应用数学和力学，第13卷第9期（1992年9月）
Applied Mathematics and Mechanics

应用数学和力学编委会编
重庆出版社出版

权回归与一般影响度量*

杨　虎

（重庆交通学院，1989年3月1日收到）

摘　要

权回归模型是近年来文献中较多涉及的一种．本文从回归诊断的角度进一步研究了它的若干问题，如扰动、影响度量及估计效率等．由于这类模型的一般性，因而本文的结果均可看作回归诊断中较成熟的一些结论[2][3][7]的推广．

关键词　异方差　回归诊断　效率　广义学生化残差　影响度量

一、引　言

考虑线性加权回归模型

$$WY=WX\beta+\varepsilon,\ \ \varepsilon\sim(0,\sigma^2I) \tag{1.1}$$

这里Y为$n\times1$观测向量，X为$n\times p$设计阵，β为$p\times1$参数向量，ε为$n\times1$随机误差向量．符号$\varepsilon\sim(0,\sigma^2I)$表示$E(\varepsilon)=0$且$\mathrm{cov}(\varepsilon)=\sigma^2I$，其中$\sigma$为常数，记$W=\mathrm{diag}\ (w_1,w_2,\cdots,w_n)$，$w_i(i=1,2,\cdots,n)$均为常数．很明显当$W=I$时(1.1)式即为一般Gauss-Markov模型[1]．(1.1)中β的最小二乘估计(the least square estimate：简记为LSE)为

$$\hat{\beta}_W=(X'W^2X)^{-1}X'W^2Y \tag{1.2}$$

当然它也可看成异方差线性模型

$$Y=X\beta+\varepsilon\qquad \varepsilon\sim(0,\sigma^2W^{-2}I) \tag{1.3}$$

的广义最小二乘估计[1]（the generalized least square estimate：简记为GLSE），特别地当$W=I$时，$\hat{\beta}_I=(X'X)^{-1}X'Y$就是Gauss-Markov模型参数的LSE，通常我们感兴趣的是预测值$\hat{Y}_W=WX\hat{\beta}_W$，以及它对于权阵$W$扰动的敏感程度，亦即我们需要知道在$W$存在扰动$\Delta W$时，$\hat{Y}_W$的扰动$\Delta\hat{Y}_W$的受影响情况，通常，如果取$W=I$，$\Delta W=\Omega=\mathrm{diag}(0,0,\cdots,0,-1,0,\cdots,0)$，($\Omega$的第$i$个对角元为$-1$，其余全为零)，则$\Delta\hat{Y}_I$就表示了Gauss-Markov模型中观测向量的拟合值$\hat{Y}_I$的扰动．此外，我们用$\hat{\sigma}^2_W$表示(1.1)中σ的LSE，$\|\cdot\|^2$表示欧氏范数，则$\|\Delta\hat{Y}_I\|/p\hat{\sigma}^2_I$就是Cook距离[2]，它度量了剔除第$i$个数据对回归分析的影响．与此相类似的度量还有$A$-$P$统稳定性度量[4]，一般来讲，通过适当地选择$\Delta W$，

* 钱伟长推荐．

还可得到度量数据集影响的相应的一些度量。因此研究模型(1.1)的影响度量问题显得非常必要。

本文我们考虑了这种扰动的更一般的形式，它可以作为某种一般性的影响度量用于判决各种扰动的影响以更好地进行回归诊断。

二、预测对于权扰动的敏感性

记$\delta_W=WY-WX\hat{\beta}_W$为(1.1)的普通残差向量，我们有如下结果：

定理2.1 设ΔW为W^2的扰动，$\Delta\hat{Y}_W$为$\hat{Y}_W$的扰动，则

$$\Delta\hat{Y}_W=H_WW^{-1}\Delta W(W+HW^{-1}\Delta W)^{-1}\delta_W \tag{2.1}$$

其中 $H_W=WX(X'W^2X)^{-1}X'W'$.

证明 应用如下的代数事实

$$(A+BCB')^{-1}=A^{-1}-A^{-1}BC(I+B'A^{-1}BC)^{-1}B'A^{-1} \tag{2.2}$$

我们可得

$$\begin{aligned}\hat{\beta}_{W,\Delta W}&=(X'(W^2+\Delta W)X)^{-1}X(W^2+\Delta W)Y\\&=(X'W^2X+X'\Delta WX)^{-1}X'(W^2+\Delta W)Y\\&=[(X'W^2X)^{-1}-(X'W^2X)^{-1}X'\Delta W(I\\&\quad+X(X'W^2X)^{-1}X'\Delta W)^{-1}X(X'W^2X)^{-1}]X'(W^2+\Delta W)Y\\&=\hat{\beta}_W+(X'W^2X)^{-1}X'\Delta WY-(X'W^2X)^{-1}X'\Delta W(I\\&\quad+X(X'W^2X)^{-1}X'\Delta W)^{-1}X(X'W^2X)^{-1}X'(W^2+\Delta W)Y.\end{aligned}$$

所以

$$\begin{aligned}\Delta\hat{\beta}_W&=\hat{\beta}_{W,\Delta W}-\hat{\beta}_W\\&=(X'W^2X)^{-1}X'\Delta W(I+X(X'W^2X)^{-1}X'\Delta W)^{-1}Y\\&\quad-(X'W^2X)^{-1}X'\Delta W(I+X(X'W^2X)^{-1}X'\\&\quad\Delta W)^{-1}X(X'W^2X)^{-1}X'W^2Y\\&=(X'W^2X)^{-1}X'\Delta W(W+H_WW^{-1}\Delta W)^{-1}\delta_W.\end{aligned}$$

由$\Delta\hat{Y}_W=WX\Delta\hat{\beta}_W$，明所欲证。

定义1 我们用$\|\Delta\hat{Y}_W\|^2/n\hat{\sigma}_W^2$作为扰动$\Delta W$的影响度量，记为$U_{W,\Delta W}$。

当$U_{W,\Delta W}$较大时，扰动ΔW有较大的影响，我们在实际中应找出这种扰动的原因及实际背景，并尽量避免这种扰动.很明显，$U_{W,\Delta W}$是 Cook 距离[2]的更一般的形式，当$W=I$，$\Delta W=\Omega$时，$U_{I,\Omega}$与$D(X'X,p\hat{\sigma}_I^2)$仅相差一个常数，因此我们立即得到如下的结果.

定理2.2 当$W=I$，$\Delta W=\Omega$时

$$U_{I,\Omega}=\frac{1}{n}r_i^2P_i(X'X) \tag{2.3}$$

其中 $r_i^2=\delta_i^2/(1-h_{ii})\hat{\sigma}_I^2$ 为学生化残差，$P_i(X'X)$为X_i的势(参见文献[5-7])，这里X_i'为设计阵X的第i行.

证明 当$W=I$时，H_W即为帽子矩阵[7]，设$H=(h_{ij})_{n\times n}\triangleq H_I$，则

$$H\Omega=\begin{pmatrix}0&\cdots&0&-h_{1i}&0&\cdots&0\\0&\cdots&0&-h_{2i}&0&\cdots&0\\\cdots\cdots\\0&\cdots&0&-h_{ni}&0&\cdots&0\end{pmatrix}$$

所以不难推知

$$(I+H\Omega)^{-1}=\frac{(I+H\Omega)^*}{\det(I+H\Omega)}=\begin{pmatrix}1 & 0 & \cdots & 0 & -h_{1i} & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0 & -h_{2i} & 0 & \cdots & 0\\ \cdots\cdots \\ 0 & 0 & \cdots & 0 & 1-h_{ii} & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0 & -h_{ni} & 0 & \cdots & 1\end{pmatrix}^{-1} \tag{2.4}$$

这里 $\det(A)$，A^*分别表示矩阵A的行列式和伴随矩阵，很明显

$$\det(I+W\Omega)=1-h_{ii}$$

$$(I+H\Omega)^*=\begin{pmatrix}1-h_{ii} & 0 & \cdots & 0 & (-1)^i h_{1i} & 0 & \cdots & 0\\ \cdots\cdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0\\ \cdots\cdots \\ 0 & 0 & \cdots & 0 & (-1)^{n+i-1}h_{ni} & 0 & \cdots & 1-h_{ii}\end{pmatrix}$$

用分块矩阵表示为

$$(I+H\Omega)^*=\left(\begin{pmatrix}(1-h_{ii})I_{i-1}\\ 0\end{pmatrix}\quad u\quad \begin{pmatrix}0\\ (1-h_{ii})I_{n-i}\end{pmatrix}\right) \tag{2.5}$$

其中 I_i为i阶单位阵，特别地当$i=n$时我们常略去这个下标，而$u'=(-1)^i(h_{1i},-h_{2i},\cdots,(-1)^{i-2}h_{i-1,i},1,(-1)^i h_{i+1,i},\cdots,(-1)^{n-1}h_{ni})$，所以若记$\nabla'_i=(-h_{1i},\ -h_{2i},\ \cdots,\ -h_{ni})$，$\delta'_I=(\delta_1,\delta_2,\cdots,\delta_n)$为普通残差向量，则

$$\begin{aligned}\Delta\hat{Y}_I&=H\Omega(I+H\Omega)^{-1}\delta_I\\ &=(1-h_{ii})^{-1}(0\quad \nabla_i\quad 0)\begin{pmatrix}\delta_1\\ \delta_2\\ \vdots\\ \delta_n\end{pmatrix}\\ &=(1-h_{ii})^{-1}\nabla_i\delta_i\\ &=-X(X'X)^{-1}X_i\delta_i/(1-h_{ii})\end{aligned} \tag{2.6}$$

故$U_{I,\Omega}=\|\Delta\hat{Y}_I\|^2/n\hat{\sigma}_I=h_{ii}\delta_i^2/(1-h_{ii})^2n\hat{\sigma}_I^2$，由于$X_i$的势$P_i(X'X)=h_{ii}/(1-h_{ii})$[5]，明证。

三、加权回归的LSE与效率

现在我们来看$\hat{\boldsymbol{\beta}}_{W,\Delta W}$相对于GLSE $\hat{\boldsymbol{\beta}}_W$的均方误差比效率[8]

$$\rho=\frac{\text{MSE }X\hat{\boldsymbol{\beta}}_W}{\text{MSE }X\hat{\boldsymbol{\beta}}_{W,\Delta W}} \tag{3.1}$$

其中 MSE 表示的均方误差 (mean square error)。我们很容易证明下面的两个结果

i) $\text{MSE }X\hat{\boldsymbol{\beta}}_W=\sigma^2\ \text{tr}\ X(X'W^2X)^{-1}X'$

ii) $\text{MSE }X\hat{\boldsymbol{\beta}}_{W,\Delta W}=\sigma^2\ \text{tr}\ (X'X)^{-1/2}X'W^{-2}X(X'X)^{-1/2}$

从而立即可以建立下面的定理3.1，它给出了加权回归参数估计效率的上界。

定理3.1　对线性模型(1.2)，当$W^2+\Delta W=I$时

$$\rho \leqslant \left[\frac{2\sum_{i=1}^{s} W_i W_{n-i+1}}{\sum_{i=1}^{s}(W_i+W_{n-i+1})} \right]^2 \tag{3.2}$$

其中 $s=\min(p,n-p)$。

注 当$W^2+\Delta W$正定时，条件 $W^2+\Delta W=I$ 总是可行的，或者说我们可以去掉这个条件，事实上，我们直接用$(W+W)^{-1/2}$去左乘模型(1.2)，可以得到一个新的线性模型，对于此模型，定理3.1的条件即能满足。

定义2 设$V=W^2-I=\mathrm{diag}(v_1,v_2,\cdots,v_n)$，$C=(I+HV)^{-1}$我们称

$$-\frac{\delta_I'[C'VHVC+(I-HVC)'V(I-HVC)]\delta_I}{\hat{\sigma}_I^2} \tag{3.3}$$

为广义学生化残差，记为r_V^2。

定理3.2 当$V=\Omega$时

$$r_\Omega^2=\frac{\delta_i^2}{(1-h_{ii})\hat{\sigma}_I^2} \tag{3.4}$$

这同一般文献中的学生化残差形式上一样。

证明 由(2.6)式得

$$HVC\delta_I=-X(X'X)^{-1}X_i\delta_i/(1-h_{ii}) \tag{3.5}$$

所以

$$\frac{\delta_I'C'VHVC\delta_I}{\hat{\sigma}_I^2}=\frac{h_{ii}\delta_i^2}{(1-h_{ii})^2\hat{\sigma}_I^2} \tag{3.6}$$

而

$$\begin{aligned}
&\delta_I'(I-HVC)'V(I-HVC)\delta_I/\hat{\sigma}_I^2\\
&=\left(\delta_I-\frac{1}{1-h_{ii}}V_i\delta_i\right)'\Omega\left(\delta_I-\frac{1}{1-h_{ii}}V_i\delta_i\right)/\hat{\sigma}_I^2\\
&=-\frac{1}{\hat{\sigma}_I^2}\left[\delta_i^2+2\frac{h_{ii}\delta_i^2}{1-h_{ii}}+\frac{h_{ii}^2\delta_i^2}{(1-h_{ii})^2}\right]\\
&=-\delta_i^2/(1-h_{ii})^2\hat{\sigma}_I^2
\end{aligned} \tag{3.7}$$

由(3.6)，(3.7)两式，立即推得定理的结果。

在计算$U_{W,\Delta W}$时，需要求出$\hat{\sigma}_W^2$，下面的定理给出了计算$\hat{\sigma}_W$的一种方法。

定理3.3 对加权线性回归模型(1.1)，我们有

$$\hat{\sigma}_W^2=\hat{\sigma}_I^2-r_V^2\hat{\sigma}_I^2/(n-p-1) \tag{3.8}$$

证明 因为

$$(n-p-1)\hat{\sigma}_W^2=\sum_{i=1}^{n}W_i^2(Y_i-X_i'\hat{\beta}_W)^2$$

$$=\sum_{i=1}^{n}W_i^2[\delta_i+X_i'(\hat{\beta}_I-\hat{\beta}_W)]^2$$

由定理2.1的证明可见

$$\hat{\beta}_W-\hat{\beta}_I=(X'X)^{-1}XVC\delta_I \tag{3.9}$$

从而有

$$(n-p-1)\hat{\sigma}_{W}^{2}=(n-p-1)\hat{\sigma}_{I}^{2}+2\sum_{i=1}^{p}\delta_i\nabla_i' VC\delta_I+\sum_{i=1}^{n}(\nabla_i' VC\delta_I)^2$$

$$+\sum_{i=1}^{n}U_i[\delta_i+\nabla_i' VC\delta_I]^2 \tag{3.10}$$

而

i) $\sum_{i=1}^{n}\delta_i\nabla_i'=\delta_I'H=0$

ii) $\sum_{i=1}^{n}(\nabla_i' VC\delta_I)^2=\delta_I'C'V'HVC\delta_I$

iii) $\sum_{i=1}^{n}U_i[\delta_i+\nabla_i' VC\delta_I]^2=\delta_I'(I-HVC)'V(I-HVC)\delta_I$

将这些结 果代入(3.10)式即得(3.8)。

参 考 文 献

[1] 王松桂，《线性模型的理论及其应用》，安徽教育出版社 (1987).

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

[3] Andrews, D. F. and, D. Pregibon, Finding Outliers that Matter, *J. Ray. Stat.*, Soc. B, **40**, (1978), 85—93

[4] 杨虎，强影响的分布与稳定性度量，数学研究与评论，**10** (3) (1990), 393—398.

[5] Weisberg, S., *Applied Linear Regression*, New York (1980).

[6] Huber, P., *Robust Statistics*, Wiley New York (1981)

[7] 陈希孺、王松桂，《近代回归分析——原理方法及应用》，安徽教育出版社 (1987).

[8] 杨虎，Kantorovich 不等式的延拓与均方误差比效率，应用数学，1(4)(1988), 85—90.

A Kind of General Influence Measure on the Linear Weighted Regression

Yang Hu

(Chongqing Jiaotong Institute, Chongqing)

Abstract

The linear weighted regression model is one of the models studied in many articles in recent years. Some further problems, such as disturbation, influence measure and estimate efficiency, have been discussed in this paper on the basis of the regression diagnosties. The partial conclusions of this paper are the extension of the familiar concepts in the regression diagnoties theory[2,3,7] because they are representative of this kind of model.

Key words heteroscedastic, regression diagnostic, efficiency, generalized student's residual, influence measure