STT 843: Multivariate Analysis
2. Matrix Algebra (Chapter 2.1-2.2)

Guanqun Cao

Department of Statistics and Probability
Michigan State University

Spring 2026

Outline

1. **Vectors**

2. **Matrices**

3. **Positive Definite Matrices**

Vectors

- Definition: An array $\mathbf{x}$ of n real number $x_1, x_2, \ldots, x_n$ is called a vector, and it is written as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{or} \quad \mathbf{x}^{\mathsf{T}} = [x_1, x_2, \ldots, x_n]$$

where the prime denotes the operation of transposing a column to a row.
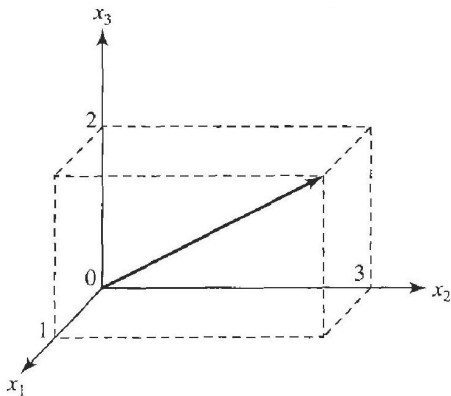
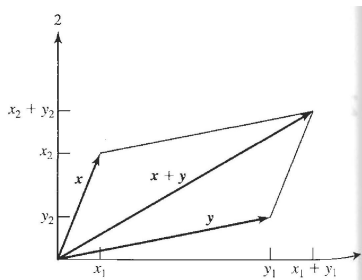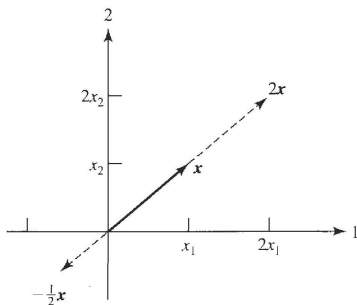Figure 2.1   The vector $x^\mathsf{T} = [1, 3, 2]$.

- Multiplying vectors by a constant c :

$$c\mathbf{X} = \mathbf{X} = \begin{bmatrix} cX_1 \\ cX_2 \\ \vdots \\ cX_n \end{bmatrix}$$

- Addition of $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

Scatter multiplication and vector addition

Length of vectors, unit vector

- When $n = 2$, $\mathbf{x} = [x1, x2]^\mathsf{T}$, the length of $\mathbf{x}$, written $L\mathbf{x}$ is defined to be

$$L_\mathbf{x} = \sqrt{x_1^2 + x_2^2}$$

- Geometrically, the length of a vector in two dimension can be viewed as the hypotenuse of a right triangle. The length of a vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^\mathsf{T}$ and $c\mathbf{x} = [cx_1, cx_2, \ldots, cx_n]^\mathsf{T}$

$$L_\mathbf{x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

$$L_{c\mathbf{x}} = \sqrt{c^2 x_1^2 + c^2 x_2^2 + \cdots + c^2 x_n^2} = |c|\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = |c| L_\mathbf{x}$$

- Choosing $c = L_\mathbf{x}^{-1}$, we obtain the unit vector $L_\mathbf{x}^{-1}\mathbf{x}$, which has length 1 and lies in the direction of $\mathbf{x}$.
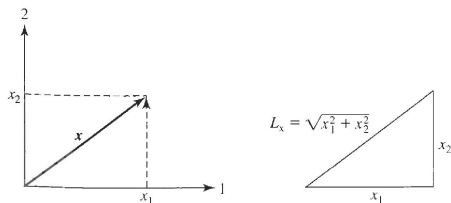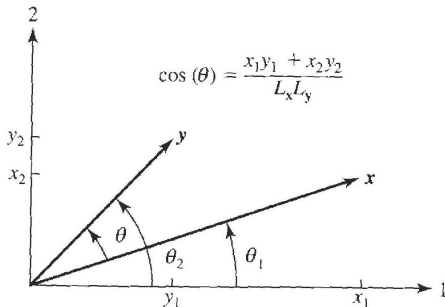
Figure 2.3 Length of $\mathbf{x} = \sqrt{x_1^2 + x_2^2}$.

Figure 2.4 The angle θ between $\mathbf{x}^{\mathsf{T}} = [x_1, x_2]$ and $\mathbf{y}^{\mathsf{T}} = [y_1, y_2]$.

Angle, inner product, perpendicular

- Consider two vectors $\mathbf{x}, \mathbf{y}$ in a plane and the angle θ between them, as in Figure 2.4. From the figure, θ can be represented as the difference between the angle θ_1 and θ_2 formed by the two vectors and the first coordinate axis. Since, by the definition,

$$\cos(\theta_1) = \frac{x_1}{L_{\mathbf{x}}}, \cos(\theta_2) = \frac{y_1}{L_{\mathbf{y}}}$$

$$\sin(\theta_1) = \frac{x_2}{L_{\mathbf{x}}}, \sin(\theta_2) = \frac{y_2}{L_{\mathbf{y}}}$$

and

$$\cos(\theta_2 - \theta_1) = \cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2)$$

- the angle θ between the two vectors is specified by

$$\cos(\theta) = \cos(\theta_2 - \theta_1) = \frac{y_1}{L_{\mathbf{y}}} \cdot \frac{x_1}{L_{\mathbf{x}}} + \frac{y_2}{L_{\mathbf{y}}} \cdot \frac{x_2}{L_{\mathbf{x}}} = \frac{x_1 y_1 + x_2 y_2}{L_{\mathbf{x}} L_{\mathbf{y}}}$$

inner product

Inner product of the two vectors **x** and **y**:

$$\mathbf{x}^\top \mathbf{y} = x_1 y_1 + x_2 y_2$$

With the definition of the inner product and $\cos(\theta)$,

$$L_{\mathbf{x}} = \sqrt{\mathbf{x}^\top \mathbf{x}}, \quad \cos(\theta) = \frac{\mathbf{x}^\top \mathbf{y}}{L_{\mathbf{x}} L_{\mathbf{y}}} = \frac{\mathbf{x}^\top \mathbf{y}}{\sqrt{\mathbf{x}^\top \mathbf{x}} \sqrt{\mathbf{y}^\top \mathbf{y}}}.$$

Example 2.1.(Calculating lengths of vectors and the angle between them)

Given the vectors $\mathbf{x}^{\mathsf{T}} = \begin{bmatrix} 1 & 3 & 2 \end{bmatrix}$ and $\mathbf{y}^{\mathsf{T}} = \begin{bmatrix} -2 & 1 & -1 \end{bmatrix}$, find $3\mathbf{x}$ and $\mathbf{x} + \mathbf{y}$. Next, determine the length of $\mathbf{x}$, the length of $\mathbf{y}$, and the angle between $\mathbf{x}$ and $\mathbf{y}$. Also, check that the length of $3\mathbf{x}$ is three times the length of $\mathbf{x}$

linearly dependency

- A pair of vectors $\mathbf{x}$ and $\mathbf{y}$ of the same dimension is said to be linearly dependent if there exist constants c_1 and c_2, both not zero, such that $c_1\mathbf{x} + c_2\mathbf{y} = 0$. A set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ is said to be linearly dependent if there exist constants $c_1, c_2, \ldots, c_k$, not all zero, such that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_k\mathbf{x}_k = 0$$

- Linear dependence implies that at least one vector in the set can be written as linear combination of the other vectors. Vector of the same dimension that are not linearly dependent are said to be linearly independent.

projection

projection (or shadow) of a vector $\mathbf{x}$ on a vector $\mathbf{y}$:

$$\frac{\left(\mathbf{x}^\top\mathbf{y}\right)}{\mathbf{y}^\top\mathbf{y}} \cdot \mathbf{y} = \frac{\left(\mathbf{x}^\top\mathbf{y}\right)}{L_\mathbf{y}}\frac{1}{L_\mathbf{y}}\mathbf{y}$$

where the vector $L_\mathbf{y}^{-1}\mathbf{y}$ has unit length. The length of the projection is

$$\text{Length of projection } = \frac{\left|\mathbf{x}^\top\mathbf{y}\right|}{L_\mathbf{y}} = L_\mathbf{x}\left|\frac{\mathbf{x}^\top\mathbf{y}}{L_\mathbf{x}L_\mathbf{y}}\right| = L_\mathbf{x}|\cos(\theta)|$$

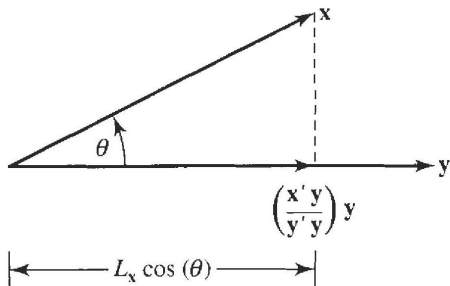where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$.

Figure 2.5   The projection of **x** on **y**.

Example 2.2 (Identifying linearly independent vectors)

Consider if the set of vectors is linearly dependent.

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Matrices

A matrix is any rectangular array of real numbers. We denote an arbitrary array of n rows and p columns

$$\mathbf{A}_{n \times p} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

Example 2.3 (Transpose of a matrix)

if

$$\mathbf{A}_{\{2\times3\}} = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 5 & 4 \end{bmatrix}$$

then

$$\mathbf{A}^{\mathsf{T}}_{\{3\times2\}} = \begin{bmatrix} 3 & 1 \\ -1 & 5 \\ 2 & 4 \end{bmatrix}$$

The product $c\mathbf{A}$ is the matrix that results from multiplying each elements of $\mathbf{A}$ by c. Thus

$$c\mathbf{A}_{\{n\times p\}} = \begin{bmatrix} ca_{11} & ca_{12} & \ldots & ca_{1p} \\ ca_{21} & ca_{22} & \ldots & ca_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{np} \end{bmatrix}$$

Example 2.4 (The sum of two matrices and multiplication of a matrix by a constant)

If

$$\mathbf{A}_{\{2\times3\}} = \begin{bmatrix} 0 & 3 & 1 \\ 1 & -1 & 1 \end{bmatrix} \quad \mathbf{B}_{\{2\times3\}} = \begin{bmatrix} 1 & -2 & -3 \\ 2 & 5 & 1 \end{bmatrix}$$

then $4\mathbf{A}$ and $\mathbf{A} + \mathbf{B}$?

The matrix product **AB** is

$A_{\{n \times k\}} B_{\{k \times p\}}$ = the $(n \times p)$ matrix whose entry in the i th row and j th column is the inner product of the i th row of **A** and the j th column of **B**.

or

$$(i, j) \text{ entry of } \mathbf{AB} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj} = \sum_{\ell=1}^{k} a_{i\ell}b_{\ell j}$$

Example 2.5 (Matrix multiplication)

If

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 5 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -2 \\ 7 \\ 9 \end{bmatrix}, \quad \text{and} \quad \mathbf{C} = \begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix}$$

then **AB** and **CA** ?

Example 2.6 (Some typical products and their dimensions)

Let

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 4 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 7 \\ -3 \\ 6 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 5 \\ 8 \\ -4 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} 2 \\ 9 \end{bmatrix}$$

Then $\mathbf{Ab}, \mathbf{bc}^{\mathsf{T}}, \mathbf{b}^{\mathsf{T}}\mathbf{c}$, and $\mathbf{d}^{\mathsf{T}}\mathbf{Ad}$?

- Square matrices will be of special importance in our development of statistical methods. A square matrix is said to be symmetric if $\mathbf{A} = \mathbf{A}^\top$ or $a_{ij} = a_{ji}$ for all i and j.

- Identity matrix $\mathbf{I}$ acts like 1 in ordinary multiplication $(1 \cdot a = a \cdot 1 = a)$,

$$\mathbf{I}_{(k \times k)} \mathbf{A}_{(k \times k)} = \mathbf{A}_{(k \times k)} \mathbf{I}_{(k \times k)} = \mathbf{A}_{(k \times k)} \quad \text{for any} \quad \mathbf{A}_{(k \times k)}$$

- The fundamental scalar relation about the existence of an inverse number a^{-1} such that $a^{-1}a = aa^{-1} = 1$ if $a \neq 0$ has the following matrix algebra extension: If there exists a matrix $\mathbf{B}$ such that

$$\mathbf{BA} = \mathbf{AB} = \mathbf{I}$$

then $\mathbf{B}$ is called the inverse of $\mathbf{A}$ and is denoted by $\mathbf{A}^{-1}$.

- Example 2.7 (The existence of a matrix inverse) For

$$\mathbf{A} = \left[\begin{array}{cc} 3 & 2 \\ 4 & 1 \end{array} \right]$$

- Diagonal matrices
- Orthogonal matrices

$$\mathbf{Q}\mathbf{Q}^\top = \mathbf{Q}^\top\mathbf{Q} = \mathbf{I} \text{ or } \mathbf{Q}^\top = \mathbf{Q}^{-1}.$$

- Eigenvalue λ with corresponding eigenvector $\mathbf{x} \neq 0$ if

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

Ordinarily, $\mathbf{x}$ is normalized so that it has length unity; that is $\mathbf{x}^\mathsf{T}\mathbf{x} = 1$.

- Let $\mathbf{A}$ be a $k \times k$ square symmetric matrix. Then $\mathbf{A}$ has k pairs of eigenvalues and eigenvectors namely

$$\lambda_1\mathbf{e}_1, \lambda_2\mathbf{e}_2, \ldots, \lambda_k\mathbf{e}_k$$

- The eigenvectors can be chosen to satisfy $1 = \mathbf{e}_1^\mathsf{T}\mathbf{e}_1 = \cdots = \mathbf{e}_k^\mathsf{T}\mathbf{e}$ and be mutually perpendicular. The eigenvectors are unique unless two or more eigenvalues are equal.

- Example 2.8 (Verifying eigenvalues and eigenvectors) Let

$$\mathbf{A} = \begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix}$$

show that $\lambda_1 = 6$ and $\lambda_2 = -4$ is its eigenvalues and the corresponding eigenvectors are $\mathbf{e}_1 = [1/\sqrt{2}, -1/\sqrt{2}]^\mathsf{T}$ and $\mathbf{e}_2 = [1/\sqrt{2}, 1/\sqrt{2}]$.

Positive Definite Matrices

- The study of variation and interrelationships in multivariate data is often based upon distances and the assumption that the data are multivariate normally distributed.

- Squared distance and the multivariate normal density can be expressed in terms of matrix products called quadratic forms.

- Consequently, it should not be surprising that quadratic forms play central role in multivariate analysis. Quadratic forms that are always nonnegative and the associated positive definite matrices.

spectral decomposition

- spectral decomposition for symmetric matrices

$$\mathbf{A}_{(k \times k)} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^\mathsf{T} + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^\mathsf{T} + \cdots + \lambda_k \mathbf{e}_k \mathbf{e}_k^\mathsf{T}$$

- $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the eigenvalues and $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k$ are the associated normalized $k \times 1$ eigenvectors. $\mathbf{e}_i^\mathsf{T} \mathbf{e}_i = 1$ for $i = 1, 2, \ldots, k$ and $\mathbf{e}_i^\mathsf{T} \mathbf{e}_j = 0$ for $i \neq j$.

Because $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ has only square terms x_i^2 and product terms $x_i x_k$, it is called a quadratic form. When a $k \times k$ symmetric matrix $\mathbf{A}$ is such that

$$0 \leq \mathbf{x}^\top \mathbf{A} \mathbf{x}$$

for all $\mathbf{x}^\top = [x_1, x_2, \ldots, x_k]$, both the matrix $\mathbf{A}$ and the quadratic form are said to be nonnegative definite. If the equality holds in the equation above only for the vector $\mathbf{x}^\top = [0, 0, \ldots, 0]$, then $\mathbf{A}$ or the quadratic form is said to be positive definite. In other words, $\mathbf{A}$ is positive definite if

$$0 < \mathbf{x}^\top \mathbf{A} \mathbf{x}$$

for all vectors $\mathbf{x} \neq 0$.

- Using the spectral decomposition, we can easily show that a $k \times k$ matrix $\mathbf{A}$ is a positive definite matrix if and only if every eigenvalue of $\mathbf{A}$ is positive. $\mathbf{A}$ is a nonnegative definite matrix if and only if all of its eigenvalues are greater than or equal to zero.

Example 2.9 (The spectral decomposition of a matrix)

Consider the symmetric matrix and find its spectral decomposition.

$$\mathbf{A} = \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix}$$