Journal of the KSQC
Vol. 19, No. 2, Dec.1991

1991년 12월 韓國品質管理學會誌 제19권 제2호

A DYNAMIC GRAPHICAL METHOD FOR REGRESSION DIAGNOSTICS

Sung H. Park You H. Kim

ABSTRACT

Recently, Cook and Weisberg(1989) presented dynamic graphics for regression diagnostics. They suggested animating graphics which could aid to understanding the effects of adding a variable to a model. In this paper, using the Cook and Weisberg's idea of animation, we propose a dynamic graphical method for residuals to display the effects of removing an observation from a model. Based on the information obtained from these animating graphics, it is possible to see the influence of outliers or influencial observations for regression diagnostics.

KEY WORDS : Regrssion diagnostics ; Dynamic graphics ; Animation ; Outliers ; Residuals.

1. INTRODUCTION

Recently, graphical techniques become an essential part of statistical methodology. One of the important graphics in regression analysis is residual plot. In regression analysis the plotting of residuals versus the independent variables of predicted value has been recommended by Draper and Smith

(1966) and Cox and Snell(1969). These plots help to detect outliers, to assess the presence of inhomogeneity of variance, and to check model adequacy. Larsen and McCleary (1972) introduced partial residual plots, and Mallows(1986) introduced augmented partial residual plots, which can detect the importance of each independent variable and assess some nonlinearity or necessary transformation of varible.

Recent surveys of graphical methods were given by Chambers, Cleveland, Kleiner, and Tukey (1983), Atkinson(1985), and Cleveland(1987). For the purpose of regression diagnostics, Cook and Weisberg(1989) recently introduced dynamic statistical graphics. They considered uses and interpretation of two recently proposed types of dynamic displays, rotation and animation, in regression diagnostics. Some of issues that they addressed by using dynamic graphics include adding predictors to a model, assessing the need to transform, and checking for interactions and normality. They used animation to show dynamic effects of adding a variable to a model and mentioned idea of simultaneously adding variables to a model.

Assume the general linear model

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

where y is the $n \times 1$ vector of observations, X is the $n \times p$ known design matrix, β is the $p \times 1$ vector of unkown regression parameters, and ε is the $n \times 1$ vector of errors, which are identically and independently distributed with $N(0, \sigma^2)$. And X consists of X_1 and X_2 where X_1 is $n \times (p-1)$ matrix, and X_2 is $n \times 1$ matrix, that is, $X = (X_1, X_2)$.

The basic idea of Cool and Weisberg (1989) is to begin with the model $y = X_1\beta_1 + \varepsilon$ and then smoothly add X_2, ending with a fit of the full model $y = X_1\beta_1 + X_2\beta_2 + \varepsilon$ where β_1 is $(p-1) \times 1$ vector and β_2 is an unknown scalar. Since the animated plot that they proposed involves only fitted values and residuals, they worked in terms of a modified version of the full model (1) given by

$$\begin{aligned} y &= X\beta + \varepsilon \\ &= X_1\beta_1 + X_2\beta_2 + \varepsilon \\ &= X_1\beta_1^* + \widetilde{X}_2\beta_2^* + \varepsilon \\ &= Z\beta^* + \varepsilon \end{aligned} \tag{2}$$

where $\widetilde{X}_2 = Q_1X_2/\| Q_1X_2 \|$ is the part of X_2 orthogonal to X_1, normalized to unit length, $Q_1 = I - P_1$, $P_1 = X_1(X_1'X_1)^{-1}X_1'$, $Z = (X_1, \widetilde{X}_2)$, and $\beta^* = \binom{\beta_1^*}{\beta_2^*}$.

Next for each $0 < \lambda \leq 1$, they estimate β^* by

$$\hat{\beta}_\lambda = (Z'Z + \frac{1-\lambda}{\lambda} bb')^{-1} Z'y \quad (3)$$

where b is a $p \times 1$ vector of zeros except for single 1 corresponding to X_2. Since

$$(Z'Z + \frac{1-\lambda}{\lambda} bb')^{-1} = \begin{pmatrix} X_1'X_1 & 0 \\ 0' & \tilde{X}_2'\tilde{X}_2 + \frac{1-\lambda}{\lambda} \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} X_1'X_1 & 0 \\ 0' & \frac{1}{\lambda} \end{pmatrix}^{-1}$$

we can obtain

$$\hat{\beta}_\lambda = \begin{bmatrix} (X_1'X_1)^{-1} X_1'y \\ \lambda \tilde{X}_2'y \end{bmatrix}$$

So as λ tends to 0, (3) corresponds to the regression of y on X_1 alone. And if $\lambda = 1$, then (3) corresponds to the ordinary least squares regression of y on X_1 and X_2. Thus as λ increases from 0 to 1, $\hat{\beta}_\lambda$ represents a smooth sequence of estimators that add X_2 to the model, and an animated plot of $e(\lambda)$ versus $\hat{y}(\lambda)$, where $e(\lambda) = y - \hat{y}(\lambda)$ and $\hat{y}(\lambda) = Z\hat{\beta}_\lambda$, gives a dynamic view of the effects of adding X_2 to the model that already includes X_1.

In this paper, using the Cook and Weisberg's idea of animation, we want to propose an animating graphical method to display the effects of removing an outlier from a model for regression diagnostic purpose.

2. REMOVING AN OBSERVATION WITH ANIMATION

Using the idea of Cook and Weisberg(1989) for adding a variable with animation, we want to view dynamic effects of removing the i^{th} observation from the model (1). First, consider the mean shift model $y = X\beta + \Upsilon_i^* u_i + \varepsilon$ where u_i is the vector of zeros except for single 1 corresponding to the i^{th} observation. We can work in terms of a modified version of the mean shift model given by

$$\begin{aligned} y &= X\beta + \Upsilon_i^* u_i + \varepsilon \\ &= Z\beta^* + \varepsilon \end{aligned} \quad (4)$$

where $\tilde{u}_i = Q_x u_i / \| Q_x u_i \|$ is the orthogonal part of u_i to X normalized to unit length, $Q_x = I - P_x$, $P_x = X(X_1'X_1)^{-1}X_1'$, $Z = (X, u_i)$, and $\beta^* = \binom{\beta}{\gamma_i^*}$. And then for each $0 < \lambda \leq 1$, we estimate β^* by

$$\hat{\beta}_\lambda = (Z'Z + \frac{1-\lambda}{\lambda} bb')^{-1} Z'y \tag{5}$$

where b is the same vector of u_i.

Now we can think some properties of $\hat{\beta}_\lambda$. First, without loss of generality, we take X and y as the form $X = \binom{X_{(i)}}{x_i'}$ and $y = \binom{y_{(i)}}{y_i}$ where x_i' is the i^{th} row vector of X, $X_{(i)}$ is the matrix X without the i^{th} row, and $y_{(i)}$ is the vector y without y_i. That is, put the i^{th} observation to the bottom and so u_i and b become vectors of zeros except for last 1. Then since

$$(Z'Z + \frac{1-\lambda}{\lambda} bb')^{-1} = \begin{pmatrix} X'X & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} (X'X)^{-1} & 0 \\ 0 & \lambda \end{pmatrix}$$

and

$$Zy = \begin{pmatrix} X_1' y \\ u_i' y \end{pmatrix},$$

we can obtain that

$$\hat{\beta}_\lambda = \begin{pmatrix} \hat{\beta} \\ \hat{\gamma}_i^* \end{pmatrix}$$

$$= \begin{pmatrix} (X_1' X_1)^{-1} \quad X_1' y \\ \lambda u_i' y \end{pmatrix}$$

and

$$\hat{y}(\lambda) = Z\hat{\beta}_\lambda$$

$$= X(X'X)^{-1}X'y + \lambda \tilde{u}_i \tilde{u}_i y.$$

Hence at $\lambda = 0, \hat{y}(\lambda) = X(X'X)^{-1}X'y$ is the predicted vector of observed values for the full model by the method of ordinary least squares. And at $\lambda = 1$, we can get the following lemma.

Lemma 1.

$$\hat{y}(\lambda=1) = X(X'X)^{-1}X'y + \tilde{u}_i\tilde{u}_i'y.$$

$$= \begin{pmatrix} X_{(i)}\hat{\beta}_{(i)} \\ y_i \end{pmatrix}$$

Proof : Using the fact

$$(X'X)^{-1} = (X'_{(i)}X_{(i)} + x_ix_i')^{-1}$$

$$= (X'_{(i)}X_{(i)})^{-1} - \frac{(X'_{(i)}X_{(i)})^{-1}x_ix'_i(X'_{(i)}X_{(i)})^{-1}}{1 + t_{ii}}$$

where

$$t_{ii} = x_i'(X'_{(i)}X_{(i)})^{-1}x_i,$$

we can obtain

$$P_X = X(X'X)^{-1}X'$$

$$= \begin{pmatrix} X_{(i)} \\ x_i' \end{pmatrix} \left((X'_{(i)}X_{(i)})^{-1} - \frac{(X'_{(i)}X_{(i)}) + x_ix_i'(X'_{(i)}X_{(i)})^{-1}}{1 + t_{ii}} \right) (X'_{(i)}\ x_i)$$

$$= \begin{pmatrix} (X_{(i)}X'_{(i)}X_{(i)})^{-1}X'_{(i)} - \frac{(X_{(i)}(X'_{(i)}X_{(i)})^{-1}x_ix_i'(X'_{(i)}X_{(i)})^{-1}X'_{(i)}}{1 + t_{ii}} & \frac{X_{(i)}(X'_{(i)}X_{(i)})^{-1}x_i}{1 + t_{ii}} \\ \frac{x_{(i)}'(X_{(i)}'X_{(i)})^{-1}X_{(i)}'}{1 + t_{ii}} & \frac{t_{ii}}{1+t_{ii}} \end{pmatrix}$$

and

$$P_{XY} = X(X'X)^{-1}X'y$$

$$\begin{pmatrix} X_{(i)}\hat{\beta}_{(i)} - \frac{1}{1 - t_{ii}}(X_{(i)}(X'_{(i)}X_{(i)})^{-1}x_ix_i'\hat{\beta}_{(i)} - X_{(i)}(X_{(i)}'X_{(i)})^{-1}x_iy_i) \\ \frac{1}{1 + t_{ii}}(x_{(i)}'\hat{\beta}_{(i)} + t_{ii}y_i) \end{pmatrix}$$

where $\hat{\beta}_{(i)} = (X'_{(i)}X_{(i)})^{-1}X'_{(i)}y_{(i)}$. And since

$$(I - P_X)\, u_i = \frac{1}{1 - t_{ii}} \begin{pmatrix} X_{(i)}(X'_{(i)}X_{(i)})^{-1}\, x_i \\ 1 \end{pmatrix}$$

and

$$\| (I - P_X)\, u_i \|^2 = \frac{1}{1 + t_{ii}},$$

we can get that

$$u_i u_i' y = \frac{1}{1 + t_{ii}} \begin{pmatrix} X_{(i)}(X'_{(i)}X_{(i)})^{-1}\, x_i x_i' \hat{\beta}_{(i)} - X_{(i)}(X'_{(i)}X_{(i)})^{-1}\, x_i y_i) \\ - x'_i \beta_{(i)} + y_i \end{pmatrix}$$

therefore,

$$X(X'X)^{-1}X'y + \tilde{u}_i\tilde{u}_i' y = \begin{pmatrix} X_{(i)}\hat{\beta}_{(i)} \\ y_i \end{pmatrix}$$

Thus as λ increases from 0 to 1, and animated plot of $e(\lambda)$ versus $\hat{y}(\lambda)$ gives a dynamic view of effects of removing the i^{th} observation from the model (1).

The following Lemma shows that the residuals $e(\lambda)$ and fitted values $\hat{y}(\lambda)$ can be computed from the residuals e, fitted values $\hat{y} = \hat{y}(\lambda = 0)$ from the full model, and the fitted values $\hat{y}(\lambda = 1)$ from the model which does not cotain the i^{th} observation.

Lemma 2.

(i) $\hat{y}(\lambda) = \hat{y}(\lambda = 0) + \lambda\,(\hat{y}(\lambda = 1) - \hat{y}(\lambda = 0))$

(ii) $e(\lambda) = e - \lambda\,(\hat{y}(\lambda = 1) - \hat{y}(\lambda = 0))$

Proof : Using the fact

$$\begin{pmatrix} X'X & X'u_i \\ u'X & 1 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} (X'X)^{-1} + (X'X)^{-1}X'u_i Hu'_i X(X'X)^{-1} & -(X'X)^{-1}X'u_i H \\ -Hu'_i X(X'X)^{-1} & H \end{pmatrix}$$

where

$$\begin{aligned} H &= (u'_i u_i - u'_i X(X'X)^{-1}X'u_i)^{-1} \\ &= (u'_i (I - P_X) u_i)^{-1} \\ &= \frac{1}{\| Q_X u_i \|^2}, \end{aligned}$$

we can show that

$$\begin{aligned} P(x, u_i) &= (X \quad u_i) \begin{pmatrix} X'X & X'u_i \\ u'_i X & u'_i u_i \end{pmatrix}^{-1} \begin{pmatrix} X' \\ u'_i \end{pmatrix} \\ &= Px + \frac{(I-P_X)\, u_i u'_i (I-P_X)}{\| Q_X u_i \|^2} \\ &= Px + \tilde{u}_i \tilde{u}'_i \end{aligned}$$

where $P(x, u_i)$ is the projection matrix onto the column space of (X, u_i). Therefore

$$\begin{aligned} \hat{y}(\lambda) &= X(X'X)^{-1}X'y + \lambda \tilde{u}_i \tilde{u}'y \\ &= \hat{y}(\lambda = 0) + \lambda(P(x, u_i) - Px)y \\ &= \hat{y}(\lambda = 0) + \lambda(\hat{y}(\lambda = 1) - \hat{y}(\lambda = 0)) \end{aligned}$$

since

$$P(x, \tilde{u}_i) = P(x, u_i).$$

And property (ii) can be proved by the fact that

$$e(\lambda) = y - \hat{y}(\lambda)$$

$$= y - \hat{y}(\lambda = 0) - \lambda(\hat{y}(\lambda = 1) - \hat{y}(\lambda = 0))$$

$$e - \lambda(\hat{y}(\lambda = 1) - \hat{y}(\lambda = 0)).$$

Because of the simplicity of Lemma 2, an animated plot of $e(\lambda)$ versus $\hat{y}(\lambda)$ as λ is varied between 0 and 1 can be easily computed.

The appropriate number of frames (values of λ) for an animated plot depends on the speed with which the computer screen can be refreshed and thus on the hardware being used. With too many frames, changes become often too small to be noticed, and as a consequence the overall trend can be missed. With too small frames, smoothness and the behavior of individual points cannot be detected.

3. EXAMPLES

In this section, we will illustrate the use of the plot $e(\lambda)$ versus $\hat{y}(\lambda)$ as an aid to understanding the dynamic effects of removing an observation from a model.

3.1 Phosphorus Data

Our first illustration is based on the phosphorus data reported in Snedecor and Cochran(1967, p. 384). An investigation of source from which corn plants obtain their phosphorus was carried out. Concentrations of phosphorus in parts per millions in each of 18 soils were measured. The variables are

X_1 = concentratiions of inorganic phosphorus in the soil,
X_2 = concentratiions of organic phosphorus in the soil, and
Y = phosphorus content of corn grown in the soil at 20℃

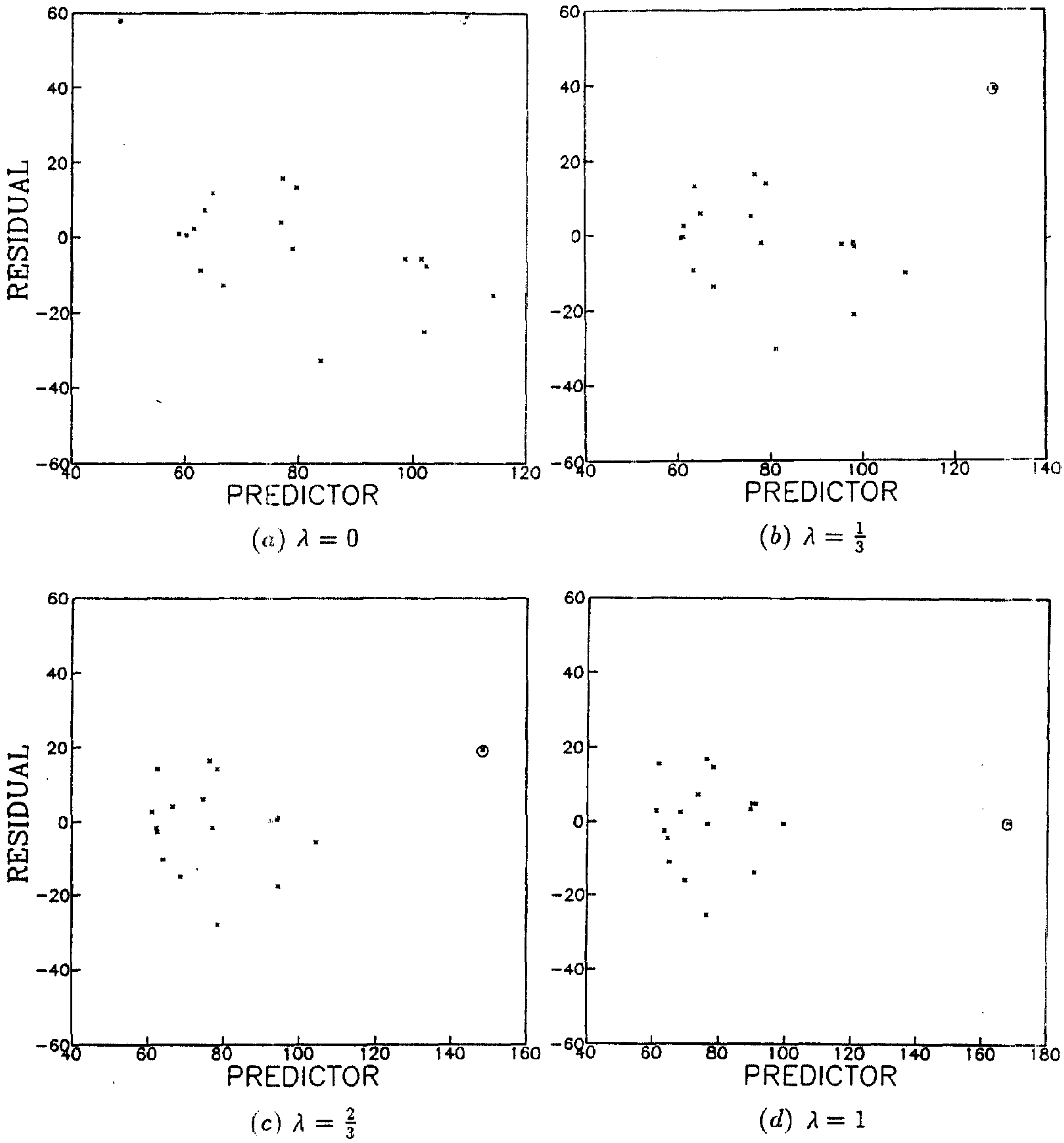
The data set together with the ordinary reisduals e_i, the diagonal terms h_{ii} of hat matrix $H = X(X'X)^{-1}X'$, the studentized residuals r_i, and Cook's distances $Cook_i$ are shown in Table 1 under the linear model assumption. From Table 1, we can conclude that the 17[th] observation, soil number 17 is the most influential observation since it has the largest residual ($|e_{17}|$ = 58.76), studentized residual ($|e_{17}|$ = 3.18), and Cook's distance ($Cook_{17}$ = 0.837675).

Figure 1 shows four frames of an animated plot of $e(\lambda)$ versus $\hat{y}(\lambda)$ for removing the 17^{th} observation. The first frame (a) is for $\lambda = 0$ and thus corresponds to the usual plot of residuals versus fitted values from the regression of y on $X = (X_1, X_2)$. And we can see in (a) the 17^{th} observation is located on the upper right corner. The second (b), third (c), and fourth (d) frames correspond to $\lambda = \frac{1}{3}, \frac{2}{3}$, and 1, respectively. So the fourth frame (d) is the usual plot of the residuals versus the fitted values from the regression of $y_{(17)}$ on $X_{(17)}$ where the subscript represents omission of the corresponding observation. We can see that as λ increases from 0 to 1, the 17^{th} observation moves to the right and down becoming the rightmost point in (b) and (d). And considering the plotting form the residual plot in (a) has an undesirable form because it does not have a random form in band between -60 and $+60$, but in (d) its form has randomness in band between -20 and $+20$. This result can be confirmed from the fact that the model fitting without the 17^{th} observation has the larger $R^2 = 0.5253$ than $R^2 = 0.4823$ of the full data model fitting. Figure 1 gives the reason why soil number 17 was omitted from the analysis of Snedecor and Cochran (1980).

Figure 2 gives four frames of an animated plot of $e(\lambda)$ versus $\hat{y}(\lambda)$ for removing the 6^{th} observation which has the largest diagonal term ($h_{66} = 0.46$) of hat matrix. But we can not find much difference in (a) ~ (d). Thus the 6^{th} observation gives not much effect to the model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$.

Table 1. Phosphorus Data and Statistics When Y is Regressed in X_1 and X_2

Soil	X_1	X_2	Y	e_i	h_{ii}	r_i	$Cook_i$
1	0.4	53	64	2.44	0.26	0.14	0.002243
2	0.4	23	60	1.04	0.19	0.06	0.000243
3	3.1	19	71	7.55	0.23	0.42	0.016711
4	0.6	34	61	0.73	0.13	0.04	0.000071
5	4.7	24	54	−12.74	0.16	−0.67	0.028762
6	1.7	65	77	12.07	0.46	0.79	0.178790
7	9.4	44	81	4.11	0.06	0.21	0.000965
8	10.1	31	93	15.99	0.10	0.81	0.023851
9	11.6	29	93	13.47	0.12	0.70	0.022543
10	12.6	58	51	−32.83	0.15	−1.72	0.178095
11	10.9	37	76	−2.97	0.06	−0.15	0.000503
12	23.1	46	96	−5.58	0.13	−0.29	0.004179
13	23.1	50	77	−24.93	0.13	−1.29	0.080664
14	21.6	44	93	−5.72	0.12	−0.29	0.003768
15	23.1	56	95	−7.45	0.15	−0.39	0.008668
16	1.9	36	54	−8.77	0.11	−0.45	0.008624
17	26.8	58	168	58.76	0.20	3.18	0.837675
18	29.9	51	99	−15.18	0.24	−0.84	0.075463



(a) $\lambda = 0$

(b) $\lambda = \frac{1}{3}$

(c) $\lambda = \frac{2}{3}$

(d) $\lambda = 1$

Figure 1. Phosphorus Data: Animated Plot for Removing the 17^{th} Observation
($\otimes$ represents the 17^{th} observation)

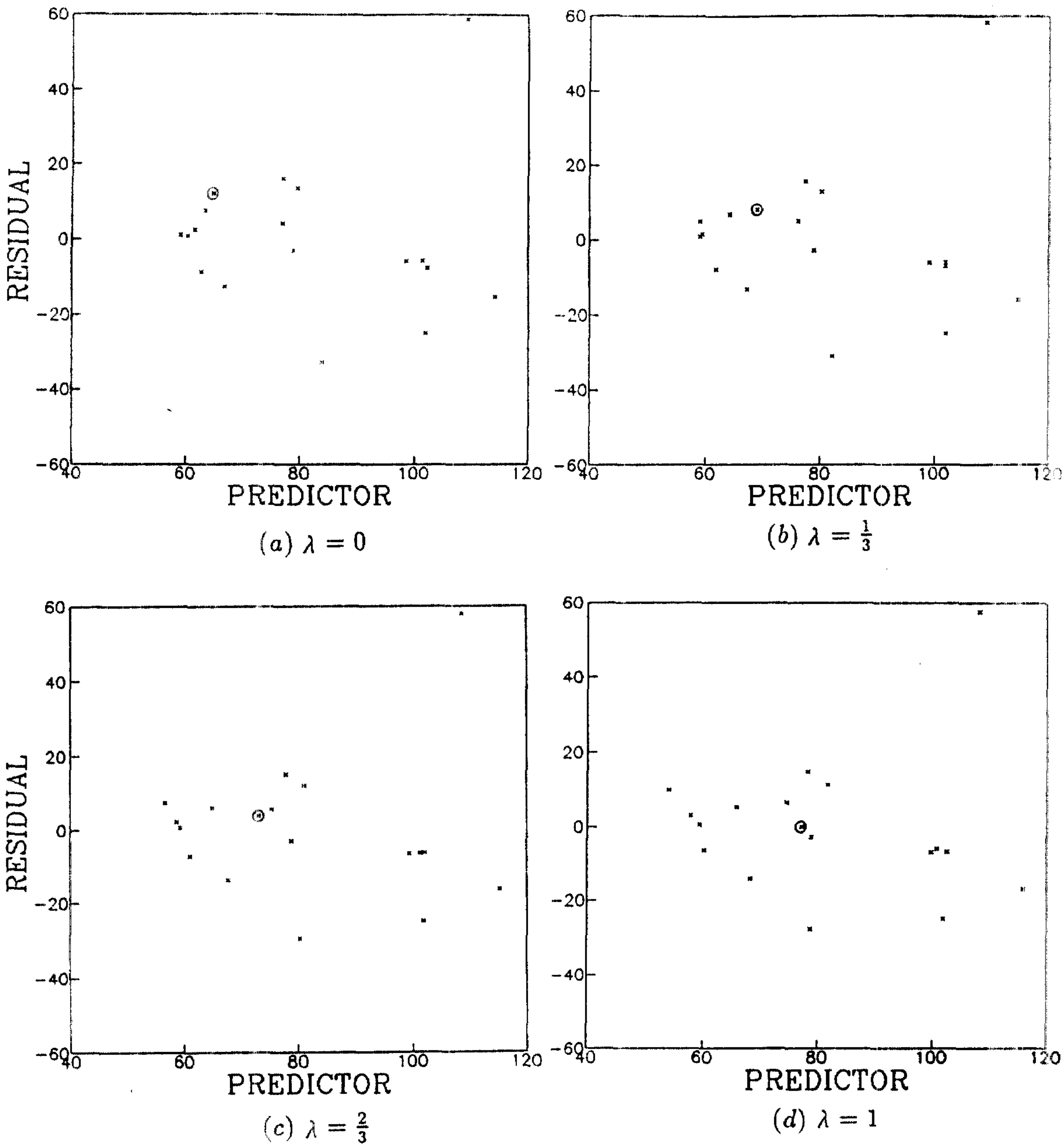


Figure 2. Phosphorus Data: Animated Plot for Removing the 6^{th} Observation
($\otimes$ represents the 6^{th} observation)

3.2 Rat data

This example is based on the rat data taken from Cook and Weisberg (1985, p.122). Varing doses of a drug were given nineteen rats. The response y is the percent of the dose that was absorbed into the rat's liver. The explanatory variables are dose X_1, liver weight X_2, and body weight X_3. The data set together with the ordinary reisduals e_i, the studentized residuals r_i, *Cook's* distances $Cook_i$, and the diagonal terms h_{ii} of hat matrix are shown in Table 2 under the linear model assumption. From Table 1, we can conclude that the 3rd observation is the high leverage point since it has the largest diagonal term (h_{33} = 0.93) although it has a relatively small residual (e_3 =0.02).

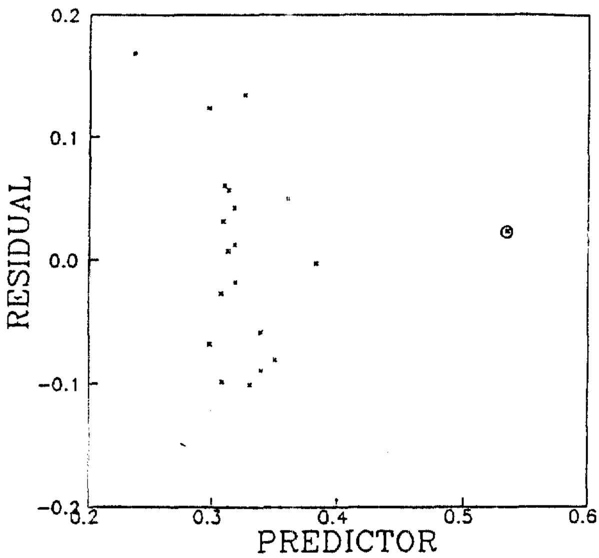
Figure 3 shows four frames of animated plot of $e(\lambda)$ versus $\hat{y}(\lambda)$ for removing the 3rd observation at $\lambda = 0, \frac{1}{3}, \frac{2}{3}, 1$. The first frame (a) is for $\lambda = 0$ and thus corresponds to the usual plot of residuals versus fitted values from the regression of Y on $X = (X_1, X_2, X_3)$. And in (a) the 3rd observation is located on the rightmost position, so we can confirm the fact that the 3rd observation is the high leverage point. As λ increases from 0 to 1, that is, as the 3rd observation is being smoothly removed, we can find that the plotting forms become more peculiar. This means that the 3rd observation is very influential. This result can be confirmed from the fact that the model fitting without the 3rd observation has smaller R^2 = 0.0211 than R^2 = 0.3639 of the full data model fitting.

Figure 4 shows four frames of animated plot for removing the 19th observation which has the largest residual (e_{19} = 0.13) and the largest studentized residual (r_{19} = 1.92). We can find much difference in Figure 4.

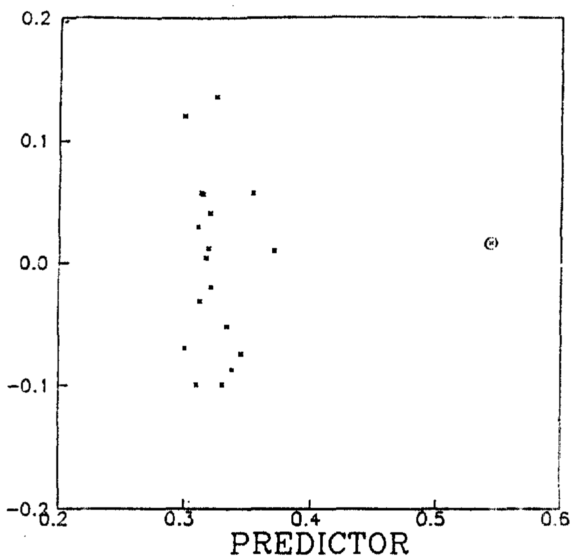
In Figure 3 and Figure 4, the observation which has the largest residual gives much effect to the animated plot rather than the high leverage point.

Table 2. Rat Data and Statistics When Y is Regressed on (X_1, X_2, X_3)

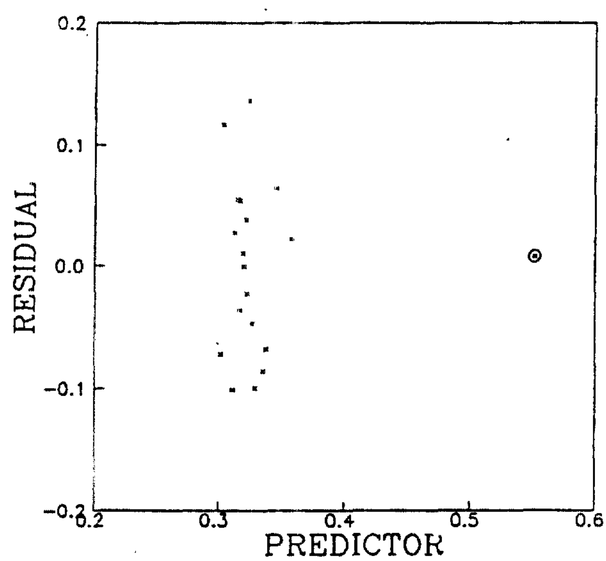
X_1	X_2	X_3	Y	e_i	r_i	$Cook_i$	h_{ii}
176	6.5	0.88	0.42	0.12	1.77	0.17	0.18
176	9.5	0.88	0.25	−0.09	−1.27	0.09	0.18
190	9.0	1.00	0.56	0.02	0.81	0.93	0.85
176	8.9	0.88	0.23	−0.10	−1.38	0.06	0.11
200	7.2	1.00	0.23	−0.07	−1.12	0.20	0.39
167	8.9	0.83	0.32	0.01	0.10	0.00	0.16
188	8.0	0/94	0.37	0.06	0.79	0.02	0.14
195	10.0	0.98	0.41	0.05	0.74	0.05	0.25
176	8.0	0.88	0.33	0.01	0.16	0.00	0.07
165	7.9	0.84	0.38	0.00	−0.04	0.00	0.12
158	6.9	0.80	0.27	−0.08	−1.11	0.04	0.12
148	7.3	0.74	0.36	0.04	0.6	0.02	0.17
149	5.2	0.75	0.21	−0.10	−1.54	0.27	0.32
163	8.4	0.81	0.28	−0.03	−0.38	0.01	0.13
170	7.2	0.85	0.34	0.03	0.43	0.00	0.08
186	6.8	0.94	0.28	−0.06	−0.86	0.05	0.22
146	7.3	0.73	0.30	−0.02	−0.26	0.00	0.20
181	9.0	0.90	0.37	0.06	0.85	0.03	0.15
149	6.4	0.75	0.46	0.13	1.92	0.20	0.18



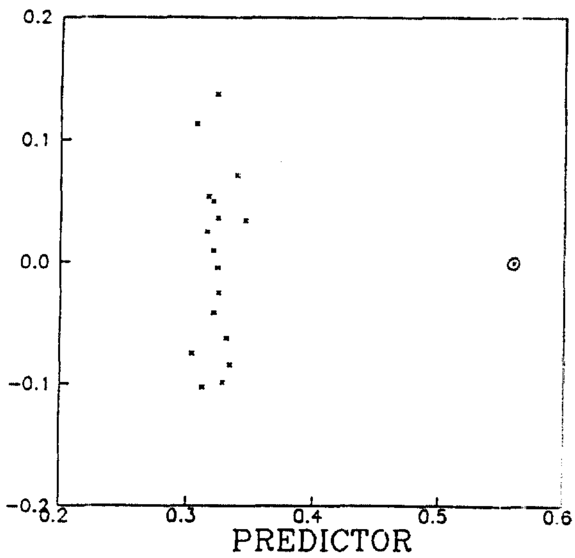
(*a*) $\lambda = 0$



(*b*) $\lambda = \frac{1}{3}$



(*c*) $\lambda = \frac{2}{3}$



(*d*) $\lambda = 1$

Figure 3. Rat Data: Animated Plot of Removing the 3^{rd} Observation
($\otimes$ represents the 3^{rd} observation)

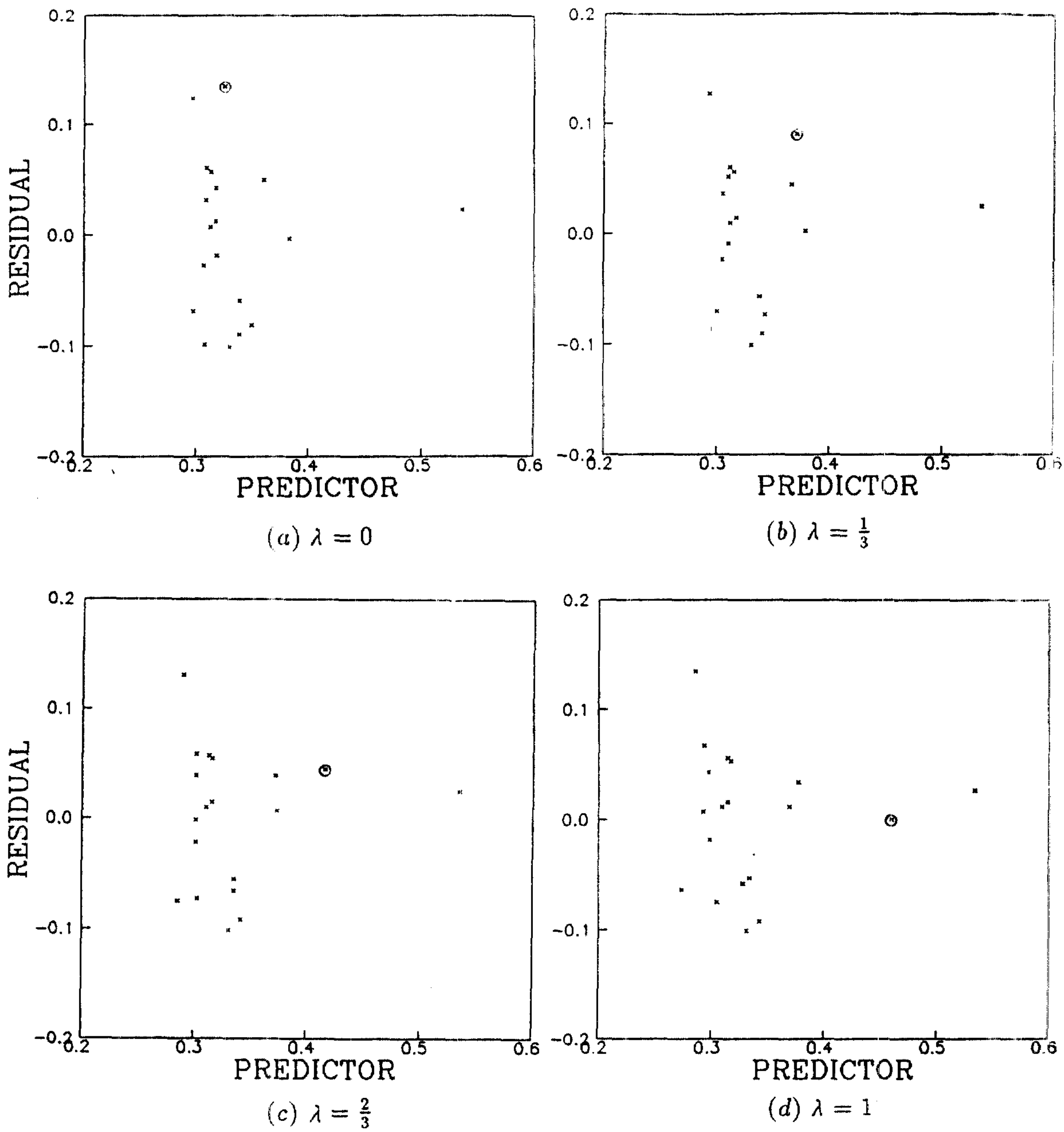


Figure 4. Rat Data: Animated Plot of Removing the 19^{th} Observation
(⊗ represents the 19^{th} observation)

REFERENCES

[1] Atkinson, A. C. (1985), *Plots, Transformations, and Regression*, Oxford, U. K. : Oxford University Press.

[2] Chambers, J. M., Cleveland, W., Kleiner, B., and Tukey, P. (1983), *Graphical Methods for Data Analysis*, Monterey, CA : Wadsworth.

[3] Cleveland, W. S. (1987), "Research in Statistical Graphics", *Journal of the American Statistical Association*, 82, 419－423.

[4] Cook, R. D., and Weisberg, S. (1982), *Residuals and Influence in Regression*, New York : Chapman & Hall.

[5] Cook, R. D., and Weisberg, S. (1989), "Regression Dignostics With Dynamic Graphics", *Technometrics*, 31, 277－291.

[6] Cox, D. R. and Snell, E. J. (1968), "A General Definition of Residuals", *Journal of the Royal Statistical Society, Ser. B*, 30, 248－275.

[7] Draper, N. R. and Smith, H. (1966), *Applied Regression Analysis*, 2nd ed., New York : Wiely.

[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] Larsen, W. A., and McCleary, S. J. (1972), "The Use of Partial Residual Plots in Regression Analysis", *Technometrics*, 14, 781－790.

[10] Mallows, C. L. (1986), "Augmented Partial Residual Plots", *Technometrics*, 28, 313－320.

[11] Snedecor, G. W., and Cochran, W. G. (1967), *Statistical Methods*, 6th ed., Ames : Iowa State University Press.