Sankhyā : The Indian Journal of Statistics
1984, Volume 46, Series A, Pt. 2 pp. 277-286.

CONVERGENCE OF GENERALIZED INVERSES WITH APPLICATIONS TO ASYMPTOTIC HYPOTHESIS TESTING

By MADAN L. PURI and CARL T. RUSSELL
Indiana University, Bloomington
and
THOMAS MATHEW
Indian Statistical Institute

SUMMARY. Let A_N, A be $m \times n$ matrices with $A_N \to A$. It is shown that $R(A_N) \to R(A)$ is necessary for the convergence of any sequence of generalized inverses $A_N^- \to A^-$, and sufficient conditions are given for the existence of a convergence sequence of g-inverses with specified row and column spaces. This generalizes a result of Stewart (1969). Applications to asymptotic hypothesis testing are discussed and an optimal property of the Moore-Penrose inverse is presented.

1. Notation and preliminary results

Boldface capital letters denote matrices, and boldface lower-case letters denote column vectors over the complex field, $\mathbf{0}_{m\times n}$ denotes the zero matrix of order $m \times n$ and will be written an "0" when the order is clear from context. $\boldsymbol{A}^*$, $\boldsymbol{A}'$, $R(\boldsymbol{A})$, $\mathcal{M}(\boldsymbol{A})$, and $O(\boldsymbol{A})$ denote, respectively, the conjugate transpose of $\boldsymbol{A}$, the transpose of $\boldsymbol{A}$ when $\boldsymbol{A}$ has real components, the rank of $\boldsymbol{A}$, the column space of $\boldsymbol{A}$, and the space orthogonal to $\mathcal{M}(\boldsymbol{A})$ with respect to the usual inner product. For two subspace $\mathcal{S}$ and $\mathcal{T}$ of the same vector space, $\mathcal{S} \cap \mathcal{T}$ denotes the intersection, and if $\mathcal{S} \cap \mathcal{T} = \{0\}$, $\mathcal{S} \oplus \mathcal{T}$ denotes the direct sum; $\delta(\mathcal{S})$ denotes the dimension of $\mathcal{S}$. $\|x\| = (\boldsymbol{x}^*\boldsymbol{x})^{\frac{1}{2}}$, $\|A\| = \sup\{\|Ax\|: \|x\| = 1\}$; $\boldsymbol{x}_N \to \boldsymbol{x}$ means $\lim_{N\to\infty} \|x_N - x\| = 0$, $\boldsymbol{A}_N \to \boldsymbol{A}$ means $\boldsymbol{A}_N\boldsymbol{x} \to \boldsymbol{A}\boldsymbol{x}$ for all $\boldsymbol{x}$ or equivalently $\|\boldsymbol{A}_N - \boldsymbol{A}\| \to 0$. The frequent references to Rao and Mitra (1971) will be indicated by RM.

Definition. $\boldsymbol{G}$ is said to be a *generalized inverse* of $\boldsymbol{A}$ if

$$\boldsymbol{AGA} = \boldsymbol{A}. \qquad \ldots \quad (1)$$

G is also called a g-inverse of $\boldsymbol{A}$ and written $\boldsymbol{G} = \boldsymbol{A}^-$. $\boldsymbol{G}$ is said to be a reflexive g-inverse of $\boldsymbol{A}$ if in addition to (1)

$$\boldsymbol{GAG} = \boldsymbol{G} \qquad \ldots \quad (2)$$

holds, in which case one writes $\boldsymbol{G} = \boldsymbol{A}_r^-$.

AMS (1980) *subject classification :* 15 A09, 62F05

Key words and phrases : Generalized inverse, Moore-Penrose inverse, Asymptotic hypothesis testing.

By Lemma 2.5.1 (RM), if $G = A^-$, then $G = A_r^-$ iff $R(G) = R(A)$. It is worthwhile noticing that a reflexive g-inverse is uniquely determined by its row and column spaces.

Lemma 1 : *If* $G_1 = A_r^-$ *and* $G_2 = A_r^-$, *then* $G_1 = G_2$ *iff* $\mathcal{M}(G_1) = \mathcal{M}(G_2)$ *and* $\mathcal{M}(G_1^*) = \mathcal{M}(G_2^*)$.

Proof : Necessity is obvious.

If $G_1 = DG_2$ and $G_2 = G_1E$, then

$$G_1 = DG_2 = DG_2AG_2 = G_1AG_1E = G_1E = G_2. \qquad \text{Q.E.D.}$$

In a sense, all g-inverses are reflexive, as the following lemma shows.

Lemma 2. (RM Theorem 2.7.1) : *Let* A *be of order* $m \times n$ *and rank* a, *and let* r *be an integer satisfying* $a \leqslant r \leqslant min(m, n)$. *Then* $G = A^-$ *iff* $G = (A+MN)_r^-$ *where* M *of order* $m \times (r-a)$ *and* N *of order* $(r-a) \times n$ *are arbitrary matrices satisfying* $R(A : M) = R(A^* : N^*) = r$.

The following simple lemma is actually the key to the proofs in the next section.

Lemma 3 : *If* $A_N \to A$, *then* $R(A_N) \geqslant R(A)$ *for* N *sufficiently large.*

Proof : Let $r = R(A)$ and let B and C be matrices such that $BAC = I_r$ (the $r \times r$ identity matrix). Then $BA_NC \to BAC$ so $|BA_NC| \to |BAC| = 1$ (where $|\cdot|$ denotes determinant). Thus $R(A_N) \geqslant r$ for N sufficiently large. Q.E.D.

2. Necessary and sufficient conditions for convergence of G-inverses

Stewart (1969) showed that if $A_N \to A$ then $A_N^+ \to A^+$ if $R(A_N) \to R(A)$ (where $(\cdot)^+$ denotes the Moore-Penrose inverse defined preceding Corollary 8). We now show that this condition is necessary for arbitrary choices of A_N^-.

Theorem 4 : *Suppose* $A_N \to A$ *and let* $G_N = A_N^-$. *In order that* G_N *converge, it is necessary that* $R(A_N) \to R(A)$.

Proof : Suppose $R(A_N) \not\to R(A)$. Then, there exists a sequence $\{A_{N_k}\}$ such that $R(A_{N_k}) > R(A)$, by Lemma 3. There exists $x_{N_k} \in \mathcal{M}(G_{N_k} A_{N_k})$ such that $Ax_{N_k} = 0$ and $\|x_{N_k}\| = 1$. Thus

$$x_{N_k} = G_{N_k}A_{N_k}x_{N_k} = G_{N_k}(A_{N_k}-A+A)x_{N_k} = G_{N_k}(A_{N_k}-A)x_{N_k}$$

so that $\quad 1 = \|x_{N_k}\| \leqslant \|G_{N_k}\| \; \|A_{N_k}-A\| \; \|x_{N_k}\| = \|G_{N_k}\| \; \|A_{N_k}-A\|.$

It follows that $\|G_{N_k}\| \geqslant \|A_{N_k}-A\|^{-1} \to \infty$, so G_{N_k} does not converge. Q.E.D.

Since neither A_N^- nor $R(A_N^-)$ is in general unique, $A_N \to A$ and $R(A_N) \to R(A)$ can not alone guarantee $A_N^- \to A^-$. (For example let

$$A_N = A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } A_N^- = \begin{pmatrix} 1 & a_N \\ b_N & c_N \end{pmatrix}, \quad A^- = \begin{pmatrix} 1 & a \\ b & c \end{pmatrix},$$

a_N, b_N, c_N, a, b, c completely arbitrary.) We show, however, that if allowable convergent row and column spaces are specified, then there is a convergent sequence of g-inverses with the specified row and column spaces. The following result of Rao and Mitra gives necessary and sufficient conditions to realize a g-inverse with row and column spaces contained in specified spaces.

Lemma 5 (RM Lemma 4.4.1) : *Given matrices* A, P, Q, *a necessary and sufficient condition for* A *to have a g-inverse of the form* $G = PCQ$ *is that* $R(QAP) = R(A)$ *in which case the only choices for* C *are* $(QAP)^-$. *An inverse with the required property is unique, if further* $R(P) = R(Q) = R(A)$.

The following criterion is also needed.

Lemma 6 : $R(QAP) = R(A)$ *iff*

$$R(A) = R(P) - \delta(\mathcal{M}(P) \cap O(A^*)) = R(Q) - \delta(\mathcal{M}(Q^*) \cap O(A)).$$

Proof : It is easy to show that $R(AP) = R(P) - \delta(\mathcal{M}(P) \cap O(A^*))$ and $R(A^*Q^*) = R(Q^*) - \delta(\mathcal{M}(Q^*) \cap O(A))$. Q.E.D.

In view of Theorem 4 and Lemma 5, the following result is the best one could hope to obtain by specifying row and column spaces for g-inverses.

Theorem 7 : *For* $N = 1, 2, \ldots$ *let* A_N *be an* $m \times n$ *matrix of rank* a_N, *let* S_N *be an* $n \times s$ *matrix of rank* r_N, *and let* T_N *be a* $t \times m$ *matrix of rank* r_N. *Let* A *be an* $m \times n$ *matrix of rank* a, *let* S *be an* $n \times s$ *matrix of rank* r, *and let* T *be a* $t \times m$ *matrix of rank* r. *Suppose that* $A_N \to A$, $S_N \to S$, $T_N \to T$, *and* $a_N \to a$. *Suppose*

$$R(T_N A_N S_N) = R(A_N) - a_N \quad \text{for } N = 1, 2, \ldots$$

and

$$R(TAS) = R(A) = a.$$

Then there exist matrices G_N, G *such that* $G_N = A_N^-$, $G = A^-$. $\mathcal{M}(G_N) = \mathcal{M}(S_N)$, $\mathcal{M}(G_N^*) = \mathcal{M}(T_N^*)$, $\mathcal{M}(G) = \mathcal{M}(S)$, $\mathcal{M}(G^*) = \mathcal{M}(T^*)$, *and* $G_N \to G$. *If* $r_N = a_N$ *for* $N = 1, 2, \ldots$, *then* G_N *and* G *are the unique reflexive g-inverses having the specified row and column spaces.*

Proof : (The basic idea behind this proof was suggested by S. K. Mitra.) By Lemma 6, $\delta(\mathcal{M}(S) \cap O(A^*)) = r - a = \delta(\mathcal{M}(T^*) \cap O(A))$. Thus there exist nonsingular matrices B and C such that $SB = (B_{10} : B_{20} : 0_{n \times (s-r)})$,

$(CT)^* = (C_{10}^* : C_{20}^* : 0_{m\times(l-r)})$, $\mathscr{M}(B_{20}) = \mathscr{M}(S) \cap O(A^*)$, $\mathscr{M}(B_{10} : B_{20}) = M(S)$, $B_{20}^* B_{10} = 0_{(r-a)\times a}$, $\mathscr{M}(C_{20}^*) = \mathscr{M}(T^*) \cap O(A)$, $\mathscr{M}(C_{10}^* : C_{20}^*) = \mathscr{M}(T^*)$, and $C_{10}C_{20}^* = 0_{a\times(r-a)}$. Considering $\{S_N B\}_{N=1}^{\infty}$ and $\{CT_N\}_{N=1}^{\infty}$, it is clear that one may assume $S = (S_{10} : S_{20} : 0_{n\times(s-r)})$, $T^* = (T^*{}_{10} : T_{20}^* : 0_{m\times(l-r)})$ where $\mathscr{M}(S_{20}) = \mathscr{M}(S) \cap O(A^*)$, $S_{20}^* S_{10} = 0_{(r-a)\times a}$, $\mathscr{M}(T_{20}^*) = \mathscr{M}(T^*) \cap O(A)$, and $T_{10}T_{20}^* = 0_{a\times(r-a)}$.

Write $S_N = (S_{1N} : S_{2N} : S_{3N})$ and $T_N^* = (T_{1N}^* : T_{2N}^* : T_{3N}^*)$, where S_{1N}, S_{2N}, S_{3N}, T_{1N}, T_{2N}, T_{3N} are of orders $n\times a$, $n\times(r-a)$, $n\times(s-r)$, $a\times m$, $(r-a)\times m$, and $(l-r)\times m$, respectively, so that $S_{1N}\to S_{10}$, $S_{2N}\to S_{20}$, $S_{3N}\to 0$, $T_{1N}\to T_{10}$, $T_{2N}\to T_{20}$, and $T_{3N}\to 0$. Since

$$a = R(TAS) = R\begin{pmatrix} T_{10}AS_{10} & 0 \\ 0 & 0 \end{pmatrix} = R(T_{10}AS_{10})$$

and
$$R(T_{20}T_{20}^* S_{20}^* S_{20}) = r-a$$

it follows that

$$R\left(\begin{pmatrix} T_{10} \\ T_{20} \end{pmatrix}(A+T_{20}^*S_{20}^*)(S_{10} : S_{20})\right) = R\begin{pmatrix} T_{10}AS_{10} & 0_{a\times(r-a)} \\ 0_{(r-a)\times a} & T_{20}T_{20}^*S_{20}^*S_{20} \end{pmatrix} = r$$

so by Lemmas 2 and 5,

$$G = (S_{10} : S_{20})\left[\begin{pmatrix} T_{10} \\ T_{20} \end{pmatrix}(A+T_{20}^*S_{20}^*)(S_{10} : S_{20})\right]^{-1}\begin{pmatrix} T_{10} \\ T_{20} \end{pmatrix}$$

is a g-inverse of A with $\mathscr{M}(G) = \mathscr{M}(S)$ and $\mathscr{M}(G^*) = \mathscr{M}(T^*)$. Moreover, $(A_N : T_{2N}^*) \to (A : T_{20}^*)$, $(A_N^* : S_{2N}) \to (A^* : S_{20})$, and $R(A_N : T_{2N}^*)$, $R(A_N^* : S_{2N}) \leqslant a_N + (r-a)$ so by Lemma 3 $R(A_N : T_{2N}^*) = R(A_N^* : S_{2N}) = r$ for N sufficiently large. Likewise

$$R\left(\begin{pmatrix} T_{1N} \\ T_{2N} \end{pmatrix}(A_N+T_{2N}^*\, S_{2N}^*)(S_{1N} : S_{2N})\right) = r$$

for N sufficiently large. Thus it follows from Lemmas 2 and 5 that for $N \geqslant N_0$ (N_0 sufficiently large)

$$G_{0N} = (S_{1N} : S_{2N})\left[\begin{pmatrix} T_{1N} \\ T_{2N} \end{pmatrix}(A_N+T_{2N}^*\, S_{2N}^*)(S_{1N} : S_{2N})\right]^{-1}\begin{pmatrix} T_{1N} \\ T_{2N} \end{pmatrix}$$

is a g-inverse of both $A_{0N} = A_N + T_{2N}^*\, S_{2N}^*$ and A_N, $\mathscr{M}(A_N) \subset \mathscr{M}(A_{0N})$ and $\mathscr{M}(A_N^*) \subset \mathscr{M}(A_{0N}^*)$, and $\mathscr{M}(G_{0N}) = \mathscr{M}(S_{1N} : S_{2N}) \subset \mathscr{M}(S_N)$ and $\mathscr{M}(G_{0N}^*) = \mathscr{M}(T_{1N}^* : T_{2N}^*) \subset \mathscr{M}(T_N^*)$. Clearly $G_{0N} \to G$.

Let $N \geqslant N_0$. By Lemmas 5 and 6,

$$\delta(\mathscr{M}(S_{1N} : S_{2N}) \cap O(A_{0N}^*)) = \delta(\mathscr{M}(T_{1N}^* : T_{2N}^*) \cap O(A_{0N})) = 0$$

and $$\delta(\mathcal{M}(S_N)\cap O(A^*_{0N}))=\delta(\mathcal{M}(T^*_N)\cap O(A_{0N}))=r_N-r.$$

If $r_N>r$, let the columns of B_N and C^*_N consist of orthonormal bases for $\mathcal{M}(S_N)\cap O(A^*_{0N})$ and $\mathcal{M}(T^*_N)\cap O(A_{0N})$, respectively. If $r_N=r$ let $B_N=0_{n\times 1}$ and $C_N=0_{1\times m}$. Define

$$G_N=G_{0N}+N^{-1}B_NC_N \quad (N\geqslant N_0).$$

Clearly $G_N\to G$. Since $O(A_{0N})\subset O(A_N)$ and $O(A^*_{0N})\subset O(A^*_N)$, $A_NG_NA_N=A_NG_{0N}A_N+N^{-1}A_NB_NC_NA_N=A_N$. Finally,

$$\mathcal{M}(G_N)=\mathcal{M}(G_{0N})\oplus\mathcal{M}(B_N)=\mathcal{M}(S_N)$$

and

$$\mathcal{M}(G^*_N)=\mathcal{M}(G^*_{0N})\oplus\mathcal{M}(C^*_N)=\mathcal{M}(T^*_N).$$

For $N<N_0$ let G_N be any g-inverse of A_N satisfying $\mathcal{M}(G_N)=\mathcal{M}(S_N)$ and $\mathcal{M}(G^*_N)=\mathcal{M}(T^*_N)$ (one may use the construction which yielded G).

G_N and G defined above satisfy all the specified requirements. If $r_N=a_N$ for all N, then by Lemma 1, G_N and G are the unique reflexive g-inverses of A_N and A having the specified row and column spaces. Q.E.D.

Since one may define the Moore-Penrose inverse A^+ of a matrix A to be the unique g-inverse of A satisfying $\mathcal{M}(A^+)=\mathcal{M}(A^*)$, $\mathcal{M}(A^{+*})=\mathcal{M}(A)$ (A^+ exists by Lemmas 6 and 5), the result of Stewart (1969) referred to earlier follows immediately from Theorems 4 and 7.

Corollary 8 : *Let $A_N\to A$. Then $A^+_N\to A^+$ iff $R(A_N)\to R(A)$.*

Another way to specify a unique g-inverse G of A is to specify square matrices E and F with

$$R(E)=R(F)=R(FAE)=R(A) \qquad \dots \quad (3)$$

and require $\mathcal{M}(G)=\mathcal{M}(E)$, $\mathcal{M}(G^*)=\mathcal{M}(F^*)$. We shall denote such g-inverses by A^-_{EF}. By Lemma 5, $A^-_{EF}=E(FAE)^+F$. If in particular E and F are diagonal matrices, then A^-_{EF} is the matrix obtained by striking from A the columns corresponding to zero diagonal entries in E and the rows corresponding to zero diagonal entries in F, inverting this reduced matrix and finally expanding again by adding zeros to obtain A^-_{EF} (cf. RM(11.2.3)). Considering a sequence $A_N\to A$ and specified E, F, it follows from Theorem 7 that the condition (3) required for existence of $(A_N)^-_{EF}$ and A^-_{EF} automatically guarantees $(A_N)^-_{EF}\to A^-_{EF}$.

3. Applications to asymptotic hypothesis testing

In asymptotic hypothesis testing one often bases a test on a quadratic form

$$Q_N = X'_N B_N X_N$$

where X_N is asymptotically p-variate normal $\mathcal{N}_p(\mu, A)$ under an appropriate null hypothesis (in which case $\mu = 0$) or a sequence of "near-by" alternative hypotheses. In general the dispersion matrix A is unknown but a sequence of consistent n.n.d. estimators $\hat{A}_N \to A$ (stochastically) is available. If $\hat{A}_N$ is nonsingular, one takes $B_N = \hat{A}_N^{-1}$. The case where A_N is singular is frequently dismissed by remarking that one can work with a largest nonsingular minor and the corresponding variates; one then assumes that these reduced matrices converge to a nonsingular matrix. This amounts to assuming that $\hat{A}_N \to A$ (stochastically) and setting $B_N = (\hat{A}_N)^-_{EF}$ for some fixed choice of a diagonal matrix $E = F$ where $R(E) = R(EA) = R(A)$. If $\hat{A}_N$ is unbiased for A, then $\mathcal{M}(A_N) = \mathcal{M}(A)$ for all N sufficiently large (for proof see Appendix) in particular, $R(\hat{A}_N) \to R(A)$ (a.s.). In general, however, $\hat{A}_N$ is not unbiased for A and no useful conditions are known to guarantee the stochastic convergence $R(\hat{A}_N) \to R(A)$.

Suppose in what follows that the assumption $R(\hat{A}_N) \xrightarrow{P} R(A\ ) = r$ is valid.* Thus one can specify (Theorem 7) convergent g-inverses $A^-_N \xrightarrow{P} A^-$ (in particular, Corollary 8 exhibits such choice), and it follows that the asymptotic distribution of $Q_N = X'_N \hat{A}^-_N X_N$ is the same as that of $Q = X'A^-X$ where X has distribution $\mathcal{N}_p(\mu, A)$. By Theorem 9.2.3 (RM), Q has chi-square distribution with $r = R(A)$ degrees of freedom and noncentrality parameter $\mu' A^- \mu$ provided either $\mu \in \mathcal{M}(A)$ or A^- is a symmetric reflexive g-inverse. Thus if $\hat{A}^-_N$ and A^- are symmetric reflexive g-inverses of $\hat{A}_N$ and A, respectively, chosen such that $\hat{A}^-_N \xrightarrow{P} A^-$ (in particular $\hat{A}^+_N$, A^+ or an appropriate choice of $(\hat{A}_N)^-_{EF}$, A^-_{EF} will do), then Q_N has asymptotically chi-square distribution with r d.f. and noncentrality parameter $\mu' A^- \mu$. Denote by χ^2_α the upper α point of the central chi-square distribution with r d.f. and assume $r < p$

Consider first the following restrictive case. Let X_N have mean m_N and dispersion matrix D_N, and suppose that

$$\mathcal{M}(m_N) = \mathcal{M}(\mu) \text{ and } \mathcal{M}(\hat{A}_N) = \mathcal{M}(D_N) = \mathcal{M}(A).$$

*The notation "$\xrightarrow{P}$" denotes stochastic convergence.

(This will be the case in particular whenever X_N is unbiased for μ and $\hat{A}_N$ is unbiased for $D_N = A$ as in the following classical situation : Let U_N have mean $N^{-\frac{1}{2}}\mu$ and dispersion matrix A, let $Y_1, \ldots, Y_N$ be N independent copies of U_N, let

$$X_N = N^{\frac{1}{2}}\bar{Y}_N, \text{ and let } \hat{A}_N = (N-1)^{-1}\sum_{i=1}^{N}(Y_i-\bar{Y}_N)(Y_i-\bar{Y}_N)'.$$

It should be observed that $R(\hat{A}_N) = R(A)$ for $N > r$. Thus $\mathcal{M}(\hat{A}_N) = \mathcal{M}(A)$ for all $N > r$.

For $N > r$, let C be a $p\times m$ matrix of rank $p-r$ satisfying $C'\hat{A}_N = 0$ and consider the test Φ_{1N} which rejects $H_0 : \mu = 0$ when $C'X_N \neq 0$ as well as when $Q_N > \chi^2_r$ and the usual test Φ_{0N} which rejects H_0 only when $Q_N > \chi^2_r$. Then Φ_{1N} has the same size as Φ_{0N} but has power 1 against alternatives $\mu \notin \mathcal{M}(A)$.

Unfortunately the preceding test Φ_{1N} is not robust against even minor deviations from $\mathcal{M}(m_N) = \mathcal{M}(\mu)$ and $\mathcal{M}(\hat{A}_N) = \mathcal{M}(D_N) = \mathcal{M}(A)$ since $P[X_N - m_N \in \mathcal{M}(D_N)] = 1$. In fact one often knows only that $X_N \xrightarrow{P} \mu$, $D_N \rightarrow A$, and $\hat{A}_N \xrightarrow{P} A$ so that Φ_{1N} is not applicable and one must hope to detect arbitrary deviations from H_0 using Φ_{0N} alone. But no choice of Φ_{0N} is sensitive against arbitrary deviations $\mu \in \mathcal{M}(A)$ since

$$\lim_{N\rightarrow\infty} P_\mu[X'_N\hat{A}^-_N X_N > \chi^2_r] = \alpha \text{ for } \mu \in O(A^-).$$

Moreover, different symmetric reflexive choices for $\hat{A}^-_N$ yield different spaces $O(A^-)$ and different noncentrality parameters $\mu' A^-\mu$. Furthermore, even for fixed $\mu \in \mathcal{M}(A)$, there is no g-inverse G of A which maximizes $\mu'G\mu$.

Example 11. Let $e_1, \ldots, e_r$ be an orthonormal basis for $\mathcal{M}(A)$, let $e_{r+1}, \ldots, e_p$ be an orthonormal basis for $O(A)$, and write matrices and vectors with respect to the basis $e_1, \ldots, e_p$. Let $1 \leqslant k \leqslant \min(r, p-r)$ and $l = r-k$. Then

$$A = \begin{pmatrix} D_k^2 & 0 & 0 \\ 0 & D_l^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where $D_k = \text{diag}(d_1, \ldots, d_k)$, $D_l = \text{diag}(d_{k+1}, \ldots, d_r)$, and $d_1, \ldots, d_r$ are positive. Also for $t \in (-\infty, \infty)$

$$A_t^- = \begin{pmatrix} D_k^{-2} & 0 & tD_k^{-2} & 0 \\ 0 & D_l^{-2} & 0 & 0 \\ t\,D_k^{-2} & 0 & t^2D_k^{-2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

is a symmetric reflexive g-inverse of A (notice that $A^+ = A_0^-$). If $\mu = (\mu_1, \ldots, \mu_p)'$, then

$$\mu' A_t^- \mu = \sum_{i=1}^{k} d_i^{-2}(\mu_i + t\mu_{r+i})^2 + \sum_{i=k+1}^{r} d_i^{-2}\mu_i^2,$$

so if $(\mu_{r+1}, \ldots, \mu_p)' \neq 0$, $\mu' A_t^- \mu$ increases without bound as $|t| \to \infty$. In particular, if for fixed $\mu \notin \mathcal{M}(A)$, e_1 and e_{r+1} are chosen so that $\mu_i = 0$ for $1 \neq i \neq r+1$, then $\mu' A_t^- \mu = d_1^{-2}(\mu_1 + t\mu_{r+1})^2$, so if G is any fixed g-inverse of A and $K > 0$ is given, one can find a symmetric reflexive g-inverse $A^-(G, K)$ such that $\mu' A^-(G, K)\mu > K\mu' G\mu$.

One should not conclude, however, that a g-inverse of the type A_t^- is generally desirable. For even though the test Φ_{0N} using A_t^- for large $|t|$ is highly sensitive agyinst certain alternatives μ far from $\mathcal{M}(A)$ (in the sense that $\sum_{i=1}^{r} \mu_i^2 / \left(\sum_{i=1}^{p} \mu_i^2\right)$ is small) such a test is necessarily relatively insensitive against certain alternatives near $\mathcal{M}(A)$; more precisely, if $\mu_1 < 0$, $t\mu_{r+1} > 0$ and $\mu_i = 0$ $(1 \neq i \neq r+1)$, then $\mu' A_t^- \mu < \mu' A^+ \mu$ holds whenever $\mu_1(t\mu_{r+1})^{-1} < -\frac{1}{2}$. (Even without making a special choice $\mu_i = 0$ $(1 \neq i \neq r+1)$, it is clear from the expression $\mu' A_t^- \mu$ that there exists μ satisfying $\mu' A_t^- \mu < \mu' A^+ \mu$ for any given t). In general, as long as $\mu \in O(A)$, the maximum gain in sensitivity for detecting μ far from $\mathcal{M}(A)$ which can be realized by choosing a reflexive g-inverse G different from $A_0^- = A^+$ is more than offset by the maximum loss in sensitivity for detecting μ near $\mathcal{M}(A)$ in the sense given in the following lemma.

Lemma 12 : *Let G be any symmetric reflexive g-inverse of A different from A^+, and let $S_\epsilon = \{\mu : \|P_A \mu\| \geqslant \epsilon$ and $\|\mu\| = 1\}$ where P_A is the orthogonal projector on $\mathcal{M}(A)$. Then there is an ϵ_0 (depending on G) such that for $0 < \epsilon \leqslant \epsilon_0$*

$$\max_{\mu \in S_\epsilon} \frac{\mu' A^+ \mu}{\mu' G \mu} > \max_{\mu \in S_\epsilon} \left\{\frac{\mu' G \mu}{\mu' A^+ \mu}\right\}$$

where $\max[a, \infty) = \infty$. In fact for $0 < \epsilon \leqslant \epsilon_0$, there is a $\mu_0 \in S_\epsilon$ such that $\mu_0' A^+ \mu_0 > \mu_0 G \mu_0 = 0$.

Proof : It suffices to prove the last statement of the lemma.

In the notation of Example 11, an arbitrary symmetric reflexive g-inverse G of A has the form

$$G = \begin{pmatrix} D^{-1} \\ B'D \end{pmatrix} (D^{-1} : DB)$$

where $D = \text{diag}(d_1, \ldots, d_r)$. Since $G \neq A^+$, $DB \neq 0$ so that also $K = \|D^{\frac{1}{2}}B\| > 0$. Let $\varepsilon_0 = K^2/(K^2+1)$. and let $0 < \varepsilon \leqslant \varepsilon_0$ so that $\varepsilon/(1-\varepsilon) \leqslant K^2$. Choose v such that $\|D^{\frac{1}{2}}Bv\|^2/\|v\|^2 > \varepsilon/(1-\varepsilon)$. and let $u = -D^{\frac{1}{2}}Bv$. Then, $\mu_0 = (\|u\|^2+\|v\|^2)^{-\frac{1}{2}}(u' : v')' \in S_\varepsilon$, since

$$\|P_A\mu_0\| = \|u\|/(\|u\|^2+\|v\|^2)^{\frac{1}{2}} > \varepsilon$$

and $\|\mu_0\| = 1$.

Furthermore, $$\mu_0' A^+ \mu_0 = \frac{\|DBv\|^2}{\|u\|^2+\|v\|^2} > 0$$

while $$\mu_0' G \mu_0 = \frac{\|-DBv+DBv\|}{\|u\|^2+\|v\|^2} = 0. \quad \text{Q.E.D.}$$

Lemma 12 shows that the Moore-Penrose inverse is in one sense optimal if one wishes to consider alternatives $\mu \notin \mathcal{M}(A)$. One the other hand, if the only alternatives of interest are $0 \neq \mu \in \mathcal{M}(A)$, the choice of g-inverse is irrelevant since $\mu' A^- \mu$ does not depend on the choice of A^- if $\mu \in \mathcal{M}(A)$.

In summary, we have seen that when an asymptotic test of $H_0 : \mu = 0$ is based on the quadratic form $Q_N = X_N' \hat{A}_N^- X_N$ in the asymptotically $\mathcal{N}_p(\mu, A)$ random variable X_N, there are many choices of $\hat{A}_N^-$ which yield an asymptotically chi-square distribution for the test statistic Q_N; one of these amounts to working with a largest nonsingular minor and the corresponding variates. The procedure for detecting deviations $\mu \notin \mathcal{M}(A)$ which works for the usual exact and asymptotic normal theory test statistics is not generally applicable and one must often rely solely on Q_N to detect arbitrary deviations from H_0. But even for fixed $\mu \notin \mathcal{M}(A)$ there is no way to choose $\hat{A}_N^-$ to achieve maximum sensitivity at μ; fortunately, the Moore-Penrose inverse does have an optimal property for detecting $0 \neq \mu \notin \mathcal{M}(A)$. Since the choice of g-inverse has no effect on the sensitivity of Q_N to deviations $0 \neq \mu \in \mathcal{M}(A)$, the Moore-Penrose inverse should be used unless one wishes to increase the sensitivity of Q_N to a particular $\mu \notin \mathcal{M}(A)$ with a corresponding (but greater unless $\mu \in C(A)$) loss of sensitivity at some other $\mu_0 \notin \mathcal{M}(A)$.

Appendix

Proposition : If $\hat{A}_N \xrightarrow{P} A$ and $\hat{A}_N$ is unbiased for A, then $\mathcal{M}(\hat{A}_N) = \mathcal{M}(A)$ for all N sufficiently large.

Proof : Since $A_N \xrightarrow{P} \hat{A}$, by Lemma 3 we obtain

$$R(\hat{A}_N) \geqslant R(A) \text{ with probability one.} \quad \ldots \text{ (A1)}$$

for all N sufficiently large. Consider a vector λ orthogonal to the columns of A, i.e., $\lambda' A = 0$. Then $\lambda' A_N \lambda$ is nonnegative random variable with $E(\lambda' \hat{A}_N \lambda) = 0$. Hence $\lambda' \hat{A}_N \lambda = 0$ with probability one or equivalently $\lambda' \hat{A}_N = 0$. Thus we get

$$\mathcal{M}(\hat{A}_N) \subset \mathcal{M}(A) \quad \ldots \text{ (A2)}$$

for all N. In particular, (A2) $\Longrightarrow R(\hat{A}_N) \leqslant R(A)$ for all N. This, together with (A1) gives $R(\hat{A}_N) = R(A)$ for all N sufficiently large. Hence (A2) gives $\mathcal{M}(\hat{A}_N) = \mathcal{M}(A)$ for all N sufficiently large.

Acknowledgement. The authors would like to express their sincere thanks to Professor S. K. Mitra for very valuable suggestions which led to a considerable improvement in the original draft of the paper.

References

Rao, C. R. and Mitra, S. K. (1971) : *Generalized Inverse of Matrices and its Applications*, John Wily, New York.

Russell, C. T. (1973a) : On the use of g-inverses in asymptotic hypothesis testing. (Abstract). *Notices Am. Math. Soc.*, **20**, A-206.

——— (1973b) : Rank tests for multivariate paired comparisons involving random variables which are not identically distributed. Ph. D. Thesis, Indiana University.

Stewart, G. W. (1969) : On the continuity of the generalized inverse. *SIAM J. Appl. Math.* **17**, 33–45.

Paper received : June, 1983.