Matrix Algebra From a Statistician's Perspective

David A. Harville

LENOVO-1023

21

Eigenvalues and Eigenvectors

The decomposition of a matrix $\mathbf{A}$ into a product of two or three matrices can (depending on the characteristics of those matrices) be a very useful first step in computing such things as the rank, the determinant, or an (ordinary or generalized) inverse (of $\mathbf{A}$) as well as a solution to a linear system having $\mathbf{A}$ as its coefficient matrix. In particular, decompositions like the QR, LDU, U′DU, and Cholesky decompositions—refer to Sections 6.4 and 14.5—in which the component matrices are diagonal or triangular, or have orthonormal rows or columns, can be very useful for computational purposes. Moreover, establishing that a matrix has a decomposition of a certain type can be instructive about the nature of the matrix. In particular, if it can be shown that a matrix $\mathbf{A}$ is expressible as $\mathbf{A} = \mathbf{R}\mathbf{R}'$ for some matrix $\mathbf{R}$, then it can be concluded—refer to Corollary 14.2.14—that $\mathbf{A}$ is nonnegative definite (and symmetric).

In the present chapter, consideration is given to a decomposition known as the spectral decomposition, which exists for symmetric matrices, and to a related decomposition known as the singular-value decomposition, which exists for all matrices. In the spectral or singular-value decomposition of a matrix $\mathbf{A}$, $\mathbf{A}$ is decomposed into $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{P}$, where $\mathbf{P}$ and $\mathbf{Q}$ are orthogonal matrices and where $\mathbf{D}$ is a diagonal matrix or, more generally, where $\mathbf{D} = \text{diag}(\mathbf{D}_1, \mathbf{0})$ for some diagonal matrix $\mathbf{D}_1$—in the case of the spectral decomposition, $\mathbf{P} = \mathbf{Q}'$.

The first part of the present chapter is devoted to scalars and (column) vectors that satisfy the definitions of what are, respectively, known as an eigenvalue and an eigenvector of a (square) matrix. And, in subsequent parts of the chapter, it is shown that the existence and construction of the spectral or singular-value decomposition of a matrix $\mathbf{A}$ is equivalent to the existence and computation of linearly independent eigenvectors of $\mathbf{A}$ or of $\mathbf{A}'\mathbf{A}$ or $\mathbf{A}\mathbf{A}'$.

The existence of the spectral or singular-value decomposition can (once it has been established) be very useful in establishing various other results on matrices, including a number of the results that were established earlier in the book without resort to the existence of this decomposition. And, once the spectral or singular-value decomposition of a matrix has been constructed, it can be very useful for computational purposes. However, there are drawbacks in using the existence of

the spectral or singular-value decomposition in "theoretical" arguments (when there are other options) and in constructing this decomposition for computational purposes. Any proof of existence implicitly or explicitly involves some relatively deep mathematics, and in general the construction of the spectral or singular-value decomposition requires iterative methods and is much more computationally intensive than the construction of various other decompositions.

Most (if not all) of the more basic theoretical results in linear statistical models and in related areas of statistics can be established without implicit or explicit use of the existence of the spectral or singular-value decomposition. And, in addressing the computational problems encountered in this area of statistics, it is typically not efficient to adopt an approach that requires the construction of the spectral or singular-vlaue decomposition. However, there are exceptions—refer, for example, to Harville and Fenech (1985).

21.1 Definitions, Terminology, and Some Basic Results

A scalar (real number) λ is said to be an *eigenvalue* of an $n \times n$ matrix $\mathbf{A} = \{a_{ij}\}$ if there exists an $n \times 1$ nonnull vector $\mathbf{x}$ such that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

or, equivalently, such that

$$(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}.$$

Clearly, λ is an eigenvalue of $\mathbf{A}$ if and only if the matrix $\mathbf{A} - \lambda\mathbf{I}$ is singular. The set comprising the different scalars that are eigenvalues of $\mathbf{A}$ is called the *spectrum* of $\mathbf{A}$.

An $n \times 1$ nonnull vector $\mathbf{x}$ is said to be an *eigenvector* of the matrix $\mathbf{A}$ if there exists a scalar (real number) λ such that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, in which case λ is (by definition) an eigenvalue of $\mathbf{A}$. For any particular eigenvector $\mathbf{x}$ (of $\mathbf{A}$), there is only one eigenvalue λ such that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, which (since $\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \Rightarrow \mathbf{x}'\mathbf{A}\mathbf{x} = \lambda\mathbf{x}'\mathbf{x}$) is

$$\lambda = \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{x}}.$$

The eigenvector $\mathbf{x}$ is said to correspond to (or belong to) this eigenvalue. Note that if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$, then, for any nonzero scalar c, the scalar multiple $c\mathbf{x}$ is also an eigenvector of $\mathbf{A}$ and $c\mathbf{x}$ corresponds to the same eigenvalue as $\mathbf{x}$.

Clearly, a nonnull vector $\mathbf{x}$ is an eigenvector (of the $n \times n$ matrix $\mathbf{A}$) corresponding to any particular eigenvalue λ if and only if it is a solution to the homogeneous linear system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{z} = \mathbf{0}$ (in $\mathbf{z}$) or, equivalently, if and only if $\mathbf{x} \in \mathfrak{N}(\mathbf{A} - \lambda\mathbf{I})$. Thus, the set comprising all of the eigenvectors (of $\mathbf{A}$) corresponding to the eigenvalue λ is the set obtained from $\mathfrak{N}(\mathbf{A} - \lambda\mathbf{I})$ by deleting the $(n \times 1)$ null vector. The linear space $\mathfrak{N}(\mathbf{A} - \lambda\mathbf{I})$ is sometimes referred to as the *eigenspace* of the eigenvalue λ, and the dimension of this space is referred to as the *geometric multiplicity* of λ. As a consequence of Lemma 11.3.1,