

Appendix: Matrix Theory

In this appendix, conformable matrices are matrices that are the correct sizes when multiplied together. All matrices in this appendix are real, though many of the results also hold for complex matrices (see Seber 2008). Because of lack of uniformity in the literature on some definitions I give the following definitions.

A symmetric $n \times n$ matrix \mathbf{A} is said to be non-negative definite (n.n.d.) if $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ for all \mathbf{x} , while if $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$ we say that \mathbf{A} is positive definite (p.d.). The matrix \mathbf{A} is said to be positive semidefinite if it is non-negative definite and there exists $\mathbf{x} \neq \mathbf{0}$ such that $\mathbf{x}'\mathbf{A}\mathbf{x} = 0$, that is \mathbf{A} is singular. A matrix \mathbf{A} is said to be negative definite if $-\mathbf{A}$ is positive definite.

A matrix \mathbf{A}^- is called a weak inverse of \mathbf{A} if $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$. (We use the term weak inverse as the term generalized inverse has different meanings in the literature.)

Trace

Theorem A.1 *If \mathbf{A} is $m \times n$ and \mathbf{B} is $n \times m$, then*

$$\text{trace}[\mathbf{AB}] = \text{trace}[\mathbf{BA}] = \text{trace}[\mathbf{B}'\mathbf{A}'] = \text{trace}[\mathbf{A}'\mathbf{B}'].$$

Proof

$$\text{trace}[\mathbf{AB}] = \sum_{i=1}^m \sum_{j=1}^n a_{ij}b_{ji} = \sum_{j=1}^n \sum_{i=1}^m b_{ji}a_{ij} = \sum_{i=1}^m \sum_{j=1}^n b'_{ij}a'_{ji} = \sum_{n=1}^n \sum_{i=1}^m a'_{ji}b'_{ij}.$$

If $m = n$ and either \mathbf{A} or \mathbf{B} is symmetric then $\text{trace}[\mathbf{AB}] = \sum_{i=1}^n a_{ij}b_{ij}$.

Rank

Theorem A.2 *If \mathbf{A} and \mathbf{B} are conformable matrices, then*

$$\text{rank}[\mathbf{AB}] \leq \text{minimum}\{\text{rank } \mathbf{A}, \text{rank } \mathbf{B}\}.$$

Proof The i th row of \mathbf{AB} is $\sum_j a_{ij}\mathbf{b}'_j$, where \mathbf{b}'_j is the j th row of \mathbf{B} . The rows of \mathbf{AB} are therefore linear combinations of the rows of \mathbf{B} so that the number of linearly independent rows of \mathbf{AB} is less than or equal to those of \mathbf{B} ; thus $\text{rank}[\mathbf{AB}] \leq \text{rank}[\mathbf{B}]$. Similarly, the columns of \mathbf{AB} are linear combinations of the columns of \mathbf{A} , so that $\text{rank}[\mathbf{AB}] \leq \text{rank}[\mathbf{A}]$.

Theorem A.3 *Let \mathbf{A} be an $m \times n$ matrix with rank r and nullity s , where the nullity is the dimension of the null space of \mathbf{A} , then*

$$r + s = \text{number of columns of } \mathbf{A}.$$

Proof Let $\alpha_1, \alpha_2, \dots, \alpha_s$ be a basis for $\mathcal{N}[\mathbf{A}]$. Enlarge this set of vectors to give a basis $\alpha_1, \alpha_2, \dots, \alpha_r, \beta_1, \beta_2, \dots, \beta_t$ for \mathbb{R}^n . Every vector in $\mathcal{C}[\mathbf{A}]$ can be expressed in the form

$$\begin{aligned} \mathbf{Ax} &= \mathbf{A} \left(\sum_{i=1}^s a_i \alpha_i + \sum_{j=1}^t b_j \beta_j \right) \\ &= \sum_{j=1}^t b_j \mathbf{A}\beta_j \\ &= \sum_{j=1}^t b_j \gamma_j, \quad \text{say.} \end{aligned}$$

Now suppose that $\sum_{j=1}^t c_j \gamma_j = \mathbf{0}$, then

$$\mathbf{A} \left(\sum_{j=1}^t c_j \beta_j \right) = \sum_{j=1}^t c_j \gamma_j = \mathbf{0}$$

and $\sum c_j \beta_j \in \mathcal{N}[\mathbf{A}]$. This is only possible if the c_j 's are all zero so that the γ_j are linearly independent. Since every vector \mathbf{Ax} in $\mathcal{C}[\mathbf{A}]$ can be expressed in terms of the γ_j 's, the γ_j 's form a basis for $\mathcal{C}[\mathbf{A}]$; thus $t = s$. Since $s + t = n$, our proof is complete.

Theorem A.4 *Let \mathbf{A} be any matrix.*

- (i) *The rank of \mathbf{A} is unchanged when \mathbf{A} is pre- or post-multiplied by a non-singular matrix.*
(ii) $\text{rank}[\mathbf{A}'\mathbf{A}] = \text{rank}[\mathbf{A}]$. *Since $\text{rank}[\mathbf{A}'] = \text{rank}[\mathbf{A}]$ this implies that $\text{rank}[\mathbf{A}\mathbf{A}'] = \text{rank}[\mathbf{A}]$.*

Proof

- (i) If \mathbf{Q} is a conformable non-singular matrix, then by A.2

$$\text{rank}[\mathbf{A}] \leq \text{rank}[\mathbf{A}\mathbf{Q}] \leq \text{rank}[\mathbf{A}\mathbf{Q}\mathbf{Q}^{-1}] = \text{rank}[\mathbf{A}]$$

so that $\text{rank}[\mathbf{A}] = \text{rank}[\mathbf{A}\mathbf{Q}]$ etc.

- (ii) $\mathbf{A}\mathbf{x} = \mathbf{0}$ implies that $\mathbf{A}'\mathbf{A}\mathbf{x} = \mathbf{0}$. Conversely, if $\mathbf{A}'\mathbf{A}\mathbf{x} = \mathbf{0}$ then $\mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x} = 0$, which implies $\mathbf{A}\mathbf{x} = \mathbf{0}$. Hence the null spaces of \mathbf{A} and $\mathbf{A}'\mathbf{A}$ are the same. Since \mathbf{A} and $\mathbf{A}'\mathbf{A}$ have the same number of columns, it follows from A.3 that \mathbf{A} and $\mathbf{A}'\mathbf{A}$ have the same ranks. Similarly, replacing \mathbf{A} by \mathbf{A}' and using $\text{rank}[\mathbf{A}] = \text{rank}[\mathbf{A}']$ we have $\text{rank}[\mathbf{A}'] = \text{rank}[\mathbf{A}\mathbf{A}']$, and the result follows.

Theorem A.5 *If \mathbf{A} is $n \times p$ of rank p and \mathbf{B} is $p \times r$ of rank r , then $\mathbf{A}\mathbf{B}$ has rank r .*

Proof We note that $n \geq p \geq r$. From A.4(ii), $\mathbf{A}'\mathbf{A}$ and $\mathbf{B}'\mathbf{B}$ are nonsingular. Multiplying $\mathbf{A}\mathbf{B}\mathbf{x} = \mathbf{0}$ on the left by $(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ gives us $\mathbf{x} = \mathbf{0}$ so that the columns of $n \times r$ matrix $\mathbf{A}\mathbf{B}$ are linearly independent. Hence $\mathbf{A}\mathbf{B}$ has rank r .

Eigenvalues

Theorem A.6 *For conformable matrices, the nonzero eigenvalues of $\mathbf{A}\mathbf{B}$ are the same as those of $\mathbf{B}\mathbf{A}$.*

Proof Let λ be a nonzero eigenvalue of $\mathbf{A}\mathbf{B}$. Then there exists $\mathbf{u} (\neq \mathbf{0})$ such that $\mathbf{A}\mathbf{B}\mathbf{u} = \lambda\mathbf{u}$, that is $\mathbf{B}\mathbf{A}\mathbf{B}\mathbf{u} = \lambda\mathbf{B}\mathbf{u}$. Hence $\mathbf{B}\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, where $\mathbf{v} = \mathbf{B}\mathbf{u} \neq \mathbf{0}$ (as $\mathbf{A}\mathbf{B}\mathbf{u} \neq \mathbf{0}$), and λ is an eigenvalue of $\mathbf{B}\mathbf{A}$. The argument reverses by interchanging the roles of \mathbf{A} and \mathbf{B} .

Theorem A.7 (Spectral Decomposition Theorem) *Let \mathbf{A} be any $n \times n$ symmetric matrix. Then there exists an orthogonal matrix \mathbf{T} such that $\mathbf{T}'\mathbf{A}\mathbf{T} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, where the λ_i are the eigenvalues of \mathbf{A} . [For further details relating to this theorem see Seber (2008: 16.44).]*

Proof Most matrix books give a proof of this important result.

Theorem A.8 *If \mathbf{A} is an $n \times n$ positive-definite matrix and \mathbf{B} is a symmetric $n \times n$ matrix, then there exists a non-singular matrix \mathbf{V} such that $\mathbf{V}'\mathbf{A}\mathbf{V} = \mathbf{I}_n$ and $\mathbf{V}'\mathbf{B}\mathbf{V} = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_n)$, where the γ_i are the roots of $|\gamma\mathbf{A} - \mathbf{B}| = 0$, (i.e., are the eigenvalues of $\mathbf{A}^{-1}\mathbf{B}$ (or $\mathbf{B}\mathbf{A}^{-1}$ or $\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}$)).*

Proof There exists an orthogonal \mathbf{T} such that $\mathbf{T}'\mathbf{A}\mathbf{T} = \mathbf{\Lambda}$, the diagonal matrix of (positive) eigenvalues of \mathbf{A} . Let $\mathbf{\Lambda}^{1/2}$ be the square root of $\mathbf{\Lambda}$, that is has diagonal elements $\lambda_i^{1/2}$, and let $\mathbf{R} = \mathbf{T}\mathbf{\Lambda}^{-1/2}$. Then $\mathbf{R}'\mathbf{A}\mathbf{R} = \mathbf{\Lambda}^{-1/2}\mathbf{T}'\mathbf{A}\mathbf{T}\mathbf{\Lambda}^{-1/2} = \mathbf{I}_n$. As $\mathbf{C} = \mathbf{R}'\mathbf{B}\mathbf{R}$ is symmetric, there exists an orthogonal matrix \mathbf{S} such that $\mathbf{S}'\mathbf{C}\mathbf{S} = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_n) = \mathbf{\Gamma}$, say, where the diagonal elements of $\mathbf{\Gamma}$ are the eigenvalues of \mathbf{C} . Setting $\mathbf{V} = \mathbf{R}\mathbf{S}$ we have $\mathbf{V}'\mathbf{A}\mathbf{V} = \mathbf{S}'\mathbf{R}'\mathbf{A}\mathbf{R}\mathbf{S} = \mathbf{I}_n$ and $\mathbf{V}'\mathbf{B}\mathbf{V} = \mathbf{S}'\mathbf{C}\mathbf{S} = \mathbf{\Gamma}$, where the γ_i are the roots of

$$0 = |\gamma\mathbf{I}_n - \mathbf{R}'\mathbf{B}\mathbf{R}| = |\gamma\mathbf{R}'\mathbf{A}\mathbf{R} - \mathbf{R}'\mathbf{B}\mathbf{R}| = |\mathbf{R}||\gamma\mathbf{A} - \mathbf{B}||\mathbf{R}'| = |\gamma\mathbf{A} - \mathbf{B}|,$$

that is of $|\gamma\mathbf{I}_n - \mathbf{A}^{-1}\mathbf{B}| = 0$. Using A.9(ii), we then apply A.6 to $\mathbf{A}^{-1/2}\mathbf{A}^{-1/2}\mathbf{B}$, which has the same eigenvalues as $\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}$ to complete the proof.

Non-negative Definite Matrices

Theorem A.9 *Let \mathbf{A} be an $n \times n$ matrix of rank r ($r \leq n$).*

- (i) \mathbf{A} is non-negative (positive) definite if and only if all its eigenvalues are non-negative (positive).
- (ii) If \mathbf{A} is non-negative (positive) definite, then exists a non-negative (positive) definite matrix $\mathbf{A}^{1/2}$ such that $\mathbf{A} = (\mathbf{A}^{1/2})^2$.
- (iii) \mathbf{A} is non-negative definite if and only if $\mathbf{A} = \mathbf{R}\mathbf{R}'$ where \mathbf{R} is $n \times n$ of rank r . This result is also true if we replace \mathbf{R} by an $n \times r$ matrix of rank r . If \mathbf{A} is positive definite then $r = n$ and \mathbf{R} is nonsingular.
- (iv) If \mathbf{A} is an $n \times n$ non-negative (positive) definite matrix and \mathbf{C} is an $n \times p$ matrix of rank p , then $\mathbf{C}'\mathbf{B}\mathbf{C}$ is non-negative (positive) definite.
- (v) If \mathbf{A} is non-negative definite and $\mathbf{C}'\mathbf{A}\mathbf{C} = \mathbf{0}$, then $\mathbf{A}\mathbf{C} = \mathbf{0}$; in particular, $\mathbf{C}'\mathbf{C} = \mathbf{0}$ implies that $\mathbf{C} = \mathbf{0}$.
- (vi) If \mathbf{A} is positive definite then so is \mathbf{A}^{-1} .
- (vii) If \mathbf{A} is $n \times p$ of rank p , then $\mathbf{A}'\mathbf{A}$ is nonsingular and therefore positive definite.
- (viii) If the elements of $n \times n$ matrix $\mathbf{A}(\boldsymbol{\theta})$ are continuous functions of $\boldsymbol{\theta}$ and $\mathbf{A}(\boldsymbol{\theta}_0)$ is positive definite, then it will be positive definite in a neighborhood of $\boldsymbol{\theta}_0$.
- (ix) If \mathbf{A} is non-negative definite (n.n.d.), then $\text{trace}[\mathbf{A}]$ is the non-negative sum of the eigenvalues of \mathbf{A} .

Proof

- (i) Since \mathbf{A} is symmetric there exists an orthogonal matrix \mathbf{T} such that $\mathbf{T}'\mathbf{A}\mathbf{T} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = \mathbf{\Lambda}$, where the λ_i are the eigenvalues of \mathbf{A} . Now \mathbf{A} is n.n.d. if and only if $\mathbf{x}'\mathbf{T}'\mathbf{A}\mathbf{T}\mathbf{x} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_n x_n^2 \geq 0$, if and only if the λ_i are nonnegative as we can set $\mathbf{x} = \mathbf{e}_i$ for each i , where \mathbf{e}_i has one for the i th element and zeros elsewhere.
- (ii) From the previous proof,

$$\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}' = \mathbf{T}\mathbf{\Lambda}^{1/2}\mathbf{T}'\mathbf{T}\mathbf{\Lambda}^{1/2}\mathbf{T}' = (\mathbf{A}^{1/2})^2,$$

where $\mathbf{\Lambda}^{1/2} = \text{diag}(\lambda_1^{1/2}, \lambda_2^{1/2}, \dots, \lambda_d^{1/2})$ is n.n.d. and $\mathbf{A}^{1/2} = \mathbf{T}\mathbf{\Lambda}^{1/2}\mathbf{T}'$ is n.n.d. by (iii).

- (iii) Since \mathbf{A} is positive semidefinite of rank r , we have from the proof of (i) that $\mathbf{T}'\mathbf{A}\mathbf{T} = \mathbf{\Lambda}$, where the eigenvalues λ_i are all positive for $i = 1, 2, \dots, r$, say, and zero for the rest. Let $\mathbf{\Lambda}^{1/2} = \text{diag}(\lambda_1^{1/2}, \lambda_2^{1/2}, \dots, \lambda_r^{1/2}, 0, \dots, 0)$. Then $\mathbf{A} = \mathbf{T}\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\mathbf{T}' = \mathbf{R}\mathbf{R}'$, where $\mathbf{R} = \mathbf{T}\mathbf{\Lambda}^{1/2}$ has rank r . Conversely, if $\mathbf{A} = \mathbf{R}\mathbf{R}'$, then $\text{rank}[\mathbf{R}] = r = \text{rank}[\mathbf{R}\mathbf{R}'] = \text{rank}[\mathbf{A}]$, and $\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{R}\mathbf{R}'\mathbf{x} = \mathbf{y}'\mathbf{y} \geq 0$, where $\mathbf{y} = \mathbf{R}'\mathbf{x}$. Hence \mathbf{A} is positive semidefinite of rank r .

We can replace $\mathbf{R} = \mathbf{T}\mathbf{\Lambda}^{1/2}$ by the $n \times r$ matrix $\mathbf{T}_r\mathbf{\Lambda}_r^{1/2}$, where $\mathbf{\Lambda}_r^{1/2} = \text{diag}(\lambda_1^{1/2}, \dots, \lambda_r^{1/2})$ and \mathbf{T}_r consists of the first r columns of \mathbf{T} .

- (iv) We note that $\mathbf{y}'\mathbf{C}'\mathbf{R}\mathbf{R}'\mathbf{C}\mathbf{y} = \mathbf{z}'\mathbf{z} \geq 0$, where $\mathbf{z} = \mathbf{R}'\mathbf{C}\mathbf{y}$. If \mathbf{R} is nonsingular, $\mathbf{z} = \mathbf{0}$ if and only if $\mathbf{y} = \mathbf{0}$ as \mathbf{C} has full column rank.
- (v) We have from (iii) that $\mathbf{A} = \mathbf{R}\mathbf{R}'$ so that $\mathbf{0} = \mathbf{C}'\mathbf{R}\mathbf{R}'\mathbf{C} = \mathbf{B}'\mathbf{B}$ ($\mathbf{B} = \mathbf{R}'\mathbf{C}$), which implies that $\mathbf{b}_i'\mathbf{b}_i = 0$ and $\mathbf{b}_i = \mathbf{0}$ for every column \mathbf{b}_i of \mathbf{B} . Hence $\mathbf{B} = \mathbf{0}$ and $\mathbf{A}\mathbf{C} = \mathbf{R}\mathbf{R}'\mathbf{C} = \mathbf{R}\mathbf{B} = \mathbf{0}$.
- (vi) Using (iii),

$$\mathbf{A}^{-1} = (\mathbf{R}\mathbf{R}')^{-1} = \mathbf{R}'^{-1}\mathbf{R}^{-1} = \mathbf{R}'^{-1}\mathbf{R}^{-1} = \mathbf{S}'\mathbf{S},$$

say, where \mathbf{S} is nonsingular. Hence \mathbf{A}^{-1} is positive definite.

- (vii) If $\mathbf{y} = \mathbf{A}\mathbf{x}$, then $\mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x} = \mathbf{y}'\mathbf{y} \geq 0$ and $\mathbf{A}'\mathbf{A}$ is positive semi-definite. However by A.4(ii), the $p \times p$ matrix $\mathbf{A}'\mathbf{A}$ has rank p and is therefore nonsingular and positive definite.
- (viii) It is well-known that a matrix is positive definite if and only if all its leading minor determinants are positive (for a proof see Seber and Lee (2003, 461–462)). Now at θ_0 the i th leading minor determinant of $\mathbf{A}(\theta)$ is positive, so by continuity it will be positive in a neighborhood N_i of θ_0 . Hence all the m leading minor determinants will be positive in the neighborhood $\mathcal{N} = \cap_{i=1}^m N_i$, and $\mathbf{A}(\theta)$ will be positive definite in \mathcal{N} .
- (ix) This follows from the proof of (ii), with \mathbf{T} orthogonal, that

$$\text{trace}[\mathbf{A}] = \text{trace}[\mathbf{T}\mathbf{\Lambda}\mathbf{T}'] = \text{trace}[\mathbf{T}'\mathbf{T}\mathbf{\Lambda}] = \text{trace}[\mathbf{\Lambda}] = \sum_i \lambda_i \geq 0,$$

by A.1 and (i).

Theorem A.10 *Let f be the matrix function*

$$f(\mathbf{\Sigma}) = \log |\mathbf{\Sigma}| + \text{trace}[\mathbf{\Sigma}^{-1}\mathbf{A}].$$

If the $d \times d$ matrix \mathbf{A} is positive definite, then, subject to $\mathbf{\Sigma}$ being positive definite, $f(\mathbf{\Sigma})$ is minimized uniquely at $\mathbf{\Sigma} = \mathbf{A}$.

Proof Let $\lambda_1, \lambda_2, \dots, \lambda_d$ be the eigenvalues of $\mathbf{\Sigma}^{-1}\mathbf{A}$, that is of $\mathbf{\Sigma}^{-1/2}\mathbf{A}\mathbf{\Sigma}^{-1/2}$ (by A.8). Since the latter matrix is positive definite (by A.9(iv)), the λ_i are positive.

Also, since the determinant of a symmetric matrix is the product of its eigenvalues

$$|(\Sigma^{-1}\mathbf{A})| = |\Sigma^{-1}||\mathbf{A}| = |\Sigma^{-1/2}\mathbf{A}\Sigma^{-1/2}| = \prod_i \lambda_i.$$

Hence

$$\begin{aligned} f(\Sigma) - f(\mathbf{A}) &= \log |\Sigma\mathbf{A}^{-1}| + \text{trace}[\Sigma^{-1}\mathbf{A}] - \text{trace}\mathbf{I}_d \\ &= -\log |\Sigma^{-1/2}\mathbf{A}\Sigma^{-1/2}| + \text{trace}[\Sigma^{-1/2}\mathbf{A}\Sigma^{-1/2}] - d \\ &= -\log \prod_i \lambda_i + \sum_i \lambda_i - d \\ &= \sum_{i=1}^d (-\log \lambda_i + \lambda_i - 1) \geq 0, \end{aligned}$$

as $\log x \leq x - 1$ for $x > 0$. Equality occurs when each λ_i is unity, that is when $\Sigma = \mathbf{A}$.

Identifiability Conditions

Theorem A.11 *Let \mathbf{X} be an $n \times p$ matrix of rank r , and \mathbf{H} a $t \times p$ matrix. Then the equations $\boldsymbol{\theta} = \mathbf{X}\boldsymbol{\beta}$ and $\mathbf{H}\boldsymbol{\beta} = \mathbf{0}$ have a unique solution for $\boldsymbol{\beta}$ for every $\boldsymbol{\theta} \in \mathcal{C}[\mathbf{X}]$ if and only if*

- (i) $\mathcal{C}[\mathbf{X}'] \cap \mathcal{C}[\mathbf{H}'] = \mathbf{0}$, and
- (ii) $\text{rank}[\mathbf{G}] = \text{rank} \begin{bmatrix} \mathbf{X} \\ \mathbf{H} \end{bmatrix} = p$.

Proof (Scheffé 1959: 17) We first of all find necessary and sufficient conditions for $\boldsymbol{\beta}$ to exist. Now $\boldsymbol{\beta}$ will exist if and only if

$$\boldsymbol{\phi} = \begin{pmatrix} \boldsymbol{\theta} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{X} \\ \mathbf{H} \end{pmatrix} \boldsymbol{\beta} = \mathbf{G}\boldsymbol{\beta} \in \mathcal{C}[\mathbf{G}] \quad \text{for every } \boldsymbol{\theta} \in \mathcal{C}[\mathbf{X}].$$

This statement is equivalent to: every vector perpendicular to $\mathcal{C}[\mathbf{G}]$ is perpendicular to $\boldsymbol{\phi}$ for every $\boldsymbol{\theta} \in \mathcal{C}[\mathbf{X}]$. Let $\mathbf{a}' = (\mathbf{a}'_X, \mathbf{a}'_H)$ be any $n + t$ dimensional vector. Then

$$\begin{aligned} \mathbf{G}'\mathbf{a} = \mathbf{0} &\Rightarrow \boldsymbol{\phi}'\mathbf{a} = 0 \quad \text{if and only if} \\ \mathbf{X}'\mathbf{a}_X + \mathbf{H}'\mathbf{a}_H = \mathbf{0} &\Rightarrow \boldsymbol{\theta}'\mathbf{a}_X = 0 \quad \text{for every } \boldsymbol{\theta} \in \mathcal{C}[\mathbf{X}] \text{ if and only if} \\ \mathbf{X}'\mathbf{a}_X + \mathbf{H}'\mathbf{a}_H = \mathbf{0} &\Rightarrow \mathbf{X}'\mathbf{a}_X = \mathbf{0} \text{ and hence } \mathbf{H}'\mathbf{a}_H = \mathbf{0}. \end{aligned}$$

Thus $\boldsymbol{\beta}$ will exist if and only if no linear combination of the rows of \mathbf{X} is a linear combination of the rows of \mathbf{H} except $\mathbf{0}$, or $\mathcal{C}[\mathbf{X}'] \cap \mathcal{C}[\mathbf{H}'] = \mathbf{0}$.

If β is to be unique, then the columns of \mathbf{G} must be linearly independent so that $\text{rank}[\mathbf{G}] = p$.

We note that the theorem implies that $\text{rank}[\mathbf{H}]$ must be $p - r$ for identifiability, so we usually have $t = p - r$, with the rows of \mathbf{H} linearly independent.

Idempotent Matrices

Theorem A.12 Let $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$ be a sequence of $n \times n$ symmetric matrices such that $\sum_{i=1}^m \mathbf{A}_i = \mathbf{I}_n$. Then the following conditions are equivalent:

- (i) $\sum_{i=1}^m r_i = n$, where $r_i = \text{rank}[\mathbf{A}_i]$.
- (ii) $\mathbf{A}_i \mathbf{A}_j = \mathbf{0}$ for all $i, j, i \neq j$.
- (iii) $\mathbf{A}_i^2 = \mathbf{A}_i$ for $i = 1, 2, \dots, m$.

Proof We first show that (i) implies (ii) and (iii). Since

$$\mathbf{y} = \mathbf{I}_n \mathbf{y} = \mathbf{A}_1 \mathbf{y} + \mathbf{A}_2 \mathbf{y} + \dots + \mathbf{A}_m \mathbf{y}, \quad (\text{A.1})$$

(i) implies that $\mathbb{R}^n = \mathcal{C}[\mathbf{A}_1] \oplus \dots \oplus \mathcal{C}[\mathbf{A}_m]$. Let $\mathbf{y} \in \mathcal{C}[\mathbf{A}_j]$. Then the unique expression of \mathbf{y} in the above form is

$$\mathbf{y} = \mathbf{0} + \dots + \mathbf{y} + \dots + \mathbf{0}. \quad (\text{A.2})$$

Since Eqs. (A.1) and (A.2) must be equivalent as \mathbf{y} has a unique decomposition into components in mutually exclusive subspaces, we have $\mathbf{A}_i \mathbf{y} = \mathbf{0}$ (all $i, i \neq j$) and $\mathbf{A}_j \mathbf{y} = \mathbf{y}$ when $\mathbf{y} \in \mathcal{C}[\mathbf{A}_j]$. In particular, for any \mathbf{x} , we have by putting $\mathbf{y} = \mathbf{A}_j \mathbf{x}$ that $\mathbf{A}_i \mathbf{A}_j \mathbf{x} = \mathbf{0}$ and $\mathbf{A}_j^2 \mathbf{x} = \mathbf{A}_j \mathbf{x}$. Hence (ii) and (iii) are true.

That (ii) implies (iii) is trivial; we simply multiply $\sum_k \mathbf{A}_k = \mathbf{I}_n$ by \mathbf{A}_i .

If (iii) is true so that each \mathbf{A}_i is idempotent, then $\text{rank}[\mathbf{A}_i] = \text{trace}[\mathbf{A}_i]$ (by A.13) and

$$\begin{aligned} n &= \text{trace}[\mathbf{I}_n] \\ &= \text{trace}\left[\sum_i \mathbf{A}_i\right] \\ &= \sum_i \text{trace}[\mathbf{A}_i] \\ &= \sum_i \text{rank}[\mathbf{A}_i] \\ &= \sum_i r_i, \end{aligned}$$

so that (iii) implies (i). This completes the proof.

Theorem A.13 If \mathbf{A} is an idempotent matrix (not necessarily symmetric) of rank r , then its eigenvalues are 0 or 1 and $\text{trace}[\mathbf{A}] = \text{rank}[\mathbf{A}] = r$.

Proof As $\mathbf{A}^2 - \mathbf{A} = \mathbf{0}$, $\lambda^2 - \lambda = 0$ is the minimal polynomial. Hence its eigenvalues are 0 or 1 and \mathbf{A} is diagonalizable. Therefore there exists a nonsingular matrix \mathbf{R} such that

$$\mathbf{R}^{-1}\mathbf{A}\mathbf{R} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

since the rank is unchanged when pre- or post-multiplying by a nonsingular matrix (A.4(i)). Hence

$$\text{trace}[\mathbf{A}] = \text{trace}[\mathbf{A}\mathbf{R}\mathbf{R}^{-1}] = \text{trace}[\mathbf{R}^{-1}\mathbf{A}\mathbf{R}] = r.$$

When \mathbf{A} is also symmetric we see from Theorem 1.4 in Sect. 1.5 that \mathbf{R} is replaced by an orthogonal matrix.

Weak (Generalized) Inverse

Theorem A.14 *If \mathbf{A} is any matrix with weak inverse \mathbf{A}^- , then $\mathbf{A}\mathbf{A}^-$ is idempotent and $\text{trace}[\mathbf{A}\mathbf{A}^-] = \text{rank}[\mathbf{A}\mathbf{A}^-] = \text{rank}[\mathbf{A}]$.*

Proof $(\mathbf{A}\mathbf{A}^-)\mathbf{A}^- = \mathbf{A}\mathbf{A}^-$, so that $\mathbf{A}\mathbf{A}^-$ is idempotent. Now from A.4(ii),

$$\text{rank}[\mathbf{A}] = \text{rank}[\mathbf{A}\mathbf{A}^- \mathbf{A}] \leq \text{rank}[\mathbf{A}\mathbf{A}^-] \leq \text{rank}[\mathbf{A}].$$

Hence $\text{rank}[\mathbf{A}\mathbf{A}^-] = \text{rank}[\mathbf{A}] = \text{trace}[\mathbf{A}\mathbf{A}^-]$, by A.13.

Theorem A.15

- (i) $(\mathbf{A}^-)'$ is a weak inverse of \mathbf{A}' , which we can then describe symbolically as $(\mathbf{A}^-)' = (\mathbf{A}')^-$. (Technically \mathbf{A}^- is not unique as it represents a family of matrices.)
- (ii) If \mathbf{X} is $n \times p$ of rank $r < p$ and $(\mathbf{X}'\mathbf{X})^-$ is any weak inverse of $\mathbf{X}'\mathbf{X}$, then we have $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ is the unique projection matrix onto $\mathcal{C}[\mathbf{X}]$, so that it is symmetric and idempotent.
- (iii) If \mathbf{G} is defined in A.11, then $(\mathbf{G}'\mathbf{G})^{-1}$ is a weak inverse of $\mathbf{X}'\mathbf{X}$.

Proof

- (i) This is proved by taking the transpose of $\mathbf{A} = \mathbf{A}\mathbf{A}^- \mathbf{A}$.
- (ii) Let $\Omega = \mathcal{C}[\mathbf{X}]$ and let $\theta = \mathbf{X}\beta \in \Omega$. Given the normal equations $\mathbf{X}'\mathbf{X}\beta = \mathbf{X}'\mathbf{y}$, these have a solution $\hat{\beta} = (\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{y}$. If $\hat{\theta} = \mathbf{X}\hat{\beta}$, then

$$\begin{aligned} \hat{\theta}'(\mathbf{y} - \hat{\theta}) &= \hat{\beta}'\mathbf{X}'(\mathbf{y} - \mathbf{X}\hat{\beta}) \\ &= \hat{\beta}'(\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\hat{\beta}) \\ &= 0. \end{aligned}$$

We therefore have an orthogonal decomposition of $\mathbf{y} = \hat{\boldsymbol{\theta}} + \mathbf{y} - \hat{\boldsymbol{\theta}}$ such that $\hat{\boldsymbol{\theta}} \in \Omega$ and $(\mathbf{y} - \hat{\boldsymbol{\theta}}) \perp \Omega$. Since $\hat{\boldsymbol{\theta}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ and the orthogonal projection is unique, we must have $\mathbf{P}_\Omega = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

- (iii) Using the normal equations $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$ and adding $\mathbf{H}\hat{\boldsymbol{\beta}} = \mathbf{0}$ we have $(\mathbf{X}'\mathbf{X} + \mathbf{H}'\mathbf{H})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$ so that $\mathbf{0} = \mathbf{H}\hat{\boldsymbol{\beta}} = \mathbf{H}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{X}'\mathbf{y}$ for all \mathbf{y} . Hence $\mathbf{H}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{X}' = \mathbf{0}$ and

$$\mathbf{X}'\mathbf{X}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{X}'\mathbf{X} = (\mathbf{X}'\mathbf{X} + \mathbf{H}'\mathbf{H})(\mathbf{G}'\mathbf{G})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{X}'\mathbf{X}.$$

Theorem A.16 Let $\mathbf{y} \sim N_n[\mathbf{0}, \Sigma]$, where Σ is a nonnegative-definite matrix of rank s . If Σ^- is any weak inverse of Σ , then $\mathbf{y}'\Sigma^-\mathbf{y} \sim \chi_s^2$.

Proof If $\mathbf{z} \sim N_n[\mathbf{I}_s, \mathbf{0}]$, then \mathbf{y} has the same distribution as $\Sigma^{1/2}\mathbf{z}$ (cf. A.9(ii)) since $\text{Var}[\mathbf{y}] = (\Sigma^{1/2})^2 = \Sigma$. Now $\mathbf{y}'\mathbf{A}\mathbf{y} = \mathbf{z}'\Sigma^{1/2}\mathbf{A}\Sigma^{1/2}\mathbf{z}$ is χ_r^2 if $\Sigma^{1/2}\mathbf{A}\Sigma^{1/2}$ is idempotent (Theorem 1.10 in Sect. 1.9), where

$$r = \text{trace}[\Sigma^{1/2}\mathbf{A}\Sigma^{1/2}] = \text{trace}[\Sigma\mathbf{A}]$$

(since $\text{trace}[\mathbf{CD}] = \text{trace}[\mathbf{DC}]$), that is if

$$\Sigma^{1/2}\mathbf{A}\Sigma^{1/2}\Sigma^{1/2}\mathbf{A}\Sigma^{1/2} = \Sigma^{1/2}\mathbf{A}\Sigma^{1/2}. \quad (\text{A.3})$$

Multiplying the above equation on the left and right by $\Sigma^{1/2}$ we get

$$\Sigma\mathbf{A}\Sigma\mathbf{A}\Sigma = \Sigma\mathbf{A}\Sigma. \quad (\text{A.4})$$

We now show that Eqs. (A.3) and (A.4) are equivalent conditions.

Let $\mathbf{B} = \mathbf{A}\Sigma\mathbf{A} - \mathbf{A}$, then we need to show that $\Sigma\mathbf{B}\Sigma = \mathbf{0}$ implies that the matrix $\mathbf{D} = \Sigma^{1/2}\mathbf{B}\Sigma^{1/2} = \mathbf{0}$. Now \mathbf{D} is symmetric and given $\Sigma\mathbf{B}\Sigma = \mathbf{0}$,

$$\begin{aligned} \text{trace}[\mathbf{D}^2] &= \text{trace}[\Sigma^{1/2}\mathbf{B}\Sigma^{1/2}\Sigma^{1/2}\mathbf{B}\Sigma^{1/2}] \\ &= \text{trace}[\Sigma\mathbf{B}\Sigma\mathbf{B}] \\ &= \mathbf{0}. \end{aligned}$$

However $\text{trace}[\mathbf{D}^2] = \sum_i \sum_j d_{ij}^2 = 0$ implies that $\mathbf{D} = \mathbf{0}$.

We now set $\mathbf{A} = \Sigma^-$, then

$$\Sigma\mathbf{A}\Sigma\mathbf{A}\Sigma = (\Sigma\Sigma^-\Sigma)\Sigma^-\Sigma = \Sigma\Sigma^-\Sigma = \Sigma\mathbf{A}\Sigma,$$

and the condition for idempotency is satisfied. We note that $r = \text{trace}[\Sigma\mathbf{A}] = \text{trace}[\Sigma\Sigma^-]$ and, from A.14,

$$\text{rank}[\Sigma\Sigma^-] = \text{rank}[\Sigma] = s = \text{trace}[\Sigma\Sigma^-].$$

Hence $r = s$ and $\mathbf{y}'\Sigma^{-1}\mathbf{y}$ is χ_s^2 .

Inverse of a Partitioned Matrix

Theorem A.17 *If \mathbf{A} and \mathbf{C} are symmetric matrices and all inverses exist, then*

$$\begin{pmatrix} \mathbf{A} & \mathbf{B}' \\ \mathbf{B} & \mathbf{C} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{F}^{-1} & -\mathbf{F}^{-1}\mathbf{G}' \\ -\mathbf{G}\mathbf{F}^{-1} & \mathbf{C}^{-1} + \mathbf{G}\mathbf{F}^{-1}\mathbf{G}' \end{pmatrix},$$

where $\mathbf{F} = \mathbf{A} - \mathbf{B}'\mathbf{C}^{-1}\mathbf{B}$ and $\mathbf{G} = \mathbf{C}^{-1}\mathbf{B}$.

Proof The result is proved by confirming that the matrix multiplied on the left by its inverse is the identity matrix.

Theorem A.18 *If \mathbf{A} is positive definite and all inverses exist, then*

$$\begin{pmatrix} \mathbf{A} & \mathbf{B}' \\ \mathbf{B} & \mathbf{0} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}'(\mathbf{B}\mathbf{A}^{-1}\mathbf{B}')^{-1}\mathbf{B}\mathbf{A}^{-1} & \mathbf{A}^{-1}\mathbf{B}'(\mathbf{B}\mathbf{A}^{-1}\mathbf{B}')^{-1} \\ (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}')^{-1}\mathbf{B}\mathbf{A}^{-1} & -(\mathbf{B}\mathbf{A}^{-1}\mathbf{B}')^{-1} \end{pmatrix}$$

Proof The matrix times its inverse is the identity matrix.

Theorem A.19 *Let \mathbf{A} be an $r \times r$ positive definite matrix, \mathbf{B} be an $s_1 \times r$ matrix of rank s_1 , \mathbf{C} be an $s_2 \times r$ matrix of rank s_2 , and $\mathcal{C}[\mathbf{B}'] \cap \mathcal{C}[\mathbf{C}'] = \mathbf{0}$ so that $\text{rank}[(\mathbf{B}', \mathbf{C}')] = s_1 + s_2$. If*

$$\mathbf{Z} = \begin{pmatrix} \mathbf{A} & \mathbf{B}' & \mathbf{C}' \\ \mathbf{B} & \mathbf{0} & \mathbf{0} \\ \mathbf{C} & \mathbf{0} & \mathbf{0} \end{pmatrix},$$

then

$$\mathbf{Z}^{-1} = \begin{pmatrix} \mathbf{P} & \mathbf{Q}' & \mathbf{R}' \\ \mathbf{Q} & -\mathbf{Q}\mathbf{A}\mathbf{Q}' & -\mathbf{Q}\mathbf{A}\mathbf{R}' \\ \mathbf{R} & -\mathbf{R}\mathbf{A}\mathbf{Q}' & \mathbf{R}\mathbf{A}\mathbf{R}' \end{pmatrix},$$

where

$$\mathbf{P} = \mathbf{M} - \mathbf{M}\mathbf{C}'(\mathbf{C}\mathbf{M}\mathbf{C}')^{-1}\mathbf{C}\mathbf{M},$$

$$\mathbf{Q} = (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}')^{-1}\mathbf{B}\mathbf{A}^{-1}[\mathbf{I}_r - \mathbf{C}'(\mathbf{C}\mathbf{M}\mathbf{C}')^{-1}\mathbf{C}\mathbf{M}],$$

$$\mathbf{R} = (\mathbf{C}\mathbf{M}\mathbf{C}')^{-1}\mathbf{C}\mathbf{M}, \quad \text{and}$$

$$\mathbf{M} = \mathbf{M}' = \mathbf{A}^{-1}[\mathbf{I}_r - \mathbf{B}'(\mathbf{B}\mathbf{A}^{-1}\mathbf{B}')^{-1}\mathbf{B}\mathbf{A}^{-1}].$$

Proof Let

$$\mathbf{Z}^{-1} = \begin{pmatrix} \mathbf{P} & \mathbf{Q}' & \mathbf{R}' \\ \mathbf{Q} & \mathbf{T} & \mathbf{U}' \\ \mathbf{R} & \mathbf{U} & \mathbf{V} \end{pmatrix},$$

then from $\mathbf{ZZ}^{-1} = \mathbf{I}_{r+s_1+s_2}$, we have a system of nine matrix equations, namely

$$\mathbf{AP} + \mathbf{B}'\mathbf{Q} + \mathbf{C}'\mathbf{R} = \mathbf{I}_r, \quad \mathbf{BP} = \mathbf{0}, \quad \mathbf{CP} = \mathbf{0}, \quad (\text{A.5})$$

$$\mathbf{AQ}' + \mathbf{B}'\mathbf{T} + \mathbf{C}'\mathbf{U} = \mathbf{0}, \quad \mathbf{BQ}' = \mathbf{I}_{s_1}, \quad \mathbf{CQ}' = \mathbf{0}, \quad (\text{A.6})$$

$$\mathbf{AR}' + \mathbf{B}'\mathbf{U}' + \mathbf{C}'\mathbf{V} = \mathbf{0}, \quad \mathbf{BR}' = \mathbf{0}, \quad \mathbf{CR}' = \mathbf{I}_{s_2}. \quad (\text{A.7})$$

Now from Eq. (A.5)

$$\mathbf{P} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}'\mathbf{Q} - \mathbf{A}^{-1}\mathbf{C}'\mathbf{R} \quad (\text{A.8})$$

and using $\mathbf{BP} = \mathbf{0}$ gives us

$$\mathbf{Q} = (\mathbf{BA}^{-1}\mathbf{B}')^{-1}\mathbf{BA}^{-1}(\mathbf{I}_r - \mathbf{C}'\mathbf{R}).$$

Substituting back into Eq. (A.8) and using $\mathbf{CP} = \mathbf{0}$ leads to

$$\mathbf{CM} - \mathbf{CMC}'\mathbf{R} = \mathbf{0}. \quad (\text{A.9})$$

Since \mathbf{A}^{-1} is positive definite, there exists a nonsingular $r \times r$ matrix \mathbf{L} such that $\mathbf{A}^{-1} = \mathbf{L}'\mathbf{L}$ (A.9(iii)). Now \mathbf{LB}' is $r \times s_1$ of rank s_1 so that

$$\begin{aligned} \mathbf{CMC}' &= \mathbf{CL}'[\mathbf{I}_r - \mathbf{LB}'(\mathbf{BL}'\mathbf{LB}')^{-1}\mathbf{BL}']\mathbf{LC}' \\ &= \mathbf{CL}'(\mathbf{I}_r - \mathbf{P}_{\mathcal{C}[\mathbf{LB}']})\mathbf{LC}' \\ &= \mathbf{CL}'\mathbf{P}_{\mathcal{N}[\mathbf{BL}']}\mathbf{LC}' \end{aligned}$$

by Theorem 1.1 in Sect. 1.2. Now Theorem 4.4 in Sect. 4.2 states that if \mathbf{A} is $q \times n$ of rank q then $\text{rank}[\mathbf{P}_{\Omega}\mathbf{A}] = q$ if and only if $\mathcal{C}[\mathbf{A}'] \cap \Omega^\perp = \mathbf{0}$. If $\Omega = \mathcal{N}[\mathbf{BL}']$ and $\mathbf{A} = \mathbf{CL}'$, an $s_2 \times r$ matrix of rank s_2 , then

$$\mathcal{C}[\mathbf{A}'] \cap \Omega^\perp = \mathcal{C}[\mathbf{L}'\mathbf{C}] \cap \mathcal{C}[\mathbf{L}'\mathbf{B}] = \mathbf{0},$$

since $\text{rank}[\mathbf{B}', \mathbf{C}']$ is unchanged by premultiplying by \mathbf{L}' , a nonsingular matrix. Hence $\text{rank}[\mathbf{P}_{\mathcal{N}[\mathbf{BL}']}\mathbf{LC}'] = s_2$. As $\mathbf{P}_{\mathcal{N}[\mathbf{BL}']}$ is symmetric and idempotent, $\mathbf{CL}'\mathbf{P}_{\mathcal{N}[\mathbf{BL}']}\mathbf{P}_{\mathcal{N}[\mathbf{BL}']}\mathbf{LC}'$ is $s_2 \times s_2$ of rank s_2 , and is therefore nonsingular, so that \mathbf{CMC}' has an inverse. From Eq. (A.9)

$$\mathbf{R} = (\mathbf{C}\mathbf{M}\mathbf{C}')^{-1}\mathbf{C}\mathbf{M},$$

and from Eq. (A.6)

$$\begin{aligned}\mathbf{0} &= \mathbf{Q}\mathbf{A}\mathbf{Q}' + \mathbf{Q}\mathbf{B}'\mathbf{T} + \mathbf{Q}\mathbf{C}'\mathbf{U} \\ &= \mathbf{Q}\mathbf{A}\mathbf{Q}' + \mathbf{I}_{s_1}\mathbf{T}\end{aligned}$$

so that $\mathbf{T} = -\mathbf{Q}\mathbf{A}\mathbf{Q}'$. From premultiplying Eq. (A.7) by \mathbf{Q} , and then premultiplying (A.7) by \mathbf{R} , we obtain $\mathbf{U}' = -\mathbf{Q}\mathbf{A}\mathbf{R}'$. Since $\mathbf{R}\mathbf{C}' = \mathbf{I}_{s_2}$, $\mathbf{V} = \mathbf{R}\mathbf{A}\mathbf{R}'$.

Differentiation

Theorem A.20 *If $d/d\beta$ denotes the column vector with i th element $d/d\beta_i$, then:*

- (i) $d(\mathbf{a}'\beta)/d\beta = \mathbf{a}$.
- (ii) $d(\beta'\mathbf{A}\beta)/d\beta = 2\mathbf{A}\beta$.

Proof

- (i) $d\sum_i a_i\beta_i/d\beta_i = a_i$.
- (ii)

$$\begin{aligned}d(\beta'\mathbf{A}\beta)/d\beta_i &= d\left(\sum_i a_{ii}\beta_i^2 + \sum_i \sum_{j:j\neq i} a_{ij}\beta_i\beta_j\right)/d\beta_i \\ &= 2a_{ii}\beta_i + \sum_{j:j\neq i} (a_{ij} + a_{ji})\beta_j \\ &= 2\sum_j a_{ij}\beta_j.\end{aligned}$$

Inequalities

Theorem A.21

- (i) *If \mathbf{D} is positive definite, then for any \mathbf{a}*

$$\sup_{\mathbf{x}:\mathbf{x}\neq\mathbf{0}} \left\{ \frac{(\mathbf{a}'\mathbf{x})^2}{\mathbf{x}'\mathbf{D}\mathbf{x}} \right\} = \mathbf{a}'\mathbf{D}^{-1}\mathbf{a}.$$

(ii) If \mathbf{M} and \mathbf{N} are positive definite, then

$$\sup_{\mathbf{x}, \mathbf{y}, \mathbf{x} \neq \mathbf{0}, \mathbf{y} \neq \mathbf{0}} \left\{ \frac{(\mathbf{x}'\mathbf{L}\mathbf{y})^2}{\mathbf{x}'\mathbf{M}\mathbf{x} \cdot \mathbf{y}'\mathbf{N}\mathbf{y}} \right\} = \theta_{\max},$$

where θ_{\max} is the largest eigenvalue of $\mathbf{M}^{-1}\mathbf{L}\mathbf{N}^{-1}\mathbf{L}'$, and of $\mathbf{N}^{-1}\mathbf{L}'\mathbf{M}^{-1}\mathbf{L}$.

Proof Proofs are given by Seber (1984: 527).

References

- Scheffé, H. (1959). *The analysis of variance*. New York: Wiley.
- Seber, G. A. F. (1984). *Multivariate observations*. New York: Wiley. Also reproduced in paperback by Wiley in 2004.
- Seber, G. A. F. (2008). *A matrix handbook for statisticians*. New York: Wiley.
- Seber, G. A. F., & Lee, A. J. (2003). *Linear regression analysis* (2nd ed.). New York: Wiley.

Index

A

- Analysis of covariance, examples of, 24–25
- Analysis of variance (ANOVA)
 - examples of, 22–24, 67
 - orthogonality in, 81
 - table, 50
 - table for randomized block design, 95
- Autocorrelation, 39

B

- Basis, vector space, 3

C

- Cochran's theorem
 - multivariate, 137
 - univariate, 48
- Column space, 2
- Concentrated likelihood, 123
- Concomitant variable, 104
- Confidence intervals, 57
- Constraint-equation hypothesis, 149
- Contrasts, 55
- Covariance of two vectors, 6
- Critical region, 61

D

- Dimension, 3

E

- Eigenvalues, 191–192
- Enlarging the model
 - examples of, 103

- missing observations and, 111–116
 - randomized block design and, 109
 - regression extensions, 107
- Estimable functions, 34, 54
 - Expectation of a random matrix, 6
 - Expected information matrix, 150

F

- Factorial experiment, 55
- Fitted values, 38
- Freedom-equation hypothesis, 175–179, 186–188

G

- Gauss-Markov theorem
 - multivariate, 133
 - univariate, 34
- Generalized inverse, 33, 132
- Generalized linear hypothesis, 143
- Growth curves, 144

H

- Hat matrix, 38
- Hotelling's test statistic, 144–146
- Hypothesis sum of squares, 50

I

- Idempotent matrix, 195–196
 - eigenvalues of, 5
 - non-central chi-square and, 14
 - orthogonal projection and, 28, 29
 - symmetric, 5

Identifiability conditions, 18, 23, 32, 33, 161
 multivariate, 132
 Independent quadratic forms, 16
 Indicator variables, 22
 Inequalities, 200
 Information matrix
 definition of, 150
 positive definite, 152, 176–177
 positive semidefinite, 161–171, 177–179
 Internally studentized residual, 39
 Intersection of vector spaces, 3

L

Lagrange multiplier, 17, 32, 44, 110, 157
 multiplier test, 127, 166–171
 zero for identifiability conditions, 33
 Least squares equations, 30
 Least squares estimate, 27
 Length of a vector, 1
 Less than full rank model, 32–34
 Likelihood ratio test, 47, 127, 185
 Linear hypothesis, 21–26
 definition of, 25
 Linear independence, 3
 Linear model approximation, 155
 Linear regression
 examples of, 21–22, 29, 30, 52
 orthogonal hypotheses and, 78–81
 Linear vector space, 2

M

Maximum likelihood estimation, 44–45
 Minimum variance estimator, 34
 Missing observations, 111–116
 analysis of covariance for, 115
 randomized block design and, 114
 Moment generating function, 9
 Multinomial distribution, 181
 Multiple confidence intervals, 147
 Multiple correlation coefficient, 42
 Multivariate maximum likelihood estimates,
 136
 Multivariate models, 129–147
 Multivariate normal distribution, 8
 Multivariate randomized block design, 129,
 141

N

Nested test method, 73
 Non-central chi-square distribution, 9, 48, 160
 Non-central F -distribution, 10, 49

Non-centrality parameter, 9
 Non-linear regression
 asymptotic theory for, 122
 examples of, 117
 large sample tests for, 127
 Non-negative definite matrix, 5, 192
 Normal equations, 30, 33
 Nullity, 3
 Null space, 2

O

Orthogonal complement, 2
 Orthogonal decomposition, 74
 Orthogonal hypotheses
 definition of, 76
 multivariate, 143
 nonlinear, 171–174
 p -factor layouts and, 89–93
 randomized block design and,
 93–96
 regression models and, 78–81
 two-factor layouts and, 81–89
 Orthogonal matrix, 192
 Orthogonal polynomials, 80
 Orthogonal projection, 27, 48
 Orthogonal transformation, 56
 Orthonormal basis, 3, 27, 30, 48
 Orthonormal contrasts, 55

P

Partitioned matrix, 198–200
 p -factor layouts, 89–93
 Pitman's limiting power, 154
 Polynomial regression, 21
 Positive-definite matrix, 189, 192, 193, 200
 Positive semidefinite matrix, 189
 Power of the F -test, 61–63
 Projection matrix, 4

Q

Quadratically balanced design, 36
 Quadratically balanced test, 67
 Quadratic form
 expectation of, 11
 variance of, 12

R

Randomized block design, 93–96
 Rank, 190–191
 Residuals, 38

Residual sum of squares, 35
Robustness of F -test, 64–69
Row space, 2

S

Score test, 127, 184–186
Separate test method, 76
Serial correlation effect, 39
Simultaneous confidence intervals,
58
Sum of two vector spaces, 3
Systemic bias, 39

T

Test for independence, 186
Trace, 1, 189

Two-factor layouts, 81–89
Two-way ANOVA table, 88

U

Uniformly most powerful test, 61

V

Variance-covariance matrix, 6
Variance estimation, 35
Vector differentiation, 30, 200

W

Wald test, 51–55, 127, 165, 184–186
Weak inverse, 33, 197
Wishart distribution, 135