Appendix A
Matrix Algebra

There are numerous books on matrix algebra that contain results useful for the discussion of linear models. See for instance books by Graybill (1961); Mardia, Kent, and Bibby (1979); Searle (1982); Rao (1973a); Rao and Mitra (1971); and Rao and Rao (1998), to mention a few. We collect in this Appendix some of the important results for ready reference. Proofs are generally omitted. References to original sources are given wherever necessary.

A.1 Overview

Definition A.1 *An $m \times n$-matrix A is a rectangular array of elements in m rows and n columns.*

In the context of the material treated in the book and in this Appendix, the elements of a matrix are taken as real numbers. We indicate an $m \times n$-matrix by writing $A : m \times n$ or $\underset{m,n}{A}$.

Let a_{ij} be the element in the ith row and the jth column of A. Then A may be represented as

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})\,.$$

A matrix with $n = m$ rows and columns is called a square matrix. A square matrix having zeros as elements below (above) the diagonal is called an upper (lower) triangular matrix.

Definition A.2 *The transpose $A' : n \times m$ of a matrix $A : m \times n$ is given by interchanging the rows and columns of A. Thus*

$$A' = (a_{ji})\,.$$

Then we have the following rules:

$$(A')' = A\,, \quad (A + B)' = A' + B'\,, \quad (AB)' = B'A'\,.$$

Definition A.3 *A square matrix is called symmetric if $A' = A$.*

Definition A.4 *An $m \times 1$ matrix a is said to be an m-vector and written as a column*

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

Definition A.5 *A $1 \times n$-matrix a' is said to be a row vector*

$$a' = (a_1, \cdots, a_n).$$

$A : m \times n$ may be written alternatively in a partitioned form as

$$A = (a_{(1)}, \ldots, a_{(n)}) = \begin{pmatrix} a'_1 \\ \vdots \\ a'_m \end{pmatrix}$$

with

$$a_{(j)} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, \quad a_i = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix}.$$

Definition A.6 *The $n \times 1$ row vector $(1, \cdots, 1)'$ is denoted by $1'_n$ or $1'$.*

Definition A.7 *The matrix $A : m \times m$ with $a_{ij} = 1$ (for all i,j) is given the symbol J_m, that is,*

$$J_m = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \vdots & 1 \end{pmatrix} = 1_m 1'_m\,.$$

Definition A.8 *The n-vector*

$$e_i = (0, \cdots, 0, 1, 0, \cdots, 0)'$$

with the ith component as 1 and all the others as 0, is called the ith unit vector.

Definition A.9 *A $n \times n$ (square) matrix with elements 1 on the main diagonal and zeros off the diagonal is called the identity matrix I_n.*

Definition A.10 *A square matrix $A : n \times n$ with zeros in the off diagonal is called a diagonal matrix. We write*

$$A = \operatorname{diag}(a_{11}, \cdots, a_{nn}) = \operatorname{diag}(a_{ii}) = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}.$$

Definition A.11 *A matrix A is said to be partitioned if its elements are arranged in submatrices.*

Examples are

$$\underset{m,n}{A} = (\underset{m,r}{A_1}, \underset{m,s}{A_2}) \quad \text{with} \quad r + s = n$$

or

$$\underset{m,n}{A} = \begin{pmatrix} \underset{r,n-s}{A_{11}} & \underset{r,s}{A_{12}} \\ \underset{m-r,n-s}{A_{21}} & \underset{m-r,s}{A_{22}} \end{pmatrix}.$$

For partitioned matrices we get the transposess as

$$A' = \begin{pmatrix} A_1' \\ A_2' \end{pmatrix}, \quad A' = \begin{pmatrix} A_{11}' & A_{21}' \\ A_{12}' & A_{22}' \end{pmatrix},$$

respectively.

A.2 Trace of a Matrix

Definition A.12 *Let $a_{11}, \ldots, a_{nn}$ be the elements on the main diagonal of a square matrix $A : n \times n$. Then the trace of A is defined as the sum*

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}\,.$$

Theorem A.13 *Let A and B be square $n \times n$ matrices, and let c be a scalar factor. Then we have the following rules:*

(i) $\operatorname{tr}(A \pm B) = \operatorname{tr}(A) \pm \operatorname{tr}(B)$;

(ii) $\operatorname{tr}(A') = \operatorname{tr}(A)$;

(iii) $\operatorname{tr}(cA) = c \operatorname{tr}(A)$;

(iv) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ *(here A and B can be rectangular matrices of the form $A : m \times n$ and $B : n \times m$);*

(v) $\operatorname{tr}(AA') = \operatorname{tr}(A'A) = \sum_{i,j} a_{ij}^2$;

(vi) *If* $a = (a_1, \ldots, a_n)'$ *is an* n*-vector, then its squared norm may be written as*

$$\|a\|^2 = a'a = \sum_{i=1}^{n} a_i^2 = \mathrm{tr}(aa').$$

Note, that rules (iv) and (v) also hold for the case $A : n \times m$ and $B : m \times n$.

A.3 Determinant of a Matrix

Definition A.14 *Let* $n > 1$ *be a positive integer. The determinant of a square matrix* $A : n \times n$ *is defined by*

$$|A| = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} |M_{ij}| \quad \textit{(for any } j\textit{, } j \textit{ fixed),}$$

with $|M_{ij}|$ *being the minor of the element* a_{ij}*.* $|M_{ij}|$ *is the determinant of the remaining* $(n-1) \times (n-1)$ *matrix when the* i*th row and the* j*th column of* A *are deleted.* $A_{ij} = (-1)^{i+j}|M_{ij}|$ *is called the cofactor of* a_{ij}*.*

Examples:

For $n = 2$: $|A| = a_{11}a_{22} - a_{12}a_{21}$.

For $n = 3$ (first column ($j = 1$) fixed):

$$\begin{aligned}
A_{11} &= (-1)^2 \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = (-1)^2 M_{11} \\
A_{21} &= (-1)^3 \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} = (-1)^3 M_{21} \\
A_{31} &= (-1)^4 \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = (-1)^4 M_{31} \\
\Rightarrow |A| &= a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31}.
\end{aligned}$$

Note: As an alternative one may fix a row and develop the determinant of A according to

$$|A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} |M_{ij}| \quad \text{(for any } i\text{, } i \text{ fixed).}$$

Definition A.15 *A square matrix* A *is said to be regular or nonsingular if* $|A| \neq 0$*. Otherwise* A *is said to be singular.*

Theorem A.16 *Let* A *and* B *be* $n \times n$ *square matrices, and* c *be a scalar. Then we have*

(i) $|A'| = |A|$,

(ii) $|cA| = c^n |A|$,

(iii) $|AB| = |A||B|$,

(iv) $|A^2| = |A|^2$,

(v) *If A is diagonal or triangular, then*

$$|A| = \prod_{i=1}^{n} a_{ii}\,.$$

(vi) *For* $D = \begin{pmatrix} \underset{n,n}{A} & \underset{n,m}{C} \\ \underset{m,n}{0} & \underset{m,m}{B} \end{pmatrix}$ *we have*

$$\begin{vmatrix} A & C \\ 0 & B \end{vmatrix} = |A||B|,$$

and analogously

$$\begin{vmatrix} A' & 0' \\ C' & B' \end{vmatrix} = |A||B|.$$

(vii) *If A is partitioned with $A_{11} : p \times p$ and $A_{22} : q \times q$ square and nonsingular, then*

$$\begin{aligned} \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} &= |A_{11}||A_{22} - A_{21}A_{11}^{-1}A_{12}| \\ &= |A_{22}||A_{11} - A_{12}A_{22}^{-1}A_{21}|. \end{aligned}$$

Proof: Define the following matrices

$$Z_1 = \begin{pmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{pmatrix} \quad \text{and} \quad Z_2 = \begin{pmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{pmatrix},$$

where $|Z_1| = |Z_2| = 1$ by (vi). Then we have

$$Z_1 A Z_2 = \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{pmatrix}$$

and [using (iii) and (iv)]

$$|Z_1 A Z_2| = |A| = |A_{22}||A_{11} - A_{12}A_{22}^{-1}A_{21}|.$$

(viii) $\begin{vmatrix} A & x \\ x' & c \end{vmatrix} = |A|(c - x'A^{-1}x)$ *where x is an n-vector.*

Proof: Use (vii) with A instead of A_{11} and c instead of A_{22}.

(ix) *Let $B : p \times n$ and $C : n \times p$ be any matrices and $A : p \times p$ a nonsingular matrix. Then*

$$\begin{aligned} |A + BC| &= |A||I_p + A^{-1}BC| \\ &= |A||I_n + CA^{-1}B|. \end{aligned}$$

Proof: The first relationship follows from (iii) and

$$(A + BC) = A(I_p + A^{-1}BC)$$

immediately. The second relationship is a consequence of (vii) applied to the matrix

$$\begin{vmatrix} I_p & -A^{-1}B \\ C & I_n \end{vmatrix} = |I_p||I_n + CA^{-1}B| = |I_n||I_p + A^{-1}BC|.$$

(x) $|A + aa'| = |A|(1 + a'A^{-1}a)$, *if A is nonsingular.*

(xi) $|I_p + BC| = |I_n + CB|$, *if $B : p \times n$ and $C : n \times p$.*

A.4 Inverse of a Matrix

Definition A.17 *A matrix $B : n \times n$ is said to be an inverse of $A : n \times n$ if $AB = I$. If such a B exists, it is denoted by A^{-1}. It is easily seen that A^{-1} exists if and only if A is nonsingular. It is easy to establish that if A^{-1} exists; then $AA^{-1} = A^{-1}A = I$.*

Theorem A.18 *If all the inverses exist, we have*

(i) $(cA)^{-1} = c^{-1}A^{-1}$.

(ii) $(AB)^{-1} = B^{-1}A^{-1}$.

(iii) *If $A : p \times p$, $B : p \times n$, $C : n \times n$ and $D : n \times p$ then*

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}.$$

(iv) *If $1 + b'A^{-1}a \neq 0$, then we get from (iii)*

$$(A + ab')^{-1} = A^{-1} - \frac{A^{-1}ab'A^{-1}}{1 + b'A^{-1}a}.$$

(v) $|A^{-1}| = |A|^{-1}$.

Theorem A.19 (Inverse of a partitioned matrix) *For partitioned regular A*

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix},$$

where $E : (n_1 \times n_1)$, $F : (n_1 \times n_2)$, $G : (n_2 \times n_1)$ and $H : (n_2 \times n_2)$ $(n_1 + n_2 = n)$ are such that E and $D = H - GE^{-1}F$ are regular, the partitioned inverse is given by

$$A^{-1} = \begin{pmatrix} E^{-1}(I + FD^{-1}GE^{-1}) & -E^{-1}FD^{-1} \\ -D^{-1}GE^{-1} & D^{-1} \end{pmatrix} = \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix}.$$

Proof: Check that the product of A and A^{-1} reduces to the identity matrix, that is,

$$AA^{-1} = A^{-1}A = I.$$

A.5 Orthogonal Matrices

Definition A.20 *A square matrix $A : n \times n$ is said to be orthogonal if $AA' = I = A'A$. For orthogonal matrices, we have*

(i) $A' = A^{-1}$.

(ii) $|A| = \pm 1$.

(iii) *Let $\delta_{ij} = 1$ for $i = j$ and 0 for $i \neq j$ denote the Kronecker symbol. Then the row vectors a_i and the column vectors $a_{(i)}$ of A satisfy the conditions*

$$a_i a_j' = \delta_{ij}\,, \quad a_{(i)}' a_{(j)} = \delta_{ij}\,.$$

(iv) *AB is orthogonal if A and B are orthogonal.*

Theorem A.21 *For $A : n \times n$ and $B : n \times n$ symmetric matrices, there exists an orthogonal matrix H such that $H'AH$ and $H'BH$ become diagonal if and only if A and B commute, that is,*

$$AB = BA.$$

A.6 Rank of a Matrix

Definition A.22 *The rank of $A : m \times n$ is the maximum number of linearly independent rows (or columns) of A. We write* $\operatorname{rank}(A) = p$.

Theorem A.23 (Rules for ranks)

(i) $0 \leq \operatorname{rank}(A) \leq \min(m, n)$.

(ii) $\operatorname{rank}(A) = \operatorname{rank}(A')$.

(iii) $\operatorname{rank}(A + B) \leq \operatorname{rank}(A) + \operatorname{rank}(B)$.

(iv) $\operatorname{rank}(AB) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}$.

(v) $\operatorname{rank}(AA') = \operatorname{rank}(A'A) = \operatorname{rank}(A) = \operatorname{rank}(A')$.

(vi) *For nonsingular $B : m \times m$ and $C : n \times n$, we have* $\operatorname{rank}(BAC) = \operatorname{rank}(A)$.

(vii) *For $A : n \times n$,* $\operatorname{rank}(A) = n$ *if and only if A is nonsingular.*

(viii) *If $A = \operatorname{diag}(a_i)$, then* $\operatorname{rank}(A)$ *equals the number of the $a_i \neq 0$.*

A.7 Range and Null Space

Definition A.24

(i) *The range* $\mathcal{R}(A)$ *of a matrix* $A : m \times n$ *is the vector space spanned by the column vectors of* A*, that is,*

$$\mathcal{R}(A) = \left\{ z : z = Ax = \sum_{i=1}^{n} a_{(i)} x_i , \quad x \in \mathbb{R}^n \right\} \subset \mathbb{R}^m ,$$

where $a_{(1)}, \ldots, a_{(n)}$ *are the column vectors of* A*.*

(ii) *The null space* $\mathcal{N}(A)$ *is the vector space defined by*

$$\mathcal{N}(A) = \{ x \in \mathbb{R}^n \quad \text{and} \quad Ax = 0 \} \subset \mathbb{R}^n .$$

Theorem A.25

(i) $\operatorname{rank}(A) = \dim \mathcal{R}(A)$, *where* $\dim \mathcal{V}$ *denotes the number of basis vectors of a vector space* $\mathcal{V}$*.*

(ii) $\dim \mathcal{R}(A) + \dim \mathcal{N}(A) = n$.

(iii) $\mathcal{N}(A) = \{\mathcal{R}(A')\}^{\perp}$. *(*$\mathcal{V}^{\perp}$ *is the orthogonal complement of a vector space* $\mathcal{V}$ *defined by* $\mathcal{V}^{\perp} = \{x : x'y = 0 \quad \forall\, y \in \mathcal{V}\}$*.)*

(iv) $\mathcal{R}(AA') = \mathcal{R}(A)$.

(v) $\mathcal{R}(AB) \subseteq \mathcal{R}(A)$ *for any* A *and* B*.*

(vi) *For* $A \geq 0$ *and any* B, $\mathcal{R}(BAB') = \mathcal{R}(BA)$.

A.8 Eigenvalues and Eigenvectors

Definition A.26 *If* $A : p \times p$ *is a square matrix, then*

$$q(\lambda) = |A - \lambda I|$$

is a pth order polynomial in λ*. The* p *roots* $\lambda_1, \ldots, \lambda_p$ *of the characteristic equation* $q(\lambda) = |A - \lambda I| = 0$ *are called eigenvalues or characteristic roots of* A*.*

The eigenvalues possibly may be complex numbers. Since $|A - \lambda_i I| = 0$, $A - \lambda_i I$ is a singular matrix. Hence, there exists a nonzero vector $\gamma_i \neq 0$ satisfying $(A - \lambda_i I)\gamma_i = 0$, that is,

$$A\gamma_i = \lambda_i \gamma_i .$$

γ_i is called the (right) eigenvector of A for the eigenvalue λ_i. If λ_i is complex, then γ_i may have complex components. An eigenvector γ with real components is called standardized if $\gamma'\gamma = 1$.

Theorem A.27

(i) *If x and y are nonzero eigenvectors of A for λ_i, and α and β are any real numbers, then $\alpha x + \beta y$ also is an eigenvector for λ_i, that is,*

$$A(\alpha x + \beta y) = \lambda_i(\alpha x + \beta y).$$

Thus the eigenvectors for any λ_i span a vector space, which is called the eigenspace of A for λ_i.

(ii) *The polynomial $q(\lambda) = |A - \lambda I|$ has the normal form in terms of the roots*

$$q(\lambda) = \prod_{i=1}^{p} (\lambda_i - \lambda).$$

Hence, $q(0) = \prod_{i=1}^{p} \lambda_i$ and

$$|A| = \prod_{i=1}^{p} \lambda_i.$$

(iii) *Matching the coefficients of λ^{n-1} in $q(\lambda) = \prod_{i=1}^{p}(\lambda_i - \lambda)$ and $|A - \lambda I|$ gives*

$$tr(A) = \sum_{i=1}^{p} \lambda_i.$$

(iv) *Let $C : p \times p$ be a regular matrix. Then A and CAC^{-1} have the same eigenvalues λ_i. If γ_i is an eigenvector for λ_i, then $C\gamma_i$ is an eigenvector of CAC^{-1} for λ_i.*

Proof: As C is nonsingular, it has an inverse C^{-1} with $CC^{-1} = I$. We have $|C^{-1}| = |C|^{-1}$ and

$$\begin{aligned} |A - \lambda I| &= |C||A - \lambda C^{-1}C||C^{-1}| \\ &= |CAC^{-1} - \lambda I|. \end{aligned}$$

Thus, A and CAC^{-1} have the same eigenvalues. Let $A\gamma_i = \lambda_i \gamma_i$, and multiply from the left by C:

$$CAC^{-1}C\gamma_i = (CAC^{-1})(C\gamma_i) = \lambda_i(C\gamma_i).$$

(v) *The matrix $A + \alpha I$ with α a real number has the eigenvalues $\tilde{\lambda}_i = \lambda_i + \alpha$, and the eigenvectors of A and $A + \alpha I$ coincide.*

(vi) *Let λ_1 denote any eigenvalue of $A : p \times p$ with eigenspace H of dimension r. If k denotes the multiplicity of λ_1 in $q(\lambda)$, then*

$$1 \leq r \leq k.$$

Remarks:

(a) For symmetric matrices A, we have $r = k$.

(b) If A is not symmetric, then it is possible that $r < k$. Example:

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A \neq A' \\ |A - \lambda I| &= \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2 = 0. \end{aligned}$$

The multiplicity of the eigenvalue $\lambda = 0$ is $k = 2$.

The eigenvectors for $\lambda = 0$ are $\gamma = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and generate an eigenspace of dimension 1.

(c) If for any particular eigenvalue λ, $\dim(H) = r = 1$, then the standardized eigenvector for λ is unique (up to the sign).

Theorem A.28 *Let $A : n \times p$ and $B : p \times n$ with $n \geq p$ be any two matrices. Then from Theorem A.16 (vii),*

$$\begin{vmatrix} -\lambda I_n & -A \\ B & I_p \end{vmatrix} = (-\lambda)^{n-p} |BA - \lambda I_p| = |AB - \lambda I_n|.$$

Hence the n eigenvalues of AB are equal to the p eigenvalues of BA plus the eigenvalue 0 with multiplicity $n - p$. Suppose that $x \neq 0$ is an eigenvector of AB for any particular $\lambda \neq 0$. Then $y = Bx$ is an eigenvector of BA for this λ and we have $y \neq 0$, too.

Corollary: A matrix $A = aa'$ with a as a nonnull vector has all eigenvalues 0 except one, with $\lambda = a'a$ and the corresponding eigenvector a.

Corollary: The nonzero eigenvalues of AA' are equal to the nonzero eigenvalues of $A'A$.

Theorem A.29 *If A is symmetric, then all the eigenvalues are real.*

A.9 Decomposition of Matrices

Theorem A.30 (Spectral decomposition theorem) *Any symmetric matrix $A : (p \times p)$ can be written as*

$$A = \Gamma \Lambda \Gamma' = \sum \lambda_i \gamma_{(i)} \gamma'_{(i)} ,$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$ is the diagonal matrix of the eigenvalues of A, and $\Gamma = (\gamma_{(1)}, \ldots, \gamma_{(p)})$ is the orthogonal matrix of the standardized eigenvectors $\gamma_{(i)}$.

Theorem A.31 *Suppose A is symmetric and $A = \Gamma \Lambda \Gamma'$. Then*

(i) *A and Λ have the same eigenvalues (with the same multiplicity).*

(ii) *From $A = \Gamma\Lambda\Gamma'$ we get $\Lambda = \Gamma' A \Gamma$.*

(iii) *If $A : p \times p$ is a symmetric matrix, then for any integer n, $A^n = \Gamma\Lambda^n\Gamma'$ and $\Lambda^n = \mathrm{diag}(\lambda_i^n)$. If the eigenvalues of A are positive, then we can define the rational powers*

$$A^{\frac{r}{s}} = \Gamma\Lambda^{\frac{r}{s}}\Gamma' \quad \textit{with} \quad \Lambda^{\frac{r}{s}} = \mathrm{diag}(\lambda_i^{\frac{r}{s}})$$

for integers $s > 0$ and r. Important special cases are (when $\lambda_i > 0$)

$$A^{-1} = \Gamma\Lambda^{-1}\Gamma' \quad \textit{with} \quad \Lambda^{-1} = \mathrm{diag}(\lambda_i^{-1});$$

the symmetric square root decomposition of A (when $\lambda_i \geq 0$)

$$A^{\frac{1}{2}} = \Gamma\Lambda^{\frac{1}{2}}\Gamma' \quad \textit{with} \quad \Lambda^{\frac{1}{2}} = \mathrm{diag}(\lambda_i^{\frac{1}{2}})$$

and if $\lambda_i > 0$

$$A^{-\frac{1}{2}} = \Gamma\Lambda^{-\frac{1}{2}}\Gamma' \quad \textit{with} \quad \Lambda^{-\frac{1}{2}} = \mathrm{diag}(\lambda_i^{-\frac{1}{2}}).$$

(iv) *For any square matrix A, the rank of A equals the number of nonzero eigenvalues.*

Proof: According to Theorem A.23 (vi) we have $\mathrm{rank}(A) = \mathrm{rank}(\Gamma\Lambda\Gamma')$ $= \mathrm{rank}(\Lambda)$. But $\mathrm{rank}(\Lambda)$ equals the number of nonzero λ_i's.

(v) *A symmetric matrix A is uniquely determined by its distinct eigenvalues and the corresponding eigenspaces. If the distinct eigenvalues λ_i are ordered as $\lambda_1 \geq \cdots \geq \lambda_p$, then the matrix Γ is unique (up to sign).*

(vi) *$A^{\frac{1}{2}}$ and A have the same eigenvectors. Hence, $A^{\frac{1}{2}}$ is unique.*

(vii) *Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$ be the nonzero eigenvalues and $\lambda_{k+1} = \cdots = \lambda_p = 0$. Then we have*

$$A = (\Gamma_1 \Gamma_2) \begin{pmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Gamma_1' \\ \Gamma_2' \end{pmatrix} = \Gamma_1 \Lambda_1 \Gamma_1'$$

with $\Lambda_1 = \mathrm{diag}(\lambda_1, \cdots, \lambda_k)$ and $\Gamma_1 = (\gamma_{(1)}, \cdots, \gamma_{(k)})$, whereas $\Gamma_1'\Gamma_1 = I_k$ holds so that Γ_1 is column-orthogonal.

(viii) *A symmetric matrix A is of rank 1 if and only if $A = aa'$ where $a \neq 0$.*

Proof: If $\mathrm{rank}(A) = \mathrm{rank}(\Lambda) = 1$, then $\Lambda = \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix}$, $A = \lambda\gamma\gamma' = aa'$ with $a = \sqrt{\lambda}\gamma$. If $A = aa'$, then by Theorem A.23 (v) we have $\mathrm{rank}(A) = \mathrm{rank}(a) = 1$.

Theorem A.32 (Singular-value decomposition of a rectangular matrix) *Let A : $n \times p$ be a rectangular matrix of rank r. Then we have*

$$\underset{n,p}{A} = \underset{n,r}{U} \underset{r,r}{L} \underset{r,p}{V'}$$

with $U'U = I_r$, $V'V = I_r$, and $L = \operatorname{diag}(l_1, \cdots, l_r)$, $l_i > 0$.

For a proof, see Rao (1973, p. 42).

Theorem A.33 *If $A : p \times q$ has $\operatorname{rank}(A) = r$, then A contains at least one nonsingular (r, r)-submatrix X, such that A has the so-called normal presentation*

$$\underset{p,q}{A} = \begin{pmatrix} \underset{r,r}{X} & \underset{r,q-r}{Y} \\ \underset{p-r,r}{Z} & \underset{p-r,q-r}{W} \end{pmatrix}.$$

All square submatrices of type $(r + s, r + s)$ with $(s \geq 1)$ are singular.

Proof: As $\operatorname{rank}(A) = \operatorname{rank}(X)$ holds, the first r rows of (X, Y) are linearly independent. Then the $p-r$ rows (Z, W) are linear combinations of (X, Y); that is, there exists a matrix F such that

$$(Z, W) = F(X, Y).$$

Analogously, there exists a matrix H satisfying

$$\begin{pmatrix} Y \\ W \end{pmatrix} = \begin{pmatrix} X \\ Z \end{pmatrix} H.$$

Hence we get $W = FY = FXH$, and

$$\begin{aligned} A = \begin{pmatrix} X & Y \\ Z & W \end{pmatrix} &= \begin{pmatrix} X & XH \\ FX & FXH \end{pmatrix} \\ &= \begin{pmatrix} I \\ F \end{pmatrix} X(I, H) \\ &= \begin{pmatrix} X \\ FX \end{pmatrix} (I, H) = \begin{pmatrix} I \\ F \end{pmatrix} (X, XH). \end{aligned}$$

As X is nonsingular, the inverse X^{-1} exists. Then we obtain $F = ZX^{-1}$, $H = X^{-1}Y$, $W = ZX^{-1}Y$, and

$$\begin{aligned} A = \begin{pmatrix} X & Y \\ Z & W \end{pmatrix} &= \begin{pmatrix} I \\ ZX^{-1} \end{pmatrix} X(I, X^{-1}Y) \\ &= \begin{pmatrix} X \\ Z \end{pmatrix} (I, X^{-1}Y) \\ &= \begin{pmatrix} I \\ ZX^{-1} \end{pmatrix} (X\, Y). \end{aligned}$$

Theorem A.34 (Full rank factorization)

(i) *If* $A : p \times q$ *has* $\operatorname{rank}(A) = r$, *then* A *may be written as*

$$\underset{p,q}{A} = \underset{p,r}{K}\,\underset{r,q}{L}$$

with K *of full column rank* r *and* L *of full row rank* r.

Proof: Theorem A.33.

(ii) *If* $A : p \times q$ *has* $\operatorname{rank}(A) = p$, *then* A *may be written as*

$$A = M(I, H)\,, \quad \textit{where } M : p \times p \textit{ is regular.}$$

Proof: Theorem A.34 (i).

A.10 Definite Matrices and Quadratic Forms

Definition A.35 *Suppose* $A : n \times n$ *is symmetric and* $x : n \times 1$ *is any vector. Then the quadratic form in* x *is defined as the function*

$$Q(x) = x'Ax = \sum_{i,j} a_{ij} x_i x_j \,.$$

Clearly, $Q(0) = 0$.

Definition A.36 *The matrix* A *is called positive definite (p.d.) if* $Q(x) > 0$ *for all* $x \neq 0$. *We write* $A > 0$.

Note: If $A > 0$, then $(-A)$ is called negative definite.

Definition A.37 *The quadratic form* $x'Ax$ *(and the matrix* A*, also) is called positive semidefinite (p.s.d.) if* $Q(x) \geq 0$ *for all* x *and* $Q(x) = 0$ *for at least one* $x \neq 0$.

Definition A.38 *The quadratic form* $x'Ax$ *(and* A*) is called nonnegative definite (n.n.d.) if it is either p.d. or p.s.d., that is, if* $x'Ax \geq 0$ *for all* x. *If* A *is n.n.d., we write* $A \geq 0$.

Theorem A.39 *Let the* $n \times n$ *matrix* $A > 0$. *Then*

(i) A *has all eigenvalues* $\lambda_i > 0$.

(ii) $x'Ax > 0$ *for any* $x \neq 0$.

(iii) A *is nonsingular and* $|A| > 0$.

(iv) $A^{-1} > 0$.

(v) $\operatorname{tr}(A) > 0$.

(vi) *Let* $P : n \times m$ *be of* $\operatorname{rank}(P) = m \leq n$. *Then* $P'AP > 0$ *and in particular* $P'P > 0$, *choosing* $A = I$.

(vii) *Let* $P : n \times m$ *be of* $\operatorname{rank}(P) < m \leq n$. *Then* $P'AP \geq 0$ *and* $P'P \geq 0$.

Theorem A.40 *Let* $A : n\times n$ *and* $B : n\times n$ *such that* $A > 0$ *and* $B : n\times n \geq 0$. *Then*

(i) $C = A + B > 0$.

(ii) $A^{-1} - (A + B)^{-1} \geq 0$.

(iii) $|A| \leq |A + B|$.

Theorem A.41 *Let* $A \geq 0$. *Then*

(i) $\lambda_i \geq 0$.

(ii) $\operatorname{tr}(A) \geq 0$.

(iii) $A = A^{\frac{1}{2}} A^{\frac{1}{2}}$ *with* $A^{\frac{1}{2}} = \Gamma \Lambda^{\frac{1}{2}} \Gamma'$.

(iv) *For any matrix* $C : n \times m$ *we have* $C'AC \geq 0$.

(v) *For any matrix* C, *we have* $C'C \geq 0$ *and* $CC' \geq 0$.

Theorem A.42 *For any matrix* $A \geq 0$, *we have* $0 \leq \lambda_i \leq 1$ *if and only if* $(I - A) \geq 0$.

Proof: Write the symmetric matrix A in its spectral form as $A = \Gamma\Lambda\Gamma'$. Then we have

$$(I - A) = \Gamma(I - \Lambda)\Gamma' \geq 0$$

if and only if

$$\Gamma'\Gamma(I - \Lambda)\Gamma'\Gamma = I - \Lambda \geq 0.$$

(a) If $I - \Lambda \geq 0$, then for the eigenvalues of $I - A$ we have $1 - \lambda_i \geq 0$ (i.e., $0 \leq \lambda_i \leq 1$).

(b) If $0 \leq \lambda_i \leq 1$, then for any $x \neq 0$,

$$x'(I - \Lambda)x = \sum x_i^2(1 - \lambda_i) \geq 0,$$

that is, $I - \Lambda \geq 0$.

Theorem A.43 (Theobald, 1974) *Let* $D : n \times n$ *be symmetric. Then* $D \geq 0$ *if and only if* $\operatorname{tr}\{CD\} \geq 0$ *for all* $C \geq 0$.

Proof: D is symmetric, so that

$$D = \Gamma\Lambda\Gamma' = \sum \lambda_i \gamma_i \gamma_i' \,,$$

and hence

$$\begin{aligned}\operatorname{tr}\{CD\} &= \operatorname{tr}\left\{\sum \lambda_i C\gamma_i\gamma_i'\right\} \\ &= \sum \lambda_i \gamma_i' C\gamma_i\,.\end{aligned}$$

(a) Let $D \geq 0$, and, hence, $\lambda_i \geq 0$ for all i. Then $\operatorname{tr}(CD) \geq 0$ if $C \geq 0$.

(b) Let $\operatorname{tr}\{CD\} \geq 0$ for all $C \geq 0$. Choose $C = \gamma_i\gamma_i' \quad (i = 1, \ldots, n,\ i \text{ fixed})$ so that

$$\begin{aligned}0 \leq \operatorname{tr}\{CD\} &= \operatorname{tr}\{\gamma_i\gamma_i'(\sum_j \lambda_j\gamma_j\gamma_j')\} \\ &= \lambda_i \qquad (i = 1, \cdots, n)\end{aligned}$$

and $D = \Gamma\Lambda\Gamma' \geq 0$.

Theorem A.44 *Let $A : n \times n$ be symmetric with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Then*

$$\sup_x \frac{x'Ax}{x'x} = \lambda_1, \quad \inf_x \frac{x'Ax}{x'x} = \lambda_n\,.$$

Proof: See Rao (1973, p. 62).

Theorem A.45 *Let $A : n \times r = (A_1, A_2)$, with A_1 of order $n \times r_1$, A_2 of order $n \times r_2$, and $\operatorname{rank}(A) = r = r_1 + r_2$. Define the orthogonal projectors $M_1 = A_1(A_1'A_1)^{-1}A_1'$ and $M = A(A'A)^{-1}A'$. Then*

$$M = M_1 + (I - M_1)A_2(A_2'(I - M_1)A_2)^{-1}A_2'(I - M_1).$$

Proof: M_1 and M are symmetric idempotent matrices fulfilling the conditions $M_1A_1 = 0$ and $MA = 0$. Using Theorem A.19 for partial inversion of $A'A$, that is,

$$(A'A)^{-1} = \begin{pmatrix} A_1'A_1 & A_1'A_2 \\ A_2'A_1 & A_2'A_2 \end{pmatrix}^{-1}$$

and using the special form of the matrix D defined in A.19, that is,

$$D = A_2'(I - M_1)A_2,$$

straightforward calculation concludes the proof.

Theorem A.46 *Let $A : n \times m$ with $\operatorname{rank}(A) = m \leq n$ and $B : m \times m$ be any symmetric matrix. Then*

$$ABA' \geq 0 \quad \textit{if and only if } B \geq 0.$$

Proof:

(a) $B \geq 0 \Rightarrow ABA' \geq 0$ for all A.

(b) Let $\operatorname{rank}(A) = m \le n$ and assume $ABA' \ge 0$, so that $x'ABA'x \ge 0$ for all $x \in \mathbb{R}^n$.
We have to prove that $y'By \ge 0$ for all $y \in \mathbb{R}^m$. As $\operatorname{rank}(A) = m$, the inverse $(A'A)^{-1}$ exists. Setting $z = A(A'A)^{-1}y$, we have $A'z = y$ and $y'By = z'ABA'z \ge 0$ so that $B \ge 0$.

Definition A.47 *Let $A : n \times n$ and $B : n \times n$ be any matrices. Then the roots $\lambda_i = \lambda_i^B(A)$ of the equation*

$$|A - \lambda B| = 0$$

are called the eigenvalues of A in the metric of B. For $B = I$ we obtain the usual eigenvalues defined in Definition A.26 (cf. Dhrymes, 1974, p. 581).

Theorem A.48 *Let $B > 0$ and $A \ge 0$. Then $\lambda_i^B(A) \ge 0$.*

Proof: $B > 0$ is equivalent to $B = B^{\frac{1}{2}}B^{\frac{1}{2}}$ with $B^{\frac{1}{2}}$ nonsingular and unique (A.31 (iii)). Then we may write

$$0 = |A - \lambda B| = |B^{\frac{1}{2}}|^2 |B^{-\frac{1}{2}}AB^{-\frac{1}{2}} - \lambda I|$$

and $\lambda_i^B(A) = \lambda_i^I(B^{-\frac{1}{2}}AB^{-\frac{1}{2}}) \ge 0$, as $B^{-\frac{1}{2}}AB^{-\frac{1}{2}} \ge 0$.

Theorem A.49 (Simultaneous diagonalization) *Let $B > 0$ and $A \ge 0$, and denote by $\Lambda = \operatorname{diag}\left(\lambda_i^B(A)\right)$ the diagonal matrix of the eigenvalues of A in the metric of B. Then there exists a nonsingular matrix W such that*

$$B = W'W \quad \textit{and} \quad A = W'\Lambda W.$$

Proof: From the proof of Theorem A.48 we know that the roots $\lambda_i^B(A)$ are the usual eigenvalues of the matrix $B^{-\frac{1}{2}}AB^{-\frac{1}{2}}$. Let X be the matrix of the corresponding eigenvectors:

$$B^{-\frac{1}{2}}AB^{-\frac{1}{2}}X = X\Lambda,$$

that is,

$$A = B^{\frac{1}{2}}X\Lambda X'B^{\frac{1}{2}} = W'\Lambda W$$

with $W' = B^{\frac{1}{2}}X$ regular and

$$B = W'W = B^{\frac{1}{2}}XX'B^{\frac{1}{2}} = B^{\frac{1}{2}}B^{\frac{1}{2}}.$$

Theorem A.50 *Let $A > 0$ (or $A \ge 0$) and $B > 0$. Then*

$$B - A > 0 \quad \textit{if and only if} \quad \lambda_i^B(A) < 1.$$

Proof: Using Theorem A.49, we may write

$$B - A = W'(I - \Lambda)W,$$

namely,

$$x'(B - A)x \quad = \quad x'W'(I - \Lambda)Wx$$

$$\begin{aligned} &= y'(I - \Lambda)y \\ &= \sum \left(1 - \lambda_i^B(A)\right) y_i^2 \end{aligned}$$

with $y = Wx$, W regular, and hence $y \neq 0$ for $x \neq 0$. Then $x'(B - A)x > 0$ holds if and only if

$$\lambda_i^B(A) < 1.$$

Theorem A.51 *Let $A > 0$ (or $A \geq 0$) and $B > 0$. Then*

$$A - B \geq 0$$

if and only if

$$\lambda_i^B(A) \leq 1\,.$$

Proof: Similar to Theorem A.50.

Theorem A.52 *Let $A > 0$ and $B > 0$. Then*

$$B - A > 0 \quad \textit{if and only if} \quad A^{-1} - B^{-1} > 0.$$

Proof: From Theorem A.49 we have

$$B = W'W, \quad A = W'\Lambda W.$$

Since W is regular, we have

$$B^{-1} = W^{-1}W'^{-1}, \quad A^{-1} = W^{-1}\Lambda^{-1}W'^{-1},$$

that is,

$$A^{-1} - B^{-1} = W^{-1}(\Lambda^{-1} - I)W'^{-1} > 0,$$

as $\lambda_i^B(A) < 1$ and, hence, $\Lambda^{-1} - I > 0$.

Theorem A.53 *Let $B - A > 0$. Then $|B| > |A|$ and $\mathrm{tr}(B) > \mathrm{tr}(A)$. If $B - A \geq 0$, then $|B| \geq |A|$ and $\mathrm{tr}(B) \geq \mathrm{tr}(A)$.*

Proof: From Theorems A.49 and A.16 (iii), (v), we get

$$\begin{aligned} |B| &= |W'W| = |W|^2, \\ |A| &= |W'\Lambda W| = |W|^2 |\Lambda| = |W|^2 \prod \lambda_i^B(A), \end{aligned}$$

that is,

$$|A| = |B| \prod \lambda_i^B(A).$$

For $B - A > 0$, we have $\lambda_i^B(A) < 1$ (i.e., $|A| < |B|$). For $B - A \geq 0$, we have $\lambda_i^B(A) \leq 1$ (i.e., $|A| \leq |B|$). $B - A > 0$ implies $\mathrm{tr}(B - A) > 0$, and $\mathrm{tr}(B) > \mathrm{tr}(A)$. Analogously, $B - A \geq 0$ implies $\mathrm{tr}(B) \geq \mathrm{tr}(A)$.

Theorem A.54 (Cauchy-Schwarz inequality) *Let x, y be real vectors of the same dimension. Then*

$$(x'y)^2 \leq (x'x)(y'y),$$

with equality if and only if x and y are linearly dependent.

Theorem A.55 *Let x, y be real vectors and $A > 0$. Then we have the following results:*

(i) $(x'Ay)^2 \leq (x'Ax)(y'Ay)$.

(ii) $(x'y)^2 \leq (x'Ax)(y'A^{-1}y)$.

Proof:

(a) $A \geq 0$ is equivalent to $A = BB$ with $B = A^{\frac{1}{2}}$ (Theorem A.41 (iii)). Let $Bx = \tilde{x}$ and $By = \tilde{y}$. Then (i) is a consequence of Theorem A.54.

(b) $A > 0$ is equivalent to $A = A^{\frac{1}{2}}A^{\frac{1}{2}}$ and $A^{-1} = A^{-\frac{1}{2}}A^{-\frac{1}{2}}$. Let $A^{\frac{1}{2}}x = \tilde{x}$ and $A^{-\frac{1}{2}}y = \tilde{y}$; then (ii) is a consequence of Theorem A.54.

Theorem A.56 *Let $A > 0$ and T be any square matrix. Then*

(i) $\sup_{x \neq 0} \frac{(x'y)^2}{x'Ax} = y'A^{-1}y$.

(ii) $\sup_{x \neq 0} \frac{(y'Tx)^2}{x'Ax} = y'TA^{-1}T'y$.

Proof: Use Theorem A.55 (ii).

Theorem A.57 *Let $I : n \times n$ be the identity matrix and let a be an n-vector. Then*

$$I - aa' \geq 0 \quad \textit{if and only if} \quad a'a \leq 1.$$

Proof: The matrix aa' is of rank 1 and $aa' \geq 0$. The spectral decomposition is $aa' = C\Lambda C'$ with $\Lambda = \mathrm{diag}(\lambda, 0, \cdots, 0)$ and $\lambda = a'a$. Hence, $I - aa' = C(I - \Lambda)C' \geq 0$ if and only if $\lambda = a'a \leq 1$ (see Theorem A.42).

Theorem A.58 *Assume $MM' - NN' \geq 0$. Then there exists a matrix H such that $N = MH$.*

Proof (Milliken and Akdeniz, 1977) : Let M (n, r) of $\mathrm{rank}(M) = s$, and let x be any vector $\in \mathcal{R}(I - MM^-)$, implying $x'M = 0$ and $x'MM'x = 0$. As NN' and $MM' - NN'$ (by assumption) are n.n.d., we may conclude that $x'NN'x \geq 0$ and

$$x'(MM' - NN')x = -x'NN'x \geq 0,$$

so that $x'NN'x = 0$ and $x'N = 0$. Hence, $N \subset \mathcal{R}(M)$ or, equivalently, $N = MH$ for some matrix H (r, k).

Theorem A.59 *Let A be an $n \times n$-matrix and assume $(-A) > 0$. Let a be an n-vector. In case $n \geq 2$, the matrix $A + aa'$ is never n.n.d.*

Proof (Guilkey and Price, 1981) : The matrix aa' is of rank ≤ 1. In case $n \geq 2$, there exists a nonzero vector w such that $w'aa'w = 0$, implying $w'(A + aa')w = w'Aw < 0$.

A.11 Idempotent Matrices

Definition A.60 *A square matrix A is called idempotent if it satisfies*

$$A^2 = AA = A\,.$$

An idempotent matrix A is called an orthogonal projector if $A = A'$. Otherwise, A is called an oblique projector.

Theorem A.61 *Let $A : n \times n$ be idempotent with* $\text{rank}(A) = r \leq n$. *Then we have:*

(i) *The eigenvalues of A are 1 or 0.*

(ii) $\text{tr}(A) = \text{rank}(A) = r$.

(iii) *If A is of full rank n, then $A = I_n$.*

(iv) *If A and B are idempotent and if $AB = BA$, then AB is also idempotent.*

(v) *If A is idempotent and P is orthogonal, then PAP' is also idempotent.*

(vi) *If A is idempotent, then $I - A$ is idempotent and*

$$A(I - A) = (I - A)A = 0.$$

Proof:

(a) The characteristic equation

$$Ax = \lambda x$$

multiplied by A gives

$$AAx = Ax = \lambda Ax = \lambda^2 x.$$

Multiplication of both equations by x' then yields

$$x'Ax = \lambda x'x = \lambda^2 x'x,$$

that is,

$$\lambda(\lambda - 1) = 0.$$

(b) From the spectral decomposition

$$A = \Gamma \Lambda \Gamma'\,,$$

we obtain

$$\text{rank}(A) = \text{rank}(\Lambda) = \text{tr}(\Lambda) = r,$$

where r is the number of characteristic roots with value 1.

(c) Let $\operatorname{rank}(A) = \operatorname{rank}(\Lambda) = n$, then $\Lambda = I_n$ and

$$A = \Gamma\Lambda\Gamma' = I_n \,.$$

(a)–(c) follow from the definition of an idempotent matrix.

A.12 Generalized Inverse

Definition A.62 *Let A be an $m \times n$-matrix. Then a matrix $A^- : n \times m$ is said to be a generalized inverse of A if*

$$AA^-A = A$$

holds (see Rao (1973), p. 24).

Theorem A.63 *A generalized inverse always exists although it is not unique in general.*

Proof: Assume $\operatorname{rank}(A) = r$. According to the singular-value decomposition (Theorem A.32), we have

$$\underset{m,n}{A} = \underset{m,r}{U}\,\underset{r,r}{L}\,\underset{r,n}{V'}$$

with $U'U = I_r$ and $V'V = I_r$ and

$$L = \operatorname{diag}(l_1, \cdots, l_r), \quad l_i > 0.$$

Then

$$A^- = V \begin{pmatrix} L^{-1} & X \\ Y & Z \end{pmatrix} U'$$

(X, Y and Z are arbitrary matrices of suitable dimensions) is a g-inverse of A. Using Theorem A.33, namely,

$$A = \begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$$

with X nonsingular, we have

$$A^- = \begin{pmatrix} X^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

as a special g-inverse.

Definition A.64 (Moore-Penrose inverse) *A matrix A^+ satisfying the following conditions is called the Moore-Penrose inverse of A:*

(i) $AA^+A = A\,,$

(ii) $A^+AA^+ = A^+\,,$

(iii) $(A^+A)' = A^+A$,

(iv) $(AA^+)' = AA^+$.

A^+ is unique.

Theorem A.65 *For any matrix $A : m \times n$ and any g-inverse $A^- : m \times n$, we have*

(i) *A^-A and AA^- are idempotent.*

(ii) $\text{rank}(A) = \text{rank}(AA^-) = \text{rank}(A^-A)$.

(iii) $\text{rank}(A) \leq \text{rank}(A^-)$.

Proof:

(a) Using the definition of g-inverse,

$$(A^-A)(A^-A) = A^-(AA^-A) = A^-A.$$

(b) According to Theorem A.23 (iv), we get

$$\text{rank}(A) = \text{rank}(AA^-A) \leq \text{rank}(A^-A) \leq \text{rank}(A),$$

that is, $\text{rank}(A^-A) = \text{rank}(A)$. Analogously, we see that $\text{rank}(A) = \text{rank}(AA^-)$.

(c) $\text{rank}(A) = \text{rank}(AA^-A) \leq \text{rank}(AA^-) \leq \text{rank}(A^-)$.

Theorem A.66 *Let A be an $m \times n$-matrix. Then*

(i) *A regular $\Rightarrow A^+ = A^{-1}$.*

(ii) $(A^+)^+ = A$.

(iii) $(A^+)' = (A')^+$.

(iv) $\text{rank}(A) = \text{rank}(A^+) = \text{rank}(A^+A) = \text{rank}(AA^+)$.

(v) *A an orthogonal projector $\Rightarrow A^+ = A$.*

(vi) $\text{rank}(A) : m \times n = m \Rightarrow A^+ = A'(AA')^{-1}$ *and* $AA^+ = I_m$.

(vii) $\text{rank}(A) : m \times n = n \Rightarrow A^+ = (A'A)^{-1}A'$ *and* $A^+A = I_n$.

(viii) *If $P : m \times m$ and $Q : n \times n$ are orthogonal $\Rightarrow$ $(PAQ)^+ = Q^{-1}A^+P^{-1}$.*

(ix) $(A'A)^+ = A^+(A')^+$ *and* $(AA')^+ = (A')^+A^+$.

(x) $A^+ = (A'A)^+A' = A'(AA')^+$.

For further details see Rao and Mitra (1971).

Theorem A.67 (Baksalary, Kala, and Klaczynski (1983)) *Let $M : n \times n \geq 0$ and $N : m \times n$ be any matrices. Then*

$$M - N'(NM^+N')^+N \geq 0$$

if and only if

$$\mathcal{R}(N'NM) \subset \mathcal{R}(M).$$

Theorem A.68 *Let A be any square $n \times n$-matrix and a be an n-vector with $a \notin \mathcal{R}(A)$. Then a g-inverse of $A + aa'$ is given by*

$$\begin{aligned}(A + aa')^- &= A^- - \frac{A^-aa'U'U}{a'U'Ua} \\ &\quad - \frac{VV'aa'A^-}{a'VV'a} + \phi\frac{VV'aa'U'U}{(a'U'Ua)(a'VV'a)},\end{aligned}$$

with A^- any g-inverse of A and

$$\phi = 1 + a'A^-a, \quad U = I - AA^-, \quad V = I - A^-A.$$

Proof: Straightforward by checking $AA^-A = A$.

Theorem A.69 *Let A be a square $n \times n$-matrix. Then we have the following results:*

(i) *Assume a, b are vectors with $a, b \in \mathcal{R}(A)$, and let A be symmetric. Then the bilinear form $a'A^-b$ is invariant to the choice of A^-.*

(ii) *$A(A'A)^-A'$ is invariant to the choice of $(A'A)^-$.*

Proof:

(a) $a, b \in \mathcal{R}(A) \Rightarrow a = Ac$ and $b = Ad$. Using the symmetry of A gives

$$\begin{aligned}a'A^-b &= c'A'A^-Ad \\ &= c'Ad.\end{aligned}$$

(b) Using the rowwise representation of A as $A = \begin{pmatrix} a_1' \\ \vdots \\ a_n' \end{pmatrix}$ gives

$$A(A'A)^-A' = (a_i'(A'A)^-a_j).$$

Since $A'A$ is symmetric, we may conclude then (i) that all bilinear forms $a_i'(A'A)a_j$ are invariant to the choice of $(A'A)^-$, and hence (ii) is proved.

Theorem A.70 *Let $A : n \times n$ be symmetric, $a \in \mathcal{R}(A)$, $b \in \mathcal{R}(A)$, and assume $1 + b'A^+a \neq 0$. Then*

$$(A + ab')^+ = A^+ - \frac{A^+ab'A^+}{1 + b'A^+a}.$$

Proof: Straightforward, using Theorems A.68 and A.69.

Theorem A.71 *Let $A : n \times n$ be symmetric, a be an n-vector, and $\alpha > 0$ be any scalar. Then the following statements are equivalent:*

(i) $\alpha A - aa' \geq 0$.

(ii) $A \geq 0$, $a \in \mathcal{R}(A)$, *and* $a'A^-a \leq \alpha$, *with* A^- *being any g-inverse of* A.

Proof:

(i) $\Rightarrow$ (ii): $\alpha A - aa' \geq 0 \Rightarrow \alpha A = (\alpha A - aa') + aa' \geq 0 \Rightarrow A \geq 0$. Using Theorem A.31 for $\alpha A - aa' \geq 0$, we have $\alpha A - aa' = BB$, and, hence,

$$\begin{aligned} \alpha A = BB + aa' &= (B, a)(B, a)'. \\ \Rightarrow \quad & \mathcal{R}(\alpha A) = \mathcal{R}(A) = \mathcal{R}(B, a) \\ \Rightarrow \quad & a \in \mathcal{R}(A) \\ \Rightarrow \quad & a = Ac \quad \text{with} \quad c \in \mathbb{R}^n \\ \Rightarrow \quad & a'A^-a = c'Ac. \end{aligned}$$

As $\alpha A - aa' \geq 0 \quad \Rightarrow$

$$x'(\alpha A - aa')x \geq 0$$

for any vector x, choosing $x = c$, we have

$$\begin{aligned} \alpha c'Ac - c'aa'c &= \alpha c'Ac - (c'Ac)^2 \geq 0, \\ &\Rightarrow c'Ac \leq \alpha. \end{aligned}$$

(ii) $\Rightarrow$ (i): Let $x \in \mathbb{R}^n$ be any vector. Then, using Theorem A.54,

$$\begin{aligned} x'(\alpha A - aa')x &= \alpha x'Ax - (x'a)^2 \\ &= \alpha x'Ax - (x'Ac)^2 \\ &\geq \alpha x'Ax - (x'Ax)(c'Ac) \end{aligned}$$

$$\Rightarrow \quad x'(\alpha A - aa')x \geq (x'Ax)(\alpha - c'Ac).$$

In (ii) we have assumed $A \geq 0$ and $c'Ac = a'A^-a \leq \alpha$. Hence, $\alpha A - aa' \geq 0$.

Note: This theorem is due to Baksalary and Kala (1983). The version given here and the proof are formulated by G. Trenkler.

Theorem A.72 *For any matrix A we have*

$$A'A = 0 \quad \textit{if and only if } A = 0.$$

Proof:

(a) $A = 0 \Rightarrow A'A = 0$.

(b) Let $A'A = 0$, and let $A = (a_{(1)}, \cdots, a_{(n)})$ be the columnwise presentation. Then

$$A'A = (a'_{(i)} a_{(j)}) = 0,$$

so that all the elements on the diagonal are zero: $a'_{(i)} a_{(i)} = 0 \Rightarrow a_{(i)} = 0$ and $A = 0$.

Theorem A.73 *Let $X \neq 0$ be an $m \times n$-matrix and A an $n \times n$-matrix. Then*

$$X'XAX'X = X'X \quad \Rightarrow \quad XAX'X = X \quad \textit{and} \quad X'XAX' = X'.$$

Proof: As $X \neq 0$ and $X'X \neq 0$, we have

$$\begin{aligned} X'XAX'X - X'X = (X'XA - I)X'X &= 0 \\ \Rightarrow (X'XA - I) &= 0 \end{aligned}$$

$$\begin{aligned} \Rightarrow 0 &= (X'XA - I)(X'XAX'X - X'X) \\ &= (X'XAX' - X')(XAX'X - X) = Y'Y, \end{aligned}$$

so that (by Theorem A.72) $Y = 0$, and, hence $XAX'X = X$.

Corollary: Let $X \neq 0$ be an $m \times n$-matrix and A and b $n \times n$-matrices. Then

$$AX'X = BX'X \quad \Leftrightarrow \quad AX' = BX'.$$

Theorem A.74 (Albert's theorem)
Let $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ be symmetric. Then

(i) *$A \geq 0$ if and only if*

(a) $A_{22} \geq 0$,
(b) $A_{21} = A_{22} A_{22}^- A_{21}$,
(c) $A_{11} \geq A_{12} A_{22}^- A_{21}$,

((b) and (c) are invariant of the choice of A_{22}^-).

(ii) *$A > 0$ if and only if*

(a) $A_{22} > 0$,
(b) $A_{11} > A_{12} A_{22}^{-1} A_{21}$.

Proof (Bekker and Neudecker, 1989) :

(i) Assume $A \geq 0$.

(a) $A \geq 0 \quad \Rightarrow \quad x'Ax \geq 0$ for any x. Choosing $x' = (0', x_2')$ $\Rightarrow x'Ax = x_2' A_{22} x_2 \geq 0$ for any $x_2 \Rightarrow A_{22} \geq 0$.

(b) Let $B' = (0, I - A_{22} A_{22}^-) \Rightarrow$

$$\begin{aligned} B'A &= \big((I - A_{22}A_{22}^-)A_{21}, A_{22} - A_{22}A_{22}^-A_{22}\big) \\ &= \big((I - A_{22}A_{22}^-)A_{21}, 0\big) \end{aligned}$$

and $B'AB = B'A^{\frac{1}{2}}A^{\frac{1}{2}}B = 0$. Hence, by Theorem A.72 we get $B'A^{\frac{1}{2}} = 0$.

$$\begin{aligned} \Rightarrow B'A^{\frac{1}{2}}A^{\frac{1}{2}} = B'A &= 0. \\ \Rightarrow (I - A_{22}A_{22}^{-})A_{21} &= 0. \end{aligned}$$

This proves (b).

(c) Let $C' = (I, -(A_{22}^{-}A_{21})')$. $A \geq 0 \Rightarrow$

$$\begin{aligned} 0 \leq C'AC &= A_{11} - A_{12}(A_{22}^{-})'A_{21} - A_{12}A_{22}^{-}A_{21} \\ &\quad + A_{12}(A_{22}^{-})'A_{22}A_{22}^{-}A_{21} \\ &= A_{11} - A_{12}A_{22}^{-}A_{21}\,. \end{aligned}$$

(Since A_{22} is symmetric, we have $(A_{22}^{-})' = A_{22}$.)

Now assume (a), (b), and (c). Then

$$D = \begin{pmatrix} A_{11} - A_{12}A_{22}^{-}A_{21} & 0 \\ 0 & A_{22} \end{pmatrix} \geq 0,$$

as the submatrices are n.n.d. by (a) and (b). Hence,

$$A = \begin{pmatrix} I & A_{12}(A_{22}^{-}) \\ 0 & I \end{pmatrix} D \begin{pmatrix} I & 0 \\ A_{22}^{-}A_{21} & I \end{pmatrix} \geq 0.$$

(ii) Proof as in (i) if A_{22}^{-} is replaced by A_{22}^{-1}.

Theorem A.75 *If $A : n \times n$ and $B : n \times n$ are symmetric, then*

(i) $0 \leq B \leq A$ *if and only if*

(a) $A \geq 0$,
(b) $B = AA^{-}B$,
(c) $B \geq BA^{-}B$.

(ii) $0 < B < A$ *if and only if* $0 < A^{-1} < B^{-1}$.

Proof: Apply Theorem A.74 to the matrix $\begin{pmatrix} B & B \\ B & A \end{pmatrix}$.

Theorem A.76 *Let A be symmetric and $c \in \mathcal{R}(A)$. Then the following statements are equivalent:*

(i) $\operatorname{rank}(A + cc') = \operatorname{rank}(A)$.

(ii) $\mathcal{R}(A + cc') = \mathcal{R}(A)$.

(iii) $1 + c'A^{-}c \neq 0$.

Corollary 1: Assume (i) or (ii) or (iii) holds; then

$$(A + cc')^{-} = A^{-} - \frac{A^{-}cc'A^{-}}{1 + c'A^{-}c}$$

for any choice of A^{-}.

Corollary 2: Assume (i) or (ii) or (iii) holds; then

$$\begin{aligned} c'(A+cc')^- c &= c'A^- c - \frac{(c'A^- c)^2}{1+c'A^- c} \\ &= 1 - \frac{1}{1+c'A^- c}\,. \end{aligned}$$

Moreover, as $c \in \mathcal{R}(A+cc')$, the results are invariant for any special choices of the g-inverses involved.

Proof: $c \in \mathcal{R}(A) \Leftrightarrow AA^- c = c \Rightarrow$

$$\mathcal{R}(A+cc') = \mathcal{R}(AA^-(A+cc')) \subset \mathcal{R}(A).$$

Hence, (i) and (ii) become equivalent. Proof of (iii): Consider the following product of matrices:

$$\begin{pmatrix} 1 & 0 \\ c & A+cc' \end{pmatrix} \begin{pmatrix} 1 & -c \\ 0 & I \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -A^- c & I \end{pmatrix} = \begin{pmatrix} 1+c'A^- c & -c \\ 0 & A \end{pmatrix}.$$

The left-hand side has the rank

$$1 + \operatorname{rank}(A+cc') = 1 + \operatorname{rank}(A)$$

(see (i) or (ii)). The right-hand side has the rank $1 + \operatorname{rank}(A)$ if and only if $1 + c'A^- c \neq 0$.

Theorem A.77 *Let $A : n \times n$ be a symmetric and nonsingular matrix and $c \notin \mathcal{R}(A)$. Then we have*

(i) $c \in \mathcal{R}(A+cc')$.

(ii) $\mathcal{R}(A) \subset \mathcal{R}(A+cc')$.

(iii) $c'(A+cc')^- c = 1$.

(iv) $A(A+cc')^- A = A$.

(v) $A(A+cc')^- c = 0$.

Proof: As A is assumed to be nonsingular, the equation $Al = 0$ has a nontrivial solution $l \neq 0$, which may be standardized as $(c'l)^{-1} l$ such that $c'l = 1$. Then we have $c = (A+cc')l \in \mathcal{R}(A+cc')$, and hence (i) is proved. Relation (ii) holds as $c \notin \mathcal{R}(A)$. Relation (i) is seen to be equivalent to

$$(A+cc')(A+cc')^- c = c.$$

Then (iii) follows:

$$\begin{aligned} c'(A+cc')^- c &= l'(A+cc')(A+cc')^- c \\ &= l'c = 1\,, \end{aligned}$$

which proves (iii). From

$$\begin{aligned} c &= (A+cc')(A+cc')^{-}c \\ &= A(A+cc')^{-}c + cc'(A+cc')^{-}c \\ &= A(A+cc')^{-}c + c\,, \end{aligned}$$

we have (v).

(iv) is a consequence of the general definition of a g-inverse and of (iii) and (iv):

$$\begin{aligned} A+cc' &= (A+cc')(A+cc')^{-}(A+cc') \\ &= A(A+cc')^{-}A \\ &\quad + cc'(A+cc')^{-}cc' \quad [= cc' \text{ using (iii)}] \\ &\quad + A(A+cc')^{-}cc' \quad [= 0 \text{ using (v)}] \\ &\quad + cc'(A+cc')^{-}A \quad [= 0 \text{ using (v)}]. \end{aligned}$$

Theorem A.78 *We have $A \geq 0$ if and only if*

(i) $A + cc' \geq 0$.

(ii) $(A+cc')(A+cc')^{-}c = c$.

(iii) $c'(A+cc')^{-}c \leq 1$.

Assume $A \geq 0$; then

(a) $c = 0 \Leftrightarrow c'(A+cc')^{-}c = 0$.

(b) $c \in \mathcal{R}(A) \Leftrightarrow c'(A+cc')^{-}c < 1$.

(c) $c \notin \mathcal{R}(A) \Leftrightarrow c'(A+cc')^{-}c = 1$.

Proof: $A \geq 0$ is equivalent to

$$0 \leq cc' \leq A + cc'.$$

Straightforward application of Theorem A.75 gives (i)–(iii).

Proof of (a): $A \geq 0 \Rightarrow A + cc' \geq 0$. Assume

$$c'(A+cc')^{-}c = 0\,,$$

and replace c by (ii) $\Rightarrow$

$$\begin{aligned} &c'(A+cc')^{-}(A+cc')(A+cc')^{-}c = 0 \Rightarrow \\ &(A+cc')(A+cc')^{-}c = 0 \end{aligned}$$

as $(A+cc') \geq 0$. Assuming $c = 0 \Rightarrow c'(A+cc')c = 0$.

Proof of (b): Assume $A \geq 0$ and $c \in \mathcal{R}(A)$, and use Theorem A.76 (Corollary 2) $\Rightarrow$

$$c'(A+cc')^{-}c = 1 - \frac{1}{1 + c'A^{-}c} < 1.$$

The opposite direction of (b) is a consequence of (c).

Proof of (c): Assume $A \geq 0$ and $c \notin \mathcal{R}(A)$, and use Theorem A.77 (iii) $\Rightarrow$

$$c'(A + cc')^- c = 1.$$

The opposite direction of (c) is a consequence of (b).

Note: The proofs of Theorems A.74–A.78 are given in Bekker and Neudecker (1989).

Theorem A.79 *The linear equation $Ax = a$ has a solution if and only if*

$$a \in \mathcal{R}(A) \quad or \quad AA^- a = a$$

for any g-inverse A.

If this condition holds, then all solutions are given by

$$x = A^- a + (I - A^- A)w\,,$$

where w is an arbitrary m-vector. Further, $q'x$ has a unique value for all solutions of $Ax = a$ if and only if $q'A^- A = q'$, or $q \in \mathcal{R}(A')$.

For a proof, see Rao (1973, p. 25).

A.13 Projectors

Consider the range space $\mathcal{R}(A)$ of the matrix $A : m \times n$ with rank r. Then there exists $\mathcal{R}(A)^{\perp}$, which is the orthogonal complement of $\mathcal{R}(A)$ with dimension $m - r$. Any vector $x \in \mathbb{R}^m$ has the unique decomposition

$$x = x_1 + x_2\,, \quad x_1 \in \mathcal{R}(A)\,, \quad \text{and} \quad x_2 \in \mathcal{R}(A)^{\perp}\,,$$

of which the component x_1 is called the orthogonal projection of x on $\mathcal{R}(A)$. The component x_1 can be computed as Px, where

$$P = A(A'A)^- A'\,,$$

which is called the projection operator on $\mathcal{R}(A)$. Note that P is unique for any choice of the g-inverse $(A'A)^-$.

Theorem A.80 *For any $P : n \times n$, the following statements are equivalent:*

(i) *P is an orthogonal projection operator.*

(ii) *P is symmetric and idempotent.*

For proofs and other details, the reader is referred to Rao (1973) and Rao and Mitra (1971).

Theorem A.81 *Let X be a matrix of order $T \times K$ with rank $r < K$, and $U : (K - r) \times K$ be such that $\mathcal{R}(X') \cap \mathcal{R}(U') = \{0\}$. Then*

(i) $X(X'X + U'U)^{-1}U' = 0$.

(ii) $X'X(X'X + U'U)^{-1}X'X = X'X$; *that is,* $(X'X + U'U)^{-1}$ *is a g-inverse of* $X'X$.

(iii) $U'U(X'X + U'U)^{-1}U'U = U'U$; *that is,* $(X'X + U'U)^{-1}$ *is also a g-inverse of* $U'U$.

(iv) $U(X'X + U'U)^{-1}U'u = u$ *if* $u \in \mathcal{R}(U)$.

Proof: Since $X'X + U'U$ is of full rank, there exists a matrix A such that

$$\begin{aligned} & (X'X + U'U)A = U' \\ \Rightarrow \quad & X'XA = U' - U'UA \quad \Rightarrow \quad XA = 0 \text{ and } U' = U'UA \end{aligned}$$

since $\mathcal{R}(X')$ and $\mathcal{R}(U')$ are disjoint.

Proof of (i):

$$X(X'X + U'U)^{-1}U' = X(X'X + U'U)^{-1}(X'X + U'U)A = XA = 0\,.$$

Proof of (ii):

$$\begin{aligned} & X'X(X'X + U'U)^{-1}(X'X + U'U - U'U) \\ = \quad & X'X - X'X(X'X + U'U)^{-1}U'U = X'X\,. \end{aligned}$$

Result (iii) follows on the same lines as result (ii).

Proof of (iv):

$$U(X'X + U'U)^{-1}U'u = U(X'X + U'U)^{-1}U'Ua = Ua = u$$

since $u \in \mathcal{R}(U)$.

A.14 Functions of Normally Distributed Variables

Let $x' = (x_1, \cdots, x_p)$ be a p-dimensional random vector. Then x is said to have a p-dimensional normal distribution with expectation vector μ and covariance matrix $\Sigma > 0$ if the joint density is

$$f(x; \mu, \Sigma) = \{(2\pi)^p|\Sigma|\}^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x - \mu)'\Sigma^{-1}(x - \mu)\right\}.$$

In such a case we write $x \sim N_p(\mu, \Sigma)$.

Theorem A.82 *Assume* $x \sim N_p(\mu, \Sigma)$, *and* $A : p \times p$ *and* $b : p \times 1$ *nonstochastic. Then*

$$y = Ax + b \sim N_q(A\mu + b, A\Sigma A') \quad \textit{with}\ q = \operatorname{rank}(A).$$

Theorem A.83 *If* $x \sim N_p(0, I)$, *then*

$$x'x \sim \chi^2_p$$

(central χ^2*-distribution with* p *degrees of freedom).*

Theorem A.84 *If* $x \sim N_p(\mu, I)$, *then*

$$x'x \sim \chi^2_p(\lambda)$$

has a noncentral χ^2*-distribution with noncentrality parameter*

$$\lambda = \mu'\mu = \sum_{i=1}^{p} \mu_i^2.$$

Theorem A.85 *If* $x \sim N_p(\mu, \Sigma)$, *then*

(i) $x'\Sigma^{-1}x \sim \chi^2_p(\mu'\Sigma^{-1}\mu)$.

(ii) $(x-\mu)'\Sigma^{-1}(x-\mu) \sim \chi^2_p$.

Proof: $\Sigma > 0 \quad \Rightarrow \quad \Sigma = \Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}}$ with $\Sigma^{\frac{1}{2}}$ regular and symmetric. Hence, $\Sigma^{-\frac{1}{2}}x = y \sim N_p(\Sigma^{-\frac{1}{2}}\mu, I) \Rightarrow$

$$x'\Sigma^{-1}x = y'y \sim \chi^2_p(\mu'\Sigma^{-1}\mu)$$

and

$$(x-\mu)'\Sigma^{-1}(x-\mu) = (y - \Sigma^{-\frac{1}{2}}\mu)'(y - \Sigma^{-\frac{1}{2}}\mu) \sim \chi^2_p.$$

Theorem A.86 *If* $Q_1 \sim \chi^2_m(\lambda)$ *and* $Q_2 \sim \chi^2_n$, *and* Q_1 *and* Q_2 *are independent, then*

(i) *The ratio*

$$F = \frac{Q_1/m}{Q_2/n}$$

has a noncentral $F_{m,n}(\lambda)$*-distribution.*

(ii) *If* $\lambda = 0$, *then* $F \sim F_{m,n}$ *(the central* F*-distribution).*

(iii) *If* $m = 1$, *then* $\sqrt{F}$ *has a noncentral* $t_n(\sqrt{\lambda})$*-distribution or a central* t_n*-distribution if* $\lambda = 0$.

Theorem A.87 *If* $x \sim N_p(\mu, I)$ *and* $A : p \times p$ *is a symmetric, idempotent matrix with* $\operatorname{rank}(A) = r$, *then*

$$x'Ax \sim \chi^2_r(\mu'A\mu).$$

Proof: We have $A = P\Lambda P'$ (Theorem A.30) and without loss of generality (Theorem A.61 (i)) we may write $\Lambda = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, that is, $P'AP = \Lambda$ with P orthogonal. Let $P = (\underset{p,r}{P_1} \quad \underset{p,(p-r)}{P_2})$ and

$$P'x = y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} P_1'x \\ P_2'x \end{pmatrix}.$$

Therefore

$$\begin{aligned} y &\sim N_p(P'\mu, I_p) \quad \text{(Theorem A.82)} \\ y_1 &\sim N_r(P_1'\mu, I_r) \\ \text{and } y_1'y_1 &\sim \chi_r^2(\mu' P_1 P_1'\mu) \quad \text{(Theorem A.84).} \end{aligned}$$

As P is orthogonal, we have

$$\begin{aligned} A &= (PP')A(PP') = P(P'AP)P \\ &= (P_1 \; P_2) \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P_1' \\ P_2' \end{pmatrix} = P_1 P_1' \,, \end{aligned}$$

and therefore

$$x'Ax = x'P_1P_1'x = y_1'y_1 \sim \chi_r^2(\mu' A\mu).$$

Theorem A.88 *Let $x \sim N_p(\mu, I)$, $A : p \times p$ be idempotent of rank r, and $B : n \times p$ be any matrix. Then the linear form Bx is independent of the quadratic form $x'Ax$ if and only if $BA = 0$.*

Proof: Let P be the matrix as in Theorem A.87. Then $BPP'AP = BAP = 0$, as $BA = 0$ was assumed. Let $BP = D = (D_1, D_2) = (BP_1, BP_2)$, then

$$BPP'AP = (D_1, D_2) \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = (D_1, 0) = (0, 0),$$

so that $D_1 = 0$. This gives

$$Bx = BPP'x = Dy = (0, D_2) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = D_2 y_2 \,,$$

where $y_2 = P_2'x$. Since P is orthogonal and hence regular, we may conclude that all the components of $y = P'x$ are independent $\Rightarrow$ $Bx = D_2y_2$ and $x'Ax = y_1'y_1$ are independent.

Theorem A.89 *Let $x \sim N_p(0, I)$ and A and B be idempotent $p \times p$-matrices with $\operatorname{rank}(A) = r$ and $\operatorname{rank}(B) = s$. Then the quadratic forms $x'Ax$ and $x'Bx$ are independently distributed if and only if $BA = 0$.*

Proof: If we use P from Theorem A.87 and set $C = P'BP$ (C symmetric), we get with the assumption $BA = 0$,

$$\begin{aligned} CP'AP &= P'BPP'AP \\ &= P'BAP = 0 \,. \end{aligned}$$

Using

$$\begin{aligned} C &= \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} B(P_1' \; P_2') \\ &= \begin{pmatrix} C_1 & C_2 \\ C_2' & C_3 \end{pmatrix} = \begin{pmatrix} P_1BP_1' & P_1BP_2' \\ P_2BP_1' & P_2BP_2' \end{pmatrix} , \end{aligned}$$

this relation may be written as

$$CP'AP = \begin{pmatrix} C_1 & C_2 \\ C_2' & C_3 \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} C_1 & 0 \\ C_2' & 0 \end{pmatrix} = 0 .$$

Therefore, $C_1 = 0$ and $C_2 = 0$,

$$\begin{aligned} x'Bx &= x'(PP')B(PP')x \\ &= x'P(P'BP)P'x \\ &= x'PCP'x \\ &= (y_1', y_2') \begin{pmatrix} 0 & 0 \\ 0 & C_3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y_2' C_3 y_2 . \end{aligned}$$

As shown in Theorem A.87, we have $x'Ax = y_1'y_1$, and therefore the quadratic forms $x'Ax$ and $x'Bx$ are independent.

A.15 Differentiation of Scalar Functions of Matrices

Definition A.90 *If $f(X)$ is a real function of an $m \times n$-matrix $X = (x_{ij})$, then the partial differential of f with respect to X is defined as the $m \times n$-matrix of partial differentials $\partial f / \partial x_{ij}$:*

$$\frac{\partial f(X)}{\partial X} = \begin{pmatrix} \frac{\partial f}{\partial x_{11}} & \cdots & \frac{\partial f}{\partial x_{1n}} \\ \vdots & & \vdots \\ \frac{\partial f}{\partial x_{m1}} & \cdots & \frac{\partial f}{\partial x_{mn}} \end{pmatrix} .$$

Theorem A.91 *Let x be an n-vector and A be a symmetric $n \times n$-matrix. Then*

$$\frac{\partial}{\partial x} x'Ax = 2Ax.$$

Proof:

$$\begin{aligned} x'Ax &= \sum_{r,s=1}^{n} a_{rs} x_r x_s , \\ \frac{\partial f}{\partial x_i} x'Ax &= \sum_{\substack{s=1 \\ (s \neq i)}}^{n} a_{is} x_s + \sum_{\substack{r=1 \\ (r \neq i)}}^{n} a_{ri} x_r + 2a_{ii} x_i \\ &= 2 \sum_{s=1}^{n} a_{is} x_s \quad (\text{as } a_{ij} = a_{ji}) \\ &= 2a_i' x \quad (a_i' \text{: } i\text{th row vector of } A). \end{aligned}$$

According to Definition A.90, we get

$$\frac{\partial x'Ax}{\partial x} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} (x'Ax) = 2 \begin{pmatrix} a_1' \\ \vdots \\ a_n' \end{pmatrix} x = 2Ax.$$

Theorem A.92 *If x is an n-vector, y is an m-vector, and C an $n \times m$-matrix, then*

$$\frac{\partial}{\partial C} x'Cy = xy'.$$

Proof:

$$\begin{aligned} x'Cy &= \sum_{r=1}^{m} \sum_{s=1}^{n} x_s c_{sr} y_r , \\ \frac{\partial}{\partial c_{k\lambda}} x'Cy &= x_k y_\lambda \quad \text{(the } (k,\lambda)\text{th element of } xy'), \\ \frac{\partial}{\partial C} x'Cy &= (x_k y_\lambda) = xy' . \end{aligned}$$

Theorem A.93 *Let x be a K-vector, A a symmetric $T \times T$-matrix, and C a $T \times K$-matrix. Then*

$$\frac{\partial x'C'}{\partial C} x'C'ACx = 2ACxx' .$$

Proof: We have

$$\begin{aligned} x'C' &= \left(\sum_{i=1}^{K} x_i c_{1i}, \cdots, \sum_{i=1}^{K} x_i c_{Ti} \right) , \\ \frac{\partial}{\partial c_{k\lambda}} &= (0, \cdots, 0, x_\lambda, 0, \cdots, 0) \quad (x_\lambda \text{ is an element of the } k\text{th column}). \end{aligned}$$

Using the product rule yields

$$\frac{\partial}{\partial c_{k\lambda}} x'C'ACx = \left(\frac{\partial}{\partial c_{k\lambda}} x'C' \right) ACx + x'C'A \left(\frac{\partial}{\partial c_{k\lambda}} Cx \right) .$$

Since

$$x'C'A = \left(\sum_{t=1}^{T} \sum_{i=1}^{K} x_i c_{ti} a_{t1}, \cdots, \sum_{t=1}^{T} \sum_{i=1}^{K} x_i c_{ti} a_{Tt} \right) ,$$

we get

$$\begin{aligned} x'C'A \left(\frac{\partial}{\partial c_{k\lambda}} Cx \right) &= \sum_{t,i} x_i x_\lambda c_{ti} a_{kt} \\ &= \sum_{t,i} x_i x_\lambda c_{ti} a_{tk} \qquad \text{(as } A \text{ is symmetric)} \end{aligned}$$

$$= \left(\frac{\partial}{\partial c_{k\lambda}} x'C'\right) ACx.$$

But $\sum_{t,i} x_i x_\lambda c_{ti} a_{tk}$ is just the (k, λ)-th element of the matrix $ACxx'$.

Theorem A.94 *Assume $A = A(x)$ to be an $n \times n$-matrix, where its elements $a_{ij}(x)$ are real functions of a scalar x. Let B be an $n \times n$-matrix, such that its elements are independent of x. Then*

$$\frac{\partial}{\partial x} \operatorname{tr}(AB) = \operatorname{tr}\left(\frac{\partial A}{\partial x} B\right).$$

Proof:

$$\begin{aligned} \operatorname{tr}(AB) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} b_{ji}, \\ \frac{\partial}{\partial x}\operatorname{tr}(AB) &= \sum_{i}\sum_{j} \frac{\partial a_{ij}}{\partial x} b_{ji} \\ &= \operatorname{tr}\left(\frac{\partial A}{\partial x} B\right), \end{aligned}$$

where $\partial A/\partial x = (\partial a_{ij}/\partial x)$.

Theorem A.95 *For the differentials of the trace we have the following rules:*

	y	$\partial y/\partial X$
(i)	$\operatorname{tr}(AX)$	A'
(ii)	$\operatorname{tr}(X'AX)$	$(A + A')X$
(iii)	$\operatorname{tr}(XAX)$	$X'A + A'X'$
(iv)	$\operatorname{tr}(XAX')$	$X(A + A')$
(v)	$\operatorname{tr}(X'AX')$	$AX' + X'A$
(vi)	$\operatorname{tr}(X'AXB)$	$AXB + A'XB'$

Differentiation of Inverse Matrices

Theorem A.96 *Let $T = T(x)$ be a regular matrix, such that its elements depend on a scalar x. Then*

$$\frac{\partial T^{-1}}{\partial x} = -T^{-1}\frac{\partial T}{\partial x} T^{-1}.$$

Proof: We have $T^{-1}T = I$, $\partial I/\partial x = 0$, and

$$\frac{\partial (T^{-1}T)}{\partial x} = \frac{\partial T^{-1}}{\partial x} T + T^{-1}\frac{\partial T}{\partial x} = 0.$$

Theorem A.97 *For nonsingular X, we have*

$$\frac{\partial \operatorname{tr}(AX^{-1})}{\partial X} = -(X^{-1}AX^{-1})',$$

$$\frac{\partial\,\mathrm{tr}(X^{-1}AX^{-1}B)}{\partial X} \quad = \quad -(X^{-1}AX^{-1}BX^{-1} + X^{-1}BX^{-1}AX^{-1})'\,.$$

Proof: Use Theorems A95 and A96 and the product rule.

Differentiation of a Determinant

Theorem A.98 *For a nonsingular matrix Z, we have*

(i) $\frac{\partial}{\partial Z}|Z| = |Z|(Z')^{-1}$.

(ii) $\frac{\partial}{\partial Z}log|Z| = (Z')^{-1}$.

A.16 Miscellaneous Results, Stochastic Convergence

Theorem A.99 (Kronecker product) *Let $A : m \times n = (a_{ij})$ and $B : p \times q = (b_{rs})$ be any matrices. Then the Kronecker product of A and B is defined as*

$$\underset{mp,nq}{C} = \underset{m,n}{A} \otimes \underset{p,q}{B} = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ \vdots & \vdots & & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix},$$

and the following rules hold:

(i) $c(A \otimes B) = (cA) \otimes B = A \otimes (cB)$ *(c a scalar),*

(ii) $A \otimes (B \otimes C) = (A \otimes B) \otimes C$,

(iii) $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$,

(iv) $(A \otimes B)' = A' \otimes B'$.

Theorem A.100 (Chebyschev's inequality) *For any n-dimensional random vector X and a given scalar $\epsilon > 0$, we have*

$$P\{|X| \geq \epsilon\} \leq \frac{\mathrm{E}\,|X|^2}{\epsilon^2}\,.$$

Proof: Let $F(x)$ be the joint distribution function of $X = (x_1, \ldots, x_n)$. Then

$$\begin{aligned} \mathrm{E}\,|x|^2 &= \int |x|^2 dF(x) \\ &= \int_{\{x:|x|\geq\epsilon\}} |x|^2 dF(x) + \int_{\{x:|x|<\epsilon\}} |x|^2 dF(x) \\ &\geq \epsilon^2 \int_{\{x:|x|\geq\epsilon\}} dF(x) = \epsilon^2 P\{|x| \geq \epsilon\}\,. \end{aligned}$$

Definition A.101 *Let* $\{x(t)\}$, $t = 1, 2, \ldots$ *be a multivariate stochastic process.*

(i) *Weak convergence: If*

$$\lim_{t\to\infty} P\{|x(t) - \tilde{x}| \geq \delta\} = 0\,,$$

where $\delta > 0$ *is any given scalar and* $\tilde{x}$ *is a finite vector, then* $\tilde{x}$ *is called the probability limit of* $\{x(t)\}$, *and we write*

$$p\lim x = \tilde{x}\,.$$

(ii) *Strong convergence: Assume that* $\{x(t)\}$ *is defined on a probability space* (Ω, Σ, P). *Then* $\{x(t)\}$ *is said to be strongly convergent to* $\tilde{x}$, *that is,*

$$\{x(t)\} \to \tilde{x} \quad \textit{almost surely (a.s.)}$$

if there exists a set $T \in \Sigma$, $P(T) = 0$, *and* $x_\omega(t) \to \tilde{x}_\omega$, *as* $T \to \infty$, *for each* $\omega \in \Omega - T$

Theorem A.102 (Slutsky's theorem) *Using Definition A.101, we have*

(i) *if* $p\lim x = \tilde{x}$, *then* $\lim_{t\to\infty} \mathrm{E}\{x(t)\} = \bar{\mathrm{E}}(x) = \tilde{x}$,

(ii) *if* c *is a vector of constants, then* $p\lim c = c$,

(iii) *(Slutsky's theorem) if* $p\lim x = \tilde{x}$ *and* $y = f(x)$ *is any continuous vector function of* x, *then* $p\lim y = f(\tilde{x})$,

(iv) *if* A *and* B *are random matrices, then the following limits exist:*

$$p\lim(AB) = (p\lim A)(p\lim B)$$

and

$$p\lim(A^{-1}) = (p\lim A)^{-1}\,,$$

(v) *if* $p\lim\left[\sqrt{T}(x(t) - Ex(t))\right]'\left[\sqrt{T}(x(t) - Ex(t))\right] = V$, *then the asymptotic covariance matrix is*

$$\bar{\mathrm{V}}(x, x) = \bar{\mathrm{E}}\left[x - \bar{\mathrm{E}}(x)\right]'\left[x - \bar{\mathrm{E}}(x)\right] = T^{-1}V\,.$$

Definition A.103 *If* $\{x(t)\}, t = 1, 2, \ldots$ *is a multivariate stochastic process satisfying*

$$\lim_{t\to\infty} \mathrm{E}\,|x(t) - \tilde{x}|^2 = 0\,,$$

then $\{x(t)\}$ *is called convergent in the quadratic mean, and we write*

$$\mathrm{l.i.m.}\, x = \tilde{x}\,.$$

Theorem A.104 *If* $\mathrm{l.i.m.}\, x = \tilde{x}$, *then* $p\lim x = \tilde{x}$.

Proof: Using Theorem A.100 we get

$$0 \leq \lim_{t\to\infty} P(|x(t) - \tilde{x}| \geq \epsilon) \leq \lim_{t\to\infty} \frac{\mathrm{E}\,|x(t) - \tilde{x}|^2}{\epsilon^2} = 0\,.$$

Theorem A.105 *If* l.i.m. $(x(t) - \mathrm{E}\,x(t)) = 0$ *and* $\lim_{t\to\infty} \mathrm{E}\,x(t) = c$, *then* $p\lim x(t) = c$.

Proof:

$$\begin{aligned}
&\lim_{t\to\infty} P(|x(t) - c| \geq \epsilon) \\
&\quad \leq \epsilon^{-2} \lim_{t\to\infty} \mathrm{E}\,|x(t) - c|^2 \\
&\quad = \epsilon^{-2} \lim_{t\to\infty} \mathrm{E}\,\big|x(t) - \mathrm{E}\,x(t) + \mathrm{E}\,x(t) - c\big|^2 \\
&\quad = \epsilon^{-2} \lim_{t\to\infty} \mathrm{E}\,\big|x(t) - \mathrm{E}\,x(t)\big|^2 + \epsilon^{-2} \lim_{t\to\infty} \big|\mathrm{E}\,x(t) - c\big|^2 \\
&\qquad + 2\epsilon^{-2} \lim_{t\to\infty} \big\{\big(\mathrm{E}\,x(t) - c)'(x(t) - \mathrm{E}\,x(t)\big)\big\} \\
&\quad = 0\,.
\end{aligned}$$

Theorem A.106 l.i.m. $x = c$ *if and only if*

$$\text{l.i.m.}\,\big(x(t) - \mathrm{E}\,x(t)\big) = 0 \text{ and } \lim_{t\to\infty} \mathrm{E}\,x(t) = c\,.$$

Proof: As in Theorem A.105, we may write

$$\begin{aligned}
\lim_{t\to\infty} \mathrm{E}\,\big|x(t) - c\big|^2 &= \lim_{t\to\infty} \mathrm{E}\,\big|x(t) - \mathrm{E}\,x(t)\big|^2 + \lim_{t\to\infty} \big|\mathrm{E}\,x(t) - c\big|^2 \\
&\quad + 2 \lim_{t\to\infty} \mathrm{E}\,\big(\mathrm{E}\,x(t) - c\big)'\big(x(t) - \mathrm{E}\,x(t)\big) = 0\,.
\end{aligned}$$

Theorem A.107 *Let $x(t)$ be an estimator of a parameter vector θ. Then we have the result*

$$\lim_{t\to\infty} \mathrm{E}\,x(t) = \theta \quad \text{if} \quad \text{l.i.m.}\,(x(t) - \theta) = 0\,.$$

That is, $x(t)$ is an asymptotically unbiased estimator for θ if $x(t)$ converges to θ in the quadratic mean.

Proof: Use Theorem A.106.

Theorem A.108 *Let $V : p \times p$ and n.n.d. and $X : p \times m$ matrices. Then one choice of the g-inverse of*

$$\begin{pmatrix} V & X \\ X' & 0 \end{pmatrix}$$

is

$$\begin{pmatrix} C_1 & C_2 \\ C_2' & -C_4 \end{pmatrix}$$

where, with $T = V + XX'$,

$$\begin{aligned} C_1 &= T - T^- X(X'T^- X)^- X'T^- \\ C_2' &= (X'T^- X)^- X'T^- \\ -C_4 &= (X'T^- X)^-(X'T^- X - I) \end{aligned}$$

For details, see Rao (1989).