Matrix Algebra

A **matrix** is a rectangular array of numbers, denoted

$$\mathbf{A} = [a_{ik}] = [\mathbf{A}]_{ik} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nK} \end{bmatrix}.$$

Square matrix

- A **symmetric matrix** is one in which $a_{ik} = a_{ki}$ for all i and k.
- A **diagonal matrix** is a square matrix whose only nonzero elements appear on the **main diagonal**, that is, moving from upper left to lower right.
- A **scalar matrix** is a diagonal matrix with the same value in all diagonal elements.
- An **identity matrix** is a scalar matrix with ones on the diagonal. This matrix is always denoted **I**. A subscript is sometimes included to indicate its size, or **order**. For example, $\mathbf{I}_4$ indicates a 4×4 identity matrix.
- A **triangular matrix** is one that has only zeros either above or below the main diagonal. If the zeros are above the diagonal, the matrix is **lower triangular**.

Transposition

The **transpose** of a matrix $\mathbf{A}$, denoted $\mathbf{A}'$, is obtained by creating the matrix whose *kth* row is the *kth* column of the original matrix.[2] Thus, if $\mathbf{B} = \mathbf{A}'$, then each column of $\mathbf{A}$ will appear as the corresponding row of $\mathbf{B}$. If $\mathbf{A}$ is $n \times K$, then $\mathbf{A}'$ is $K \times n$.

The definition of a symmetric matrix implies that

$$\text{if (and only if) } \mathbf{A} \text{ is symmetric, then } \mathbf{A} = \mathbf{A}'.$$

It also follows from the definition that for any $\mathbf{A}$,

$$(\mathbf{A}')' = \mathbf{A}.$$

Finally, the transpose of a column vector, $\mathbf{a}$, is a row vector:

$$\mathbf{a}' = [a_1 \quad a_2 \quad \cdots \quad a_n].$$

Multiplication

Matrices are multiplied by using the **inner product**. The inner product, or **dot product**, of two vectors, **a** and **b**, is a scalar and is written

$$\mathbf{a}'\mathbf{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n = \Sigma_{j=1}^{n} a_jb_j.$$

Matrix multiplication

For an $n \times K$ matrix $\mathbf{A}$ and a $K \times M$ matrix $\mathbf{B}$, the product matrix, $\mathbf{C} = \mathbf{AB}$, is an $n \times M$ matrix whose ikth element is the inner product of row i of $\mathbf{A}$ and column k of $\mathbf{B}$. Thus, the product matrix $\mathbf{C}$ is

[3]A simple way to check the conformability of two matrices for multiplication is to write down the dimensions of the operation, for example, $(n \times K)$ times $(K \times M)$. The inner dimensions must be equal; the result has dimensions equal to the outer values.

$$\mathbf{C} = \mathbf{AB} \Rightarrow c_{ik} = \mathbf{a}_i'\mathbf{b}_k.$$

The product of a matrix and a vector is written

$$\mathbf{c} = \mathbf{A}\mathbf{b}.$$

The number of elements in **b** must equal the number of columns in **A**; the result is a vector with number of elements equal to the number of rows in **A**. For example,

$$\begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

We can interpret this in two ways. First, it is a compact way of writing the three equations

$$\begin{aligned} 5 &= 4a + 2b + 1c, \\ 4 &= 2a + 6b + 1c, \\ 1 &= 1a + 1b + 0c. \end{aligned}$$

Second, by writing the set of equations as

$$\begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} = a \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} + b \begin{bmatrix} 2 \\ 6 \\ 1 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

Some general rules for matrix multiplication are as follows

- **Associative law:** $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.
- **Distributive law:** $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$.
- **Transpose of a product:** $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.
- **Transpose of an extended product:** $(\mathbf{ABC})' = \mathbf{C}'\mathbf{B}'\mathbf{A}'$.

SUMS OF VALUES

Denote by $\mathbf{i}$ a vector that contains a column of ones. Then,

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n = \mathbf{i}'\mathbf{x}.$$

If all elements in $\mathbf{x}$ are equal to the same constant a, then $\mathbf{x} = a\mathbf{i}$ and

$$\sum_{i=1}^{n} x_i = \mathbf{i}'(a\mathbf{i}) = a(\mathbf{i}'\mathbf{i}) = na.$$

For any constant a and vector $\mathbf{x}$,

$$\sum_{i=1}^{n} a x_i = a \sum_{i=1}^{n} x_i = a\mathbf{i}'\mathbf{x}.$$

If $a = 1/n$, then we obtain the arithmetic mean,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{n}\mathbf{i}'\mathbf{x},$$

from which it follows that

$$\sum_{i=1}^{n} x_i = \mathbf{i}'\mathbf{x} = n\bar{x}.$$

IDEMPOTENT MATRIX

$$\mathbf{i}\bar{x} = \mathbf{i}\frac{1}{n}\mathbf{i}'\mathbf{x} = \begin{bmatrix} \bar{x} \\ \bar{x} \\ \vdots \\ \bar{x} \end{bmatrix} = \frac{1}{n}\mathbf{i}\mathbf{i}'\mathbf{x}.$$

The matrix $(1/n)\mathbf{i}\mathbf{i}'$ is an $n \times n$ matrix with every element equal to $1/n$. The set of values in deviations form is

$$\begin{bmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \cdots \\ x_n - \bar{x} \end{bmatrix} = [\mathbf{x} - \mathbf{i}\bar{x}] = \left[\mathbf{x} - \frac{1}{n}\mathbf{i}\mathbf{i}'\mathbf{x}\right].$$

Because $\mathbf{x} = \mathbf{I}\mathbf{x}$,

$$\left[\mathbf{x} - \frac{1}{n}\mathbf{i}\mathbf{i}'\mathbf{x}\right] = \left[\mathbf{I}\mathbf{x} - \frac{1}{n}\mathbf{i}\mathbf{i}'\mathbf{x}\right] = \left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right]\mathbf{x} = \mathbf{M}^0\mathbf{x}.$$

Two properties of $\mathbf{M}^0$ are useful at this point. First, because all off-diagonal elements of $\mathbf{M}^0$ equal $-1/n$, $\mathbf{M}^0$ is symmetric. Second, as can easily be verified by multiplication, $\mathbf{M}^0$ is equal to its square; $\mathbf{M}^0\mathbf{M}^0 = \mathbf{M}^0$.

DEFINITION A.1 Idempotent Matrix

An ***idempotent*** *matrix,* $\mathbf{M}$*, is one that is equal to its square, that is,* $\mathbf{M}^2 = \mathbf{MM} = \mathbf{M}$. *If* $\mathbf{M}$ *is a symmetric idempotent matrix (all of the idempotent matrices we shall encounter are symmetric), then* $\mathbf{M}'\mathbf{M} = \mathbf{M}$ as well.

GEOMETRY OF MATRICES

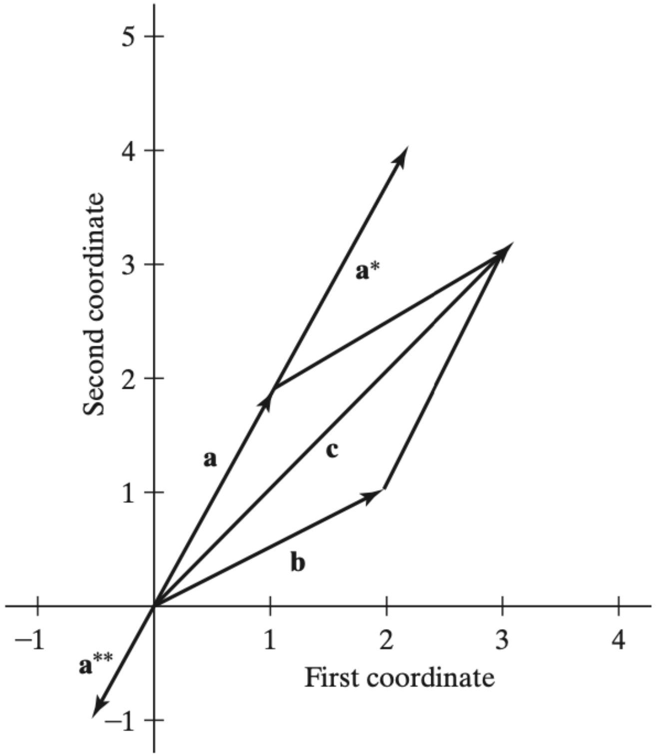
- VECTOR SPACES

The K elements of a column vector

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_K \end{bmatrix}$$

can be viewed as the coordinates of a point in a K-dimensional space, as shown in Figure A.1 for two dimensions, or as the definition of the line segment connecting the origin and the point defined by $\mathbf{a}$.

FIGURE A.1 Vector Space.



GEOMETRY OF MATRICES

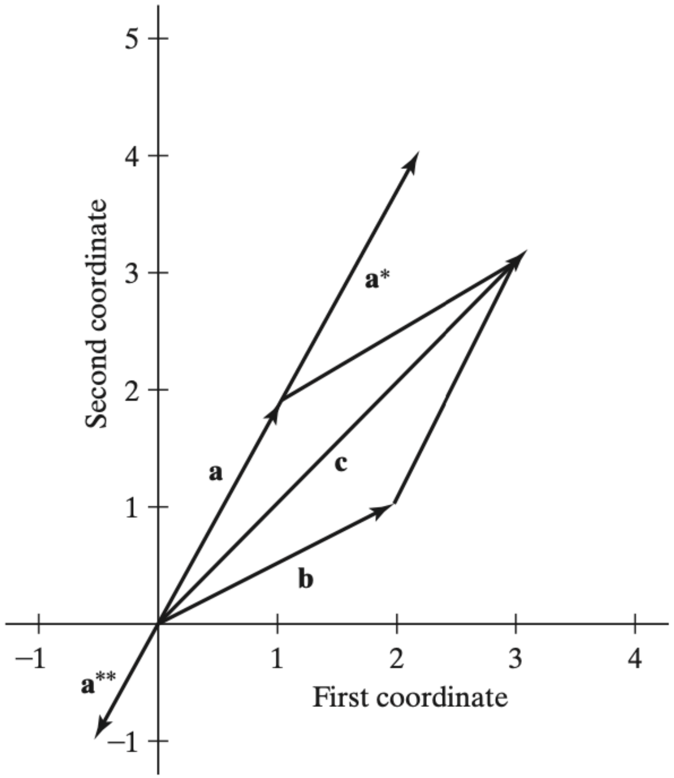
- VECTOR SPACES

scalar multiplication and **addition**

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{a}^* = 2\mathbf{a} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad \mathbf{a}^{**} = -\frac{1}{2}\mathbf{a} = \begin{bmatrix} -\frac{1}{2} \\ -1 \end{bmatrix}.$$

$$\mathbf{c} = \mathbf{a} + \mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

FIGURE A.1 Vector Space.



GEOMETRY OF MATRICES

- VECTOR SPACES

The two-dimensional plane is the set of all vectors with two real-valued coordinates. We label this set $\mathbb{R}^2$ ("R two," not "R squared"). It has two important properties.

- $\mathbb{R}^2$ *is closed under scalar multiplication;* every scalar multiple of a vector in $\mathbb{R}^2$ is also in $\mathbb{R}^2$.
- $\mathbb{R}^2$ *is closed under addition;* the sum of any two vectors in the plane is always a vector in $\mathbb{R}^2$.

*A **vector space** is any set of vectors that is closed under scalar multiplication and addition.*

*A **vector space** is any set of vectors that is closed under scalar multiplication and addition.*

GEOMETRY OF MATRICES

- VECTOR SPACES

*A set of vectors is **linearly independent** if and only if the only solution* $(\alpha_1, \ldots, \alpha_K)$ *to*

$$\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \cdots + \alpha_K\mathbf{a}_K = \mathbf{0}$$

is

$$\alpha_1 = \alpha_2 = \cdots = \alpha_K = 0.$$

the preceding implies the following equivalent definition of a basis.

*The set of all linear combinations of a set of vectors is the vector space that is **spanned** by those vectors.*

*The **column space** of a matrix is the vector space that is spanned by its column vectors.*

GEOMETRY OF MATRICES

- VECTOR SPACES

$$\mathbf{A} = \begin{bmatrix} 1 & 5 & 6 \\ 2 & 6 & 8 \\ 7 & 1 & 8 \end{bmatrix}.$$

It contains three vectors from R^3 , but the third is the sum of the first two, so the column space of this matrix cannot have three dimensions. Nor does it have only one, because the three columns are not all scalar multiples of one another. Hence, it has two, and the column space of this matrix is a two-dimensional subspace of R^3

GEOMETRY OF MATRICES

- RANK OF A MATRIX

*The **column rank** of a matrix is the dimension of the vector space that is spanned by its column vectors.*

What is the column rank of **A?** Can you see that it is 2?

$$\mathbf{A} = \begin{bmatrix} 1 & 5 & 6 \\ 2 & 6 & 8 \\ 7 & 1 & 8 \end{bmatrix}.$$

Consider Matrix **B**

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 1 & 5 \\ 6 & 4 & 5 \\ 3 & 1 & 4 \end{bmatrix}.$$

GEOMETRY OF MATRICES

- RANK OF A MATRIX

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 1 & 5 \\ 6 & 4 & 5 \\ 3 & 1 & 4 \end{bmatrix}.$$

Each column of **B** is a vector in $\mathbb{R}^4$, so the column space of **B** is a three-dimensional subspace of $\mathbb{R}^4$.

Consider, instead, the set of vectors obtained by using the *rows* of **B** instead of the columns. The new matrix would be

$$\mathbf{C} = \begin{bmatrix} 1 & 5 & 6 & 3 \\ 2 & 1 & 4 & 1 \\ 3 & 5 & 5 & 4 \end{bmatrix}.$$

GEOMETRY OF MATRICES

- RANK OF A MATRIX

$$\mathbf{C} = \begin{bmatrix} 1 & 5 & 6 & 3 \\ 2 & 1 & 4 & 1 \\ 3 & 5 & 5 & 4 \end{bmatrix}.$$

This matrix is composed of four column vectors from $\mathbb{R}^3$.

Rank of a Matrix

THEOREM A.1 Equality of Row and Column Rank
*The **column rank** and **row rank** of a matrix are equal. By the definition of row rank and its counterpart for column rank, we obtain the corollary, the **row space** and **column space** of a matrix have the same dimension.* **(A-42)**

$$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}') \leq \min\,(\text{number of rows, number of columns}).$$

GEOMETRY OF MATRICES

- DETERMINANT OF A MATRIX

The determinant of a matrix is nonzero if and only if it has full rank.

It is most convenient to begin with a diagonal matrix

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ & & & \cdots & \\ 0 & 0 & 0 & \cdots & d_K \end{bmatrix}.$$

$$|\mathbf{D}| = d_1 d_2 \ldots d_K = \prod_{k=1}^{K} d_k.$$

$$|c\mathbf{D}| = c^K |\mathbf{D}|.$$

For 2×2 matrices, the computation of the determinant is

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc.$$

A 3X3, however, might be computed on occasion; if so, the following shortcut known as Sarrus's rule will

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{23}a_{32}.$$

GEOMETRY OF MATRICES

- DETERMINANT OF A MATRIX

SOLUTION OF A SYSTEM OF LINEAR EQUATIONS

Consider the set of n linear equations

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

A homogeneous system is of the form $\mathbf{A}\mathbf{x} = \mathbf{0}$.

A nonhomogeneous system of equations is of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$, *where* $\mathbf{b}$ *is a nonzero vector.*

SOLUTION OF A SYSTEM OF LINEAR EQUATIONS - INVERSE MATRICES

To solve the system $\mathbf{Ax} = \mathbf{b}$ for $\mathbf{x}$, something akin to division by a matrix is needed.

Suppose that we could find a square matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}$.

$$\mathbf{BAx} = \mathbf{Ix} = \mathbf{x} = \mathbf{Bb}.$$

SOLUTION OF A SYSTEM OF LINEAR EQUATIONS - INVERSE MATRICES

If the matrix **B** exists, then it is the **inverse** of **A**, denoted

$$\mathbf{B} = \mathbf{A}^{-1}.$$

From the definition,

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

In addition, by premultiplying by **A**, postmultiplying by $\mathbf{A}^{-1}$, and then canceling terms, we find

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

as well.

CHARACTERISTIC ROOTS AND VECTORS

A useful set of results for analysing a square matrix **A** arises from the solutions to the set of equations

$$\mathbf{Ac} = \lambda\mathbf{c}.$$

$$\mathbf{Ac} = \lambda\mathbf{Ic},$$

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

This equation is a homogeneous system that has a nonzero solution only if the matrix $(\mathbf{A} - \lambda\mathbf{I})$ is singular or has a zero determinant. Therefore, if λ is a solution, then

$$|\mathbf{A} - \lambda\mathbf{I}| = 0.$$

CHARACTERISTIC ROOTS AND VECTORS

With λ in hand, the characteristic vectors are derived from the original problem,

$$\mathbf{Ac} = \lambda\mathbf{c},$$

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

A $K \times K$ symmetric matrix has K distinct characteristic vectors, $\mathbf{c}_1, \mathbf{c}_2, \ldots \mathbf{c}_K$. The corresponding characteristic roots, $\lambda_1, \lambda_2, \ldots, \lambda_K$, although real, need not be distinct.

the K-characteristic roots in the same order, in a diagonal matrix,

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & \lambda_K \end{bmatrix}.$$

CHARACTERISTIC ROOTS AND VECTORS

Then, the full set of equations

$$\mathbf{A}\mathbf{c}_k = \lambda_k \mathbf{c}_k$$

is contained in

$$\mathbf{AC} = \mathbf{C}\boldsymbol{\Lambda}.$$

Because the vectors are orthogonal and $\mathbf{c}_i'\mathbf{c}_i = 1$, we have

$$\mathbf{C}'\mathbf{C} = \begin{bmatrix} \mathbf{c}_1'\mathbf{c}_1 & \mathbf{c}_1'\mathbf{c}_2 & \cdots & \mathbf{c}_1'\mathbf{c}_K \\ \mathbf{c}_2'\mathbf{c}_1 & \mathbf{c}_2'\mathbf{c}_2 & \cdots & \mathbf{c}_2'\mathbf{c}_K \\ & & \vdots & \\ \mathbf{c}_K'\mathbf{c}_1 & \mathbf{c}_K'\mathbf{c}_2 & \cdots & \mathbf{c}_K'\mathbf{c}_K \end{bmatrix} = \mathbf{I}.$$

CHARACTERISTIC ROOTS AND VECTORS

*The **diagonalization** of a matrix* $\mathbf{A}$ *is*

$$\mathbf{C}'\mathbf{A}\mathbf{C} = \mathbf{C}'\mathbf{C}\boldsymbol{\Lambda} = \mathbf{I}\boldsymbol{\Lambda} = \boldsymbol{\Lambda}.$$

The spectral decomposition of $\mathbf{A}$ *is*

$$\mathbf{A} = \mathbf{C}\boldsymbol{\Lambda}\mathbf{C}' = \sum_{k=1}^{K} \lambda_k \mathbf{c}_k \mathbf{c}_k'.$$

THEOREM A.3 Rank of a Product

For any matrix $\mathbf{A}$ *and nonsingular matrices* $\mathbf{B}$ *and* $\mathbf{C}$, *the rank of* $\mathbf{BAC}$ *is equal to the rank of* $\mathbf{A}$. ***Proof:*** *By (A-45),* $\text{rank}(\mathbf{BAC}) = \text{rank}[(\mathbf{BA})\mathbf{C}] = \text{rank}(\mathbf{BA})$. *By (A-43),* $\text{rank}(\mathbf{BA}) = \text{rank}(\mathbf{A}'\mathbf{B}')$, *and applying (A-45) again,* $\text{rank}(\mathbf{A}'\mathbf{B}') = \text{rank}(\mathbf{A}')$ *because* $\mathbf{B}'$ *is nonsingular if* $\mathbf{B}$ *is nonsingular [once again, by (A-43)]. Finally, applying (A-43) again to obtain* $\text{rank}(\mathbf{A}') = \text{rank}(\mathbf{A})$ *gives the result.*

THEOREM A.4 Rank of a Symmetric Matrix

The rank of a symmetric matrix is the number of nonzero characteristic roots it contains.

TRACE OF A MATRIX

The **trace** of a square K x K matrix is the sum of its diagonal elements

THEOREM A.7 Trace of a Matrix

The trace of a matrix equals the sum of its characteristic roots.

TRACE OF A MATRIX

Some easily proven results are

$$\operatorname{tr}(c\mathbf{A}) = c(\operatorname{tr}(\mathbf{A})),$$

$$\operatorname{tr}(\mathbf{A}') = \operatorname{tr}(\mathbf{A}),$$

$$\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B}),$$

$$\operatorname{tr}(\mathbf{I}_K) = K.$$

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA}).$$

Determinant of a Matrix

THEOREM A.8 Determinant of a Matrix

The determinant of a matrix equals the product of its characteristic roots.

$$\mathbf{C}'\mathbf{A}\mathbf{C} = \mathbf{\Lambda},$$
$$|\mathbf{C}'\mathbf{A}\mathbf{C}| = |\mathbf{\Lambda}|.$$

$$|\mathbf{C}'\mathbf{A}\mathbf{C}| = |\mathbf{C}'| \cdot |\mathbf{A}| \cdot |\mathbf{C}| = |\mathbf{C}'| \cdot |\mathbf{C}| \cdot |\mathbf{A}| = |\mathbf{C}'\mathbf{C}| \cdot |\mathbf{A}| = |\mathbf{I}| \cdot |\mathbf{A}| = 1 \cdot |\mathbf{A}|$$

$$= |\mathbf{A}|$$
$$= |\mathbf{\Lambda}|.$$

QUADRATIC FORMS AND DEFINITE MATRICES

Many optimization problems involve double sums of the form

$$q = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j a_{ij}.$$

This **quadratic form** can be written

$$q = \mathbf{x}'\mathbf{A}\mathbf{x}$$

where $\mathbf{A}$ is a symmetric matrix. In general, q may be positive, negative, or zero; it depends on $\mathbf{A}$ and $\mathbf{x}$. There are some matrices, however, for which q will be positive regardless of $\mathbf{x}$, and others for which q will always be negative (or nonnegative or nonpositive).

QUADRATIC FORMS AND DEFINITE MATRICES

For a given matrix $\mathbf{A}$,

1. If $\mathbf{x}'\mathbf{A}\mathbf{x} > (<)\ 0$ for all nonzero $\mathbf{x}$, then $\mathbf{A}$ is **positive (negative) definite.**
2. If $\mathbf{x}'\mathbf{A}\mathbf{x} \geq (\leq)\ 0$ for all nonzero $\mathbf{x}$, then $\mathbf{A}$ is **nonnegative definite** or **positive semidefinite** (nonpositive definite).

Recall that a symmetric matrix can be decomposed into

$$\mathbf{A} = \mathbf{C}\boldsymbol{\Lambda}\mathbf{C}'.$$

Therefore, the quadratic form can be written as

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{C}\boldsymbol{\Lambda}\mathbf{C}'\mathbf{x}.$$

Let $\mathbf{y} = \mathbf{C}'\mathbf{x}$. Then

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{y}'\boldsymbol{\Lambda}\mathbf{y} = \sum_{i=1}^{n} \lambda_i y_i^2.$$

QUADRATIC FORMS AND DEFINITE MATRICES

THEOREM A.11 Definite Matrices
Let **A** *be a symmetric matrix. If all the characteristic roots of* **A** *are positive (negative), then* **A** *is* **positive definite (negative definite)**. *If some of the roots are zero, then* **A** *is* **nonnegative (nonpositive) definite** *if the remainder are positive (negative). If* **A** *has both negative and positive roots, then* **A** *is* **indefinite.**