研究简报

两类新的广义 Kantorovich 不等式及其应用

杨 虎

(重庆交通学院,重庆 630074)

关键词 广义 Kantorovich 不等式、条件数、相对效率

本文用谱范数和欧氏范数作为度量，给出了分别属于谱范数类和欧氏范数类的一系列广义 Kantorovich 不等式(Generalized Kantorovich inequality,简记为 GKI)，与之对应的是文献[1—4]中首先给出的行列式类、商迹类和迹商类 GKI,它们分别给出了 $\det(X'AXX'A^{-1}X)$, $\operatorname{tr}(X'AXX'A^{-1}X)$ 与 $\operatorname{tr}(X'AX)/\operatorname{tr}(X'A^{-1}X)^{-1}$ 的上界,这里 X 为 $n\times p$ 阶矩阵,满足 $X'X=I$, A 为 n 阶正定阵. 本文对这两类不等式分别给出了 $\|X'AXX'A^{-1}X\|_2$ 与 $\|X'AXX'A^{-1}X\|_E$ 的上界.

考虑 Kantorovich 不等式

$$x'Axx'A^{-1}x\leqslant\frac{(\lambda_1+\lambda_n)^2}{4\lambda_1\lambda_n},\tag{1}$$

其中 A 为 n 阶正定阵，x 为 n 维向量，满足 $x'x=1$，λ_1，λ_n 分别为 A 的最大和最小特征根. 如果将向量 x 推广成 $n\times p$ 列满秩矩阵 X,并考虑适当的度量就得到了 GKI，这些不等式均为(1)式的延拓,也就是说当 X 退化为向量时，它们就是(1)式,故这里的 GKI 都包含(1)式作为它们的特例. 使用不同的度量为我们提供了区分 GKI 的标准，因此，我们把现有的 GKI 分为三类:

i）行列式类. 用行列式度量的 GKI 为数众多,它的基本形式由 Bloomfield 和 Watson[1] 与 Knott[2]同时得到,即

$$\det(X'AXX'A^{-1}X)\leqslant\prod_{i=1}^{t}\frac{(\lambda_i+\lambda_{n-i+1})^2}{4\lambda_i\lambda_{n-i+1}}.\tag{2}$$

ii）商迹类. 其基本形式由 Khatri 和 Rao[3]提出,即

$$\operatorname{tr}(X'AXX'A^{-1}X)\leqslant\sum_{i=1}^{p}\frac{(\lambda_i+\lambda_{n-i+1})^2}{4\lambda_i\lambda_{n-i+1}},\ n\geqslant 2p.\tag{3}$$

iii）迹商类. 由本文作者[4]提出,其基本形式为

$$\operatorname{tr}(X'AX)/\operatorname{tr}(X'A^{-1}X)^{-1}\leqslant\left(\frac{\sum_{i=1}^{t}(\lambda_i+\lambda_{n-i+1})}{2\sum_{i=1}^{t}\sqrt{\lambda_i\lambda_{n-i+1}}}\right)^2.\tag{4}$$

本文 1989 年 7 月 3 日收到. 1989 年 9 月 21 日收到修改稿.

以上各式中 A 为 n 阶正定阵，X 为 $n\times p$ 列满秩矩阵，满足 $X'X=I$，$s=\min(p,\ n-p)$，$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$ 表示 A 的特征根，很明显，当 $p=1$ 时，以上各式均回到(1)式。目前，这三类 GKI 已得到深入的研究，特别是 Khatri 和 Rao[5]得到了 GKI 的很一般的形式。

下面我们先给出谱范数类和欧氏范数类 GKI 的基本形式。

定理 1 设 X 为 $n\times p$ 列满秩矩阵，满足 $X'X=I$，A 为 n 阶正定阵，λ_1，λ_n 分别为 A 的最大和最小特征根，τ_1，τ_2 为 $X'AX$ 与 $X'A^{-1}X$ 的条件数，则谱范数类 GKI 的基本形式为

$$\|X'AXX'A^{-1}X\|_2\leqslant\sqrt{\tau_1\tau_2}\,\frac{(\lambda_1+\lambda_n)^2}{4\lambda_1\lambda_n}. \tag{5}$$

显然，由于 $p=1$ 时 $\tau_1=\tau_2=1$，故(1)式为(5)式的特例。

定理 2 在(5)式同样的条件下，欧氏范数类 GKI 的基本形式为

$$\|X'AXX'A^{-1}X\|_E\leqslant\sqrt{\tau_1\tau_2}\sum_{i=1}^{p}\frac{(\lambda_i+\lambda_{n-i+1})^2}{4\lambda_i\lambda_{n-i+1}},\quad n\geqslant 2p, \tag{6}$$

其中 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$ 为 A 的特征根，容易看出(6)式亦为(1)式的延拓，且当 $n<2p$ 时，由文献[5]我们类似地有

$$\|X'AXX'A^{-1}X\|_E\leqslant\sqrt{\tau_1\tau_2}\sum_{i=1}^{n-p}\frac{(\lambda_i+\lambda_{n-i+1})^2}{4\lambda_i\lambda_{n-i+1}}+2p-n. \tag{7}$$

类似于其它三类 GKI 相应的推广形式，我们有如下的一些推论：

推论 1 设 X 为 $n\times p$ 列满秩矩阵，A 为 n 阶正定阵，且 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n>0$ 为它的特征根，τ_0、τ_1、τ_2 分别为 $X'X$、$X'AX$ 和 $X'A^{-1}X$ 的条件数，$n\geqslant 2p$，则

$$\|X'A^{-1}X(X'X)^{-1}X'AX(X'X)^{-1}\|_2\leqslant\sqrt{\tau_0^3\tau_1\tau_2}\,\frac{(\lambda_1+\lambda_n)^2}{4\lambda_1\lambda_n}, \tag{8}$$

$$\|X'A^{-1}X(X'X)^{-1}X'AX(X'X)^{-1}\|_E\leqslant\sqrt{\tau_0^3\tau_1\tau_2}\sum_{i=1}^{p}\frac{(\lambda_i+\lambda_{n-i+1})^2}{4\lambda_i\lambda_{n-i+1}}. \tag{9}$$

推论 2 设 X 为 $n\times p$ 列满秩矩阵，A、B 为 n 阶正定矩阵，$\mu_1\geqslant\mu_2\geqslant\cdots\geqslant\mu_n$ 为 $|A-\mu B|=0$ 的根，τ_0、τ_1、τ_2 分别为 $X'B^{-1}X$、$X'A^{-1}X$ 和 $X'B^{-1}AB^{-1}X$ 的条件数，$n\geqslant 2p$，则

$$\|X'A^{-1}X(X'B^{-1}X)^{-1}X'B^{-1}AB^{-1}X(X'B^{-1}X)^{-1}\|_2\leqslant\sqrt{\tau_0^3\tau_1\tau_2}\,\frac{(\mu_1+\mu_n)^2}{4\mu_1\mu_n}, \tag{10}$$

$$\|X'A^{-1}X(X'B^{-1}X)^{-1}X'B^{-1}AB^{-1}X(X'B^{-1}X)^{-1}\|_E\leqslant\sqrt{\tau_0^3\tau_1\tau_2}\sum_{i=1}^{p}\frac{(\mu_i+\mu_{n-i+1})^2}{4\mu_i\mu_{n-i+1}}. \tag{11}$$

推论 3 设 X 为 $n\times p$ 列满秩矩阵，A、B 为 n 阶非奇异对称阵，且使 $AB=BA$ 为正定阵，$\mu_1\geqslant\mu_2\geqslant\cdots\geqslant\mu_n$ 为 $|A-\mu B|=0$ 的根，τ_0、τ_1、τ_2 分别为 $XABX$、$X'A^2X$ 和 $X'B^2X$ 的条件数，$n\geqslant 2p$，则

$$\|X'A^2X(X'ABX)^{-1}X'B^2X(X'ABX)^{-1}\|_2\leqslant\sqrt{\tau_0^3\tau_1\tau_2}\,\frac{(\mu_1+\mu_n)^2}{4\mu_1\mu_n}, \tag{12}$$

$$\|X'A^2X(X'ABX)^{-1}X'B^2X(X'ABX)^{-1}\|_E\leqslant\sqrt{\tau_0^3\tau\ \tau_2}\sum_{i=1}^{p}\frac{(\mu_i+\mu_{n-i+1})^2}{4\mu_i\mu_{n-i+1}}. \tag{13}$$

考虑线性模型

$$Y = X\beta + \varepsilon, \tag{14}$$

Y 为 n 维观测向量，X 为 $n\times p$ 列满秩设计阵，β 为参数向量，ε 为 n 维随机误差向量，满足 $E(\varepsilon)=0$，$\mathrm{Cov}(\varepsilon)=\Sigma$(正定)，$\beta$ 的最佳线性无偏估计为

$$\tilde{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y, \tag{15}$$

而 β 的最小二乘估计为 $\hat{\beta}=(X'X)^{-1}X'Y$，一般来说，我们用 $\hat{\beta}$ 代替 $\tilde{\beta}$ 作为 β 的估计要受一些损失，在统计中，常用 $\hat{\beta}$ 相对于 $\tilde{\beta}$ 的效率来作为这种损失的度量。文献[4]中给出了四种相对效率的度量，它们的下界可以由前面中给出的三类 GKI 得到。类似地，我们定义两种新的相对效率：

$$\rho_5 = \|(\mathrm{Cov}\tilde{\beta})^{-1}\mathrm{Cov}\hat{\beta}\|_2^{-1}, \tag{16}$$

$$\rho_6 = \|(\mathrm{Cov}\tilde{\beta})^{-1}\mathrm{Cov}\hat{\beta}\|_E^{-1}. \tag{17}$$

由于 $\mathrm{Cov}\tilde{\beta}=(X'\Sigma^{-1}X)^{-1}$，$\mathrm{Cov}\hat{\beta}=(X'X)^{-1}X'\Sigma X(X'X)^{-1}$，可见(16)和(17)式的下限可以分别通过推论 1 中给出的两种 GKI (8) 和(9)式得出.

参　考　文　献

[1] Bloomfield, P., Watson, G. S., *Biometrika*, **62** (1975), 121—128.

[2] Knott, M., *Biometrika*, 62 (1975), 129—132.

[3] Khatri, C. G., Rao, C. R., *J. Multi. Anal.*, **11**(1981), 498—505.

[4] 杨虎，应用数学，**1**(1988)，4：85—90.

[5] Khatri, C. G., Rao, C. R., *Indian J. Statist.*, Ser. A, **44**(1982), 91—102.