Matrix Algebra

Adriano Z. Zambom

Matrix Algebra

- 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

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and the transpose} \quad \mathbf{x}^T = \mathbf{x}' = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}$$

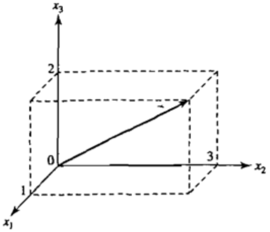

Figure 2.1 The vector $x' = [1, 3, 2]$.

Vector Operations

- Multiply by a constant c (stretch or contract the size of the vector):

$$c\mathbf{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$$

Vector Operations

- Multiply by a constant c (stretch or contract the size of the vector):

$$c\mathbf{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$$

- add vectors

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

Vector Operations

- Length of a vector **x**:

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

- Note that $L_{c\mathbf{x}} =$

Vector Operations

- Length of a vector $\mathbf{x}$:

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

- Note that $L_{c\mathbf{x}} = |c|L_{\mathbf{x}}$

Vector Operations

- Length of a vector **x**:

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

- Note that $L_{c\mathbf{x}} = |c| L_{\mathbf{x}}$
- Angles:

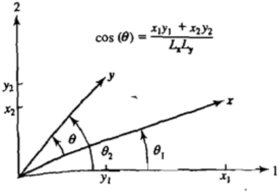

Figure 2.4 The angle θ between $\mathbf{x}' = [x_1, x_2]$ and $\mathbf{y}' = [y_1, y_2]$.

Vector Operations

- Inner product:

$$\mathbf{x}^T\mathbf{y} = \mathbf{x}'\mathbf{y} = x_1y_1 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i$$

Vector Operations

- Inner product:

$$\mathbf{x}^T\mathbf{y} = \mathbf{x}'\mathbf{y} = x_1y_1 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i$$

- Note that $L_{\mathbf{x}} = \sqrt{\mathbf{x}'\mathbf{x}}$

Vector Operations

- Inner product:

$$\mathbf{x}^T\mathbf{y} = \mathbf{x}'\mathbf{y} = x_1y_1 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i$$

- Note that $L_{\mathbf{x}} = \sqrt{\mathbf{x}'\mathbf{x}}$
- And hence $\cos(\theta) = \frac{\mathbf{x}'\mathbf{y}}{L_{\mathbf{x}}L_{\mathbf{y}}} = \frac{\mathbf{x}'\mathbf{y}}{\sqrt{\mathbf{x}'\mathbf{x}}\sqrt{\mathbf{y}'\mathbf{y}}}$

Vector Operations

- Inner product:

$$\mathbf{x}^T\mathbf{y} = \mathbf{x}'\mathbf{y} = x_1 y_1 + \ldots + x_n y_n = \sum_{i=1}^{n} x_i y_i$$

- Note that $L_{\mathbf{x}} = \sqrt{\mathbf{x}'\mathbf{x}}$
- And hence $\cos(\theta) = \frac{\mathbf{x}'\mathbf{y}}{L_{\mathbf{x}} L_{\mathbf{y}}} = \frac{\mathbf{x}'\mathbf{y}}{\sqrt{\mathbf{x}'\mathbf{x}}\sqrt{\mathbf{y}'\mathbf{y}}}$
- Exercise: Find the length of $\mathbf{x} = [1, 2, 4]$ and $\mathbf{y} = [-2, -1, 0]$. Find $\cos(\theta)$.

Vector Operations

- **Definition**: Two vectors $\mathbf{x}$ and $\mathbf{y}$ are **linearly dependent** if there exists constants c_1 and c_2 such that

$$c_1\mathbf{x}_1 + c_2\mathbf{y} = 0.$$

- **linear dependence** means that one vector can be written as a linear combination of the other.

Vector Operations

- **Definition**: Two vectors $\mathbf{x}$ and $\mathbf{y}$ are **linearly dependent** if there exists constants c_1 and c_2 such that

$$c_1\mathbf{x}_1 + c_2\mathbf{y} = 0.$$

- **linear dependence** means that one vector can be written as a linear combination of the other.
- **Definition**: A set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p$ are **linearly dependent** if there exists constants $c_1, c_2, \ldots, c_p$ not all zero such that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_p\mathbf{x}_p = 0.$$

Vector Operations

Example:

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

Setting

$$c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + c_3 \mathbf{x}_3 = \mathbf{0}$$

implies that

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ 2c_1 \qquad - 2c_3 &= 0 \\ c_1 - c_2 + c_3 &= 0 \end{aligned}$$

with the unique solution $c_1 = c_2 = c_3 = 0$. As we cannot find three constants c_1, c_2, and c_3, *not all zero*, such that $c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + c_3 \mathbf{x}_3 = \mathbf{0}$, the vectors $\mathbf{x}_1, \mathbf{x}_2$, and $\mathbf{x}_3$ are *linearly independent*. ■

Vector Operations

- **Definition**: The **projection** of $\mathbf{x}$ on $\mathbf{y}$ is

$$P_{\mathbf{x}}(\mathbf{y}) = \frac{\mathbf{x}'\mathbf{y}}{\mathbf{y}'\mathbf{y}}\mathbf{y} = \frac{\mathbf{x}'\mathbf{y}}{L_{\mathbf{y}}L_{\mathbf{y}}}\mathbf{y}$$

- Find the expression for the length of $P_{\mathbf{x}}(\mathbf{y})$ as a function of θ.

$$L_{P_{\mathbf{x}}(\mathbf{y})} =$$

Vector Operations

- **Definition**: The **projection** of $\mathbf{x}$ on $\mathbf{y}$ is

$$P_{\mathbf{x}}(\mathbf{y}) = \frac{\mathbf{x}'\mathbf{y}}{\mathbf{y}'\mathbf{y}}\mathbf{y} = \frac{\mathbf{x}'\mathbf{y}}{L_{\mathbf{y}}L_{\mathbf{y}}}\mathbf{y}$$

- Find the expression for the length of $P_{\mathbf{x}}(\mathbf{y})$ as a function of θ.

$$L_{P_{\mathbf{x}}(\mathbf{y})} = \frac{|\mathbf{x}'\mathbf{y}|}{L_{\mathbf{y}}} = L_{\mathbf{x}}\left|\frac{\mathbf{x}'\mathbf{y}}{L_{\mathbf{x}}L_{\mathbf{y}}}\right| = L_{\mathbf{x}}|\cos(\theta)|$$

Matrices

Matrices

- A matrix A is a rectangular array of real numbers

$$\mathbf{X}_{\{nxp\}} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{np} \end{bmatrix}$$

Matrices

- A matrix A is a rectangular array of real numbers

$$\mathbf{X}_{\{nxp\}} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{np} \end{bmatrix}$$

- Transpose of A is

$$\mathbf{X}^T_{\{nxp\}} = X_{\{pxn\}} = \begin{bmatrix} x_{11} & x_{21} & x_{31} & \dots & x_{n1} \\ x_{12} & x_{22} & x_{32} & \dots & x_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{1p} & x_{2p} & x_{3p} & \dots & x_{np} \end{bmatrix}$$

Matrices

- Multiply by a constant

$$c\mathbf{X}_{\{nxp\}} = \begin{bmatrix} cx_{11} & cx_{12} & cx_{13} & \dots & cx_{1p} \\ cx_{21} & cx_{22} & cx_{23} & \dots & cx_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ cx_{n1} & cx_{n2} & cx_{n3} & \dots & cx_{np} \end{bmatrix}$$

Matrices

- Multiply by a constant

$$c\mathbf{X}_{\{nxp\}} = \begin{bmatrix} cx_{11} & cx_{12} & cx_{13} & \dots & cx_{1p} \\ cx_{21} & cx_{22} & cx_{23} & \dots & cx_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ cx_{n1} & cx_{n2} & cx_{n3} & \dots & cx_{np} \end{bmatrix}$$

- Sum of Matrices

$$c\mathbf{X} + d\mathbf{Y} = \begin{bmatrix} cx_{11} + dy_{11} & cx_{12} + dy_{12} & cx_{13} + dy_{13} & \dots & cx_{1p} + dy_{1p} \\ cx_{21} + dy_{21} & cx_{1}22 + dy_{22} & cx_{23} + dy_{23} & \dots & cx_{2p} + dy_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ cx_{n1} + dy_{n1} & cx_{n2} + dy_{n2} & cx_{n3} + dy_{n3} & \dots & cx_{np} + dy_{np} \end{bmatrix}$$

Matrices

- Product

Matrices

- Product
  $\mathbf{A}_{\{nxp\}}$ and $\mathbf{B}$ is of dimension?

Matrices

- Product
  $\mathbf{A}_{\{nxp\}}$ and $\mathbf{B}$ is of dimension? $\{pxk\}$

Matrices

- Product

  $\mathbf{A}_{\{nxp\}}$ and $\mathbf{B}$ is of dimension? $\{pxk\}$

$$(i,j) \text{ entry of } \mathbf{AB} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj} = \sum_{\ell=1}^{k} a_{i\ell}b_{\ell j} \qquad (2\text{-}10)$$

When $k = 4$, we have four products to add for each entry in the matrix $\mathbf{AB}$. Thus,

$$\underset{(n\times 4)(4\times p)}{\mathbf{A}\ \mathbf{B}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & a_{i4} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2p} \\ b_{31} & \cdots & b_{3j} & \cdots & b_{3p} \\ b_{41} & \cdots & b_{4j} & \cdots & b_{4p} \end{bmatrix}$$

Matrices

- Product

  $\mathbf{A}_{\{nxp\}}$ and $\mathbf{B}$ is of dimension? $\{pxk\}$

$$(i,j) \text{ entry of } \mathbf{AB} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj} = \sum_{\ell=1}^{k} a_{i\ell}b_{\ell j} \qquad (2\text{-}10)$$

When $k = 4$, we have four products to add for each entry in the matrix $\mathbf{AB}$. Thus,

$$\underset{(n\times 4)(4\times p)}{\mathbf{A}\ \ \mathbf{B}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & a_{i4} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2p} \\ b_{31} & \cdots & b_{3j} & \cdots & b_{3p} \\ b_{41} & \cdots & b_{4j} & \cdots & b_{4p} \end{bmatrix}$$

- Exercise: find $\mathbf{AB}$ and $\mathbf{B}^T\mathbf{A}$ for

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 1 \\ -3 & 0 & -3 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \\ 0 & 3 \end{bmatrix}$$

Matrices

- **Definition**: A **symmetric** matrix is a matrix such that it is equal to its transpose, that is, $\mathbf{A} = \mathbf{A}^T$.

Matrices

- **Definition**: A **symmetric** matrix is a matrix such that it is equal to its transpose, that is, $\mathbf{A} = \mathbf{A}^T$.
- The **identity matrix** I is he matrix with 1 in the diagonal and 0 in the off diagonal, that is

$$I_{\{pxp\}} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

- Property: $I\mathbf{A} = \mathbf{A}I = \mathbf{A}$ for any pxp matrix $\mathbf{A}$

Matrices

- **Definition**: A **symmetric** matrix is a matrix such that it is equal to its transpose, that is, $\mathbf{A} = \mathbf{A}^T$.
- The **identity matrix** I is he matrix with 1 in the diagonal and 0 in the off diagonal, that is

$$I_{\{pxp\}} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

- Property: $I\mathbf{A} = \mathbf{A}I = \mathbf{A}$ for any pxp matrix $\mathbf{A}$
- In R:

Matrices

- **Definition**: If there exists a matrix $\mathbf{B}$ such that

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

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

Matrices

- **Definition**: If there exists a matrix $\mathbf{B}$ such that

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

  then $\mathbf{B}$ is called the inverse of $\mathbf{A}$ and denoted by $\mathbf{B} = \mathbf{A}^{-1}$.
- A **orthogonal matrix** is a square matrix such that

$$\mathbf{AA}^T = \mathbf{A}^T\mathbf{A} = I \text{ equivalently if } \mathbf{A}^T = \mathbf{A}^{-1}$$

Matrices

- **Definition**: A matrix **A** is said to have an eigenvalue λ, with corresponding eigenvector $\mathbf{v} \neq 0$, if

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

Matrices

- **Definition**: A matrix $\mathbf{A}$ is said to have an eigenvalue λ, with corresponding eigenvector $\mathbf{v} \neq 0$, if

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

- For a symmetric *pxp* matrix $\mathbf{A}$, there are p eigenvalues $\lambda_1, \ldots, \lambda_p$ and p eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$

Matrices

- **Definition**: A matrix $\mathbf{A}$ is said to have an eigenvalue λ, with corresponding eigenvector $\mathbf{v} \neq 0$, if

$$\mathbf{Av} = \lambda \mathbf{v}$$

- For a symmetric *pxp* matrix $\mathbf{A}$, there are p eigenvalues $\lambda_1, \ldots, \lambda_p$ and p eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$
- The eigenvectors can be chosen to satisfy $1 = \mathbf{e}_1^T \mathbf{e}_1 = \ldots, \mathbf{e}_p^T \mathbf{e}_p$, where $\mathbf{e}_i$ is the normalized $\mathbf{v}_i$ (divide by $\mathbf{v}_i^T \mathbf{v}_i$), and be mutually perpendicular. The eigenvectors are unique unless two or more eigenvalues are equal.

Matrices

- **Definition**: A matrix $\mathbf{A}$ is said to have an eigenvalue λ, with corresponding eigenvector $\mathbf{v} \neq 0$, if

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

- For a symmetric *pxp* matrix $\mathbf{A}$, there are p eigenvalues $\lambda_1, \ldots, \lambda_p$ and p eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$
- The eigenvectors can be chosen to satisfy $1 = \mathbf{e}_1^T\mathbf{e}_1 = \ldots, \mathbf{e}_p^T\mathbf{e}_p$, where $\mathbf{e}_i$ is the normalized $\mathbf{v}_i$ (divide by $\mathbf{v}_i^T\mathbf{v}_i$), and be mutually perpendicular. The eigenvectors are unique unless two or more eigenvalues are equal.
- in R: "eigen"

Matrices

- **Theorem** Given a square matrix $\mathbf{A}$ and a scalar λ, the following statements are equivalent:
  a) λ is an eigenvalue of $\mathbf{A}$,
  b) the matrix $\mathbf{A} - \lambda I$ is singular,
  c) $det(\mathbf{A} - \lambda I) = 0$.

Matrices

- **Theorem** Given a square matrix **A** and a scalar λ, the following statements are equivalent:
  a) λ is an eigenvalue of **A**,
  b) the matrix $\mathbf{A} - \lambda I$ is singular,
  c) $det(\mathbf{A} - \lambda I) = 0$.
  Definition. $det(\mathbf{A} - \lambda I) = 0$ is called the characteristic equation of the matrix **A**.
  Eigenvalues λ of **A** are roots of the characteristic equation.
  Associated eigenvectors of **A** are nonzero solutions of the equation $(\mathbf{A} - \lambda I)x = 0$.

Example. $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

Characteristic equation: $\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0.$

$(2-\lambda)^2 - 1 = 0 \implies \lambda_1 = 1, \ \lambda_2 = 3.$

$$(A - I)\mathbf{x} = \mathbf{0} \iff \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\iff \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff x + y = 0.$$

The general solution is $(-t, t) = t(-1, 1), \ t \in \mathbb{R}$. Thus $\mathbf{v}_1 = (-1, 1)$ is an eigenvector associated with the eigenvalue 1. The corresponding eigenspace is the line spanned by $\mathbf{v}_1$.

$$(A - 3I)\mathbf{x} = \mathbf{0} \iff \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\iff \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff x - y = 0.$$

The general solution is $(t, t) = t(1, 1), \quad t \in \mathbb{R}$.

Thus $\mathbf{v}_2 = (1, 1)$ is an eigenvector associated with the eigenvalue 3. The corresponding eigenspace is the line spanned by $\mathbf{v}_2$.

Positive Definite Matrices

Positive Definite Matrices

- The **spectral decomposition** of a *pxp* symmetric matrix is

$$\mathbf{A} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \ldots + \lambda_p \mathbf{e}_p \mathbf{e}_p^T,$$

where $\lambda_1, \ldots, \lambda_p$ are the eigenvalues and $\mathbf{e}_1, \ldots, \mathbf{e}_p$ the corresponding eigenvectors (vectors of length p)

Positive Definite Matrices

- The **spectral decomposition** of a *pxp* symmetric matrix is

$$\mathbf{A} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \ldots + \lambda_p \mathbf{e}_p \mathbf{e}_p^T,$$

  where $\lambda_1, \ldots, \lambda_p$ are the eigenvalues and $\mathbf{e}_1, \ldots, \mathbf{e}_p$ the corresponding eigenvectors (vectors of length p)
- Note that this is a weighted sum of the matrices $\mathbf{e}_i \mathbf{e}_i^T$

Positive Definite Matrices

- The **spectral decomposition** of a *pxp* symmetric matrix is

$$\mathbf{A} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \ldots + \lambda_p \mathbf{e}_p \mathbf{e}_p^T,$$

  where $\lambda_1, \ldots, \lambda_p$ are the eigenvalues and $\mathbf{e}_1, \ldots, \mathbf{e}_p$ the corresponding eigenvectors (vectors of length p)
- Note that this is a weighted sum of the matrices $\mathbf{e}_i \mathbf{e}_i^T$

- **Definition**: $\mathbf{x}^T \mathbf{A} \mathbf{x}$ is called a quadratic form.

Positive Definite Matrices

- If a $p \times p$ symmetric matrix A is such that

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

for all $\mathbf{x}$, both the matrix and the quadratic form are said to be **nonnegative definite.**

Positive Definite Matrices

- If a $p \times p$ symmetric matrix A is such that

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

  for all $\mathbf{x}$, both the matrix and the quadratic form are said to be **nonnegative definite.**

- If a $p \times p$ symmetric matrix A is such that

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

  for all $\mathbf{x}$, both the matrix and the quadratic form are said to be **positive definite.**

Positive Definite Matrices

- If a $p \times p$ symmetric matrix A is such that
$$0 \leq \mathbf{x}^T \mathbf{A} \mathbf{x}$$
for all $\mathbf{x}$, both the matrix and the quadratic form are said to be **nonnegative definite.**
- If a $p \times p$ symmetric matrix A is such that
$$0 < \mathbf{x}^T \mathbf{A} \mathbf{x}$$
for all $\mathbf{x}$, both the matrix and the quadratic form are said to be **positive definite.**
- Using the spectral decomposition: a *kxk* matrix A is a positive definite matrix if and only if every eigenvalue of A is positive.

Positive Definite Matrices

- If a $p \times p$ symmetric matrix A is such that
$$0 \leq \mathbf{x}^T \mathbf{A}\mathbf{x}$$
for all $\mathbf{x}$, both the matrix and the quadratic form are said to be **nonnegative definite.**
- If a $p \times p$ symmetric matrix A is such that
$$0 < \mathbf{x}^T \mathbf{A}\mathbf{x}$$
for all $\mathbf{x}$, both the matrix and the quadratic form are said to be **positive definite.**
- Using the spectral decomposition: a *kxk* matrix A is a positive definite matrix if and only if every eigenvalue of A is positive.
- A is a nonnegative definite matrix if and only if all of its eigenvalues are greater than or equal to zero.

Positive Definite Matrices

- Remember the spectral decomposition: For a positive definite matrix $\mathbf{A}$:

$$\mathbf{A} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \ldots + \lambda_p \mathbf{e}_p \mathbf{e}_p^T = \sum_{i=1}^{p} \lambda_i \mathbf{e}_i \mathbf{e}_i^T = \mathbf{P} \Lambda \mathbf{P}^T,$$

where $\mathbf{P}_{\{pxp\}} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$ is an orthogonal matrix and Λ is the diagonal matrix

$$\Lambda = \begin{bmatrix} \lambda_{11} & 0 & \ldots & 0 \\ 0 & \lambda_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{pp} \end{bmatrix}$$

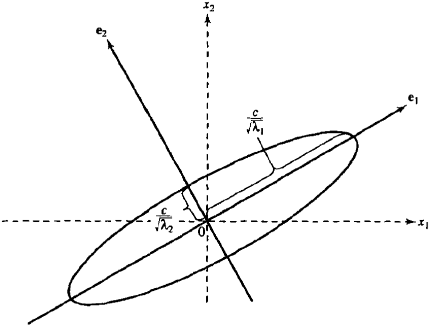

Figure 2.6 Points a constant distance c from the origin ($p = 2, 1 \leq \lambda_1 < \lambda_2$).

If $p > 2$, the points $\mathbf{x}' = [x_1, x_2, \ldots, x_p]$ a constant distance $c = \sqrt{\mathbf{x}'\mathbf{A}\mathbf{x}}$ from the origin lie on hyperellipsoids $c^2 = \lambda_1(\mathbf{x}'\mathbf{e}_1)^2 + \cdots + \lambda_p(\mathbf{x}'\mathbf{e}_p)^2$, whose axes are given by the eigenvectors of $\mathbf{A}$. The half-length in the direction $\mathbf{e}_i$ is equal to $c/\sqrt{\lambda_i}$, $i = 1, 2, \ldots, p$, where $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues of $\mathbf{A}$.

Positive Definite Matrices

- So $\mathbf{P}\Lambda\mathbf{P}^T$, then

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

Positive Definite Matrices

- So $\mathbf{P}\Lambda\mathbf{P}^T$, then

$$\mathbf{A}^{-1} = \mathbf{P}\Lambda^{-1}\mathbf{P}^T = \sum_{i=1}^{p} \frac{1}{\lambda_i}\mathbf{e}_i\mathbf{e}_i^T$$

since

$$(\mathbf{P}\Lambda^{-1}\mathbf{P}^T)\mathbf{P}\Lambda\mathbf{P}^T = \mathbf{P}\Lambda\mathbf{P}^T(\mathbf{P}\Lambda^{-1}\mathbf{P}^T) = \mathbf{P}\mathbf{P}^T = I$$

Positive Definite Matrices

- So $\mathbf{P}\Lambda\mathbf{P}^T$, then

$$\mathbf{A}^{-1} = \mathbf{P}\Lambda^{-1}\mathbf{P}^T = \sum_{i=1}^{p} \frac{1}{\lambda_i}\mathbf{e}_i\mathbf{e}_i^T$$

  since

$$(\mathbf{P}\Lambda^{-1}\mathbf{P}^T)\mathbf{P}\Lambda\mathbf{P}^T = \mathbf{P}\Lambda\mathbf{P}^T(\mathbf{P}\Lambda^{-1}\mathbf{P}^T) = \mathbf{P}\mathbf{P}^T = I$$

In R: is.positive.definite(x, tol=1e-8)

Positive Definite Matrices

- Consider now $\Lambda^{1/2}$ with diagonal elements $\sqrt{\lambda_i}$.

$$\mathbf{A}^{1/2} = \sum_{i=1}^{p} \sqrt{\lambda_i}\mathbf{e}_i\mathbf{e}_i^T = \mathbf{P}\Lambda^{1/2}\mathbf{P}^T$$

- Properties:
- $(\mathbf{A}^{1/2})^T = \mathbf{A}^{1/2}$

Positive Definite Matrices

- Consider now $\Lambda^{1/2}$ with diagonal elements $\sqrt{\lambda_i}$.

$$\mathbf{A}^{1/2} = \sum_{i=1}^{p} \sqrt{\lambda_i}\mathbf{e}_i\mathbf{e}_i^T = \mathbf{P}\Lambda^{1/2}\mathbf{P}^T$$

- Properties:
- $(\mathbf{A}^{1/2})^T = \mathbf{A}^{1/2}$
- $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$

Positive Definite Matrices

- Consider now $\Lambda^{1/2}$ with diagonal elements $\sqrt{\lambda_i}$.

$$\mathbf{A}^{1/2} = \sum_{i=1}^{p} \sqrt{\lambda_i}\mathbf{e}_i\mathbf{e}_i^T = \mathbf{P}\Lambda^{1/2}\mathbf{P}^T$$

- Properties:
- $(\mathbf{A}^{1/2})^T = \mathbf{A}^{1/2}$
- $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$
- $(\mathbf{A}^{1/2})^{-1} = \mathbf{P}\mathbf{A}^{-1/2}\mathbf{P}^T$

Positive Definite Matrices

- Consider now $\Lambda^{1/2}$ with diagonal elements $\sqrt{\lambda_i}$.

$$\mathbf{A}^{1/2} = \sum_{i=1}^{p} \sqrt{\lambda_i}\mathbf{e}_i\mathbf{e}_i^T = \mathbf{P}\Lambda^{1/2}\mathbf{P}^T$$

- Properties:
- $(\mathbf{A}^{1/2})^T = \mathbf{A}^{1/2}$
- $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$
- $(\mathbf{A}^{1/2})^{-1} = \mathbf{P}\mathbf{A}^{-1/2}\mathbf{P}^T$
- $\mathbf{A}^{1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1/2}\mathbf{A}^{1/2} = I$

Positive Definite Matrices

- Consider now $\Lambda^{1/2}$ with diagonal elements $\sqrt{\lambda_i}$.

$$\mathbf{A}^{1/2} = \sum_{i=1}^{p} \sqrt{\lambda_i}\mathbf{e}_i\mathbf{e}_i^T = \mathbf{P}\Lambda^{1/2}\mathbf{P}^T$$

- Properties:
- $(\mathbf{A}^{1/2})^T = \mathbf{A}^{1/2}$
- $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$
- $(\mathbf{A}^{1/2})^{-1} = \mathbf{P}\mathbf{A}^{-1/2}\mathbf{P}^T$
- $\mathbf{A}^{1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1/2}\mathbf{A}^{1/2} = I$
- $\mathbf{A}^{-1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1}$

See Homework file