ANIMAL BREEDING NOTES

CHAPTER 2

LINEAR DEPENDENCE, MATRIX INVERSES, AND CONSISTENCY OF LINEAR EQUATIONS

Linear dependence

Let $\{y_1, y_2, ..., y_n\}$ be a set of n m × 1 vectors. This set of n vectors is **linearly dependent** if there is a set of scalars $\{c_1, ..., c_n\}$ not all zero, such that

$$c_1y_1 + c_2y_2 + ... + c_ny_n = 0.$$

Contrarily, if the only set of scalars for which the above sum is $\{0,0, ...,0\}$, the set of vectors $\{y_1, ..., y_n\}$ is **linearly independent**.

Remarks:

1) Any set of vectors containing the zero vector is linearly dependent.
2) Any subset of a linearly independent set of vectors is linearly independent.
3) If a set contains more than m m × 1 vectors, it is linearly dependent.

Examples:

1) $y_1 = \begin{bmatrix} 0 \\ -2 \\ 6 \end{bmatrix}, \quad y_2 = \begin{bmatrix} 5 \\ 10 \\ 8 \end{bmatrix}, \quad y_3 = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$

$$c_1\begin{bmatrix} 0 \\ -2 \\ 6 \end{bmatrix} + c_2\begin{bmatrix} 5 \\ 10 \\ 8 \end{bmatrix} + c_3\begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5c_2 \\ -2c_1 + 10c_2 + 3c_3 \\ 6c_1 + 8c_2 + c_3 \end{bmatrix}$$

The set $\{y_1, y_2, y_3\}$ is linearly independent because $\sum_{i=1}^{3} c_1 y_i = 0$ only if $c_1 = c_2 = c_3 = 0$.

2) $x_1 = \begin{bmatrix} 0 \\ -2 \\ 6 \end{bmatrix}$, $x_2 = \begin{bmatrix} 5 \\ 10 \\ 8 \end{bmatrix}$, $x_3 = \begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix}$

The set $\{x_1, x_2, x_3\}$ is linearly dependent because $x_1 + 2x_3 = 0$.

Rank of a matrix: The columns of a matrix $A_{m\times n}$ can be considered as a set of vectors (i.e., column vectors). Similarly, the rows of a matrix $A_{m\times n}$ constitute a set of row vectors. The rank of a matrix $A_{m\times n}$ is the number of linearly independent column vectors (column rank) or row vectors (row rank). The row rank and the column rank of a matrix are equal.

Remarks:

1) The rank of $A_{m\times n}$, $m \neq n \leq \min(m, n)$.

   [**Note:** rank (A) = rank (A')]

2) The rank of $AB \leq \min$ (rank A, rank B).

3) The rank of a square matrix is equal to or less than its order.

4) The rank of $(A \oplus B)$ = rank of A + rank of B.

5) The following statements are equivalent for a nonsingular (square) matrix A_n:

   a) $Ax = 0 \Rightarrow x = 0$, and

   b) $|A| \neq 0$.

6) For D = diagonal matrix, rank (D) = number of nonzero elements. In particular, rank $(I_n) = n$.

Examples:

1) The matrix $A_3 = \begin{bmatrix} 0 & 5 & 0 \\ -2 & 10 & 3 \\ 6 & 8 & 1 \end{bmatrix}$ has rank = 3 (i.e., it is nonsingular) because:

   a) $Ax = 0 \Rightarrow x = [0\ \ 0\ \ 0]'$

   b) $|A| = 5\begin{vmatrix} -2 & 3 \\ 6 & 1 \end{vmatrix}(-1)^3 = 100$

2) The matrix $B_3 = \begin{bmatrix} 0 & 5 & 0 \\ -2 & 10 & 1 \\ 6 & 8 & -3 \end{bmatrix}$ has rank = 2 (i.e., it is singular) because:

   a) $Bx = 0$ for $x = [1\ \ 0\ \ 2]'$

   b) $|B| = 5\begin{vmatrix} -2 & 1 \\ 6 & -3 \end{vmatrix}(-1)^3 = -5(0) = 0$

   c) If the first or the third columns are ignored the remaining columns are linearly independent.

Inverse of a matrix: The matrix B such that AB = BA = I is called the inverse of A and it is denoted by A^{-1}. The inverse is defined only for square matrices.

Remarks:

1) The matrix A has an inverse if it is nonsingular, i.e.,

$$A^{-1} \text{ exists } \Rightarrow \begin{cases} Ax = 0 \Rightarrow x = 0 \\ |A| \neq 0 \end{cases}$$

2) A^{-1} is unique.

3) $(A^{-1})^{-1} = A$

4) $(A')^{-1} = (A^{-1})'$

5) If A is symmetric (i.e., $A' = A$), then A^{-1} is also symmetric (i.e., $(A^{-1})' = A^{-1}$).

6) If A and B are nonsingular, then $(AB)^{-1} = B^{-1}A^{-1}$

7) If $A^{-1} = A'$ then $AA' = I \Rightarrow A \equiv$ orthogonal matrix.

8) $(A \oplus B)^{-1} = A^{-1} \oplus B^{-1}$

9) $(A * B)^{-1} = A^{-1} * B^{-1}$

10) $D = \text{diag}\ \{d_{ii}\} \ \Rightarrow\ D^{-1} = \left\{\dfrac{1}{d_{ii}}\right\}$.

Computation of the inverse of a matrix

$$A^{-1} = |A|^{-1}\ \text{adj}\ (A)$$

where

$|A|$ = Determinant of the matrix A

adj (A) = Transposed matrix of cofactors of the elements of A

= **adjugate** or **adjoint** of A

Example:

a) $A_{2\times 2} = \begin{bmatrix} 6 & 2 \\ 4 & 5 \end{bmatrix} \Rightarrow A^{-1} = \dfrac{1}{22}\begin{bmatrix} 5 & -2 \\ -4 & 6 \end{bmatrix}$

b) $A_{3\times 3} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 6 \end{bmatrix}$

$$|A| = 1\begin{vmatrix} 2 & 1 \\ 4 & 0 \end{vmatrix}(-1)^{3+1} + 6\begin{vmatrix} 3 & 2 \\ 2 & 4 \end{vmatrix}(-1)^{3+3} = -4 + 48 = 44$$

$$\Rightarrow \quad A^{-1} = \frac{1}{44}\begin{bmatrix} \begin{vmatrix} 4 & 0 \\ 0 & 6 \end{vmatrix} & -\begin{vmatrix} 2 & 0 \\ 1 & 6 \end{vmatrix} & \begin{vmatrix} 2 & 4 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} 2 & 1 \\ 0 & 6 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 1 & 6 \end{vmatrix} & -\begin{vmatrix} 3 & 2 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} 2 & 1 \\ 4 & 0 \end{vmatrix} & -\begin{vmatrix} 3 & 1 \\ 2 & 0 \end{vmatrix} & \begin{vmatrix} 3 & 2 \\ 2 & 4 \end{vmatrix} \end{bmatrix}$$

$$\Rightarrow \quad A^{-1} = \frac{1}{44}\begin{bmatrix} 24 & -12 & -4 \\ -12 & 17 & 2 \\ -4 & 2 & 8 \end{bmatrix}$$

$$\Rightarrow AA^{-1} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 6 \end{bmatrix} \frac{1}{44}\begin{bmatrix} 24 & -12 & -4 \\ -12 & 17 & 2 \\ -4 & 2 & 8 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Inversion of a matrix by partitioning

If a nonsingular matrix A is too large to be directly inverted in a computer, its inverse (i.e., A^{-1}) could be obtained by partitioning A into four submatrices, i.e.,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{where } A_{11} \text{ and } A_{22} \text{ are square matrices.}$$

Let

$$A^{-1} = \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix}.$$

Note that:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

$$\Rightarrow \quad A_{11}A^{11} + A_{12}A^{21} = I \qquad [1]$$

$$A_{11}A^{12} + A_{12}A^{22} = 0 \qquad [2]$$

$$A_{21}A^{11} + A_{22}A^{21} = 0 \qquad [3]$$

$$A_{21}A^{12} + A_{22}A^{22} = I \qquad [4]$$

Assuming that A_{11} and A_{22} are nonsingular, using [1], [2], [3], and [4], we get:

from [2]: $A^{12} = -A_{11}^{-1} A_{12}A^{22}$ [5]

from [3]: $A^{21} = -A_{22}^{-1} A_{21}A^{11}$ [6]

Substituting [6] in [1] and [5] in [4] we get:

$$A_{11}A^{11} + A_{12}(-A_{22}^{-1}A_{21}A^{11}) = I$$

$$(A_{11} - A_{12}A_{22}^{-1}A_{21})A^{11} = I$$

$$\Rightarrow \quad A^{11} = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} \qquad [7]$$

and

$$A_{21}(-A_{11}^{-1}A_{12}A^{22}) + A_{22}A^{22} = I$$

$$\Rightarrow \quad A^{22} = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \qquad [8]$$

This approach requires inverting **four** matrices: A_{11}, A_{22}, $(A_{11} - A_{12}AA_{21})$ and $(A_{22} - A_{21}AA_{12})$.

A procedure that requires inverting only **two** matrices is as follows. Recall that:

$$\begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

Thus,

$$\Rightarrow \quad A^{11}A_{11} + A^{12}A_{21} = I \qquad [9]$$

$$A^{11}A_{12} + A^{12}A_{22} = 0 \quad [10]$$

$$A^{21}A_{11} + A^{22}A_{21} = 0 \quad [11]$$

$$A^{21}A_{12} + A^{22}A_{22} = I \quad [12]$$

From [10]:

$$A^{12} = -A^{11}A_{12}\,A_{22}^{-1} \quad [13]$$

From [12]:

$$A^{22} = (I - A^{21}A_{12})\,A_{22}^{-1}$$

$$A^{22} = A_{22}^{-1} - A^{21}A_{12}\,A_{22}^{-1} \quad [14]$$

$$\Rightarrow \quad \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix} = \begin{bmatrix} \left(A_{11} - A_{12}A_{22}^{-1}A_{21}\right)^{-1} & -A^{11}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}A^{11} & A_{22}^{-1} - A^{21}A_{12}A_{22}^{-1} \end{bmatrix} \quad [15]$$

Remark: Matrix [15] requires the existence of A_{22}^{-1}.

Similarly, from [11] and [9], we get:

$$A^{21} = -A^{22}A_{21}A_{11}^{-1} \quad [16]$$

and

$$A^{11} = (I - A^{12}A_{21})A_{11}^{-1}$$

$$A^{11} = A_{11}^{-1} - A^{12}A_{21}A_{11}^{-1} \quad [17]$$

$$\Rightarrow \quad \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix} = \begin{bmatrix} A_{11}^{-1} - A^{12}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}A^{22} \\ -A^{22}A_{21}A_{11}^{-1} & \left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right)^{-1} \end{bmatrix} \quad [18]$$

Remark: Matrix A_{11}^{-1} must exist if A^{-1} is to be computed using [18].

Expressions for obtaining the inverse of a symmetric matrix by partitioning are similar to [15] and [18], with A_{12}' and $A^{12}{}'$ substituted for A_{21} and A^{21}.

Example:

$$A_{3\times3} = \left[\begin{array}{c|cc} 3 & 2 & 1 \\ \hline 2 & 4 & 0 \\ 1 & 0 & 6 \end{array}\right]$$

Let

$$A_{11} = \begin{bmatrix} 3 \end{bmatrix},\ A_{12} = \begin{bmatrix} 2 & 1 \end{bmatrix},\ A_{21} = A_{12}' = \begin{bmatrix} 2 \\ 1 \end{bmatrix},\ \text{and}\ A_{22} = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}.$$

By matrix formulae [15],

$$A^{11} = \left[\begin{bmatrix} 3 \end{bmatrix} - \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right]^{-1}$$

$$A^{11} = \left[\begin{bmatrix} 3 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right]^{-1} = \left[3 - \frac{7}{6} \right]^{-1} = \left[\frac{6}{11} \right]$$

$$A^{12} = -\left[\frac{6}{11} \right] \begin{bmatrix} \frac{1}{2} & \frac{1}{6} \end{bmatrix} = -\begin{bmatrix} \frac{3}{11} & \frac{1}{11} \end{bmatrix} = A^{12\prime}$$

$$A^{22} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{6} \end{bmatrix} - \begin{bmatrix} -\frac{3}{11} \\ -\frac{1}{11} \end{bmatrix} \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{6} \end{bmatrix}$$

$$A^{22} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{6} \end{bmatrix} + \begin{bmatrix} \frac{6}{11} & \frac{3}{11} \\ \frac{2}{11} & \frac{1}{11} \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{6} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{6} \end{bmatrix} + \begin{bmatrix} \frac{3}{22} & \frac{1}{22} \\ \frac{1}{22} & \frac{1}{66} \end{bmatrix}$$

$$A^{22} = \begin{bmatrix} \frac{17}{44} & \frac{1}{22} \\ \frac{1}{22} & \frac{2}{11} \end{bmatrix}$$

$$\Rightarrow A^{-1} = \left[\begin{array}{c|cc} \frac{6}{11} & -\frac{3}{11} & -\frac{1}{11} \\ \hline -\frac{3}{11} & \frac{17}{44} & \frac{1}{22} \\ -\frac{1}{11} & \frac{1}{22} & \frac{2}{11} \end{array}\right] = \frac{1}{44}\left[\begin{array}{c|cc} 24 & -12 & -4 \\ \hline -12 & 17 & 2 \\ -4 & 2 & 8 \end{array}\right]$$

Elementary operators

Elementary operators are square matrices derived from the identity matrix. The rank of the matrix resulting from multiplying an elementary operator by a matrix $A_{m\times n}$ is the same as the rank of A.

The elementary operators are:

a) E_{ij} is I with rows i and j interchanged, e.g.,

$$E_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

b) $R_{ii}(\lambda)$ is I with λ substituted for 1 in the i^{th} diagonal element, e.g.,

$$R_{11}(2) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

c) $P_{ij}(\lambda)$ is I with λ replacing for 0 in the ij^{th} location for $i \neq j$, e.g.,

$$P_{12}(2) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Effect of elementary operators on $A_{m\times n}$

Pre-multiplication of A by an elementary operator affects the **rows** of A. Post-multiplication of A by an elementary operator affects the **columns** of A.

Pre-multiplication of A by:

a) E_{ij} interchanges the i^{th} and j^{th} rows, e.g.,

$$\underset{E_{13}}{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}} \underset{A}{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}} = \underset{E_{13}A}{\begin{bmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix}}$$

b) $R_{ii}(\lambda)$ multiples the i^{th} row by λ, e.g.,

$$\underset{R_{11}(2)}{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \underset{A}{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}} = \underset{R_{11}(2)A}{\begin{bmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}}$$

c) $P_{ij}(\lambda)$ adds λ times the j^{th} row to the i^{th} row of A, e.g.,

$$\underset{P_{12}(2)}{\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \underset{A}{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}} = \underset{P_{12}(2)A}{\begin{bmatrix} 9 & 12 & 15 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}}$$

Determinants, transposes and inverses of elementary operators

Determinant		Transpose	Inverse
$\lvert E_{ij} \rvert = -1$	$\lvert E_{ij}A \rvert = -\lvert A \rvert$	$E_{ij}' = E_{ij}$	$E_{ij}^{-1} = E_{ij}$
$\lvert R_{ii}(\lambda) \rvert = \lambda$	$\lvert R_{ii}(\lambda)A \rvert = \lambda \lvert A \rvert$	$R_{ii}'(\lambda) = R_{ii}(\lambda)$	$[R_{ii}(\lambda)]^{-1} = R_{ii}\left(\frac{1}{\lambda}\right)$
$\lvert P_{ij}(\lambda) \rvert = 1$	$\lvert P_{ij}(\lambda)A \rvert = \lvert A \rvert$	$[P_{ij}(\lambda)] = P_{ji}(\lambda)$	$[P_{ij}(\lambda)]^{-1} = P_{ij}(-\lambda)$

Reduction of a matrix to its equivalent diagonal form

Equivalence: two matrices are equivalent if one can be derived from the other by multiplying it by a series of elementary operators, i.e., a matrix B_{mn} is equivalent to a matrix A_{mn} if

$$P_u \ldots P_2 P_1 A Q_1 Q_2 \ldots Q_v = B$$

where the P_i, i=1, ..., u and Q_j, j = 1, ..., v are elementary operators. Let

$$P = P_u \ldots P_2 P_1 \text{ and } Q = Q_1 Q_2 \ldots Q_v$$

where

P and $\{P_i\}$ are m×n and Q and $\{Q_j\}$ are n×m nonsingular matrices. Thus,

$$A = P^{-1} B Q^{-1}$$

where

$$P^{-1} = P_1^{-1} \ldots P_{u-1}^{-1} P_u^{-1}$$

$$Q^{-1} = Q_v^{-1} \ldots Q_2^{-1} Q_1^{-1}$$

which are also elementary operators. Thus, B is equivalent to A and A is equivalent to B. In addition, because multiplication of matrices by elementary operators does not change their rank,

rank (A) = rank (B).

Equivalent diagonal form

$$D = \begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix}$$

Note: the rank of a matrix A_{mn} can be obtained by obtaining an equivalent matrix B whose subdiagonal elements are zero.

Examples:

1) Reduction of A_{33} to D

$$A = \begin{bmatrix} 8 & 4 & 2 \\ 4 & 6 & 1 \\ 2 & 1 & 4 \end{bmatrix}$$

$$P_2P_1 = \begin{vmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ -\frac{1}{4} & 0 & 1 \end{vmatrix} \Rightarrow P_2P_1A = \begin{bmatrix} 8 & 4 & 2 \\ 0 & 4 & 0 \\ 0 & 0 & \frac{7}{2} \end{bmatrix} \Rightarrow \text{rank}(A) = 3$$

2) Reduction of A_{33} to D

$$A = \begin{bmatrix} 6 & 4 & 2 \\ 4 & 4 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

$$P_2P_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{2}{3} & 1 & 0 \\ -\frac{1}{3} & 0 & 1 \end{bmatrix} \Rightarrow P_2P_1A = \begin{bmatrix} 6 & 4 & 2 \\ 0 & \frac{4}{3} & -\frac{4}{3} \\ 0 & -\frac{4}{3} & \frac{4}{3} \end{bmatrix}$$

$$\underset{P_3P_2P_1A}{} = \underset{P_3}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}} \underset{P_2P_1A}{\begin{bmatrix} 6 & 4 & 2 \\ 0 & \frac{4}{3} & -\frac{4}{3} \\ 0 & -\frac{4}{3} & \frac{4}{3} \end{bmatrix}} = \underset{PA}{\begin{bmatrix} 6 & 4 & 2 \\ 0 & \frac{4}{3} & -\frac{4}{3} \\ 0 & 0 & 0 \end{bmatrix}}$$

Because A is symmetric $Q_i = P_i'$; thus, $Q_1Q_2Q_3 = P_1'P_2'P_3'$,

$$\Rightarrow P_3P_2P_1AP_1'P_2'P_3' = \underset{PA}{\begin{bmatrix} 6 & 4 & 2 \\ 0 & \frac{4}{3} & \frac{4}{3} \\ 0 & 0 & 0 \end{bmatrix}} \underset{P'}{\begin{bmatrix} 1 & -\frac{2}{3} & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}} = \underset{D}{\begin{bmatrix} 6 & 0 & 0 \\ 0 & \frac{4}{3} & 0 \\ 0 & 0 & 0 \end{bmatrix}}$$

In this example, the rank(A) = 2.

Generalized inverse of a matrix (Searle, 1966)

A matrix G is said to be a generalized inverse of matrix A if AGA = A. Matrix G is not unique.

There is an infinite number of matrices G that satisfy the condition AGA = A.

Computing G (Searle, 1971)

a) Consider D = PAQ, where P and Q are products of elementary operators. The matrix D is

$$D = \begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix}$$

Define

$$D^- = \begin{bmatrix} D_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

Then, a g-inverse of matrix A is $G = QD^-P$, because $AGA = A$.

Proof:

Note that

$$DD^-D = D \qquad \text{and} \qquad D^-DD^- = D^-$$

Thus,

$$\begin{aligned} AGA &= (P^{-1}DQ^{-1})(QD^-P)(P^{-1}DQ^{-1}) \\ &= P^{-1}DID^-IDQ^{-1} \\ &= P^{-1}DQ^{-1} \\ &= A \end{aligned}$$

Also,

$$\begin{aligned} GAG &= (QD^-P)(P^{-1}DQ^{-1})(QD^-P) \\ &= QD^-IDID^-P \\ &= QD^-P \\ &= G \end{aligned}$$

b) Consider a matrix A of rank = r. If A can be partitioned in such a way that its leading principal minor is nonsingular, i.e.,

$$A_{m\times n} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} is r×r and r (A_{11}) = r. Then, a g-inverse of A is:

$$G = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

because AGA = A.

Proof:

$$\begin{aligned} AGA &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & 0 \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \end{aligned}$$

Note that A can be factored as follows: first, partition $A = \begin{bmatrix} F \\ KF \end{bmatrix}$ where $F_{r\times n}$ are r independent rows and KF are m−r linear combinations of the first r rows. If A has r independent rows, it also has r independent columns, where r ≤ m ≤ n. Assuming the set of r independent columns are the first ones, A can be further partitioned as follows: $A = \begin{bmatrix} A_{11} & B \\ KA_{11} & KB \end{bmatrix}$ where A_{11} is an r×r matrix and B is an r×(n−r) matrix, where the columns of $\begin{bmatrix} B \\ KB \end{bmatrix}$ are linear combinations of the r independent columns $\begin{bmatrix} A_{11} \\ KA_{11} \end{bmatrix}$, i.e.,

$$\begin{bmatrix} B \\ KB \end{bmatrix} = \begin{bmatrix} A_{11} \\ KA_{11}L \end{bmatrix}.$$

$$\Rightarrow A = \begin{bmatrix} A_{11} & A_{11}L \\ KA_{11} & KA_{11}L \end{bmatrix}$$

$$\Rightarrow AGA = \begin{bmatrix} A_{11} & A_{11}L \\ KA_{11} & KA_{11}L \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{11}L \\ KA_{11} & KA_{11}L \end{bmatrix}$$

$$\Rightarrow AGA = \begin{bmatrix} I & 0 \\ K & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{11}L \\ KA_{11} & KA_{11}L \end{bmatrix}$$

$$\Rightarrow AGA = \begin{bmatrix} A_{11} & A_{11}L \\ KA_{11} & KA_{11}L \end{bmatrix} = A \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

or, by substituting $A_{11}L$ for A_{12} in $A_{21}A_{11}^{-1}A_{12}$ we get

$$\begin{aligned} A_{21}A_{11}^{-1}A_{12} &= KA_{11}A_{11}^{-1}A_{11}L \\ &= KA_{11}L \\ &\equiv A_{22} \end{aligned}$$

Also, $GAG = G$. Thus,

$$GAG = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{11}L \\ KA_{11} & KA_{11}L \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

$$GAG = \begin{bmatrix} I & L \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

$$GAG = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} = G$$

or,

$$GAG = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

$$GAG = \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

$$GAG = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

c) **Symmetric matrices**. A computational algorithm for a symmetric matrix A_{mn} of rank = r is:

1) Find any nonsingular principal minor of order r, e.g., A_{ii}.
2) Invert A_{ii}, i.e., obtain A_{11}^{-1}.
3) Replace each element of A_{ii} in A by the corresponding element of A_{11}^{-1}.
4) Substitute zeroes for all the other elements of A.
5) The resulting matrix is a g-inverse of A.

Examples:

a) **Method 1**: $G = QD^{-}P$

$$A_{3x3} = \begin{bmatrix} 6 & 4 & 2 \\ 4 & 4 & 0 \\ 2 & 0 & 2 \end{bmatrix}, \ P = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{2}{3} & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}, \ Q = P'$$

$$QD^{-} = \begin{bmatrix} 1 & -\frac{2}{3} & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{6} & 0 & 0 \\ 0 & \frac{3}{4} & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} & -\frac{1}{2} & 0 \\ 0 & \frac{3}{4} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$QD^{-}P = \begin{bmatrix} \frac{1}{6} & -\frac{1}{2} & 0 \\ 0 & \frac{3}{4} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -\frac{2}{3} & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{3}{4} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, AGA = A and GAG = G.

b)**Method 2**: Principal minor.

Note that if $A_{33} = \begin{vmatrix} 6 & 4 \\ 4 & 4 \end{vmatrix}$ is used, then

$$G = \begin{bmatrix} A_{33}^{-1} & 0 \\ 0 & 0 \end{bmatrix} = QD^{-}P \text{ above. Thus,}$$

$$A^{-} = \begin{bmatrix} \begin{bmatrix} 6 & 4 \\ 4 & 4 \end{bmatrix}^{-1} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{8}\begin{bmatrix} 4 & -4 \\ -4 & 6 \end{bmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 0 \end{bmatrix}$$

$$A^{-} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{3}{4} & 0 \\ 0 & 0 & 0 \end{bmatrix} = QD^{-}P$$

The easiest to obtain G is the one based on A_{11}, where $A_{11} = \begin{vmatrix} 4 & 0 \\ 0 & 2 \end{vmatrix}$. Here,

$$G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$$

$$AG = \begin{bmatrix} 6 & 4 & 2 \\ 4 & 4 & 0 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$AGA = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & 4 & 2 \\ 4 & 4 & 0 \\ 2 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 4 & 2 \\ 4 & 4 & 0 \\ 2 & 0 & 2 \end{bmatrix} = A$$

$$GA = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 6 & 4 & 2 \\ 4 & 4 & 0 \\ 2 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$\text{GAG} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} = \text{G}$$

Thus, AGA = A and GAG = G also.

Generalized inverses for partitioned matrices (Searle, 1971, pp 16-28)

Substitute −'s for −1's in the superscripts of formulas [15] and [18] for partitioned matrices that have a unique inverse. Thus,

from equation [15],

$$\begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix} = \begin{bmatrix} \left(A_{11} - A_{12}A_{22}^{-}A_{21}\right)^{-} & -A^{11}A_{12}A_{22}^{-} \\ -A_{22}^{-}A_{21}A^{11} & A_{22} - A^{21}A_{12}A_{22}^{-} \end{bmatrix}$$

and from equation [18],

$$\begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix} = \begin{bmatrix} A_{11}^{-} - A^{12}A_{21}A_{11}^{-} & -A_{11}^{-}A_{12}A^{22} \\ -A^{22}A_{21}A_{11}^{-} & \left(A_{22} - A_{21}A_{11}^{-}A_{12}\right)^{-} \end{bmatrix}.$$

Linear Equations

Consistency: a set of linear equations Ax = y is consistent if, and only if, the linear relationships that exist among rows of the matrix A also exist among the elements of vector y.

Theorems:

1) The equations Ax = y are consistent if, and only if, the rank of the augmented matrix [A y] is

equal to the rank of A.

Proof:

$\Rightarrow$ If $Ax = y$ is consistent, then $[A\ y] = \begin{bmatrix} A_1 & y_1 \\ LA_1 & Ly_1 \end{bmatrix} \Rightarrow$ the same number of linearly independent rows exists for A and $[A\ y] \Rightarrow \text{rank}(A) = \text{rank}([A\ y])$.

$\Leftarrow$ If $\text{rank}(A) = \text{rank}([A\ y])$, then $\text{rank}\begin{bmatrix} A_1 \\ LA_1 \end{bmatrix} = \text{rank}\begin{bmatrix} A_1 & y_1 \\ LA_1 & Ly_1 \end{bmatrix}$ for some K.

Clearly this is true only if K = L, which implies that $Ax = y$ is a consistent set of equations.

2) If a matrix A has p rows and rank r, $p > r$, and if $D = PAQ$ is an equivalent diagonal form of A, then the equations $Ax = y$ are consistent if, and only if, the last $p-r$ elements of Py are zero.

Proof:

$\Rightarrow$ Given: $Ax = y$ consistent.

$$D = PAQ$$

Show: last $p-r$ elements of $Py = 0$.

$$Ax = y$$

$$PAx = Py$$

But

$$PAx = \begin{bmatrix} A_r x \\ 0 \end{bmatrix} = Py \Rightarrow \text{last } p-r \text{ elements of } Py = 0.$$

$\Leftarrow$ Given: last $p-r$ elements of $Py = 0$.

Show: $Ax = y$ is consistent.

If $Py = 0$ for the last $p-r$ elements, then the set of equations

$$\begin{bmatrix} A_r \\ 0 \end{bmatrix} x = Py$$

is consistent. Because P is a matrix of elementary operators P^{-1} exists, so

$P^{-1}\begin{bmatrix} A_r \\ 0 \end{bmatrix} x = Ax = P^{-1}P\,y = y = Ax = y$ is also a set of consistent equations. This is so

because:

a) $\begin{bmatrix} A_r \\ 0 \end{bmatrix} x = Py$ is a set of consistent equations, and

b) the same set of linear relationships (i.e., P^{-1}) were applied to the rows of $\begin{bmatrix} A_r \\ 0 \end{bmatrix} x$ and to the elements of Py.

3) A solution to the consistent set of equations Ax = y is given by x = Gy if, and only if, AGA = A.

Proof:

$\Rightarrow$ Given x = Gy, show that AGA = A.

$$x = Gy \;\Rightarrow\; Ax = AGy$$

But y = Ax,

$\Rightarrow \quad Ax = AGAx$

$\Rightarrow \quad A = AGA$

$\Leftarrow$ Given AGA = A, show that x = Gy.

$$AGAx = Ax$$

But

$$Ax = y$$

$\Rightarrow \quad AGy \quad = \; y$

or

$A(Gy) \; = \; y \;\Rightarrow\; x = Gy$ is a solution of the system of equations $Ax = y$.

Note: If A is square and full rank x is **the** vector of solutions to $Ax = y$.

4) Let A be a matrix of n columns, z be any $n\times 1$ vector and define $H \equiv GA$. Then, if $A = AGA$, x° is a solution to the consistent set of equations $Ax = y$, where

$$x^\circ \; = \; Gy + (H - I)z$$

Proof:

$$x^\circ \; = \; Gy + (H-I)z$$

$$Ax^\circ \; = \; AGy + A(GA-I)z$$

$$Ax^\circ \; = \; AGy + (AGA-A)z$$

But,

$$AGA \; = \; A,$$

$$\Rightarrow \quad Ax^\circ \; = \; AGy$$

$$\Rightarrow \quad Ax^\circ \; = \; y, \text{ by theorem 3) above.}$$

5) Given a set of consistent equations $Ax = y$ and a matrix G such that $AGA = A$, define $H \equiv GA$. Then, a linear combination of the elements of **any** solution x°, e.g., $k'x^\circ$, is unique if, and only if, $k'H = k'$.

Proof: from theorem 4), $x^\circ = Gy + (H - I)z$. Thus,

$$k'x^\circ = k'Gy + k'(H-I)z$$

which is independent of the arbitrary vector z if $k'H = k'$. If so, the value of $k'x^\circ$ is $k'Gy$ for **any** x°. To see that this statement is true, consider any two solution vectors, i.e., x_i and x_j, thus

$$x_i = G_i y$$

$$\Rightarrow x_i = G_i Ax, \text{ because } y = Ax.$$

Similarly,

$$x_j = G_j Ax$$

But

$$Ax_i = AG_iAx = Ax = y$$

$$Ax_j = AG_jAx = Ax = y$$

$$\Rightarrow \quad m'Ax_i = m'Ax = m'y$$

$$m'Ax_j = m'Ax = m'y$$

for some vector m. Thus, letting $k' = m'A$, we have that

$$k'x_i = k'x_j = k'x = m'y$$

$\Rightarrow$ no matter which matrix G we use to obtain a solution vector x°, a linear combination of the elements of x° satisfying the relationship $k'H = k'$ is unique.

Also, note that

$$k'H = m'AGA = m'A$$

$\Rightarrow$ a necessary and sufficient condition for $k'x^\circ$ to be unique is for k' to be a linear combination of the rows of A.

References

Goldberger, A. S. 1964. Econometric Theory. John Wiley and Sons, Inc., NY.

Searle, S. R. 1966. Matrix Algebra for the Biological Sciences (including Applications in Statistics). John Wiley and Sons, Inc., NY.

Searle, S. R. 1971. Linear Models. John Wiley and Sons, Inc., NY.

Scheffé, H. 1959. The analysis of variance. John Wiley and Sons, Inc., NY.