \tag{18}$$

or,

$$\beta_r^{OPT} = \frac{\gamma'_r Y_r}{1 + x'_r\gamma_r} \tag{19}$$

where,

$$Y_r = Y - (x_1\beta_1 + x_2\beta_2 + x_{r-1}\beta_r) \tag{20}$$

The algorithm illustrated is extended as follows:

α_r is already defined, the component of the r[th] column of the matrix X_p, which is perpendicular to all its preceding column vectors $\alpha_1, \alpha_2, \ldots, \alpha_{r-1}$ with $r = 1, 2, \ldots, p$. The component of *Y*, which is perpendicular to all the vectors $x_1, x_2, \ldots, x_p$ may be obtained as

$$\alpha_{p+1} = M_p Y \tag{21}$$

In summary, the optimal set of scalars $\beta_1, \beta_2, \ldots, \beta_r$ can be calculated as:

$$Y_r = Y - \alpha_{r-1} \tag{22}$$

$$\beta_r^{OPT} = \frac{\gamma'_r Y_r}{1 + x'_r\gamma_r} \tag{23}$$

$$Y_r = Y - x_r\,\beta_r^{OPT} \tag{24}$$

$$\beta_{r-1}^{OPT} = \frac{\gamma'_{r-1} Y_{r-1}}{1 + x'_{r-1}\gamma_{r-1}} \quad (25)$$

2.1 Determining the Regression Coefficient:

The precision of estimated regression equation has to be checked in predicting the values of the dependent variable Y. The measure, called the coefficient of determination R^2, is a scale-free summary of degree to which the variables x_1, x_2, …,x_p predict the dependent variable Y (Lawson and Hanson, 1974).

Now, the total variation is defined as the total sum of the squared deviations (TSS) of the vector Y about its mean to obtain. (Greene, 2002) discussed that a linear relationship $X\beta = Y$ is assumed and TSS is factored into two parts in which the first part is the sum of squares explained by the regression (SSR) and the second part is the sum of squares that could not be explained by the regression equation known as the residual sum of squares (RSS). Hence, it is defined as

$$\text{TSS} = \text{SSR} + \text{RSS} \quad (26)$$

Now, R^2 is defined as

$$R^2 = \frac{SSR}{TSS} = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS} \quad (27)$$

The RSS is computed as

$$\text{RSS} = \alpha'_{p+1}\alpha_{p+1} \quad (28)$$

Where the last residual vector α_{p+1} is defined as

$$\alpha_{p+1} = Y - (x_1\beta_1 + x_2\beta_2 + \ldots + x_p\beta_p) \quad (29)$$

This yields

$$R^2 = \frac{Y'Y - f_p(Y)}{Y'Y} = 1 - \frac{\alpha'_{p+1}\alpha_{p+1}}{Y'Y} \quad (30)$$

2.2 Examining the Model Fit

To measure the reliability for the model, the goodness of fit between the data and the model has to be measured. The F-test is applied to check whether the regression coefficients $\beta_1, \beta_2, \ldots, \beta_p$ considered together vary significantly from 0. The hypotheses are set up for testing as follows.

$$H_0: \beta = 0 \text{ versus } H_1: \beta \neq 0$$

Column 1 in Table 1 demonstrates the sources of variations as regression and the residual and column 2 shows respective degrees of freedom. The SS, MSS and F-ratio are given in columns 2, 3 and 4 of this table respectively. Now, the computed value of F-ratio is compared to the theoretical value taken from the F-table for any chosen level of significance and a decision rule is applied to reject or accept the hypothesis that the model fits the data.

Table 1: ANOVA for the F-test

Source	Df	SS	MSS	F-ratio
Regression	$p-1$	$Y'Y - \alpha'_{p+1}\alpha_{p+1}$	$\frac{Y'Y - \alpha'_{p+1}\alpha_{p+1}}{p-1}$	$\frac{(Y'Y - \alpha'_{p+1}\alpha_{p+1})/(p-1)}{(\alpha'_{p+1}\alpha_{p+1})/(m-p)}$
Error	$m-p$	$\alpha'_{p+1}\alpha_{p+1}$	$\frac{\alpha'_{p+1}\alpha_{p+1}}{m-p}$	
Total	$m-1$			

The F-statistic can be computed to test the hypothesis as

$$F_{Cal}(p-1, m-p) = \frac{R^2/(p-1)}{(1-R^2)/(m-p)} = \frac{(Y'Y - \alpha'_{p+1}\alpha_{p+1})/(p-1)}{(\alpha'_{p+1}\alpha_{p+1})/(m-p)} \quad (31)$$

where m = No. of observations $p-1$ and $m-p$ = degrees of freedom

2.3 Examining the Significance of the Explanatory Variables

Now, it will be examined about the contribution of each individual coefficient β_i significantly to the regression equation or not. Therefore, for testing the hypothesis for given coefficient, hypothesis is to be set up as

$H_0: \beta_i = 0$ versus $H_1: \beta_i \neq 0$

The usual statistical test for this is the t-statistic. Thus

$$t_{cal} = \frac{\beta_i}{SE(\beta_i)} \quad (32)$$

In order to construct the test, SE of the estimator should be known. So, the t-statistic is derived indirectly by utilizing the fact that the exact distribution of t_p^2 -statistic with p-df is F-distribution with (1, m) df, viz., F(1, p). The SSE for the full model together with p-predictor is subtracted from the RSS of a subset model with $p-1$ predictors which give the reduction in RSS by including predictor I = RSS_{p-1} - RSS_p

$$\text{Hence, } F(1, m-p) = \frac{RSS_{p-1} - RSS_p}{RSS_p/(m-p)} \quad (33)$$

where $\frac{RSS_p}{m-p}$ is the variance of the full model.

$$t_i = \sqrt{\frac{RSS_{p-1} - RSS_p}{RSS_p/(m-p)}} \quad (34)$$

Substituting (188) in (194) gives

$$t_i = \sqrt{\frac{\alpha'_{p-1}\alpha_{p-1} - \alpha'_p\alpha_p}{\frac{\alpha'_p\alpha_p}{m-p}}} \quad (35)$$

where t_i = calculated t-value for the coefficient of β_i, α_{p-1} = residual vector from the model without predictor i, α_p = residual vector from the model with all the predictors.

Table 2 provides the steps demonstrated above by partitioning the variance and constructing the ANOVA table.

Table 2: ANOVA for the t-test

Source	Df	SS	MSS	t-ratio
Full model	P	RSS ^{p}		
Model without predictor β^i	p - 1	RSS $^{p-1}$		
Predictor	1	RSS $^{p-1}$ - RSS ^{p}	RSS $^{p-1}$ - RSS ^{p}	$\frac{RSS_{p-1} - RSS_p}{RSS_p/(m-p)}$
Error (Full Model)	m - p	RSS ^{p}	$\frac{RSS_p}{m-p}$	

3. Illustrative Example

To estimate the effect of type of plant on the weight of the maize fruit, four types of maize plants given the same condition has been recorded the following weight of its fruit at harvest as

Wt. of the Plants	Type1	Type 2	Type 3	Type 4
	62 71 83 90	80 75 45	62 75	60
Total	306	200	137	60

In order to estimate the effect of the type of plant on the weight of plant it is assumed that the observation Y_{ij} is the sum of four parts as shown below:

$$Y_{i,j} = v + \alpha_I + e_{i,j}$$

where,

v is the population mean of the weight of plant

α_I is the effect of the type I on weight

$e_{i,j}$ is the random error term peculiar to the observation.

The observations are written down in terms of the equation as follows:

62 = $Y_{11} = v + \alpha_I + e_{11}$

71 = $Y_{12} = v + \alpha_I + e_{12}$

83 = $Y_{13} = v + \alpha_I + e_{13}$

90 = $Y_{14} = v + \alpha_I + e_{14}$

80 = $Y_{21} = v + \alpha_I + e_{21}$

75 = $Y_{22} = v + \alpha_I + e_{22}$

45 = $Y_{23} = v + \alpha_I + e_{23}$

62 = $Y_{31} = v + \alpha_I + e_{31}$

75 = $Y_{32} = v + \alpha_I + e_{32}$

60 = $Y_{41} = v + \alpha_I + e_{41}$

In matrix form, they are represented as:

$$\begin{bmatrix}62\\71\\83\\90\\80\\75\\45\\62\\75\\60\end{bmatrix} = \begin{bmatrix}Y_{11}\\Y_{12}\\Y_{13}\\Y_{14}\\Y_{21}\\Y_{22}\\Y_{23}\\Y_{31}\\Y_{32}\\Y_{41}\end{bmatrix} = \begin{bmatrix}1&1&0&0&0\\1&1&0&0&0\\1&1&0&0&0\\1&1&0&0&0\\1&0&1&0&0\\1&0&1&0&0\\1&0&1&0&0\\1&0&0&1&0\\1&0&0&1&0\\1&0&0&0&1\end{bmatrix}\begin{bmatrix}\vartheta\\\alpha_1\\\alpha_2\\\alpha_3\\\alpha_4\end{bmatrix} + \begin{bmatrix}e_1\\e_2\\e_3\\e_4\\e_5\\e_6\\e_7\\e_8\\e_9\\e_{10}\end{bmatrix}$$

$Y = X\beta + e$

Y is the vector of observations, X is the incidence matrix and β is the vector of parameters to be considered here.

The normal equations corresponding to the model, $Y = X\beta + e$ can be derived by LS to give,

$$X'X\beta = X'Y$$

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0` \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 10 & 4 & 3 & 2 & 1 \\ 4 & 4 & 0 & 0 & 0 \\ 3 & 0 & 3 & 3 & 0 \\ 2 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 62 \\ 71 \\ 83 \\ 90 \\ 80 \\ 75 \\ 45 \\ 62 \\ 75 \\ 60 \end{bmatrix} = \begin{bmatrix} 703 \\ 306 \\ 200 \\ 137 \\ 60 \end{bmatrix}$$

Matrix $X'X$ has determinant equal to zero and hence not of full rank. Therefore, matrix $(X'X)^{-1}$ does not exist. Hence the equation cannot be express as

$$\beta = (X'X)^{-1} X'Y$$

This implies that the normal equation has no unique solution. To get one of the solutions, the procedure will be done as follows:

$$M_o = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

α_1, M_1, α_2, M_2, α_3 and M_3 are calculated below.

M_o is updated to M_1 as shown below,

$$\alpha_1' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$M_1 = I - \frac{x_1 x_1'}{x_1' x_1} =$$

$$\begin{bmatrix} 0.9 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 \\ -0.1 & 0.9 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 \\ -0.1 & -0.1 & 0.9 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 \\ -0.1 & -0.1 & -0.1 & 0.9 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 \\ -0.1 & -0.1 & -0.1 & -0.1 & 0.9 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 \\ -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & 0.9 & -0.1 & -0.1 & -0.1 & -0.1 \\ -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & 0.9 & -0.1 & -0.1 & -0.1 \\ -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & 0.9 & -0.1 & -0.1 \\ -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & 0.9 & -0.1 \\ -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & 0.9 \end{bmatrix}$$

$$\alpha_2' = \begin{bmatrix} 0.6 & 0.6 & 0.6 & 0.6 & -0.4 & -0.4 & -0.4 & -0.4 & -0.4 & -0.4 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} 0.75 & -0.25 & -0.25 & -0.25 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.25 & 0.75 & -0.25 & -0.25 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.25 & -0.25 & 0.75 & -0.25 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.25 & -0.25 & -0.25 & 0.75 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.83 & -0.17 & -0.17 & -0.17 & -0.17 & -0.17 \\ 0 & 0 & 0 & 0 & -0.17 & 0.83 & -0.17 & -0.17 & -0.17 & -0.17 \\ 0 & 0 & 0 & 0 & -0.17 & -0.17 & 0.83 & -0.17 & -0.17 & -0.17 \\ 0 & 0 & 0 & 0 & -0.17 & -0.17 & -0.17 & 0.83 & -0.17 & -0.17 \\ 0 & 0 & 0 & 0 & -0.17 & -0.17 & -0.17 & -0.17 & 0.83 & -0.17 \\ 0 & 0 & 0 & 0 & -0.17 & -0.17 & -0.17 & -0.17 & -0.17 & 0.83 \end{bmatrix}$$

$$\alpha_3' = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0.5 & -0.5 & -0.5 & -0.5 \end{bmatrix}$$

$$M_3 =$$

$$\begin{bmatrix} 0.75 & -0.25 & -0.25 & -0.25 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.25 & 0.75 & -0.25 & -0.25 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.25 & -0.25 & 0.75 & -0.25 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.25 & -0.25 & -0.25 & 0.75 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.67 & -0.33 & -0.33 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.33 & 0.67 & -0.33 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.33 & -0.33 & 0.67 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.67 & -0.33 & -0.33 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.33 & 0.67 & -0.33 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.33 & -0.33 & 0.67 \end{bmatrix}$$

$$\alpha_4' = \begin{bmatrix} -14.5 & -5.5 & 6.5 & 13.5 & 13.33 & 8.3 & -21.67 & -3.66 & 9.33 & -5.67 \end{bmatrix}$$

$$Y_5' = \begin{bmatrix} 76.5 & 76.5 & 76.5 & 76.5 & 66.7 & 66.7 & 66.7 & 65.7 & 65.7 & 65.7 \end{bmatrix}$$

R_1 is computed as shown,

$$R_1 = (x_1^{+})' x_1^{+} = \left(\frac{x_1}{x_1' x_1} \right)' \left(\frac{x_1'}{x_1' x_{11}} \right)$$

Therefore, $R_1 =$

$$\begin{bmatrix} 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \end{bmatrix}$$

R_2 is calculated,

$$\gamma_2 = R_1 x_2 = \begin{bmatrix} 0.04 \\ 0.04 \\ 0.04 \\ 0.04 \\ 0.04 \\ 0.04 \\ 0.04 \\ 0.04 \\ 0.04 \\ 0.04 \end{bmatrix}$$

$$R_2 = \frac{R_1 - \gamma_3 \gamma_3'}{1 + x_3' \gamma_3}$$

$R_2 =$

$$\begin{bmatrix} 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 \\ 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 \\ 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 \\ 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 \\ 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 \\ 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 \\ 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.00872 & 0.0072 & 0.0072 & 0.0072 & 0.0072 \\ 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 \\ 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 \\ 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 & 0.0072 \end{bmatrix}$$

Similarly, γ_3, R_3, γ_4, R_4 and γ_5 are calculated. Thus

$$\gamma_3 = \begin{bmatrix} 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \end{bmatrix}, \gamma_4 = \begin{bmatrix} 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \\ 0.0217 \end{bmatrix} \text{ and } \gamma_5 \begin{bmatrix} 0.2436 \\ 0.2436 \\ 0.2436 \\ 0.2436 \\ 0.2436 \\ 0.2436 \\ 0.2436 \\ 0.2436 \\ 0.2436 \\ 0.2436 \end{bmatrix}$$

$R_3 =$

$$\begin{bmatrix} 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 \\ 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 \\ 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 \\ 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 \\ 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 \\ 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 \\ 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 \\ 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 \\ 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 \\ 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 & 0.0063 \end{bmatrix}$$

$R_4 =$

$$\begin{bmatrix} 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 \\ 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 \\ 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 \\ 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 \\ 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 \\ 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 \\ 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 \\ 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 \\ 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 \\ 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 & 0.2436 \end{bmatrix}$$

Now, $\beta_5, \beta_4, \beta_3, \beta_2$ *and* β_1 will be calculated as:

$$Y_5 = Y - \alpha_4 = \begin{bmatrix} 76.50 \\ 76.50 \\ 76.50 \\ 76.50 \\ 66.67 \\ 66.67 \\ 66.67 \\ 66.67 \\ 66.67 \\ 66.67 \end{bmatrix}$$

$$\beta_5^{OPT} = \frac{\gamma_5' Y_5}{1 + x_5' \gamma_5} = 171.2823$$

$$Y_4 = Y_5 - \beta_5^{OPT} x_5$$

$$Y_4 = \begin{bmatrix} 76.50 \\ 76.50 \\ 76.50 \\ 76.50 \\ 66.66 \\ 66.66 \\ 66.66 \\ 65.66 \\ 65.66 \\ -72.06 \end{bmatrix}$$

$$\beta_4^{OPT} = \frac{\gamma_4' Y_4}{1 + x_4' \gamma_4} = 7.0067$$

$$Y_3 = Y_4 - \beta_4^{OPT} x_4$$

$$Y_3 = \begin{bmatrix} 76.50 \\ 76.50 \\ 76.50 \\ 76.50 \\ 66.66 \\ 66.66 \\ 66.66 \\ 58.65 \\ 58.65 \\ -72.65 \end{bmatrix}$$

$$\beta_3^{OPT} = \frac{\gamma_3' Y_3}{1 + x_3' \gamma_3} = 11.2429$$

$$Y_2 = Y_3 - \beta_3^{OPT} x_3$$

$$Y_2 = \begin{bmatrix} 76.50 \\ 76.50 \\ 76.50 \\ 65.25 \\ 55.42 \\ 55.42 \\ 66.66 \\ 58.65 \\ 58.65 \\ -72.05 \end{bmatrix}$$

$$\beta_2^{OPT} = \frac{\gamma_2' Y_2}{1 + x_2' \gamma_2} = 17.8459$$

$$Y_1 = Y_2 - \beta_2^{OPT} x_2 = \begin{bmatrix} 58.6540 \\ 58.6540 \\ 58.6540 \\ 47.4111 \\ 55.4237 \\ 55.4237 \\ 66.6666 \\ 58.6598 \\ 58.6598 \\ -72.0594 \end{bmatrix}$$

$$\beta_1^{OPT} = \frac{\alpha_1' b_1}{a_1' \alpha_1} = 44.6147$$

$$\hat{\beta} = \begin{bmatrix} 44.61 \\ 17.84 \\ 11.29 \\ 7.006 \\ 171.28 \end{bmatrix}$$

Table 3: ANOVA

Source	DF	SS	MS	F
Regression	3	312.6	104.2	0.496
Residual	6	1259.5	209.9	
Total	9	1572.1		

R^2 = 312.6/1572.1 = 0.199

The total variation explained by the model is 19.9%, the overall model is not significant, meaning that the weight of the maize fruit does not depend on the type of maize plant.

4. Conclusion

In this paper, an application of dynamic programming has been considered as an alternative approach for the solution of least squares regression problems. The output of the algorithm based on dynamic programming can be used to find out the shortest length solutions of least squares problems in which the matrix may be less than of full rank. DP technique can be used to compute the coefficients even when multicollinearity is present among the explanatory variables under the investigation.

References

[1] Bellman R. (1957), “Dynamic Programming”, Princeton Landmarks in Mathematics.

[2] Bellman R. and Kalaba R. E. (1965), “Dynamic Programming and Modern Control Theory”, Academic Press, New York.

[3] Greene, W. H. (2002), “Econometric Analysis”, 4th ed., Prentice-Hall, New Jersey.

[4] Itiki, C., Kalaba, R.E., Natsuyama, H.H. (1999). Estimation problem of parameter of a lower limb model. Proc. 4th Int. Symp. on artificial Life, Computers and Mathematics with Applications, 37, 1 -7.

[5] Kalaba, R.E., Natsuyama, H., Uneo, S. (1999)., Regression Analysis ûia Dynamic Programming, I: Theory, 30th ISCIE International Symposium on Stochastic Systems and Their Applications, Kyoto, Japan, 1998.

[6] Kalaba, R.E., Natsuyama, H., Uneo, S. (1999). The estimation of parameter in time dependent transport problems: Dynamic Programming and associative memories. Computers and Mathematics with Applications, 37, 41 - 45.

[7] Kalaba, R.E., Johnson, J., Natsuyama, H. (2005). Statistical measures for least squares using the αQβR algorithm, Journal of Optimization Theory and Applications, 127 (3), 515 - 522.

[8] Lawson, C. and Hanson, R. (1974), “Solving Least Square Problems”, Prentice-Hall Englewood Cliff, NJ.