1. WORKING WITH MATRICES AND VECTOR

Chapter 1 - 3 of Rancher & Schaalje

Defn 1.1: A column of real numbers is called a **vector**.

Examples:

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \mathbf{a} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} \quad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Since **Y** has n elements it is said to have **order** (or dimension) n.

Defn 1.2: A rectangular array of elements with m rows and k columns is called an $m \times k$ **matrix.**

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix}$$

This matrix is said to be of **order** (or dimension) $m \times k$ where

m is the row order (dimension)

k is the column order (dimension)

Examples:

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 0 & 4 & 5 \end{bmatrix}$$

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$$

<u>Defn 1.3:</u> **Matrix addition:** If A and B are both $m \times k$ matrices, then

$$
\begin{aligned}
C &= A + B \\
&= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mk} \end{bmatrix} \\
&= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1k}+b_{1k} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2k}+b_{2k} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mk}+b_{mk} \end{bmatrix}
\end{aligned}
$$

Notation: $C_{m \times k} = \{c_{ij}\}$ where $c_{ij} = a_{ij} + b_{ij}$

Defn 1.4: **Matrix subtraction**

If A and B are $m \times k$ matrices, then $C = A - B$ is defined by

$$C = \{c_{ij}\} \text{ where } c_{ij} = a_{ij} - b_{ij} \; .$$

Examples:

$$\begin{bmatrix} 3 & 6 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 7 & -4 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ -1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}$$

Defn 1.5: **Scalar multiplication**

Let a be a scalar and $B = \{b_{ij}\}$ be an $m \times k$ matrix, then

$$aB = Ba = \{a\,b_{ij}\}$$

Example:

$$2\begin{bmatrix} 2 & -1 & 3 \\ 0 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 4 & -2 & 6 \\ 0 & 8 & -4 \end{bmatrix}$$

Defn 1.6: **Transpose**

The transpose of the $m \times k$ matrix $A = \{a_{ij}\}$ is the $k \times m$ matrix with elements $\{a_{ji}\}$. The transpose of A is denoted by A^T (or A').

Example:

$$A = \begin{bmatrix} 1 & 4 \\ 3 & 0 \\ -2 & 6 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 3 & -2 \\ 4 & 0 & 6 \end{bmatrix}$$

Defn 1.7: If a matrix has the same number of rows and columns it is called a **square matrix**.

$$A_{k \times k} = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}$$

is said to have order (or dimension) k.

Defn 1.8: A square matrix $A = \{a_{ij}\}$ is **symmetric** if $A = A^T$, that is, if $a_{ij} = a_{ji}$ for all (i, j).

Examples:

$$A = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$$

$$B = \begin{bmatrix} 4 & 3 & 2 & 1 \\ 3 & 5 & 0 & -2 \\ 2 & 0 & 3 & -1 \\ 1 & -2 & -1 & 2 \end{bmatrix}$$

Defn 1.9: **Inner product** (crossproduct) of two vectors of order n

$$\mathbf{a}^T\mathbf{Y} = [a_1, a_2, \cdots a_n] \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

$$= a_1\, Y_1 + a_2\, Y_2 + \cdots + a_n\, Y_n$$

$$= \sum_{j=1}^{n} a_j\, Y_j$$

Note that $\mathbf{a}^T\, \mathbf{Y} = \mathbf{Y}^T\, \mathbf{a}$.

Defn 1.10: **Euclidean distance** (or length of a vector)

$$\|\mathbf{Y}\| = (\mathbf{Y}^T\, \mathbf{Y})^{1/2} = \left(\sum_{j=1}^{n} Y_j^2 \right)^{1/2}$$

Defn 1.11: **Matrix multiplication**

The product of an $n \times k$ matrix A and a $k \times m$ matrix B is the $n \times m$ matrix $C = \{c_{ij}\}$ with elements

$$c_{ij} = a_{i1}\, b_{1j} + a_{i2}\, b_{2j} + \cdots + a_{ik}\, b_{kj}$$

Example:

$$A = \begin{bmatrix} 3 & 0 & -2 \\ 1 & -1 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$$

$$C = AB = \begin{bmatrix} 1 & -3 \\ 4 & 11 \end{bmatrix}$$

Defn 1.12: **Elementwise multiplication** of two matrices

$$A \# B = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{km} \end{bmatrix} \# \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{k1} & \cdots & b_{km} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}\, b_{11} & \cdots & a_{1m}\, b_{1m} \\ \vdots & & \vdots \\ a_{k1}\, b_{k1} & \cdots & a_{km}\, b_{km} \end{bmatrix}$$

Example

$$\begin{bmatrix} 3 & 1 \\ 2 & 4 \\ 0 & 6 \end{bmatrix} \# \begin{bmatrix} 1 & -5 \\ -3 & 4 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ -6 & 16 \\ 0 & 12 \end{bmatrix}$$

Defn 1.13: **Kronecker product** of two matrices

$$A_{k\times m}\otimes B_{n\times s}=\begin{bmatrix} a_{11}\,B & a_{12}\,B & \cdots & a_{1m}\,B \\ a_{21}\,B & a_{22}\,B & \cdots & a_{2m}\,B \\ \vdots & \vdots & & \vdots \\ a_{k1}\,B & a_{k2}\,B & \cdots & a_{km}\,B \end{bmatrix}$$

Examples:

$$\mathbf{a}\otimes\mathbf{Y}=\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}\otimes\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}=\begin{bmatrix} a_1\,Y_1 \\ a_1\,Y_2 \\ a_2\,Y_1 \\ a_2\,Y_2 \\ a_3\,Y_1 \\ a_3\,Y_2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 4 \\ 0 & -2 \\ 3 & -1 \end{bmatrix} \otimes \begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 20 & 12 \\ 4 & 2 & 8 & 4 \\ 0 & 0 & -10 & -6 \\ 0 & 0 & -4 & -2 \\ 15 & 9 & -5 & -3 \\ 6 & 3 & -2 & -1 \end{bmatrix}$$

Refer the handout slide1_r1.pdf.

Defn 1.14: The **determinant** of an $n \times n$ matrix A is

$$|A| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j} |M_{ij}| \quad \text{for any row } i$$

or

$$|A| = \sum_{i=1}^{n} a_{ij}(-1)^{i+j} |M_{ij}| \quad \text{for any column } j$$

where M_{ij} is the “minor” for a_{ij} obtained by deleting the i-th row and j-th column from A.

Example:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$|A| = a_{11}(-1)^{1+1}|a_{22}| + a_{12}(-1)^{1+2}|a_{21}|$$

$$\begin{vmatrix} 7 & 2 \\ 4 & 5 \end{vmatrix} = (7)(5) - (2)(4) = 27$$

Example:

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}(-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

$$+ a_{12}(-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}(-1)^{1+3} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

then

$$
\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = (1)(-1)^2 \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix}
$$

$$
+ (2)(-1)^3 \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + (3)(-1)^4 \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix}
$$

$$
= (1)(-3) - (2)(-6) + (3)(-3) = 0
$$

Example:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{vmatrix} = -3$$

Properties of determinants:

(i) $|A^T| = |A|$

(ii) $|A|$ = product of the eigenvalues of A

(iii) $|AB| = |A|\,|B|$ when A and B are square matrices of the same order.

(iv) $\begin{vmatrix} P & 0 \\ X & Q \end{vmatrix} = |P|\,|Q|$ when P and Q are square matrices of the same order and 0 is a matrix of zeros.

(v) $|AB| = |BA|$ when the matrix product is defined

(vi) $|cA| = c^k|A|$ when c is a scalar and A is a $k \times k$ matrix

Defn 1.15: A set of n-dimensional vectors $\mathbf{Y}_1 \; \mathbf{Y}_2 \cdots \mathbf{Y}_k$ are **linearly independent** if there is no set of scalars $a_1 \; a_2 \; \cdots \; a_k$ such that

$$\mathbf{0} = \sum_{j=1}^{k} a_j \, \mathbf{Y}_j$$

and at least one a_j is non-zero.

Example:

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

are linearly independent.

Example:

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

are not linearly independent because

$$(1)\,\mathbf{Y}_1 + (1)\,\mathbf{Y}_3 + (-2)\,\mathbf{Y}_2 = \mathbf{0}$$

Any two of these vectors are linearly independent, and it is said that this set contains two linearly independent vectors.

<ins>Defn 1.16:</ins> The **row rank** of a matrix is the number of linearly independent rows, where each row is considered as a vector.

Example:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

The row rank of A is 2 because

$$(-2)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + (1)\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + (-3)\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and there are no scalars a_1 and a_2 such that

$$a_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

except for $a_1 = a_2 = 0$.

Defn 1.17: The **column rank** of a matrix is the number of linearly independent columns, with each column considered as a vector.

Example:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

has column rank 2 because

$$(-2)\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + (1)\begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix} + (1)\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and there are no scalars a_1 and a_2 such that

$$a_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

except $a_1 = a_2 = 0$.

Result 1.1: The row rank and the column rank of a matrix are equal.

Defn 1.18 The **rank** of a matrix is either the row rank or the column rank of the matrix.

Defn 1.19: A square matrix $A_{k\times k}$ is **nonsingular** if its rank is equal to the number of rows (or columns).

This is equivalent to the condition

$$A_{k\times k}\,\mathbf{b}_{k\times 1} = \mathbf{0}_{k\times 1} \text{ only when } \mathbf{b} = \mathbf{0}$$

A matrix that fails to be nonsingular is called **singular**.

Result 1.2: If B and C are non-singular matrices and products with A are defined, then

$$rank(BA) = rank(AC) = rank(A).$$

Result 1.3:

$$\begin{aligned} rank(A^T A) &= rank(AA^T) \\ &= rank(A) \\ &= rank(A^T). \end{aligned}$$

Defn 1.20: The **identity matrix**, denoted by I, is a $k \times k$ matrix of the form

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

Defn 1.21: The **inverse** of a square, non-singular matrix A is the matrix, denoted by A^{-1}, such that

$$AA^{-1} = A^{-1}A = I$$

Example

$$\begin{bmatrix} 2 & 4 \\ 1 & 6 \end{bmatrix}^{-1} = \begin{bmatrix} 6/8 & -4/8 \\ -1/8 & 2/8 \end{bmatrix}$$

Result 1.4

(i) The inverse of $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is

$$A^{-1} = \frac{1}{|A|} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

(ii) In general, the (i, j) element of A^{-1} is

$$\frac{(-1)^{i+j} \, |A_{ji}|}{|A|}$$

where A_{ji} is the matrix obtained by deleting the j-th row and i-th column of A.

Result 1.5: For a $k \times k$ matrix A, the following are equivalent:

(i) A is nonsingular

(ii) $|A| \neq 0$

(iii) A^{-1} exists

Result 1.6: For $k \times k$ nonsingular matrices A and B

(i) $(A^T)^{-1} = (A^{-1})^T$

(ii) $(AB)^{-1} = B^{-1}A^{-1}$

(iii) $|A^{-1}| = 1/|A|$

(iv) A^{-1} is unique and nonsingular

(v) $(A^{-1})^{-1} = A$

(vi) If A is symmetric, than A^{-1} is symmetric

Result 1.7: Inverse of a Diagonal Matrix

$$\begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{kk} \end{bmatrix}^{-1} = \begin{bmatrix} 1/a_{11} & & & \\ & 1/a_{22} & & \\ & & \ddots & \\ & & & 1/a_{kk} \end{bmatrix}$$

Result 1.8: If B is a $k \times k$ non-singular matrix and $B + \mathbf{c}\mathbf{c}^T$ is non-singular, then

$$\left(B + \mathbf{c}\mathbf{c}^T\right)^{-1} = B^{-1} - \frac{B^{-1}\mathbf{c}\mathbf{c}^T B^{-1}}{1 + \mathbf{c}^T B^{-1}\mathbf{c}}$$

Result 1.9: Let I_n be an $n \times n$ identity matrix and let $J_n = \mathbf{1}\mathbf{1}^T$ be an $n \times n$ matrix where each element is one, then

$$(a\,I_n + b\,J_n)^{-1} = \frac{1}{a}\left(I_n - \frac{b}{a+nb}J_n\right)$$

Defn 1.22: The **trace** of a $k \times k$ matrix $A = \{a_{ij}\}$ is the sum of the diagonal elements:

$$tr(A) = \sum_{j=1}^{k} a_{jj}$$

Result 1.10 Let A and B denote $k \times k$ matrices and let c be a scalar. Then,

(i) $tr(c\,A) = c\,tr(A)$

(ii) $tr(A + B) = tr(A) + tr(B)$

(iii) $tr(AB) = tr(BA)$

(iv) $tr(B^{-1}A\,B) = tr(A)$

(v) $tr(A\,A^T) = \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}^2$

Refer the handout slide1_r2.pdf.

Defn 1.23: A square matrix A is said to be **orthogonal** if

$$A A^T = A^T A = I$$

(then $A^{-1} = A^T$)

Examples:

$$A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}, \quad A = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{bmatrix}$$

In each case the columns of A are coefficients for orthogonal contrasts.

Defn 1.24: A square matrix P is **idempotent** if $PP = P$

Example

$$P = \begin{bmatrix} 56 & 26 & -16 \\ 26 & 26 & 26 \\ -16 & 26 & 56 \end{bmatrix}$$

Example (linear regression): $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$

The least squares estimator is $\mathbf{b} = (X^T X)^{-1} X^T \mathbf{Y}$

The estimated means are $\hat{\mathbf{Y}} = X(X^T X)^{-1} X^T \mathbf{Y}$

and the residuals are $\mathbf{e} = (I - X(X^T X)^{-1} X^T)\mathbf{Y}$.

Both $X(X^T X)^{-1} X^T$ and $I - X(X^T X)^{-1} X^T$ are idempotent matrices.

Defn 1.25: Let A be a $k \times k$ matrix and let $\mathbf{Y}$ be a vector of order k, then

$$\mathbf{Y}^T A \mathbf{Y} = \sum_{i=1}^{k} \sum_{j=1}^{k} Y_i \, Y_j \, a_{ij}$$

is called a **quadratic form.**

Defn 1.26: A $k \times k$ matrix A is said to be **positive definite** if

$$\mathbf{Y}^T A \mathbf{Y} > 0$$

for any $\mathbf{Y} = (Y_1, \ldots, Y_k)^T \neq \mathbf{0}$.

Defn 1.27: A $k \times k$ matrix A is said to be **non-negative definite** (or positive semi-definite) if

$$\mathbf{Y}^T A \mathbf{Y} \geq 0$$

for any $\mathbf{Y} = (Y_1, \ldots, Y_k)^T$.

Eigenvalues and Eigenvectors

Defn 1.28: For a $k \times k$ matrix A, the scalars $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ satisfying the polynomial equation

$$|A - \lambda I| = 0$$

are called the eigenvalues (or characteristic roots) of A.

Defn 1.29: Corresponding to any eigenvalue λ_i is an eigenvector (or characteristic vector) $\mathbf{u}_i \neq \mathbf{0}$ satisfying

$$A\,\mathbf{u}_i = \lambda_i\,\boldsymbol{\mu}_i.$$

Comment: Eigenvectors are not unique

(i) If $\mathbf{u}_i$ is an eigenvector for λ_i, then $c\,\mathbf{u}_i$ is also an eigenvector for any scalar $c \neq 0$.

(ii) We will adopt the following conventions (for real symmetric matrices)

$$\begin{aligned} \mathbf{u}_i^T\mathbf{u}_i &= 1 \qquad \text{for all } i = 1, \ldots, k \\ \mathbf{u}_i^T\mathbf{u}_j &= 0 \qquad \text{for all } i \neq j \end{aligned}$$

(iii) Even with (ii), eigenvectors are not unique

- If $\mathbf{u}_i$ is an eigenvector satisfying (ii), then $-\mathbf{u}_i$ is also an eigenvector satisfying (ii).

- If $\lambda_i = \lambda_j$ then there are an infinite number of choices for $\mathbf{u}_i$ and $\mathbf{u}_j$.

Example:

$$A = \begin{bmatrix} 1.96 & 0.72 \\ 0.72 & 1.54 \end{bmatrix}$$

Eigenvalues are solutions to

$$\begin{aligned} 0 &= \left| \begin{bmatrix} 1.96 & 0.72 \\ 0.72 & 1.54 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| \\ &= \begin{vmatrix} 1.96-\lambda & 0.72 \\ 0.72 & 1.54-\lambda \end{vmatrix} \\ &= (1.96-\lambda)(1.54-\lambda)-(0.72)^2 \\ &= \lambda^2 - 3.5\lambda + 2.5 = a\lambda^2 + b\lambda + c \end{aligned}$$

Solutions to a quadratic equation:

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad \Rightarrow \quad \frac{3.5 \pm \sqrt{12.25 - 10}}{2}$$

$$\Rightarrow \quad \lambda_1 = 2.5 \text{ and } \lambda_2 = 1$$

Find the eigenvectors: $A\boldsymbol{\mu}_i = \lambda_i \boldsymbol{\mu}_i$

$$\begin{bmatrix} 1.96 & 0.72 \\ 0.72 & 1.54 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix} = 2.5 \begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix}$$

$$\Rightarrow \begin{array}{l} 1.96\,u_{11} + 0.72\,u_{12} = 2.5\,u_{11} \\ 0.72\,u_{11} + 1.54\,u_{12} = 2.5\,u_{12} \end{array}$$

$$\Rightarrow \quad u_{12} = 0.75\,u_{11}$$

then

$$\mathbf{u}_1 = \begin{bmatrix} c \\ 0.75\,c \end{bmatrix}$$

To satisfy our convention we must have

$$1 \quad = \quad \mathbf{u}_1^T \mathbf{u}_1 = c^2 + 0.5625\, c^2$$

Consequently,

$$c = 0.8 \text{ or } c = -0.8$$

then

$$\mathbf{u}_1 = \begin{bmatrix} 0.8 \\ 0.6 \end{bmatrix} \text{ or } \mathbf{u}_1 = \begin{bmatrix} -0.8 \\ -0.6 \end{bmatrix}$$

Find an eigenvector for $\lambda_2 = 1$

$$\begin{bmatrix} 1.96 & 0.72 \\ 0.72 & 1.54 \end{bmatrix} \begin{bmatrix} u_{21} \\ u_{22} \end{bmatrix} = (1) \begin{bmatrix} u_{21} \\ u_{22} \end{bmatrix}$$

$$\Rightarrow \begin{array}{l} 1.96\, u_{21} + 0.72\, u_{22} = u_{21} \\ 0.72\, u_{21} + 1.54\, u_{22} = u_{22} \end{array}$$

$$\Rightarrow \quad u_{22} = \tfrac{-4}{3}\, u_{21}$$

Then

$$\mathbf{u}_2 = \begin{bmatrix} c \\ -4/3\, c \end{bmatrix}$$

To satisfy our convention, we must have

$$1 = \mathbf{u}_2^T \mathbf{u}_2 = c^2 + \frac{16\, c^2}{9}$$

Consequently,

$$c = -0.6 \;\text{ or }\; c = 0.6$$

and

$$\mathbf{u}_2 = \begin{bmatrix} -0.6 \\ 0.8 \end{bmatrix} \quad \text{or} \quad \mathbf{u}_2 = \begin{bmatrix} 0.6 \\ -0.8 \end{bmatrix}$$

In either case, $\mathbf{u}_1^T \mathbf{u}_2 = 0$.

Result 1.11 For a $k \times k$ symmetric matrix A with elements that are real numbers

(i) every eigenvalue of A is a real number

(ii) rank(A) = number of non-zero eigenvalues

(iii) if A is non-negative definite, then $\lambda_i \geq 0$ for all $i = 1, 2, \ldots, k$

(iv) if A is positive definite then $\lambda_i > 0$ for all $i = 1, 2, \ldots, k$

(v) trace(A) = $\sum_{i=1}^{k} a_{ii} = \sum_{i=1}^{k} \lambda_i$

(vi) $|A| = \Pi_{i=1}^{k} \lambda_i$

(vii) if A is idempodent ($A\,A = A$), then the eigenvalues are either zero or one.

Result 1.12: **Spectral decomposition.**

The spectral decomposition of a $k \times k$ symmetric matrix A with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$ and eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ (with $\mathbf{u}_i^T \mathbf{u}_i = 1$ and $\mathbf{u}_i^T \mathbf{u}_j = 0$) is

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_k \mathbf{u}_k \mathbf{u}_k^T = U D U^T$$

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_k \end{bmatrix}$$

and $U = [\mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_k]$ is an orthogonal matrix.

Result 1.13: If A is a $k \times k$ symmetric nonsingular matrix with spectral decomposition

$$A = \sum_{i=1}^{k} \lambda_i \, \mathbf{u}_i \, \mathbf{u}_i^T = U \, D \, U^T$$

then

(i) $A^{-1} = \sum_{i=1}^{k} \lambda_i^{-1} \, \mathbf{u}_i \, \mathbf{u}_i^T = U \, D^{-1} \, U^T$

(ii) the square root matrix

$$A^{1/2} = \sum_{i=1}^{k} \sqrt{\lambda_i} \, \mathbf{u}_i \, \mathbf{u}_i^T$$

has the properties:

(a) $A^{1/2} \, A^{1/2} = A$, (b) $A^{1/2} \, A^{-1} \, A^{1/2} = I$, (c) $A^{1/2}$ is symmetric

(iii) The inverse square root matrix

$$
\begin{aligned}
A^{-1/2} &= \sum_{i=1}^{k} \frac{1}{\sqrt{\lambda_i}} \mathbf{u}_i \mathbf{u}_i^T \\
&= U D^{-1/2} U^T
\end{aligned}
$$

has the properties:

(a) $A^{-1/2} A^{-1/2} = A^{-1}$

(b) $A^{-1/2} A A^{-1/2} = I$

(c) $A^{-1/2}$ is symmetric

In parts (ii) and (iii), A should be positive definite to ensure that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$$

Result 1.14: Singular value decomposition

Any $p \times q$ matrix A of rank r can be expressed as

$$A = L \begin{bmatrix} \Delta & 0 \\ 0 & 0 \end{bmatrix} M^T$$

where

(i) $L_{p \times p}$ and $M_{q \times q}$ are orthogonal matrices

(ii) $\Delta_{r \times r}$ is a diagonal matrix with $\Delta^2 = \Delta\,\Delta$ containing the positive (non-zero) eigenvalues of $A^T A$ and $A\,A^T$

Note that $A^T A$ and $A\,A^T$ are non-negative definite and suitable L and M matrices can always be found but they are not unique.

Refer the handout slide1_r3.pdf.