NOTES ON MATRICES OF FULL COLUMN (ROW) RANK

Shayle R. Searle

Biometrics Unit, Cornell University, Ithaca, N. Y. 14853

BU-1361-M August 1996

ABSTRACT

A useful left (right) inverse of a full column (row) matrix is the Moore-Penrose inverse; and linear equations based on a full column (row) rank matrix have only one (many a) solution. These solutions are characterized.

Key words: left inverse, right inverse, Moore-Penrose inverse, solution to linear equations.

LEFT INVERSES

$\mathbf{A}_{r\times c}$ can have a left inverse $\mathbf{L}$, such that $\mathbf{LA} = \mathbf{I}$ only if $r \geq c$, and it does have left inverses only if it has full column rank, $r_A = c$. Then for given $\mathbf{A}$ there can be many values of $\mathbf{L}$, one of which is $\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$, the Moore-Penrose inverse of $\mathbf{A}$.

Example

For

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 4 & 1 \\ 2 & -2 \end{bmatrix}, \quad (\mathbf{A}'\mathbf{A})^{-1} = \begin{bmatrix} 24 & 6 \\ 6 & 14 \end{bmatrix}^{-1} = \frac{1}{300}\begin{bmatrix} 14 & -6 \\ -6 & 24 \end{bmatrix}$$

and

$$\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' = \frac{1}{300}\begin{bmatrix} 14 & -6 \\ -6 & 24 \end{bmatrix}\begin{bmatrix} 2 & 4 & 2 \\ 3 & 1 & -2 \end{bmatrix} = \frac{1}{30}\begin{bmatrix} 1 & 5 & 4 \\ 6 & 0 & -6 \end{bmatrix}.$$

The conditions for $\mathbf{A}^+$ being the Moore-Penrose of $\mathbf{A}$ are easily verified: first, that $\mathbf{A}^+\mathbf{A}$ is symmetric, which it is because $\mathbf{A}^+\mathbf{A} = \mathbf{I}$, then $\mathbf{AA}^+\mathbf{A} = \mathbf{A}$, $\mathbf{A}^+\mathbf{AA}^+ = \mathbf{A}^+$, and then that

$$\mathbf{AA}^+ = \begin{bmatrix} 2 & 3 \\ 4 & 1 \\ 2 & -2 \end{bmatrix}\frac{1}{30}\begin{bmatrix} 1 & 5 & 4 \\ 6 & 0 & -6 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix},$$

which is symmetric.

RIGHT INVERSES

The existence of right inverses is very much (but not entirely) the converse of the situation for left inverses. $\mathbf{A}_{r\times c}$ can have right inverses $\mathbf{R}$, with $\mathbf{AR} = \mathbf{I}$ only if $c \geq r$ and it does have them only if $\mathbf{A}$ has full row rank, $r_A = r$. One such right inverse is $\mathbf{A}^+ = \mathbf{A}'(\mathbf{AA}')^{-1}$, the Moore-Penrose inverse of $\mathbf{A}$ of full row rank.

Moore-Penrose inverse

An interesting point here is that the expressions for the Moore-Penrose inverse of $\mathbf{A}$ are not the same in the preceding two cases:

$$\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' \quad \text{for} \quad \mathbf{A} \text{ of full column rank} \tag{1}$$

and

$$\mathbf{A}^{++} = \mathbf{A}'(\mathbf{AA}')^{-1} \quad \text{for} \quad \mathbf{A} \text{ of full row rank .} \tag{2}$$

Result (2) is derived from (1) by noting that for $\mathbf{A}$ of full row rank $\mathbf{A}'$ has full column rank. Therefore (1) gives $(\mathbf{A}')^+$ as $(\mathbf{AA}')^{-1}\mathbf{A}$. But the transpose of $(\mathbf{A}')^+$ is $\mathbf{A}^+$, so giving (2). Note, too, that the two inverses $(\mathbf{A}'\mathbf{A})^{-1}$ and $(\mathbf{AA}')^{-1}$ are not the same, not even of the same order and, moreover, when one of them exists the other does not. Thus $(\mathbf{A}'\mathbf{A})^{-1}$ exists when $\mathbf{A}$ has full column but $(\mathbf{AA}')^{-1}$ does not; and vice versa for $\mathbf{A}$ of full row rank.

A general expression for the Moore-Penrose inverse $\mathbf{A}^\mathsf{M}$ of $\mathbf{A}$ is (Searle, 1982, p. 216)

$$\mathbf{A}^\mathsf{M} = \mathbf{A}'(\mathbf{AA}')^-\mathbf{A}(\mathbf{A}'\mathbf{A})^-\mathbf{A}' ,$$

where $(\mathbf{AA}')^-$ is any generalized inverse of $\mathbf{AA}'$ satisfying just the first of the Penrose conditions, $\mathbf{AA}'(\mathbf{AA}')^-\mathbf{AA}' = \mathbf{AA}'$. It is of interest to see how this reduces to $\mathbf{A}^+$ for $\mathbf{A}$ of full column rank. Write $\mathbf{A}$ as

$$\mathbf{A} = \begin{bmatrix} \mathbf{T} \\ \mathbf{KT} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{K} \end{bmatrix}\mathbf{T}$$

for some non-singular $\mathbf{T}$ and some $\mathbf{K}$. Then

$$\begin{aligned}
\mathbf{A}'(\mathbf{AA}')^-\mathbf{A} &= \mathbf{T}'[\mathbf{I} \quad \mathbf{K}'] \left(\begin{bmatrix} \mathbf{I} \\ \mathbf{K} \end{bmatrix} \mathbf{TT}'[\mathbf{I} \quad \mathbf{K}'] \right)^- \begin{bmatrix} \mathbf{I} \\ \mathbf{K} \end{bmatrix}\mathbf{T} \\
&= \mathbf{T}'[\mathbf{I} \quad \mathbf{K}'] \begin{bmatrix} (\mathbf{TT}')^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \mathbf{K} \end{bmatrix}\mathbf{T} \\
&= \mathbf{T}'(\mathbf{TT}')^{-1}\mathbf{T} = \mathbf{T}'\mathbf{T}'^{-1}\mathbf{T}^{-1}\mathbf{T} \\
&= \mathbf{I} .
\end{aligned}$$

Therefore $\mathbf{A}^\mathsf{M}$ reduces to $\mathbf{A}^+$, as it should:

$$\begin{aligned}
\mathbf{A}^\mathsf{M} &= \mathbf{I}(\mathbf{A}'\mathbf{A})^-\mathbf{A}' \\
&= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' = \mathbf{A}^+ ,
\end{aligned}$$

because $\mathbf{A}$ has full column rank. Similarly $\mathbf{A}^\mathsf{M}$ reduces to $\mathbf{A}^{++}$ when $\mathbf{A}$ has full row rank.

LINEAR EQUATIONS WITH A FULL COLUMN RANK MATRIX

Consider linear equations

$$\mathbf{Ax} = \mathbf{y}$$

for known $\mathbf{A}$, of full column rank c, and known $\mathbf{y}$. The general solution is

$$\tilde{\mathbf{x}} = \mathbf{A}^{-}\mathbf{y} + (\mathbf{I} - \mathbf{A}^{-}\mathbf{A})\mathbf{z}$$

for arbitrary $\mathbf{z}$; and, through the arbitrariness of $\mathbf{z}$, it generates all possible solutions (*loc. cit.*, Chapter 9). But, with

$$\mathbf{A} = \begin{bmatrix} \mathbf{T} \\ \mathbf{KT} \end{bmatrix}, \quad \text{an} \quad \mathbf{A}^{-} \quad \text{is} \quad \mathbf{A}^{-} = [\mathbf{T}^{-1} \quad \mathbf{0}] \, .$$

Therefore $\mathbf{A}^{-}\mathbf{A} = \mathbf{I}$ and so there is only one solution

$$\tilde{\mathbf{x}} = \mathbf{A}^{-}\mathbf{y} \, ,$$

the same for all generalized inverses $\mathbf{A}^{-}$.

One well might ask "What happens if the first c rows of $\mathbf{A}_{r \times c}$ (of full column rank) are not linearly independent?" Then $\mathbf{T}^{-1}$ would not exist. This can be circumvented by using $\mathbf{B} = \mathbf{PA}$ where $\mathbf{P}$ is a permutation matrix, and hence $\mathbf{B}$ is $\mathbf{A}$ with its rows permuted to have the first c rows of $\mathbf{B}$ be linearly independent. Then, as with $\mathbf{A}^{-}\mathbf{A} = \mathbf{I}$ in the preceding paragraph, we now have $\mathbf{B}^{-}\mathbf{B} = \mathbf{I}$. But with $\mathbf{B} = \mathbf{PA}$, $\mathbf{A} = \mathbf{P}'\mathbf{B}$ (because $\mathbf{P}$ is orthogonal) and $\mathbf{A}^{-} = \mathbf{B}^{-}\mathbf{P}$ and so

$$\mathbf{A}^{-}\mathbf{A} = \mathbf{B}^{-}\mathbf{P}\mathbf{P}'\mathbf{B} = \mathbf{B}^{-}\mathbf{B} = \mathbf{I} \, .$$

What is sometimes puzzling about the $\tilde{\mathbf{x}} = \mathbf{A}^{-}\mathbf{y}$ solution is why does $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\mathbf{A}^{-}\mathbf{y}$ equal $\mathbf{y}$? It is not because $\mathbf{A}\mathbf{A}^{-}$ is an identity matrix, since that is not so. One approach is that without knowing $\tilde{\mathbf{x}}$ we do know that $\mathbf{y} = \mathbf{Ax}$. Therefore $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\mathbf{A}^{-}\mathbf{y} = \mathbf{A}\mathbf{A}^{-}\mathbf{Ax} = \mathbf{Ax} = \mathbf{y}$.

Fortunately one does not have to rely on the existence of $\mathbf{T}^{-1}$ for calculating an $\mathbf{A}^{-}$. The full column rank property of $\mathbf{A}$ ensures the existence of $(\mathbf{A}'\mathbf{A})^{-1}$ and so $\mathbf{A}^{-} = \mathbf{A}^{+} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ is the easiest calculation and $\tilde{\mathbf{x}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{y}$ is *the* solution of $\mathbf{Ax} = \mathbf{y}$ for full column rank $\mathbf{A}$.

Example

$$\text{With} \quad \mathbf{A} = \begin{bmatrix} 9 & 2 \\ 4 & 1 \\ 7 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{Ax} = \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}$$

the solution using $\mathbf{T}^{-1}$ is

$$\tilde{\mathbf{x}} = \begin{bmatrix} \begin{pmatrix} 9 & 2 \\ 4 & 1 \end{pmatrix}^{-1} & \begin{matrix} 0 \\ 0 \end{matrix} \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 \\ -4 & 9 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix},$$

and with $\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ the solution is

$$\tilde{\mathbf{x}} = \mathbf{A}^+\mathbf{y} = \begin{bmatrix} 146 & 43 \\ 43 & 14 \end{bmatrix}^{-1} \begin{bmatrix} 9 & 4 & 7 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}$$

$$= \frac{1}{195} \begin{bmatrix} 14 & -43 \\ -43 & 146 \end{bmatrix} \begin{bmatrix} 60 \\ 15 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

as before.

LINEAR EQUATIONS WITH A FULL ROW RANK MATRIX

Consider $\mathbf{Ax} = \mathbf{y}$ with $\mathbf{A}$ of full row rank represented by $\mathbf{A}_{r \times c} = [\mathbf{R} \quad \mathbf{RQ}]$ for $\mathbf{R}$ non-singular of rank r. On partitioning $\mathbf{x}$ conformably with the partitioning of $\mathbf{A}$ we write

$$[\mathbf{R} \quad \mathbf{RQ}] \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \mathbf{y}$$

and get

$$\mathbf{Rx}_1 + \mathbf{RQx}_2 = \mathbf{y}$$

$$\tilde{\mathbf{x}}_1 = \mathbf{R}^{-1}\mathbf{y} - \mathbf{Qx}_2 \ .$$

Thus for any $\mathbf{x}_2$, the solution $\mathbf{x} = \begin{bmatrix} \tilde{\mathbf{x}}_1 \\ \mathbf{x}_2 \end{bmatrix}$ will satisfy $\mathbf{Ax} = \mathbf{y}$. Since $\mathbf{Q} = \mathbf{R}^{-1}\mathbf{RQ}$, and $\mathbf{RQ}$ is the notation for the columns of $\mathbf{A}$ beyond those of the r columns of non-singular $\mathbf{R}$, it is more useful to write

$$\mathbf{A} = [\mathbf{R} \quad \mathbf{RQ}] \quad \text{and} \quad \tilde{\mathbf{x}}_1 = \mathbf{R}^{-1}(\mathbf{y} - \mathbf{RQx}_2) \ .$$

Example

$$\mathbf{A} = \begin{bmatrix} 9 & 4 & 7 \\ 2 & 1 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{Ax} = \begin{bmatrix} 27 \\ 9 \end{bmatrix},$$

$$\tilde{\mathbf{x}}_1 = \begin{bmatrix} 9 & 4 \\ 2 & 1 \end{bmatrix}^{-1} \left[\begin{pmatrix} 27 \\ 9 \end{pmatrix} - \begin{pmatrix} 7 \\ 3 \end{pmatrix} x_2 \right]$$

$$= \begin{bmatrix} 1 & -4 \\ -2 & 9 \end{bmatrix} \begin{bmatrix} 27 - 7x_2 \\ 9 - 3x_2 \end{bmatrix} = \begin{bmatrix} -9 + 5x_2 \\ 27 - 13x_2 \end{bmatrix}.$$

Thus the solution is

$$\tilde{\mathbf{x}} = \begin{bmatrix} -9 + 5x_2 \\ 27 - 13x_2 \\ x_2 \end{bmatrix} \quad \text{with examples} \quad \begin{bmatrix} -9 \\ 27 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -4 \\ 14 \\ 1 \end{bmatrix}.$$

If $\mathbf{A}$ is such that its first r columns are not linearly independent, then resequence the columns to have $\mathbf{C} = \mathbf{AP}$ for $\mathbf{P}$ a permutation matrix, and $\mathbf{C}$ having its first r columns linearly independent. Then because $\mathbf{Ax} = \mathbf{y}$ is $\mathbf{APP'x} = \mathbf{y}$ which, with $\mathbf{z} = \mathbf{P'x}$, is $\mathbf{Cz} = \mathbf{y}$, obtain $\tilde{\mathbf{z}}$ as a solution to $\mathbf{Cz} = \mathbf{y}$ and from that get $\tilde{\mathbf{x}} = \mathbf{P}\tilde{\mathbf{z}}$ as the solution to $\mathbf{Ax} = \mathbf{y}$.

Example

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & 2 & 7 \end{bmatrix}\mathbf{x} = \begin{bmatrix} 8 \\ 18 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 2 & 7 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \mathbf{AP}$$

$$\mathbf{z}_1 = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}^{-1}\left[\begin{pmatrix} 8 \\ 18 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \end{pmatrix} z_2\right]$$

$$= \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 8 - z_2 \\ 18 - 2z_2 \end{bmatrix} = \begin{bmatrix} 2 - z_2 \\ 2 \end{bmatrix}$$

$$\tilde{\mathbf{z}} = \begin{bmatrix} 2 - z_2 \\ 2 \\ z_2 \end{bmatrix}$$

and

$$\tilde{\mathbf{x}} = \mathbf{P}\tilde{\mathbf{z}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 2 - z_2 \\ 2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 2 - z_2 \\ z_2 \\ 2 \end{bmatrix}$$

with examples $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$.

REFERENCE

Searle, Shayle R. (1982) *Matrix Algebra Useful for Statistics*, Wiley, New York.