On Bounds for Scaled Projections and Pseudoinverses

Dianne P. O'Leary*
Computer Science Department
and Institute for Advanced Computer Studies
University of Maryland
College Park, Maryland 20742

Submitted by G. W. Stewart

ABSTRACT

Let X be a matrix of full column rank, and let D be a positive definite diagonal matrix. In a recent paper, Stewart considered the weighted pseudoinverse $X_D^\dagger = (X^T DX)^{-1} X^T D$ and the associated oblique projection $P_D = XX_D^\dagger$, and gave bounds, independent of D, for the norms of these matrices. In this note, we answer a question he raised by showing that the bounds are computable.

Let X be a matrix of full column rank, and let D be a positive definite diagonal matrix. In a recent paper, Stewart [1] considered the weighted pseudoinverse $X_D^\dagger = (X^T DX)^{-1} X^T D$ and the associated oblique projection $P_D = XX_D^\dagger$. He proved two results. The first is that the spectral norms of these matrices are bounded independently of D as

$$\sup_{D \in \mathscr{D}_+} \|P_D\| \leqslant \rho^{-1}$$

and

$$\sup_{D \in \mathscr{D}_+} \|X_D^\dagger\| \leqslant \rho^{-1} \|X^\dagger\|,$$

*This work was supported by the Air Force Office of Scientific Research under Grant 87-0188.

© Elsevier Science Publishing Co., Inc., 1990
655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

where

$$\rho \overset{\text{def}}{=} \inf_{\substack{y \in \mathscr{Y} \\ x \in \mathscr{X}}} \|y - x\| > 0, \tag{1}$$

with

$$\mathscr{X} = \{x \in \mathscr{R}(X) : \|x\| = 1\}, \tag{2}$$

$$\mathscr{Y} = \{y : \exists D \in \mathscr{D}_+ \text{ such that } X^T D y = 0\}. \tag{3}$$

His second result is that if the columns of U form an orthonormal basis for $\mathscr{R}(X)$, then

$$\rho \leqslant \min \inf_+ (U_I), \tag{4}$$

where U_I denotes any submatrix formed from a nonempty set of rows of U.

In this note, we answer a question he raised by showing that

$$\rho = \min \inf_+ (U_I).$$

Since $\mathscr{X}$ and $\mathscr{Y}$ depend only on the range of X and not on its entries, we can replace X in (2) and (3) by U. Thus,

$$\mathscr{X} = \{U\alpha : \|\alpha\| = 1\}.$$

Let the sign of a scalar t be defined by

$$\operatorname{sg}(t) = \begin{cases} 1 & \text{if } \; t > 0, \\ 0 & \text{if } \; t = 0, \\ -1 & \text{if } \; t < 0, \end{cases}$$

and let the sign of a vector z be denoted by $\operatorname{sg}(z)$ and defined componentwise. Then $\mathscr{Y}$ has the property that for any vector $\hat{y} \in \mathscr{Y}$, every vector y with $\operatorname{sg}(y) = \operatorname{sg}(\hat{y})$ is also an element of $\mathscr{Y}$. This is verified by letting D be the nonnegative diagonal matrix such that $U^T D \hat{y} = 0$. Then $U^T D S y = 0$, where S is the diagonal matrix with

$$s_{ii} = \begin{cases} \hat{y}_i / y_i & \text{if } \; y_i \neq 0, \\ 1 & \text{if } \; y_i = 0. \end{cases}$$

Now,

$$\begin{aligned}\rho &= \inf_{\substack{y \in \mathscr{Y} \\ x \in \mathscr{X}}} \|y - x\| \\ &= \inf_{\hat{y} \in \mathscr{Y}} \inf_{\substack{\mathrm{sg}(y) = \mathrm{sg}(\hat{y}) \\ x \in \mathscr{X}}} \|y - x\| \\ &= \inf_{\hat{y} \in \mathscr{Y}} \inf_{\substack{\mathrm{sg}(y) = \mathrm{sg}(\hat{y}) \\ \|\alpha\| = 1}} \|y - U\alpha\|.\end{aligned}$$

In the inner infimum, for every choice of α there is a set of rows of $U\alpha$ that agree in sign with $\hat{y}$ and a set that disagree. The set of rows that disagree in sign must be nonempty; otherwise $y = U\alpha \in \mathscr{Y}$, and the infimum would be zero, which contradicts (1). Let the set of those that disagree be denoted by the subscript I. For this choice of α, the best y equals $U\alpha$ in all rows that agree in sign and has elements zero or arbitrarily close to zero in the other rows. The resulting value of $\|y - U\alpha\|$ is no less than $\|(U\alpha)_I\| = \|U_I\alpha\|$, and this value is bounded below by the smallest singular value of U_I. Thus we have shown that

$$\rho \geqslant \min \inf_{+}(U_I),$$

and combining this with Stewart's result (4) establishes the equality.

REFERENCES

1 G. W. Stewart, On Scaled Projections and Pseudoinverses, *Linear Algebra Appl.*, 112 (1989) 189–194.

Received 3 March 1989; final manuscript accepted 9 March 1989