CHAPTER 1

NOTATION

1.1 GENERAL DEFINITIONS

Vectors and matrices are denoted by boldface letters $\mathbf{a}$ and $\mathbf{A}$, respectively, and scalars are denoted by italics. Thus $\mathbf{a} = (a_i)$ is a vector with ith element a_i and $\mathbf{A} = (a_{ij})$ is a matrix with i, jth elements a_{ij}. I maintain this notation even with random variables, because using uppercase for random variables and lowercase for their values can cause confusion with vectors and matrices. In Chapters 20 and 21, which focus on random variables, we endeavor to help the reader by using the latter half of the alphabet $u, v, \ldots, z$ for random variables and the rest of the alphabet for constants.

Let $\mathbf{A}$ be an $n_1 \times n_2$ matrix. Then any $m_1 \times m_2$ matrix $\mathbf{B}$ formed by deleting any $n_1 - m_1$ rows and $n_2 - m_2$ columns of $\mathbf{A}$ is called a *submatrix* of $\mathbf{A}$. It can also be regarded as the intersection of m_1 rows and m_2 columns of $\mathbf{A}$. I shall define $\mathbf{A}$ to be a submatrix of itself, and when this is not the case I refer to a submatrix that is not $\mathbf{A}$ as a *proper submatrix* of $\mathbf{A}$. When $m_1 = m_2 = m$, the square matrix $\mathbf{B}$ is called a *principal submatrix* and it is said to be of *order m*. Its determinant, $\det(\mathbf{B})$, is called an *mth-order minor* of $\mathbf{A}$. When $\mathbf{B}$ consists of the intersection of the same numbered rows and columns (e.g., the first, second, and fourth), the minor is called a *principal minor*. If $\mathbf{B}$ consists of the intersection of the first m rows and the first m columns of $\mathbf{A}$, then it is called a *leading principal submatrix* and its determinant is called a *leading principal m-th order* minor.

A Matrix Handbook for Statisticians. By George A. F. Seber

Many matrix results hold when the elements of the matrices belong to a general field $\mathcal{F}$ of scalars. For most practitioners, this means that the elements can be real or complex, so we shall use $\mathbb{F}$ to denote either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. The expression $\mathbb{F}^n$ will denote the n-dimensional counterpart.

If $\mathbf{A}$ is complex, it can be expressed in the form $\mathbf{A} = \mathbf{B} + i\mathbf{C}$, where $\mathbf{B}$ and $\mathbf{C}$ are real matrices, and its *complex conjugate* is $\bar{\mathbf{A}} = \mathbf{B} - i\mathbf{C}$. We call $\mathbf{A}' = (a_{ji})$ the *transpose* of $\mathbf{A}$ and define the *conjugate transpose* of $\mathbf{A}$ to be $\mathbf{A}^* = \overline{\mathbf{A}}'$. In practice, we can often transfer results from real to complex matrices, and vice versa, by simply interchanging $'$ and *.

When adding or multiplying matrices together, we will assume that the sizes of the matrices are such that these operations can be carried out. We make this assumption by saying that the matrices are *conformable*. If there is any ambiguity we shall denote an $m \times n$ matrix $\mathbf{A}$ by $\mathbf{A}_{m\times n}$. A matrix partitioned into blocks is called a block matrix.

If x and y are random variables, then the symbols $\mathrm{E}(y)$, $\mathrm{var}(y)$, $\mathrm{cov}(x, y)$, and $\mathrm{E}(x \mid y)$ represent expectation, variance, covariance, and conditional expectation, respectively.

Before we give a list of all the symbols used we mention some univariate statistical distributions.

1.2 SOME CONTINUOUS UNIVARIATE DISTRIBUTIONS

We assume that the reader is familiar with the normal, chi-square, t, F, gamma, and beta univariate distributions. Multivariate vector versions of the normal and t distributions are given in Sections 20.5.1 and 20.8.1, respectively, and matrix versions of the gamma and beta are found in Section 21.9. As some noncentral distributions are referred to in the statistical chapters, we define two univariate distributions below.

1.1. (Noncentral Chi-square Distribution) The random variable x with probability density function

$$f(x) = \frac{1}{2^{\nu/2}} e^{-x^2/2} x^{(\nu/2)-1} \sum_{i=1}^{\infty} e^{-\delta/2} \left(\frac{\delta}{4}\right)^i \frac{1}{i!\,\Gamma(\frac{1}{2}\nu + i)} x^i$$

is called the *noncentral chi-square distribution* with ν degrees of freedom and noncentrality parameter δ, and we write $x \sim \chi^2_\nu(\delta)$.

(a) When $\delta = 0$, the above density reduces to the (central) chi-square distribution, which is denoted by χ^2_ν.

(b) The noncentral chi-square can be defined as the distribution of the sum of the squares of independent univariate normal variables y_i $(i = 1, 2, \ldots, n)$ with variances 1 and respective means μ_i. Thus if $\mathbf{y} \sim N_d(\boldsymbol{\mu}, \mathbf{I}_d)$, the multivariate normal distribution, then $x = \mathbf{y}'\mathbf{y} \sim \chi^2_d(\delta)$, where $\delta = \boldsymbol{\mu}'\boldsymbol{\mu}$ (Anderson [2003: 81–82]).

(c) $E(x) = \nu + \delta$.

Since $\delta > 0$, some authors set $\delta = \tau^2$, say. Others use $\delta/2$, which, because of (c), is not so memorable.

1.2. (Noncentral F-Distribution) If $x \sim \chi^2_m(\delta)$, $y \sim \chi^2_n$, and x and y are statistically independent, then $F = (x/m)/(y/n)$ is said to have a noncentral F-distribution with m and n degrees of freedom, and noncentrality parameter δ. We write $F \sim F_{m,n}(\delta)$. For a derivation of this distribution see Anderson [2003: 185]. When $\delta = 0$, we use the usual notation $F_{m,n}$ for the F-distribution.

1.3 GLOSSARY OF NOTATION

Scalars

$\mathcal{F}$	field of scalars
$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
$\mathbb{F}$	$\mathbb{R}$ or $\mathbb{C}$
$z = x + iy$	a complex number
$\overline{z} = x - iy$	complex conjugate of z
$\|z\| = (x^2 + y^2)^{1/2}$	modulus of z

Vector Spaces

$\mathbb{F}^n$	n-dimensional coordinate space
$\mathbb{R}^n$	$\mathbb{F}^n$ with $\mathbb{F} = \mathbb{R}$
$\mathbb{C}^n$	$\mathbb{F}^n$ with $\mathbb{F} = \mathbb{C}$
$\mathcal{C}(\mathbf{A})$	column space of $\mathbf{A}$, the space spanned by the columns of $\mathbf{A}$
$\mathcal{C}(\mathbf{A}')$	row space of $\mathbf{A}$
$\mathcal{N}(\mathbf{A})$	$\{\mathbf{x} : \mathbf{Ax} = \mathbf{0}\}$, null space (kernel) of $\mathbf{A}$
$\mathcal{S}(A)$	span of the set A, the vector space of all linear combinations of vectors in A
$\dim \mathcal{V}$	dimension of the vector space $\mathcal{V}$
$\mathcal{V}^\perp$	the orthogonal complement of $\mathcal{V}$
$\mathbf{x} \in \mathcal{V}$	$\mathbf{x}$ is an element of $\mathcal{V}$
$\mathcal{V} \subseteq \mathcal{W}$	$\mathcal{V}$ is a subset of $\mathcal{W}$
$\mathcal{V} \subset \mathcal{W}$	$\mathcal{V}$ is a proper subset of $\mathcal{W}$ (i.e., $\mathcal{V} \neq \mathcal{W}$)
$\mathcal{V} \cap \mathcal{W}$	intersection, $\{\mathbf{x} : \mathbf{x} \in \mathcal{V}$ and $\mathbf{x} \in \mathcal{W}\}$
$\mathcal{V} \cup \mathcal{W}$	union, $\{\mathbf{x} : \mathbf{x} \in \mathcal{V}$ and/or $\mathbf{x} \in \mathcal{W}\}$
$\mathcal{V} + \mathcal{W}$	sum, $\{\mathbf{x} + \mathbf{y} : \mathbf{x} \in \mathcal{V}, \mathbf{y} \in \mathcal{W}\}$
$\mathcal{V} \oplus \mathcal{W}$	direct sum, $\{\mathbf{x} + \mathbf{y} : \mathbf{x} \in \mathcal{V}, \mathbf{y} \in \mathcal{W}; \mathcal{V} \cap \mathcal{W} = \mathbf{0}\}$
$\langle , \rangle$	an inner product defined on a vector space
$\mathbf{x} \perp \mathbf{y}$	$\mathbf{x}$ is perpendicular to $\mathbf{y}$ (i.e., $\langle \mathbf{x}, \mathbf{y} \rangle = 0$)

Complex Matrix

$\mathbf{A} = \mathbf{B} + i\mathbf{C}$	complex matrix, with $\mathbf{B}$ and $\mathbf{C}$ real
$\overline{\mathbf{A}} = (\overline{a}_{ij}) = \mathbf{B} - i\mathbf{C}$	complex conjugate of $\mathbf{A}$
$\mathbf{A}^* = \overline{\mathbf{A}}' = (\overline{a}_{ji})$	conjugate transpose of $\mathbf{A}$
$\mathbf{A} = \mathbf{A}^*$	$\mathbf{A}$ is a Hermitian matrix
$\mathbf{A} = -\mathbf{A}^*$	$\mathbf{A}$ is a skew-Hermitian matrix
$\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$	$\mathbf{A}$ is a normal matrix

Special Symbols

sup	supremum
inf	infemum
max	maximum
min	minimum
$\rightarrow$	tends to
$\Rightarrow$	implies
$\propto$	proportional to
$\mathbf{1}_n$	the $n \times 1$ vector with unit elements
$\mathbf{I}_n$	the $n \times n$ identity matrix
$\mathbf{0}$	a vector or matrix of zeros
$\mathrm{diag}(\mathbf{d})$	$n \times n$ matrix with diagonal elements $\mathbf{d}' = (d_1, \ldots, d_n)$, and zeros elsewhere
$\mathrm{diag}(d_1, d_2, \ldots, d_n)$	same as above
diag $\mathbf{A}$	diagonal matrix ; same diagonal elements as $\mathbf{A}$
$\mathbf{A} \geq \mathbf{0}$	the elements of $\mathbf{A}$ are all non-negative
$\mathbf{A} > \mathbf{0}$	the elements of $\mathbf{A}$ are all positive
$\mathbf{A} \succeq \mathbf{0}$, n.n.d	$\mathbf{A}$ is non-negative definite ($\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$)
$\mathbf{A} \succeq \mathbf{B}$, $\mathbf{B} \preceq \mathbf{A}$	$\mathbf{A} - \mathbf{B} \succeq \mathbf{0}$
$\mathbf{A} \succ \mathbf{0}$, p.d.	$\mathbf{A}$ is positive definite ($\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$)
$\mathbf{A} \succ \mathbf{B}$, $\mathbf{B} \prec \mathbf{A}$	$\mathbf{A} - \mathbf{B} \succ \mathbf{0}$
$\mathbf{x} \ll \mathbf{y}$	$\mathbf{x}$ is (strongly) majorized by $\mathbf{y}$
$\mathbf{x} \ll_w \mathbf{y}$	$\mathbf{x}$ is weakly submajorized by $\mathbf{y}$
$\mathbf{x} \ll^w \mathbf{y}$	$\mathbf{x}$ is weakly supermajorized by $\mathbf{y}$
$\mathbf{A}' = (a_{ji})$	the transpose of $\mathbf{A}$
$\mathbf{A}^{-1}$	inverse of $\mathbf{A}$ when $\mathbf{A}$ is nonsingular
$\mathbf{A}^-$	weak inverse of $\mathbf{A}$ satisfying $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$
$\mathbf{A}^+$	Moore-Penrose inverse of $\mathbf{A}$
trace $\mathbf{A}$	sum of the diagonal elements of a square matrix $\mathbf{A}$
det $\mathbf{A}$	determinant of a square matrix $\mathbf{A}$
rank $\mathbf{A}$	rank of $\mathbf{A}$
per $\mathbf{A}$	permanent of a square matrix $\mathbf{A}$
$\mathrm{mod}(\mathbf{A})$	modulus of $\mathbf{A} = (a_{ij})$, given by $(\lvert a_{ij})\rvert)$
$\mathrm{Pf}(\mathbf{A})$	pfaffian of $\mathbf{A}$
$\rho(\mathbf{A})$	spectral radius of a square matrix $\mathbf{A}$
$\kappa_v(\mathbf{A})$	condition number of an $m \times n$ matrix, $v = 1, 2, \infty$

$\langle \mathbf{x}, \mathbf{y} \rangle$	inner product of $\mathbf{x}$ and $\mathbf{y}$
$\|\mathbf{x}\|$	a norm of vector $\mathbf{x}$ $(= \langle \mathbf{x}, \mathbf{x} \rangle^{1/2})$
$\|\mathbf{x}\|_2$	length of $\mathbf{x}$ $(= (\mathbf{x}^*\mathbf{x})^{1/2})$
$\|\mathbf{x}\|_p$	L_p vector norm of $\mathbf{x}$ $(= \sum_{i=1}^{n} \|x_i\|^p)^{1/p})$
$\|\mathbf{x}\|_\infty$	L_∞ vector norm of $\mathbf{x}$ $(= \max_i \|x_i\|)$
$\|\mathbf{A}\|_p$	a generalized matrix norm of $m \times n$ $\mathbf{A}$ $(= \sum_{i=1}^{m} \sum_{j=1}^{n} \|a_{ij}\|^p)^{1/p}, p > 1)$
$\|\mathbf{A}\|_F$	Frobenius norm of matrix $\mathbf{A}$ $(= (\sum_i \sum_j \|a_{ij}\|^2)^{1/2})$
$\|\mathbf{A}\|_{v,in}$	generalized matrix norm for $m \times n$ matrix $\mathbf{A}$ induced by a vector norm $\|\cdot\|_v$
$\|\mathbf{A}\|_{ui}$	unitarily invariant norm of $m \times n$ matrix $\mathbf{A}$
$\|\mathbf{A}\|_{oi}$	orthogonally invariant norm of $m \times n$ matrix $\mathbf{A}$
$\|\|\|\mathbf{A}\|\|\|$	matrix norm of square matrix $\mathbf{A}$
$\|\|\|\mathbf{A}\|\|\|_{v,in}$	matrix norm for a square matrix $\mathbf{A}$ induced by a vector norm $\|\cdot\|_v$
$\mathbf{A}_{m \times n}$	$m \times n$ matrix
$(\mathbf{A}, \mathbf{B})$	matrix partitioned by two matrices $\mathbf{A}$ and $\mathbf{B}$
$(\mathbf{a}_1, \ldots, \mathbf{a}_n)$	matrix partitioned by column vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product of $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A} \circ \mathbf{B}$	Hadamard (Schur) product of $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A} \odot \mathbf{B}$	Rao–Khatri product of $\mathbf{A}$ and $\mathbf{B}$
$\mathrm{vec}\, \mathbf{A}_{m \times n}$	$mn \times 1$ vector formed by writing the columns of $\mathbf{A}$ one below the other
$\mathrm{vech}\, \mathbf{A}_{m \times m}$	$\frac{1}{2}m(m+1) \times 1$ vector formed by writing the columns of the lower triangle of $\mathbf{A}$ (including the diagonal elements) one below the other
$\mathbf{I}_{(m,n)}$ or $\mathbf{K}_{nm}$	vec-permutation (commutation) matrix
$\mathbf{G}_n$ or $\mathbf{D}_n$	duplication matrix
$\mathbf{P}_n$ or $\mathbf{N}_n$	symmetrizer matrix
$\lambda(\mathbf{A})$	eigenvalue of a square matrix $\mathbf{A}$
$\sigma(\mathbf{B})$	singular value of any matrix $\mathbf{B}$