山西大学学报（自然科学版）12(2),138-142,1989
Journal of Shanxi University (Nat. Sci. Ed.)

相关性度量与估计效率

刘兰亭　张信东

摘要　在这篇文章中我们给出了最小二乘拟合值与残差之间的相关性度量和最小二乘估计相对效率的关系

关键词　相关性度量　估计效率

一　引言

考虑线性模型

$$y=X\beta+e,\ e\sim N(0,\Gamma) \tag{1}$$

其中y是$n\times 1$观测向量，X是已知的$n\times p$设计矩阵，β是未知的$p\times 1$参数向量，e是$n\times 1$随机误差向量，Γ是$n\times n$正定矩阵，即$\Gamma>0$。假定$rk(X)=p$，$n\geqslant 2p$。β的BLU估计（即最佳线性无偏估计）与LS估计（即最小二乘估计）分别为

$$\beta^*=(X^T\Gamma^{-1}X)^{-1}X^T\Gamma^{-1}y,\ \mathrm{Cov}(\beta^*)=(X^T\Gamma^{-1}X)^{-1}$$

$$\hat{\beta}=(X^TX)^{-1}X^Ty,\ \mathrm{Cov}(\hat{\beta})=(X^TX)^{-1}X^T\Gamma X(X^TX)^{-1}$$

熟知$\mathrm{Cov}(\beta^*)\leqslant\mathrm{Cov}(\hat{\beta})$。但是，在许多实际问题中，$\Gamma$是未知的，于是只能用$\hat{\beta}$来代替$\beta^*$，这就要蒙受一些损失。为此，许多作者利用

$$e(\hat{\beta})=\frac{|\mathrm{Cov}(\beta^*)|}{|\mathrm{Cov}(\hat{\beta})|} \tag{2}$$

作为LS估计的相对效率。显然，$0\leqslant e(\hat{\beta})\leqslant 1$，而且其值愈大，表示估价$\hat{\beta}$与$\beta^*$愈接近，当然所蒙受的损失就愈小。

许多专家和学者提出和研究了度量两个随机向量x和z相关程度的数量指标。例如相关系数，Hotelling和张尧庭提出的广义相关系数和Lindley（见[1]）提出的相关性量度L。对一般线性模型，王松桂（见[2]）研究了广义相关系数与估计效率的关系。本文讨论在模型(1)情形相关性度量L与估计效率的关系，并据此得出相关性度量的界。

对随机向量x和z，Lindley相关性度量定义为

1988年6月27日收到

$$L(x,z)=\iint f(s,t)\ln\frac{f(s,t)}{g(s)\cdot h(t)}\,ds\cdot dt \tag{3}$$

其中$g(s)$，$h(t)$和$f(s,t)$分别为x、z和$(x^\tau,z^\tau)^\tau$的分布密度或联合分布密度。

二 主要结果

不失一般性假定$X^\tau X=I_p$，选$n\times(n-p)$矩阵N使$(X\mid N)^\tau(X\mid N)=I_n$。记$\Sigma_1=XX^\tau\Gamma XX^\tau$，$\Sigma_2=NN^\tau\Gamma NN^\tau$，和

$$\Sigma=\begin{pmatrix}XX^\tau\Gamma XX^\tau & XX^\tau\Gamma NN^\tau\\ NN^\tau\Gamma XX^\tau & NN^\tau\Gamma NN^\tau\end{pmatrix}$$

易知$\mathrm{rk}(\Sigma_1)=p$，$\mathrm{rk}(\Sigma_2)=n-p$和$\mathrm{rk}(\Sigma)=n$

根据Rao介绍的奇异型正态分布（见[3]），我们可以得到最小二乘拟合值$XX^\tau y$、残差$NN^\tau y$和$\begin{pmatrix}XX^\tau y\\ NN^\tau y\end{pmatrix}$的分布密度分别为

$$g(s)=\frac{1}{(2\pi)^{\frac{p}{2}}\cdot(\det(X^\tau\Gamma X))^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(s-X\beta)^\tau\Sigma_1^-(s-X\beta)\right\} \tag{4}$$

$$N_1^\tau s=N_1^\tau X\beta \qquad \text{a.s.}$$

$$h(t)=\frac{1}{(2\pi)^{\frac{p}{2}}\cdot(\det(N^\tau\Gamma N))^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}t^\tau\Sigma_2^-t\right\} \tag{5}$$

$$N_2^\tau t=0 \qquad \text{a.s.}$$

$$f(s,t)=\frac{1}{(2\pi)^{\frac{n}{2}}(\det\Gamma)^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}\begin{pmatrix}s-X\beta\\ t\end{pmatrix}^\tau\Sigma^-\begin{pmatrix}s-X\beta\\ t\end{pmatrix}\right\} \tag{6}$$

$$N_3^\tau\begin{pmatrix}s\\ t\end{pmatrix}=N_3^\tau\begin{pmatrix}X\beta\\ 0\end{pmatrix}$$

其中N_1，N_2和N_3分别是$n\times(n-p)$，$n\times p$和$2n\times n$矩阵，且分别满足

$$\mu(N_1)=\mu^\perp(\Sigma_1),\quad \mu(N_2)=\mu^\perp(\Sigma_2),\quad \mu(N_3)=\mu^\perp(\Sigma)$$

$\mu(A)$表示由A的列向量张成的线性子空间，“$\perp$”表示线性子空间的正交补空间；Σ_1^-，Σ_2^-和Σ^-分别为Σ_1，Σ_2和Σ的广义逆。由于密度函数与广义逆的选取无关，故可用Σ_1和Σ_2的“$+$”号逆Σ_1^+和Σ_2^+分别代替Σ_1^-和Σ_2^-，至于Σ^-我们亦可用Σ的“$+$”号逆代替，亦可任选一个广义逆，譬如

$$\Sigma^-=\begin{pmatrix}XX^\tau\Gamma^{-1}XX^\tau & XX^\tau\Gamma^{-1}-X(X^\tau\Gamma X)^{-1}X^\tau\\ \Gamma^{-1}XX^\tau-X(X^\tau\Gamma X)^{-1}X^\tau & \Gamma^{-1}\end{pmatrix} \tag{7}$$

另外我们不难得到

$$\Sigma_1^+ = X(X^\tau\Gamma X)^{-1}X^\tau,\ \Sigma_2^+ = N(N^\tau\Gamma N)^{-1}N^\tau \tag{8}$$

为了证明简洁，定义集合$M_1=\left\{\begin{pmatrix}s\\t\end{pmatrix}: N_1^\tau s=0, N_2^\tau t=0\right\}$，$M_2=\left\{\begin{pmatrix}s\\t\end{pmatrix}: N_3^\tau\begin{pmatrix}s\\t\end{pmatrix}=0\right\}$和$M=\left\{\begin{pmatrix}s\\t\end{pmatrix}: N_1^\tau s=0,\ N_2^\tau t=0,\ N_3^\tau\begin{pmatrix}s\\t\end{pmatrix}=0\right\}$. 则$M_2\subset M_1$，从而

$$M=M_2 \tag{9}$$

为给出主要结论首先证明如下引理

引理 1 设$\rho_1^2\geqslant\rho_2^2\geqslant\cdots\geqslant\rho_p^2$是$XX^\tau y$和$NN^\tau y$之间的典型相关系数，则

$$\prod_{i=1}^{p}(1-\rho_i^2)=\det(I_p-(X^\tau\Gamma X)^{-1/2}X^\tau\Gamma N(N^\tau\Gamma N)^{-1}N^\tau\Gamma X(X^\tau\Gamma X)^{-1/2}$$

证明 因为

$$\mathrm{Cov}\begin{pmatrix}XX^\tau y\\NN^\tau y\end{pmatrix}=\Sigma\triangleq\begin{pmatrix}\Sigma_1 & \Sigma_{12}\\ \Sigma_{21} & \Sigma_2\end{pmatrix}$$

其中$\Sigma_{12}=XX^\tau\Gamma NN^\tau$，$\Sigma_{21}=\Sigma_{12}^\tau$。所以$XX^\tau y$和$NN^\tau y$之间的典型相关系数$\rho^2$应满足

$$\det(\rho^2 I_n-\Sigma_1^+\Sigma_{12}\Sigma_2^+\Sigma_{21})=0$$

而它又等价于

$$\det(\rho^2 I_n-X(X^\tau\Gamma X)^{-1}X^\tau\Gamma N(N^\tau\Gamma N)^{-1}N^\tau\Gamma XX^\tau)=0$$

由于$X(X^\tau\Gamma X)^{-1}X^\tau\Gamma N(N^\tau\Gamma N)^{-1}N^\tau\Gamma XX^\tau$与$(X^\tau\Gamma X)^{-1/2}X^\tau\Gamma N(N^\tau\Gamma N)^{-1}\times N^\tau\Gamma X(X^\tau\Gamma X)^{-1/2}$具有完全相同的非零特征根，因此$1-\rho_i^2(i=1,2,\cdots,p)$是$I_p-(X^\tau\Gamma X)^{-1/2}X^\tau\Gamma N(N^\tau\Gamma N)^{-1}N^\tau\Gamma X(X^\tau\Gamma X)^{-1/2}$的全部特征根。由此即得结论。

定理 1 对于模型(1)，最小二乘拟合值$XX^\tau y$和残差$NN^\tau y$之间的相关性度量与LS估计相对效率具有关系

$$L(XX^\tau y,NN^\tau y)=-\frac{1}{2}\ln e(\hat{\beta}) \tag{10}$$

证明 将(4)—(6)式代入(3)式，并利用(9)式

$$L(XX^\tau y,NN^\tau y)=\iint\limits_{\substack{N_1^\tau s=N_1^\tau X\beta\\ N_2^\tau t=0\\ N_3^\tau\left(\begin{smallmatrix}s\\t\end{smallmatrix}\right)=N_3^\tau\left(\begin{smallmatrix}X\beta\\0\end{smallmatrix}\right)}} f(s,t)\cdot\ln\frac{f(s,t)}{g(s)h(t)}\,dsdt$$

$$=\iint_M f(s+X\beta,t)\cdot\ln\frac{f(s+X\beta,t)}{g(s+X\beta)h(t)}\,dsdt$$

$$=\iint_{M_2}-\frac{1}{2}\ln\frac{\det(\Gamma)}{\det(X^\tau\Gamma X)\cdot\det(N^\tau\Gamma N)}f(s+X\beta,t)dsdt$$

$$-\frac{1}{2}\iint_{M_2}\begin{pmatrix}s\\t\end{pmatrix}^\tau\left[\Sigma^-\begin{pmatrix}\Sigma_1^+ & 0\\0 & \Sigma_2^+\end{pmatrix}\right]\begin{pmatrix}s\\t\end{pmatrix}f(s+X\beta,t)dsdt$$

$$= -\frac{1}{2}\ln\frac{\det(\Gamma)}{\det(X^{\tau}\Gamma X)\cdot\det(N^{\tau}\Gamma N)} - E\begin{pmatrix}s\\t\end{pmatrix}^{\tau}\left[\Sigma^{-}-\begin{pmatrix}\Sigma_1^{+} & 0\\ 0 & \Sigma_2^{+}\end{pmatrix}\right]\begin{pmatrix}s\\t\end{pmatrix}$$

这是由于$f(s+X\beta,t)$是正态分布$N(0,\Sigma)$在超平面M_2上的概率分布密度；又由于

$$E\begin{pmatrix}s\\t\end{pmatrix}^{\tau}\left[\Sigma^{-}-\begin{pmatrix}\Sigma_1^{+} & 0\\ 0 & \Sigma_2^{+}\end{pmatrix}\right]\begin{pmatrix}s\\t\end{pmatrix}$$

$$= \mathrm{tr}\left(\left[\Sigma^{-}-\begin{pmatrix}X(X^{\tau}\Gamma X)^{-1}X^{\tau} & 0\\ 0 & N(N^{\tau}\Gamma N)^{-1}N^{\tau}\end{pmatrix}\right]\Sigma\right)$$

$$= \mathrm{tr}(\Sigma^{-}\Sigma) - \mathrm{tr}\begin{pmatrix}XX^{\tau} & X(X^{\tau}\Gamma X)^{-1}X^{\tau}\Gamma NN^{\tau}\\ N(N^{\tau}\Gamma N)^{-1}N^{\tau}\Gamma XX^{\tau} & NN^{\tau}\end{pmatrix}$$

$$= \mathrm{rk}(\Sigma) - n = 0$$

于是，利用引理1

$$L(XX^{\tau}y, NN^{\tau}y) = -\frac{1}{2}\ln\frac{\det(\Gamma)}{\det(X^{\tau}\Gamma X)\cdot\det(N^{\tau}\Gamma N)}$$

$$= -\frac{1}{2}\ln[\det(I_p - (X^{\tau}\Gamma X)^{-\frac{1}{2}}X^{\tau}\Gamma N(N^{\tau}\Gamma N)^{-1}N^{\tau}\Gamma X(X\Gamma X)^{-\frac{1}{2}})]$$

$$= -\frac{1}{2}\ln\prod_{i=1}^{p}(1-\rho_i^2)$$

而Bartmann和Bloomfield已证明(见[4])

$$e(\hat{\beta}) = \prod_{i=1}^{p}(1-\rho_i^2) \tag{11}$$

故

$$L(XX^{\tau}y, NN^{\tau}y) = -\frac{1}{2}\ln e(\hat{\beta}) \qquad \text{定理证毕}$$

根据[5]((2.8)式)与定理1不难得到

推论1　在定理1条件之下，有

$$0 \leqslant L(XX^{\tau}y, NN^{\tau}y) \leqslant -\frac{1}{2}\sum_{i=1}^{p}\ln\frac{4\lambda_i\lambda_{n-i+1}}{(\lambda_i+\lambda_{n-i+1})^2} \tag{12}$$

其中$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n > 0$是$\Gamma$的特征根。

根据[1]((9.7.16)式)给出的两个非奇异正态变量之间的相关性度量，不难得到

$$L(X^{\tau}y, N^{\tau}y) = -\frac{1}{2}\ln\prod_{i=1}^{p}(1-\gamma_i^2) \tag{13}$$

其中$\gamma_1^2 \geqslant \gamma_2^2 \geqslant \cdots \geqslant \gamma_p^2$是$X^{\tau}y$和$N^{\tau}y$之间的典型相关系数。而它们又与$XX^{\tau}y$和$NN^{\tau}y$之间的典型相关系数是一致的。故可得

定理2　对于线性模型(1)，$N^{\tau}y$和$X^{\tau}y$之间的相关性度量与估计效率的关系为

$$L(X^{\tau}y,N^{\tau}y)=-\frac{1}{2}\ln e(\hat{\beta}) \tag{14}$$

类似于推论1有

推论2　在定理2条件之下，有

$$0\leqslant L(X,y,N^{\tau}y)\leqslant-\frac{1}{2}\sum_{i=1}^{p}\ln\frac{4\lambda_i\lambda_{n-i+1}}{(\lambda_i+\lambda_{n-i+1})^2} \tag{15}$$

参 考 文 献

〔1〕Kullback S. Information theory and statistics. New York: John Wiley, 1959:8

〔2〕王松桂. 广义相关系数与估计效率. 科学通报, 1985; 30(19):1521～1529

〔3〕Rao C R. Linear statistical inference and its applications. New York:John Wiley, 1973:527

〔4〕Bartmann F C, Bloomfield P. Inefficiecy and correlation. Biometrika,1981; 68(1):67～11

〔5〕Bloomfield P, Watson G S. The inefficiency of least squares. Biometrika, 1975;62(1):121～128

CORRELATION MEASURE AND ESTIMATION EFFICIENCY

Liu Lanting　　Zhang Xindong

Abstract

In this paper, We give out the correlation measure relationship between the least square fitted values and the residuals to the relative efficiency of the least square estimation.

Key words Correlation measure　Estimation efficiency