The Korean Communications
in Statistics Vol. 5, No. 1, 1998
pp. 1-18

Animation of AVP and DAVP for regression diagnostics[1]

Sung H. Park, Jae J. Kim, and Sung H. Chung[2]

Abstract

Since 1960s, in which the computer graphics system first appeared, various graphical techniques have been introduced for regression diagnostics and they have been remarkably developed. In particular, animation, one of the dynamic graphical methods which Cook and Weisberg (1989) proposed helps to show the effect of adding variables or observations to a model, or removing them from a model on the regression results. We present the added variable plots (AVP) with animation, which can be used as an optical tool of understanding the affect of some variables or observations on other variables, and the detrended added-variable plots (DAVP) with animation, through which it is possible to find out whether specific variables or observations have an effect on the nonlinearity of other variables or not.

1. Introduction

Graphics in statistics have become an essential part since Frobes (1857) introduced its concept. Expressly, the appearance of the computer graphics system in the 1960s marked a turning point in the statistical methodology, resulting in a remarkable leap of the graphical techniques (see Becker (1981) and Cleveland (1987)). Various methods for dynamic graphics, such as identification, linking, deletion, scaling, brushing, rotation, and animation are the products obtained by using its rapid computation to the best advantage. Dynamic plots based on these methods are useful in the sense that graphic is to show certain phenomenon in the two or three dimensional space to be comprehensible in comparison with flat static plots. Among dynamic methods, animation which was recently proposed by Cook and Weisberg (1989), and developed by Park and Kim (1992), is particularly useful in understanding the effects of adding variables or observations to a model or removing them from a model. The reason is that these effects can be shown with smooth changes.

In this paper we focus on the added variable plots (AVP) and detrended added variable plots (DAVP) with animation to study the impacts which some variables or observations can have on the other variables. In Section 2, we will introduce the AVP and DAVP with

1) This study was partially supported by Korean Ministry of Education through Research fund, BSRI-97-1415.

2) Department of Computer Science and Statistics, Seoul National University, Seoul, Korea

animation when some variables or observations or both of them are removed. In Section 3, one example will be given to illustrate the proposed dynamic graphics which can be useful for regression diagnostics.

The program for the analyses of the example given in the paper is written using S-plus package.

2. (Detrended) added variable plots with animation

To begin with, we introduce the AVP and the $DAVP$. Suppose that the full regression model is

$$\underline{y} = X\underline{\beta} + V\alpha + \underline{\varepsilon} \tag{1}$$

where $\underline{y}$ is an $n \times 1$ vector of observed response, X is an $n \times p$ fixed known matrix, V is an $n \times 1$ fixed known vector, $\underline{\beta}$ is a $p \times 1$ vector of unknown regression parameters, α is an unknown scalar, and $\underline{\varepsilon}$ is an $n \times 1$ vector of unobservable random errors whose distribution is usually assumed $N(\underline{0}, I\sigma^2)$. By multiplying (1) by $(I - P_X)$, where P_X is the projection matrix of X, i.e., $P_X = X(X'X)^{-1}X'$ and noting the $(I - P_X)X = 0$, we can obtain

$$\begin{aligned} \underline{e}_{y \cdot X} &= \underline{e}_{V \cdot X}\,\alpha + (I - P_X)\,\underline{\varepsilon} \\ &= \underline{e}_{V \cdot X}\,\alpha + \underline{e}. \end{aligned}$$

Here $\underline{e}_{y \cdot X} = (I - P_X)\underline{y}$ is the residual vector when $\underline{y}$ is regressed on X, $\underline{e}_{V \cdot X} = (I - P_X)V$ is the residual vector when V is regressed on X, and $\underline{e}$ is the residual vector of the full model, that is, when $\underline{y}$ is regressed on (X, V). The plot of $\underline{e}_{y \cdot X}$ versus $\underline{e}_{V \cdot X}$ is called the AVP or the partial regression plot (see Cook and Weisberg (1982), Chamber (1983), and Chatterjee and Hadi (1989)). Here $\underline{e}_{V \cdot X}$ is related to the systematic component, orthogonal to $\underline{e}$, which determines the scatter.

On the other hand, it is possible that an undue linear trend in the AVP should visually mask the defects in the form in which V enters the model (1), which includes nonlinearities and outliers. Giving a full explanation, suppose that the true model is

$$\underline{y} = X\underline{\beta} + V\alpha + U + \underline{\varepsilon}$$

where U is a model component which is unknown and nonstochastic. Then $E(\underline{e}_{y \cdot X})$ can be decomposed into two orthogonal parts

$$\begin{aligned} E(\underline{e}_{y \cdot X}) &= \underline{e}_{V \cdot X}\,\alpha + Q_X U \\ &= \underline{e}_{V \cdot X}\,\alpha^* + QU \end{aligned}$$

where $Q_X = I - P_X$, $Q = I - P_{(X, V)}$, and $\alpha^* = \alpha + \dfrac{\underline{e}'_{V \cdot X} U}{\| \underline{e}_{V \cdot X} \|^2}$.

Here $P_{(X,V)}$ is the projection matrix of (X, V) and $\frac{\underline{e}_{V\cdot X}}{\|\underline{e}_{V\cdot X}\|}$ is the part of V orthogonal to X, normalized to unit length. Precisely the expected slope of the AVP is α^* under this model and QU is the only part of U which is probably distinct as a deviation from the linear trend. Besides if QU is small relative to $\underline{e}_{V\cdot X}\,\alpha^*$, the potential for masking of U may make the visual detection of U difficult. As one device of solving this problem, the $DAVP$, $\{\underline{e}, \underline{e}_{V\cdot X}\}$, can be presented, which is obtained from removing the systematic component from the ordinate of an AVP. Since $E(\underline{e}) = QU$, often the presence of U is detected without difficulty after detrending (Cook and Weisberg (1989)).

In Section 2.1, we will study the impacts of removing some variables on the other variables remaining in the model with the animated AVP or $DAVP$. Analogously, in Section 2.2, the animated $AVP(DAVP)$ will show the effects of removing some observations on the specific variable in the model dynamically. In Section 2.3, we will also think over joint impacts of removing variables and observations smoothly on the other variables through the $AVP(DAVP)$.

2.1 Animated AVP(DAVP) in omission of multiple variables

First of all, we consider the case of omitting a variable smoothly. Assume that X_1: $n\times(p-1)$ matrix, X_2: $n\times 1$ vector, V: $n\times 1$ vector, $\underline{\beta}_1$: $(p-1)\times 1$ vector. Under these conditions, the full model can be partitioned and modified as

$$\begin{aligned}\underline{y} &= X_1\underline{\beta}_1 + X_2\beta_2 + V\theta + \underline{\varepsilon}\\ &= X_1\underline{\beta}_1^* + \tilde{X}_2\beta_2^* + \tilde{V}\theta^* + \underline{\varepsilon}\\ &= X\tilde{\underline{\beta}}^* + \tilde{V}\theta^* + \underline{\varepsilon}\\ &= Z\underline{\beta}^* + \underline{\varepsilon}\end{aligned}$$

where $\tilde{X}_2 = \frac{Q_1X_2}{\|Q_1X_2\|}$, $\tilde{V} = \frac{Q_{1,2}V}{\|Q_{1,2}V\|}$, $X = (X_1 \vdots \tilde{X}_2)$, $Z = (X \vdots \tilde{V})$, $\tilde{\underline{\beta}}^{*\prime} = (\underline{\beta}_1^{*\prime}, \beta_2^*)$, $\underline{\beta}^{*\prime} = (\tilde{\underline{\beta}}^{*\prime}, \theta^*)$. $\tilde{X}_2$ is the part of X_2 orthogonal to X_1 and $\tilde{V}$ is the part of V orthogonal to $(X_1, \tilde{X}_2)$.

Suppose that $\underline{b}$ is a $(p+1)\times 1$ vector of zeros except for a single 1 corresponding to X_2. For $0 \le \lambda < 1$ we get $\tilde{\underline{\beta}}(\lambda)$, the estimate of $\underline{\beta}^*$ (Park and Kim (1995))

And as X_2 is omitted smoothly, the fitted values can be obtained by

When X_2 is removed from the model smoothly, the projection matrix or prediction matrix

$$
\begin{aligned}
\hat{\underline{y}}(\lambda) &= Z\hat{\underline{\beta}}(\lambda) \\
&= X_1(X_1'X_1)^{-1}X_1'\underline{y}+(1-\lambda)\tilde{X}_2\tilde{X}_2'\underline{y}+\tilde{V}\tilde{V}' \\
&= \hat{\underline{y}}_1+(1-\lambda)(\hat{\underline{y}}-\hat{\underline{y}}(1))+(\hat{\underline{y}}(1)-\hat{\underline{y}}_1) \\
&= \hat{\underline{y}}-\lambda(\hat{\underline{y}}-\hat{\underline{y}}(1)).
\end{aligned}
$$

$$
\begin{aligned}
\hat{\underline{\beta}}(\lambda) &= (Z'Z+\frac{\lambda}{1-\lambda}\underline{b}\underline{b}')^{-1}Z'\underline{y} \\
&= \begin{pmatrix} (X_1'X_1)^{-1}X_1'\underline{y} \\ (1-\lambda)\tilde{X}_2'\underline{y} \\ \tilde{V}'\underline{y} \end{pmatrix}
\end{aligned}
$$

for X is formed by

$$
\begin{aligned}
P_X(\lambda) &= X(X'X+\frac{\lambda}{1-\lambda}\underline{b}_X\underline{b}_X')^{-1}X' \\
&= X\left(\begin{pmatrix} X_1'X_1 & \underline{0} \\ \underline{0}' & 1 \end{pmatrix} + \begin{pmatrix} O & \underline{0} \\ \underline{0}' & \frac{\lambda}{1-\lambda} \end{pmatrix}\right)^{-1}X' \\
&= (X_1\ \tilde{X}_2)\begin{pmatrix} (X_1'X_1)^{-1} & \underline{0} \\ \underline{0}' & 1-\lambda \end{pmatrix}\begin{pmatrix} X_1' \\ \tilde{X}_2' \end{pmatrix} \\
&= P_{X_1}+(1-\lambda)\tilde{X}_2\tilde{X}_2'
\end{aligned}
$$

where $\underline{b}_X$ is a $p\times 1$ vector of zeros except for a single 1 corresponding to X_2. Here we note that $\tilde{X}_2\tilde{X}_2' = P_X - P_{X_1}$ and $\tilde{X}_2\tilde{X}_2'\underline{y} = \hat{\underline{y}}_X - \hat{\underline{y}}_{X_1}$. On this occasion $P_X(\lambda)$ can be rewritten simply with P_X and P_{X_1}, and also $\hat{\underline{y}}_X(\lambda)$ can be represented with $\hat{\underline{y}}_X$ and $\hat{\underline{y}}_{X_1}$ as follows.

$$P_X(\lambda) = P_X - \lambda(P_X - P_{X_1})$$

and

$$
\begin{aligned}
\hat{\underline{y}}_X(\lambda) &= P_X(\lambda)\underline{y} \\
&= \hat{\underline{y}}_X - \lambda(\hat{\underline{y}}_X - \hat{\underline{y}}_{X_1})
\end{aligned}
$$

where $\hat{\underline{y}}_X$ is the fitted vector of $\underline{y}$ on only X and similarly $\hat{\underline{y}}_{X_1}$ is the fitted vector of $\underline{y}$ on only X_1.

Going on a step ahead, we can consider removing multiple variables simultaneously. The modified full model is (Park and Kim (1992))

$$
\begin{aligned}
\underline{y} &= X_1\underline{\beta}_1 + U_1\delta_1 + \cdots + U_q\delta_q + V\theta + \underline{\varepsilon} \\
&= X_1\underline{\beta}_1^* + \tilde{U}_1\delta_1^* + \cdots + \tilde{U}_q\delta_q^* + \tilde{V}\theta^* + \underline{\varepsilon} \\
&= X\underline{\beta}^* + \tilde{V}\theta^* + \underline{\varepsilon} \\
&= Z\underline{\beta}^* + \underline{\varepsilon}
\end{aligned}
$$

where $\tilde{U}_1 = \dfrac{Q_{X_1}U_1}{\|Q_{X_1}U_1\|}$, $\cdots$, $\tilde{U}_q = \dfrac{Q_{X_1,\tilde{U}_1,\cdots,\tilde{U}_{q-1}}U_q}{\|Q_{X_1,\tilde{U}_1,\cdots,\tilde{U}_{q-1}}U_q\|}$, $\tilde{V} = \dfrac{Q_{X_1,\tilde{U}_1,\cdots,\tilde{U}_q}V}{\|Q_{X_1,\tilde{U}_1,\cdots,\tilde{U}_q}V\|}$,

$X=(X_1, \tilde{U}_1,$

$\cdots, \tilde{U}_q)$, $Z=(X_1, \tilde{U}_1, \cdots, \tilde{U}_q, \tilde{V})$, $\underline{\beta}^{*\prime}=(\underline{\beta}_1^{*\prime}, \delta_1^*, \cdots, \delta_q^*, \theta^*)$, $\tilde{\underline{\beta}}^{*\prime}=(\underline{\beta}_1^{*\prime}, \delta_1^*, \cdots, \delta_q^*)$. Here Z is an $n\times(p+q)$ matrix and X is an $n\times(p-1+q)$ matrix.

When $U_1, \cdots, U_q$ are removed with the different rates, namely, $\lambda_1, \cdots, \lambda_q$, for $0\leq\lambda_i<1$, $i=1,2,\cdots,q$, $\tilde{\underline{\beta}}(\underline{\lambda})$ becomes

$$\begin{aligned}\tilde{\underline{\beta}}(\underline{\lambda}) &= \left(Z'Z+\frac{\lambda_1}{1-\lambda_1}\underline{b}_1\underline{b}_1'+\frac{\lambda_2}{1-\lambda_2}\underline{b}_2\underline{b}_2'+\cdots+\frac{\lambda_q}{1-\lambda_q}\underline{b}_q\underline{b}_q'\right)^{-1}Z'\underline{y}\\ &= \begin{pmatrix}(X_1'X_1)^{-1}X_1'\underline{y}\\(1-\lambda_1)\tilde{U}_1'\underline{y}\\ \vdots\\(1-\lambda_q)\tilde{U}_q'\underline{y}\\ \tilde{V}'\underline{y}\end{pmatrix}.\end{aligned}$$

In this case, $\underline{b}_i$, $(i=1,2,\cdots,q)$, is a $(p+q)\times 1$ vector of zeros except for a single 1 corresponding to U_i. For simplicity, assume that the smoothly omitted rates of U_1, $\cdots$, U_q are all the same, that is, $\lambda_1=\cdots=\lambda_q=\lambda$. Then for $0\leq\lambda<1$, $\tilde{\underline{\beta}}(\lambda)$ and $\hat{\underline{y}}(\lambda)$ can be rewritten simply as

$$\begin{aligned}\tilde{\underline{\beta}}(\lambda) &= \left(Z'Z+\frac{\lambda}{1-\lambda}\underline{b}_1\underline{b}_1'\frac{+\lambda}{1-\lambda}\underline{b}_2\underline{b}_2'+\cdots+\frac{\lambda}{1-\lambda}\underline{b}_q\underline{b}_q'\right)^{-1}Z'\underline{y}\\ &= \begin{pmatrix}(X_1'X_1)^{-1}X_1'\underline{y}\\(1-\lambda)\tilde{U}_1'\underline{y}\\ \vdots\\(1-\lambda)\tilde{U}_q'\underline{y}\\ \tilde{V}'\underline{y}\end{pmatrix}\end{aligned}$$

and

$$\begin{aligned}\hat{\underline{y}}(\lambda) &= Z\tilde{\underline{\beta}}(\lambda)\\ &= X_1(X_1'X_1)^{-1}X_1'\underline{y}+(1-\lambda)\tilde{U}_1\tilde{U}_1'\underline{y}+\cdots+(1-\lambda)\tilde{U}_q\tilde{U}_q'\underline{y}+\tilde{V}\tilde{V}'\underline{y}\\ &= \hat{\underline{y}}_1+(1-\lambda)(\hat{\underline{y}}-\hat{\underline{y}}(1))+(\hat{\underline{y}}(1)-\hat{\underline{y}}_1)\\ &= \hat{\underline{y}}-\lambda(\hat{\underline{y}}-\hat{\underline{y}}(1)).\end{aligned}$$

Assuming that $\underline{b}_{X_i}$ is a $(p-1+q)\times 1$ vector of zeros except for a single 1 corresponding to U_i and noting that $\tilde{U}_1\tilde{U}_1'+\cdots+\tilde{U}_q\tilde{U}_q'=P_X-P_{X_1}$ and $(\tilde{U}_1\tilde{U}_1'+\cdots+\tilde{U}_q\tilde{U}_q')\underline{y}=\hat{\underline{y}}_X-\hat{\underline{y}}_{X_1}$, $P_X(\lambda)$ can be represented only by P_X and P_{X_1}

$$\begin{aligned}P_X(\lambda) &= X\left(X'X+\frac{\lambda}{1-\lambda}\underline{b}_{X_1}\underline{b}_{X_1}'+\cdots+\frac{\lambda}{1-\lambda}\underline{b}_{X_q}\underline{b}_{X_q}'\right)^{-1}X'\\ &= P_{X_1}+(1-\lambda)(\tilde{U}_1\tilde{U}_1'+\cdots+\tilde{U}_q\tilde{U}_q')\\ &= P_X-\lambda(P_X-P_{X_1})\end{aligned}$$

And also $\hat{\underline{y}}_X(\lambda)$ can be computed only through $\hat{\underline{y}}_X$, the fitted vector of $\underline{y}$ on X, and

$\hat{\underline{y}}_{X_1}$, the fitted vector of $\underline{y}$ on X_1,

$$\begin{aligned}\hat{\underline{y}}_X(\lambda) &= P_X(\lambda)\,\underline{y} \\ &= \hat{\underline{y}}_X - \lambda(\hat{\underline{y}}_X - \hat{\underline{y}}_{X_1}).\end{aligned}$$

Thus $\underline{e}(\lambda)$, $\underline{e}_{y\cdot X}(\lambda)$, and $\underline{e}_{V\cdot X}(\lambda)$, residual vectors of $\underline{y}$ on Z, $\underline{y}$ on X, and V on X, respectively, are as follows.

$$\begin{aligned}\underline{e}(\lambda) &= \underline{y} - \hat{\underline{y}}(\lambda) \\ &= \underline{e} + \lambda(\hat{\underline{y}} - \hat{\underline{y}}(1)) \\ &= \underline{e} + \lambda(\underline{e}(1) - \underline{e}),\end{aligned}$$

$$\begin{aligned}\underline{e}_{y\cdot X}(\lambda) &= \underline{y} - \hat{\underline{y}}_X(\lambda) \\ &= \underline{e}_{y\cdot X} + \lambda(\hat{\underline{y}}_X - \hat{\underline{y}}_{X_1}) \\ &= \underline{e}_{y\cdot X} + \lambda(\underline{e}_{y\cdot X_1} - \underline{e}_{y\cdot X}),\end{aligned}$$

and

$$\begin{aligned}\underline{e}_{V\cdot X}(\lambda) &= (I - P_X(\lambda))\,V \\ &= (I - P_X + \lambda(P_X - P_{X_1}))\,V \\ &= \underline{e}_{V\cdot X} + \lambda(\underline{e}_{V\cdot X_1} - \underline{e}_{V\cdot X}).\end{aligned}$$

Therefore, we now can define the AVP (λ) and DAVP (λ), the animated AVP and DAVP, by

$$\begin{aligned}AVP(\lambda) &= \{\underline{e}_{y\cdot X}(\lambda),\ \underline{e}_{V\cdot X}(\lambda)\} \\ &= \{\underline{e}_{y\cdot X} + \lambda(\underline{e}_{y\cdot X_1} - \underline{e}_{y\cdot X}),\ \underline{e}_{V\cdot X} + \lambda(\underline{e}_{V\cdot X_1} - \underline{e}_{V\cdot X})\}\end{aligned}$$

and

$$\begin{aligned}DAVP(\lambda) &= \{\underline{e}(\lambda),\ \underline{e}_{V\cdot X}(\lambda)\} \\ &= \{\underline{e} + \lambda(\underline{e}(1) - \underline{e}),\ \underline{e}_{V\cdot X} + \lambda(\underline{e}_{V\cdot X_1} - \underline{e}_{V\cdot X})\}.\end{aligned}$$

As λ increases from 0 to 1, we can see the two dynamic effects, the supporting and the suppressing effect in the AVP (λ) and DAVP (λ). Here 'X_i supports X_j' means that X_j is significant in the presence of X_i, while X_j is not significant in the omission of X_i. Similarly 'X_i suppresses X_j' means that X_j is not significant in the presence of X_i, while X_j is significant in the omission of X_i. For example, if V gets more insignificant as $U_1, \cdots, U_q$ are smoothly removed, $U_1, \cdots, U_q$ support V. Conversely, if the smooth omission of $U_1, \cdots, U_q$ gives V more significance, V is suppressed by $U_1, \cdots, U_q$.

2.2 Animated $AVP(DAVP)$ in omission of multiple observations

First, we consider the case of omitting an observation smoothly which was well explained in references. Let $\underline{u}_i$ be the vector of zeros except for a single 1 corresponding to the i-th observation. Then the modified mean shift model can be presented by

$$\begin{aligned}\underline{y} &= X_1\underline{\beta}_1 + \underline{u}_i\delta + V\theta + \underline{\varepsilon} \\ &= X_1\underline{\beta}_1^* + \tilde{\underline{u}}_i\delta^* + \tilde{V}\theta^* + \underline{\varepsilon} \\ &= X\tilde{\underline{\beta}}^* + \tilde{V}\theta^* + \underline{\varepsilon} \\ &= Z\underline{\beta}^* + \underline{\varepsilon}\end{aligned}$$

where $\tilde{\underline{u}}_i = \dfrac{Q_X\underline{u}_i}{\|Q_X\underline{u}_i\|}$, $\tilde{V} = \dfrac{Q_{X,\tilde{\underline{u}}_i}V}{\|Q_{X,\tilde{\underline{u}}_i}V\|}$, $X = (X_1 : \tilde{\underline{u}}_i)$, $Z = (X : \tilde{V})$, $\tilde{\underline{\beta}}^{*\prime} = (\underline{\beta}_1^{*\prime}, \delta^*)$,

and $\underline{\beta}^{*\prime} = (\tilde{\underline{\beta}}^{*\prime}, \theta^*)$. For $0 < \lambda \leq 1$, $\hat{\underline{\beta}}(\lambda)$ and $\hat{\underline{y}}(\lambda)$ are obtained as

$$\begin{aligned}\hat{\underline{\beta}}(\lambda) &= (Z'Z + \frac{1-\lambda}{\lambda}\underline{b}\underline{b}')^{-1}Z'\underline{y} \\ &= \begin{pmatrix}(X_1'X_1)^{-1}X_1'\underline{y} \\ \lambda\tilde{\underline{u}}_i'\underline{y} \\ \tilde{V}'\underline{y}\end{pmatrix}\end{aligned}$$

and

$$\begin{aligned}\hat{\underline{y}}(\lambda) &= Z\hat{\underline{\beta}}(\lambda) \\ &= X_1(X_1'X_1)^{-1}X_1'\underline{y} + \lambda\tilde{\underline{u}}_i\tilde{\underline{u}}_i'\underline{y} + \tilde{V}\tilde{V}' \\ &= \hat{\underline{y}} - \lambda(\hat{\underline{y}} - \hat{\underline{y}}(\lambda=1)).\end{aligned}$$

Then we can easily show the animated projection matrix of X

$$\begin{aligned}P_X(\lambda) &= X(X'X + \frac{1-\lambda}{\lambda}\underline{b}_X\underline{b}_X')^{-1}X' \\ &= X\left(\begin{pmatrix}X_1'X_1 & \underline{0} \\ \underline{0}' & 1\end{pmatrix} + \begin{pmatrix}O & \underline{0} \\ \underline{0}' & \frac{1-\lambda}{\lambda}\end{pmatrix}\right)^{-1}X' \\ &= (X_1\ \tilde{\underline{u}}_i)\begin{pmatrix}(X_1'X_1)^{-1} & \underline{0} \\ \underline{0}' & \lambda\end{pmatrix}\begin{pmatrix}X_1' \\ \tilde{\underline{u}}_i'\end{pmatrix} \\ &= P_{X_1} + \lambda\tilde{\underline{u}}_i\tilde{\underline{u}}_i'.\end{aligned}$$

Noting that $\tilde{\underline{u}}_i\tilde{\underline{u}}_i' = P_X(1) - P_{X_1}$ and so $\tilde{\underline{u}}_i\tilde{\underline{u}}_i'\underline{y} = \hat{\underline{y}}_X(\lambda=1) - \hat{\underline{y}}_{X_1}$, $P_X(\lambda)$ can be rewritten simply and $\hat{\underline{y}}_X(\lambda)$ can be computed easily only through $\hat{\underline{y}}_{X_1}$ and $\hat{\underline{y}}_X(1)$,

$$P_X(\lambda) = P_{X_1} - \lambda(P_{X_1} - P_X(1))$$

and

$$\begin{aligned}\hat{\underline{y}}_X(\lambda) &= P_X(\lambda)\underline{y} \\ &= \hat{\underline{y}}_{X_1} - \lambda(\hat{\underline{y}}_{X_1} - \hat{\underline{y}}_X(1)).\end{aligned}$$

We now take into account the case of removing multiple observations at the same time. Suppose that be the set containing the indices of the m observations to be omitted which $I = \{i_1, i_2, \cdots, i_m\}$, $m < (n-p)$ may be thought to be the last observations and

$U_I=\{\underline{u}_{i_1}, \underline{u}_{i_2}, \cdots, \underline{u}_{i_m}\}$ be the matrix containing the corresponding indicator variables. The model is given and modified by

$$\begin{aligned} \underline{y} &= X_1\underline{\beta}_1 + U_I\underline{\delta} + V\theta + \underline{\varepsilon} \\ &= X_1\underline{\beta}_1^* + \widetilde{U}_I\underline{\delta}^* + \widetilde{V}\theta^* + \underline{\varepsilon} \\ &= X\underline{\widetilde{\beta}}^* + \widetilde{V}\theta^* + \underline{\varepsilon} \\ &= Z\underline{\beta}^* + \underline{\varepsilon} \end{aligned}$$

where $X=(X_1, \widetilde{U}_I)$, $Z=(X, \widetilde{V})$, $\underline{\widetilde{\beta}}^{*\prime}=(\underline{\beta}_1^{*\prime}, \delta^{*\prime})$, $\underline{\beta}^{*\prime}=(\underline{\widetilde{\beta}}^{*\prime}, \theta^*)$, and $\widetilde{U}_I=(\underline{\widetilde{u}}_{i_1}, \underline{\widetilde{u}}_{i_2}, \cdots, \underline{\widetilde{u}}_{i_m})$. Here $\underline{\widetilde{u}}_{i_j}=\dfrac{Q_{X, \widetilde{u}_{i_1}, \cdots, \widetilde{u}_{i_{j-1}}}\underline{u}_{i_j}}{\|Q_{X, \widetilde{u}_{i_1}, \cdots, \widetilde{u}_{i_{j-1}}}\underline{u}_{i_j}\|}$, $j\leq m$ and $\widetilde{V}=\dfrac{Q_{X_1, \widetilde{u}_{i_1}, \cdots, \widetilde{u}_{i_m}}V}{\|Q_{X_1, \widetilde{u}_{i_1}, \cdots, \widetilde{u}_{i_m}}V\|}$.

Thus $\widetilde{U}_I$ is the orthonormal basis orthogonal to the column space of X_1. For $0<\lambda\leq 1$ we can obtain $\underline{\widehat{\beta}}(\lambda)$ the estimates of $\underline{\beta}^*$

$$\begin{aligned} \underline{\widehat{\beta}}(\lambda) &= (Z'Z+\frac{1-\lambda}{\lambda}BB')^{-1}Z'\underline{y} \\ &= \begin{pmatrix} (X_1'X_1)^{-1}X_1'\underline{y} \\ \lambda\widetilde{U}_I'\underline{y} \\ \widetilde{V}'\underline{y} \end{pmatrix} \end{aligned}$$

where $B=\begin{pmatrix} O \\ I \end{pmatrix} \begin{matrix} p\times m \\ m\times m \end{matrix}$ and $\underline{\widehat{y}}(\lambda)$ are fitted values in the smooth omission of i_1-th, $\cdots$, i_m-th observations

$$\begin{aligned} \underline{\widehat{y}}(\lambda) &= Z\underline{\widehat{\beta}}(\lambda) \\ &= X_1(X_1'X_1)^{-1}X_1'\underline{y}+\lambda\widetilde{U}_I\widetilde{U}_I'\underline{y}+\widetilde{V}\widetilde{V}'\underline{y} \\ &= \underline{\widehat{y}}_1+\lambda(\underline{\widehat{y}}(1)-\underline{\widehat{y}})+(\underline{\widehat{y}}-\underline{\widehat{y}}_1) \\ &= \underline{\widehat{y}}-\lambda(\underline{\widehat{y}}-\underline{\widehat{y}}(1)). \end{aligned}$$

Noting that $\widetilde{U}_I\widetilde{U}_I'=P_X(1)-P_{X_1}$, we obtain the animated projection matrix of X

where $B_X=\begin{pmatrix} O \\ I \end{pmatrix} \begin{matrix} (p-1)\times m \\ m\times m \end{matrix}$

$$\begin{aligned} P_X(\lambda) &= X(X'X+\frac{1-\lambda}{\lambda}B_XB_X')^{-1}X' \\ &= P_{X_1}+\lambda\widetilde{U}_I\widetilde{U}_I' \\ &= P_{X_1}-\lambda(P_{X_1}-P_X(1)). \end{aligned}$$

And since $\widetilde{U}_I\widetilde{U}_I'\underline{y}=\underline{\widehat{y}}_X(1)-\underline{\widehat{y}}_{X_1}$, it follows that

$$\begin{aligned} \underline{\widehat{y}}_X(\lambda) &= P_X(\lambda)\underline{y} \\ &= \underline{\widehat{y}}_{X_1}-\lambda(\underline{\widehat{y}}_{X_1}-\underline{\widehat{y}}_X(1)). \end{aligned}$$

From these results, we can derive $\underline{e}(\lambda)$, $\underline{e}_{y\cdot X}(\lambda)$, and $\underline{e}_{V\cdot X}(\lambda)$ the animated residual vectors

$$\begin{aligned} \underline{e}(\lambda) &= \underline{y} - \hat{\underline{y}}(\lambda) \\ &= \underline{e} + \lambda(\hat{\underline{y}} - \hat{\underline{y}}(1)) \\ &= \underline{e} + \lambda(\underline{e}(1) - \underline{e}), \end{aligned}$$

$$\begin{aligned} \underline{e}_{y \cdot X}(\lambda) &= \underline{y} - \hat{\underline{y}}_X(\lambda) \\ &= \underline{e}_{y \cdot X_1} + \lambda(\hat{\underline{y}}_{X_1} - \hat{\underline{y}}_X(1)) \\ &= \underline{e}_{y \cdot X_1} + \lambda(\underline{e}_{y \cdot X}(1) - \underline{e}_{y \cdot X_1}) \end{aligned}$$

and

$$\begin{aligned} \underline{e}_{V \cdot X}(\lambda) &= (I - P_X(\lambda))V \\ &= (I - P_{X_1} + \lambda(P_{X_1} - P_X(1)))V \\ &= \underline{e}_{V \cdot X_1} + \lambda(\underline{e}_{V \cdot X}(1) - \underline{e}_{V \cdot X_1}). \end{aligned}$$

Thus the $AVP(\lambda)$ and $DAVP(\lambda)$ for smoothly removed observations are defined by

$$\begin{aligned} AVP(\lambda) &= \{\underline{e}_{y \cdot X}(\lambda),\ \underline{e}_{V \cdot X}(\lambda)\} \\ &= \{\underline{e}_{y \cdot X_1} + \lambda(\underline{e}_{y \cdot X}(1) - \underline{e}_{y \cdot X_1}),\ \underline{e}_{V \cdot X_1} + \lambda(\underline{e}_{V \cdot X}(1) - \underline{e}_{V \cdot X_1})\} \end{aligned}$$

and

$$\begin{aligned} DAVP(\lambda) &= \{\underline{e}(\lambda),\ \underline{e}_{V \cdot X}(\lambda)\} \\ &= \{\underline{e} + \lambda(\underline{e}(1) - \underline{e}),\ \underline{e}_{V \cdot X_1} + \lambda(\underline{e}_{V \cdot X}(1) - \underline{e}_{V \cdot X_1})\}. \end{aligned}$$

As λ increases from 0 to 1, we can visually understand the effect of removing multiple observations on a variable of interest with assistance of the $AVP(\lambda)$ and $DAVP(\lambda)$.

2.3 Animated $AVP(DAVP)$ in simultaneous omission of multiple variables and observations

In this section, we will concentrate upon the subject of the joint impact of the simultaneous omission of variables and observations on the other variables in a regression equation. The animated AVP or $DAVP$ will be then a diagnostic device for helping to detect the joint impact visually.

We begin with the study of the effect of omitting one variable and one observation smoothly. For this study, the following model is considered (Park and Kim (1995))

$$\begin{aligned} \underline{y} &= X_1 \underline{\beta}_1 + X_2 \beta_2 + \underline{u}_i \delta + V\theta + \underline{\varepsilon} \\ &= X_1 \underline{\beta}_1^* + \tilde{X}_2 \beta_2^* + \tilde{\underline{u}}_i \delta^* + \tilde{V}\theta^* + \underline{\varepsilon} \\ &= X \tilde{\underline{\beta}}^* + \tilde{V}\theta^* + \underline{\varepsilon} \\ &= Z \underline{\beta}^* + \underline{\varepsilon} \end{aligned}$$

where $\tilde{X}_2 = \dfrac{Q_{X_1} X_2}{\|Q_{X_1} X_2\|}$, $\tilde{\underline{u}}_i = \dfrac{Q_{X_1, \tilde{X}_2} \underline{u}_i}{\|Q_{X_1, \tilde{X}_2} \underline{u}_i\|}$, $\tilde{V} = \dfrac{Q_{X_1, \tilde{X}_2, \tilde{\underline{u}}_i} V}{\|Q_{X_1, \tilde{X}_2, \tilde{\underline{u}}_i} V\|}$, $X = (X_1 \vdots \tilde{X}_2 \vdots \tilde{\underline{u}}_i)$,

$Z = (X \vdots \tilde{V})$, $\tilde{\underline{\beta}}^{*\prime} = (\underline{\beta}_1^{*\prime}, \beta_2^*, \delta^*)$, $\underline{\beta}^{*\prime} = (\tilde{\underline{\beta}}^{*\prime}, \theta^*)$. For $0 < \lambda < 1$ we can easily obtain

$\hat{\underset{\sim}{\beta}}(\lambda)$, animated estimates of $\underset{\sim}{\beta}^{*\prime}$, and $\hat{\underset{\sim}{y}}(\lambda)$, animated fitted values by

$$\begin{aligned}\hat{\underset{\sim}{\beta}}(\lambda) &= (Z'Z+\frac{\lambda}{1-\lambda}\underset{\sim}{b}_1\underset{\sim}{b}_1'+\frac{1-\lambda}{\lambda}\underset{\sim}{b}_2\underset{\sim}{b}_2')^{-1}Z'\underset{\sim}{y} \\ &= \begin{pmatrix} (X_1'X_1)^{-1}X_1'\underset{\sim}{y} \\ (1-\lambda)\tilde{X}_2'\underset{\sim}{y} \\ \lambda\tilde{u}_i'\underset{\sim}{y} \\ \tilde{V}'\underset{\sim}{y}\end{pmatrix}\end{aligned}$$

and

$$\begin{aligned}\hat{\underset{\sim}{y}}(\lambda) &= Z\hat{\underset{\sim}{\beta}}(\lambda) \\ &= X_1(X_1'X_1)^{-1}X_1'\underset{\sim}{y}+(1-\lambda)\tilde{X}_2\tilde{X}_2'\underset{\sim}{y}+\lambda\tilde{u}_i\tilde{u}_i'\underset{\sim}{y}+\tilde{V}\tilde{V}'\underset{\sim}{y} \\ &= \hat{\underset{\sim}{y}}_1+(1-\lambda)(\hat{\underset{\sim}{y}}-\hat{\underset{\sim}{y}}_1-\tilde{V}\tilde{V}'\underset{\sim}{y})+\lambda(\hat{\underset{\sim}{y}}(1)-\hat{\underset{\sim}{y}}_1-\tilde{V}\tilde{V}'\underset{\sim}{y})+\tilde{V}\tilde{V}'\underset{\sim}{y} \\ &= \hat{\underset{\sim}{y}}-\lambda(\hat{\underset{\sim}{y}}-\hat{\underset{\sim}{y}}(1)).\end{aligned}$$

If we direct our attention to the facts that $\tilde{\underset{\sim}{u}}_i\tilde{\underset{\sim}{u}}_i' = P_X(1)-P_{X_1}$ and $\tilde{X}_2\tilde{X}_2' = P_X(0)-P_{X_1}$, it is easily derived that $\tilde{\underset{\sim}{u}}_i\tilde{\underset{\sim}{u}}_i'\underset{\sim}{y} = \hat{\underset{\sim}{y}}_X(1)-\hat{\underset{\sim}{y}}_{X_1}$ and $\tilde{X}_2\tilde{X}_2'\underset{\sim}{y} = \hat{\underset{\sim}{y}}_X(0)-\hat{\underset{\sim}{y}}_{X_1}$. Using these, we get $P_X(\lambda)$, the animated projection matrix of X

$$\begin{aligned}P_X(\lambda) &= X(X'X+\frac{\lambda}{1-\lambda}\underset{\sim}{b}_{X_1}\underset{\sim}{b}_{X_1}'+\frac{1-\lambda}{\lambda}\underset{\sim}{b}_{X_2}\underset{\sim}{b}_{X_2}')^{-1}X' \\ &= P_{X_1}+(1-\lambda)\tilde{X}_2\tilde{X}_2'+\lambda\tilde{u}_i\tilde{u}_i' \\ &= P_{X_1}+(1-\lambda)(P_X(0)-P_{X_1})+\lambda(P_X(1)-P_{X_1}) \\ &= P_X(0)-\lambda(P_X(0)-P_X(1)).\end{aligned}$$

And the fitted values with animating parameter λ is

$$\begin{aligned}\hat{\underset{\sim}{y}}_X(\lambda) &= P_X(\lambda)\underset{\sim}{y} \\ &= \hat{\underset{\sim}{y}}_X(0)-\lambda(\hat{\underset{\sim}{y}}_X(0)-\hat{\underset{\sim}{y}}_X(1)).\end{aligned}$$

Now it is of interest to study the animated plots of showing the effect of removing multiple variables and observations simultaneously. Suppose that the full model is

$$\begin{aligned}\underset{\sim}{y} &= X_1\underset{\sim}{\beta}_1+U_1\delta_1+,\cdots,+U_q\delta_q+V\theta+\underset{\sim}{\varepsilon} \\ &= X_1\underset{\sim}{\beta}_1+U\underset{\sim}{\delta}+V\theta+\underset{\sim}{\varepsilon}\end{aligned}$$

where $U=(U_1,\cdots,U_q)$ is an $n\times q$ matrix of which column vectors are variables to be deleted and $\underset{\sim}{\delta}'=(\delta_1,\cdots,\delta_q)$ is the $1\times q$ vector of their coefficients. Assuming that i_1-th, $\cdots$, i_m-th observations are omitted, we obtain the mean-shift modified model

$$
\begin{aligned}
\underline{y} &= X_1\underline{\beta}_1 + U\underline{\delta} + U_I\underline{\gamma} + V\theta + \underline{\varepsilon} \\
&= X_1\underline{\beta}_1^* + \tilde{U}\underline{\delta}^* + \tilde{U}_I\underline{\gamma}^* + \tilde{V}\theta^* + \underline{\varepsilon} \\
&= X\tilde{\underline{\beta}}^* + \tilde{V}\theta^* + \underline{\varepsilon} \\
&= Z\underline{\beta}^* + \underline{\varepsilon}
\end{aligned}
$$

where $\tilde{U}=(\tilde{U}_1,\cdots,\tilde{U}_q)$ is the orthonormal basis to X_1 and $\tilde{U}_I=(\tilde{\underline{u}}_{i_1},\cdots,\tilde{\underline{u}}_{i_m})$ is the orthonormal basis to X_1, $\tilde{U}$ and $X=(X_1,\tilde{U},\tilde{U}_I)$, $Z=(X,\tilde{V})$, $\tilde{\underline{\beta}}^{*\prime}=(\underline{\beta}_1^{*\prime},\underline{\delta}^{*\prime},\underline{\gamma}^{*\prime})$, $\underline{\beta}^{*\prime}=(\tilde{\underline{\beta}}_1^{*\prime},\theta^*)$.

For $0<\lambda<1$ we estimate $\underline{\beta}^*$ by

$$
\begin{aligned}
\hat{\underline{\beta}}(\lambda) &= \left(Z'Z+\frac{\lambda}{1-\lambda}\underline{B}_1\underline{B}_1'+\frac{1-\lambda}{\lambda}\underline{B}_2\underline{B}_2'\right)^{-1}Z'\underline{y} \\
&= \begin{pmatrix} (X_1'X_1)^{-1}X_1'\underline{y} \\ (1-\lambda)\tilde{U}'\underline{y} \\ \lambda\tilde{U}_I'\underline{y} \\ \tilde{V}'\underline{y} \end{pmatrix}.
\end{aligned}
$$

And the animated fitted values are

$$
\begin{aligned}
\hat{\underline{y}}(\lambda) &= Z\hat{\underline{\beta}}(\lambda) \\
&= X_1(X_1'X_1)^{-1}X_1'\underline{y}+(1-\lambda)\tilde{U}\tilde{U}'\underline{y}+\lambda\tilde{U}_I\tilde{U}_I'\underline{y}+\tilde{V}\tilde{V}'\underline{y} \\
&= \hat{\underline{y}}_1+(1-\lambda)(\hat{\underline{y}}-\hat{\underline{y}}_1-\tilde{V}\tilde{V}'\underline{y})+\lambda(\hat{\underline{y}}(1)-\hat{\underline{y}}_1-\tilde{V}\tilde{V}'\underline{y})+\tilde{V}\tilde{V}'\underline{y} \\
&= \hat{\underline{y}}-\lambda(\hat{\underline{y}}-\hat{\underline{y}}(1)).
\end{aligned}
$$

We note that $\tilde{U}\tilde{U}'=P_X(0)-P_{X_1}$ and $\tilde{U}_I\tilde{U}_I'=P_X(1)-P_{X_1}$. From this fact, we obtain

$$
\begin{aligned}
P_X(\lambda) &= X\left(X'X+\frac{\lambda}{1-\lambda}\underline{B}_{X_1}\underline{B}_{X_1}'+\frac{1-\lambda}{\lambda}\underline{B}_{X_2}\underline{B}_{X_2}'\right)^{-1}X' \\
&= P_{X_1}+(1-\lambda)\tilde{U}\tilde{U}'+\lambda\tilde{U}_I\tilde{U}_I' \\
&= P_{X_1}+(1-\lambda)(P_X(0)-P_{X_1})+\lambda(P_X(1)-P_{X_1}) \\
&= P_X(0)-\lambda(P_X(0)-P_X(1)).
\end{aligned}
$$

Since $\tilde{U}\tilde{U}'\underline{y}=\hat{\underline{y}}_X(0)-\hat{\underline{y}}_{X_1}$ and $\tilde{U}_I\tilde{U}_I'\underline{y}=\hat{\underline{y}}_X(\lambda=1)-\hat{\underline{y}}_{X_1}$,

$$
\begin{aligned}
\hat{\underline{y}}_X(\lambda) &= P_X(\lambda)\underline{y} \\
&= \hat{\underline{y}}_X(0)-\lambda(\hat{\underline{y}}_X(0)-\hat{\underline{y}}_X(1))
\end{aligned}
$$

can be computed only through $\hat{\underline{y}}_X(0)$ and $\hat{\underline{y}}_X(1)$. Thus $\underline{e}(\lambda)$, $\underline{e}_{y\cdot X}(\lambda)$, and $\underline{e}_{V\cdot X}(\lambda)$ are derived into

$$\begin{aligned}\underline{e}(\lambda) &= \underline{y} - \hat{\underline{y}}(\lambda)\\ &= \underline{e} + \lambda(\hat{\underline{y}} - \hat{\underline{y}}(1))\\ &= \underline{e} + \lambda(\underline{e}(1) - \underline{e}),\end{aligned}$$

$$\begin{aligned}\underline{e}_{\underline{y}\cdot X}(\lambda) &= \underline{y} - \hat{\underline{y}}_X(\lambda)\\ &= \underline{e}_{\underline{y}\cdot X}(0) + \lambda(\hat{\underline{y}}_X(0) - \hat{\underline{y}}_X(1))\\ &= \underline{e}_{\underline{y}\cdot X}(0) + \lambda(\underline{e}_{\underline{y}\cdot X}(1) - \underline{e}_{\underline{y}\cdot X}(0)),\end{aligned}$$

and

$$\begin{aligned}\underline{e}_{V\cdot X}(\lambda) &= (I - P_X(\lambda))V\\ &= (I - P_X(0) + \lambda(P_X(0) - P_X(1)))V\\ &= \underline{e}_{V\cdot X}(0) + \lambda(\underline{e}_{V\cdot X}(1) - \underline{e}_{V\cdot X}(0)).\end{aligned}$$

Here it is noticed that $\underline{e}(\lambda)$ can be computed simply from the residuals $\underline{e} = \underline{e}(0)$ from the full model and $\underline{e}(1)$ from that of $U_1, \cdots, U_q$ variables and i_1-th, $\cdots$, i_m-th observations removed. Likewise $\underline{e}_{\underline{y}\cdot X}(\lambda)$ and $\underline{e}_{V\cdot X}(\lambda)$ can be obtained without complicated calculations.

Thus we define the animated AVP and DAVP in the omission of multiple variables and observations by

$$\begin{aligned}AVP(\lambda) &= \{\underline{e}_{\underline{y}\cdot X}(\lambda),\ \underline{e}_{V\cdot X}(\lambda)\}\\ &= \{\underline{e}_{\underline{y}\cdot X}(0) + \lambda(\underline{e}_{\underline{y}\cdot X}(1) - \underline{e}_{\underline{y}\cdot X}(0)), \underline{e}_{V\cdot X}(0) + \lambda(\underline{e}_{V\cdot X}(1) - \underline{e}_{V\cdot X}(0))\}\end{aligned}$$

and

$$\begin{aligned}DAVP(\lambda) &= \{\underline{e}(\lambda),\ \underline{e}_{V\cdot X}(\lambda)\}\\ &= \{\underline{e} + \lambda(\underline{e}(1) - \underline{e}),\ \underline{e}_{V\cdot X}(0) + \lambda(\underline{e}_{V\cdot X}(1) - \underline{e}_{V\cdot X}(0))\}\end{aligned}$$

As λ increases from 0 to 1, the $AVP(\lambda)$ and $DAVP(\lambda)$ will show dynamically the joint impacts of removing multiple variables and observations on a specific variable.

3. Illustrated example

Until now we presented $AVP(\lambda)$and $DAVP(\lambda)$as a graphical tool of understanding dynamic effects of smoothly removing variables and observations separately or simultaneously from a model on a specific variable.

In this section, we will illustrate an example to show how these dynamic graphical methods can be applied in the practical aspects. As a device of obtaining the animated plots, we developed a computer program using S-plus package. We analyze the artificial data in Chatterjee and Hadi (1989) to detect the effect of omitting some variables on the other variable. The data set consists of 20 observations and 6 explanatory variables, which is presented in Table 1.

Table 1 Artificial Data

No.	x_1	x_2	x_3	x_4	x_5	x_6	y
1	1	-1	0	1.95	1	-0.59	0.14
2	2	-2	0	3.62	2	0.06	3.03
3	3	-3	0	5.66	3	-0.30	3.11
4	4	-4	0	7.08	4	-0.42	3.42
5	5	-5	0	10.41	5	-1.54	6.00
6	1	0	0	-0.26	-1	-0.06	-1.62
7	2	0	0	-0.32	-2	0.89	-0.92
8	3	0	0	0.62	-3	0.87	-0.07
9	4	0	0	-0.49	-4	4.16	0.63
10	5	0	0	-1.14	-5	4.31	0.43
11	0	1	1	-1.26	0	1.60	1.07
12	0	2	2	-0.53	0	1.85	1.92
13	0	3	3	0.15	0	2.96	3.36
14	0	4	4	0.03	0	4.39	3.95
15	0	5	5	-1.70	0	4.66	4.47
16	0	-1	1	-0.16	-1	-0.65	-3.47
17	0	-2	2	-2.29	-2	-2.31	-3.96
18	0	-3	3	-4.55	-3	-3.11	-4.68
19	0	-4	4	-3.75	-4	-2.68	-8.56
20	0	-5	5	-5.46	-5	-4.21	-9.99

The following model

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_6 x_6 + \varepsilon$$

is fitted to the data. Table 2 shows the variables by which each variable is affected in the single or joint form.

Table 2 Influential relationship of variables in Artificial Data

Supported/ Suppressed Variable (X_j)	Supporting Variable (X_i)		Suppressing Variable (X_i)	
	Single	joint	Single	joint
X_1	X_2, X_5	X_2X_3, X_2X_5, X_2X_6 X_3X_5, X_4X_5, X_5X_6		
X_2	X_1, X_5	X_1X_3, X_1X_4, X_3X_5 X_4X_5		
X_3				
X_4			X_5	X_1X_5, X_1X_6, X_2X_5 X_3X_5, X_5X_6
X_5	X_1	X_1X_6, X_2X_6		
X_6			X_1, X_2	X_1X_2, X_1X_3, X_1X_4 X_2X_3, X_2X_4, X_2X_5

In Table 2, 6 variables can be roughly divided into 2 groups. One group is the set of supported variables and the other is the set of suppressed variables. The former group contains X_1, X_2, X_5 and the latter contains X_4, X_6, X_3 is not affected by any variable. Table 2 also shows that X_1 is supported separately and jointly by X_2 and X_5, and X_6 is suppressed separately and jointly by X_1 and X_2. However X_2 is supported only separately, not jointly by X_1 and X_5. X_5 is supported jointly, but not separately by X_2 and X_6. Also we note that X_4 is suppressed jointly, but not separately by X_1 and X_6.

Now we present the $AVP(\lambda)$ and $DAVP(\lambda)$ as an aid for better examining the data structure in Table 1. Figures 1 and 2 show that X_6 becomes significant as each of X_1 and X_2 is removed smoothly. Also we can see in Figure 3 that the suppressed significance of X_6 is revived gradually as X_1 and X_2 are jointly removed with smoothness. Figure 4-6 display the opposite case, that is, one variable is supporting another. The omission of X_2 or X_6 has little effect on X_5. But X_5 becomes less and less significant as X_2 and X_6 are jointly removed and after all, not significant. By looking into the $DAVP(\lambda)$ in Figure 7, we can know that the linearity of X_5 in the full model is not excessively trended.

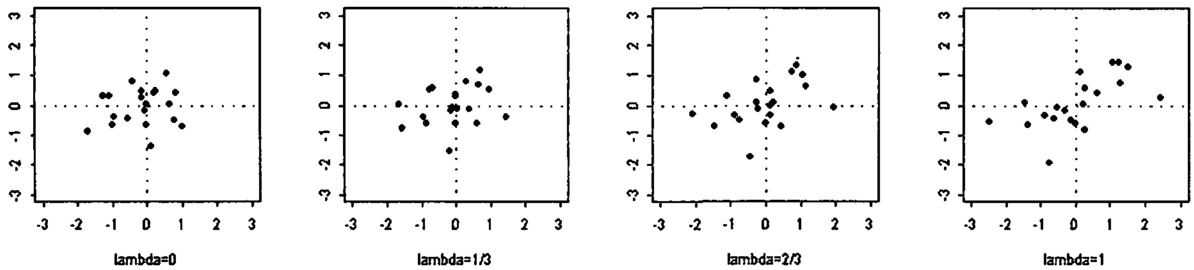

Fig. 1. $AVP(\lambda)$ of X_6 in omission of X_1.

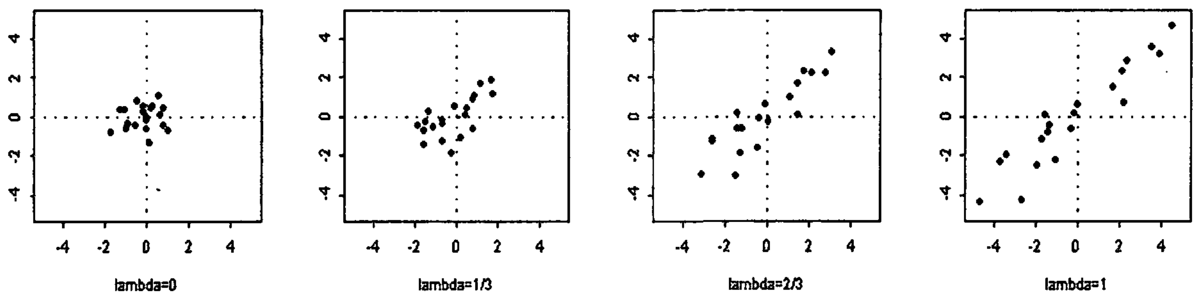

Fig. 2. $AVP(\lambda)$ of X_6 in omission of X_2.

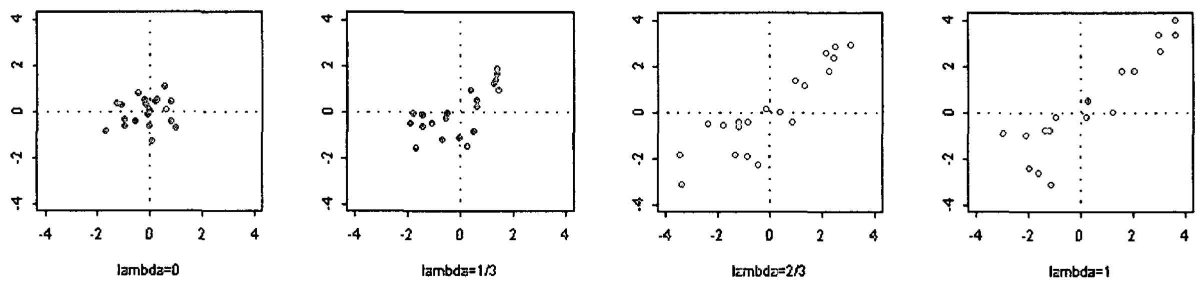

Fig. 3. $AVP(\lambda)$ of X_6 in joint omission of X_1 and X_2.

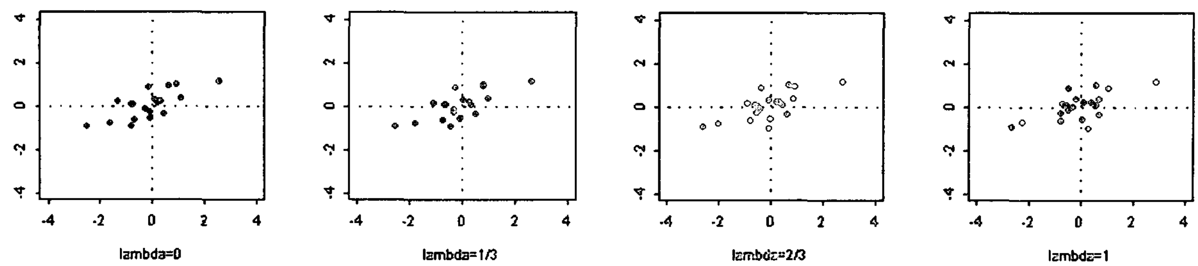

Fig. 4. $AVP(\lambda)$ of X_5 in omission of X_2.

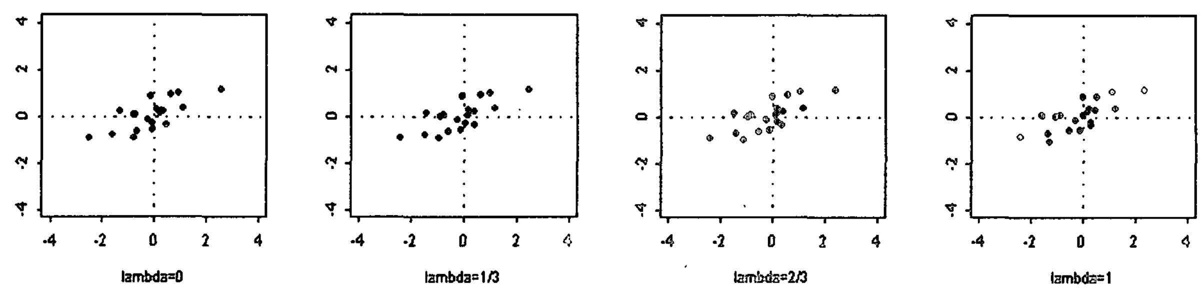

Fig. 5. $AVP(\lambda)$ of X_5 in omission of X_6.

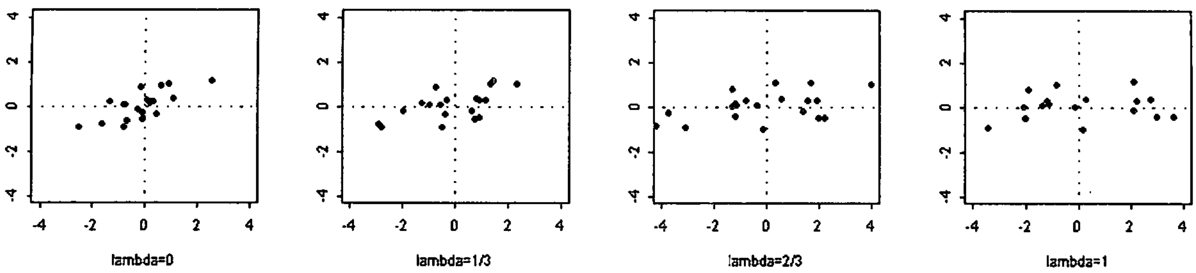

Fig. 6. $AVP(\lambda)$ of X_5 in joint omission of X_2 and X_6.

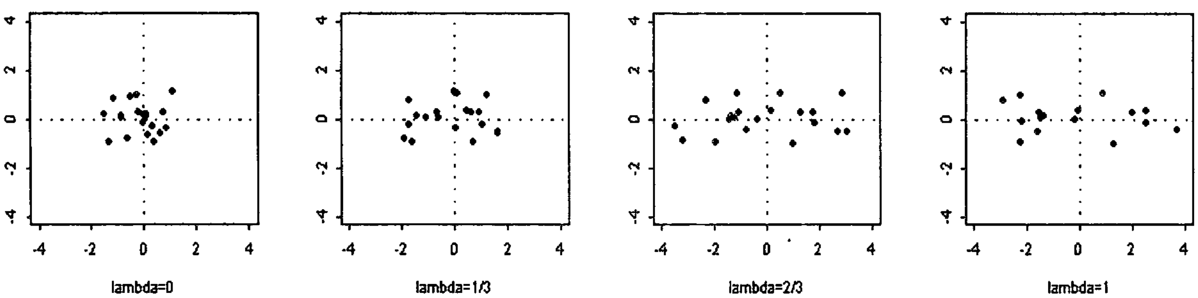

Fig. 7. $DAVP(\lambda)$ of X_5 in joint omission of X_2 and X_6.

Now we examine the joint effect of variables and observations on a specific variable by smoothly omitting both variables and observations. When we first analyze the influence of each observation on an individual variable according to the diagnostic statistics of 'Dfbetas', the 18-th observation has the opposite effects on X_5 and X_6. Figure 8 displays that as the 18-th observation gets gradually deleted, X_5 is losing the significance, that is, X_5 is supported by the 18-th observation. On the other hand, we see in Figure 9 that X_6 becomes significant together with the omission of the 18-th observation, namely, X_6 is suppressed by the 18-th observation. Figure 10 and 11 show the joint effects of smoothly omitting X_1 and the 18-th observation on X_5 and X_6. In the same manner that each of X_1 and the 18-th observation supports X_5, the joint of them also has the supporting power over X_5. Likewise, X_1 and the 18-th observation suppress X_6 in the joint form.

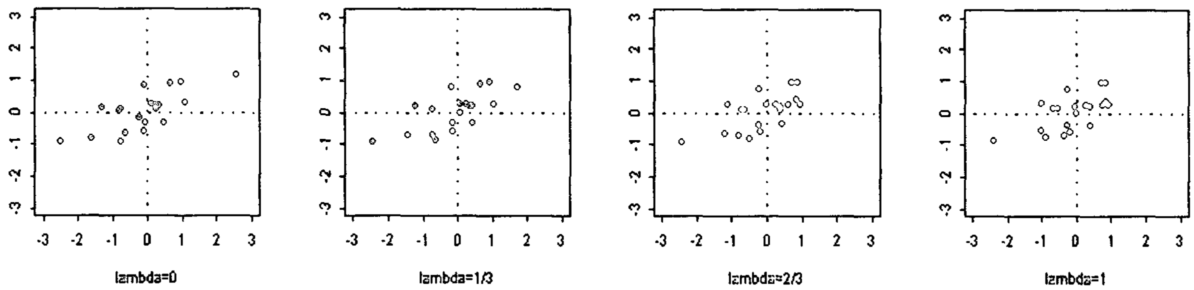

Fig. 8. $AVP(\lambda)$ of X_5 in omission of the 18-th observation.

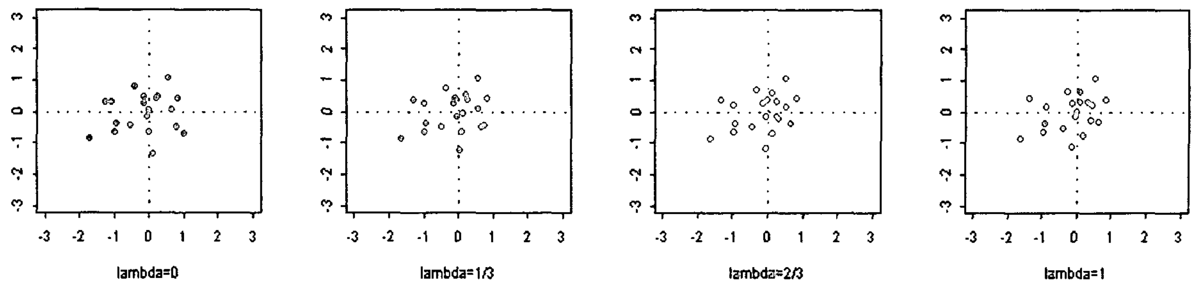

Fig. 9. $AVP(\lambda)$ of X_6 in omission of the 18-th observation.

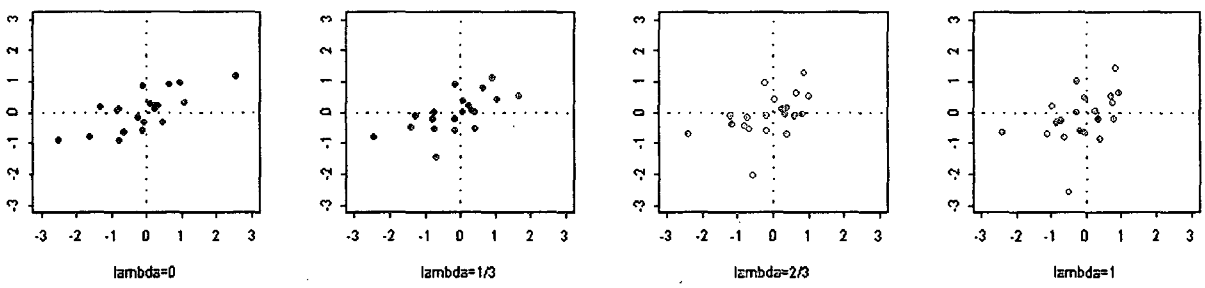

Fig. 10. $AVP(\lambda)$ of X_5 in joint omission of X_1 and the 18-th observation.

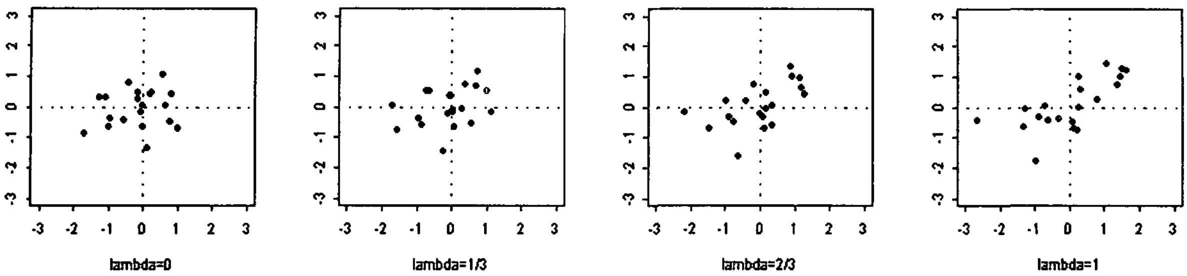

Fig. 11. $AVP(\lambda)$ of X_6 in joint omission of X_1 and the 18-th observation.

Remark. The S-program can be obtained from the authors on request.

References

[1] Becker, R. A., Cleveland, W. S., and Wilks, A. R. (1981). Dynamic graphics for data analysis (with discussion), *Statistical Science*, 2, 355-395.

[2] Belsley, D. A., Kuh, E., and Welsch, R. E. (1980). *Regression diagnostics: identifying influential data and sources of collinearity,* Wiley, New York.

[3] Chambers, J. M., Cleveland, W., Kleiner, B., and Tukey, P. (1983). *Graphical methods for data analysis,* Monterey, CA:Wadsworth.

[4] Chatterjee, S., and Hadi, A. S. (1989). *Sensitivity Analysis in Linear Regression,* John Wiley & Sons, New York.

[5] Cleveland, W. S. (1987). Research in statistical graphics, *Journal of the American Statistical Association*, 79, 807-822.

[6] Cook, R. D. and Weisberg, S. (1982). *Residuals and influence in regression,* Chapman & Hall, New York.

[7] Cook, R. D. and Weisberg, S. (1989). Regression diagnostics with dynamic graphics (with discussion), *Technometrics*, 31, 277-310.

[8] Frobes, J. D. (1857). Further experiments and remarks on the measurement of height by the boiling point of water, *Transactions of the Royal Society of Edinburgh,* 21, 135-143.

[9] Park, S. H. and Kim, S. S. (1995). Dynamic plots for displaying the roles of variables and observations in regression model, *Computational Statistics and Data Analysis,* 19, 401-418.

[10] Park, S. H. and Kim, Y. H. (1992). Regression diagnostics for removing an observation with animating graphics, *Statistical Papers,* 33, 227-240.