1. Working with Matrices and Vectors

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 $\mathbf{Y}$ 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:

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

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

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

Defn 1.3 **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:

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

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

$$a\,B = B\,a = \{a\,b_{ij}\}$$

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

$$\begin{aligned} \mathbf{a}^T\mathbf{Y} &= [a_1, a_2, \cdots a_n] \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \\ &= a_1 Y_1 + a_2 Y_2 + \cdots + a_n Y_n \\ &= \sum_{j=1}^{n} a_j Y_j \end{aligned}$$

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

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

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:

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

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


```
# This code is stored in the file
#
#          matrix.ssc
#


#------------------------------------
# Add and subtract matrices
#------------------------------------


> a <-matrix(c(3, 6, 2, 1),2,2,byrow=T)
> a
        [,1]  [,2]
  [1,]     3     6
  [2,]     2     1

> b <-matrix(c(7, -4, -3, 2),2,2,byrow=T)
> b
        [,1]  [,2]
  [1,]     7    -4
  [2,]    -3     2
```



```
> a+b
        [,1]  [,2]
  [1,]    10     2
  [2,]    -1     3

> a-b
        [,1]  [,2]
  [1,]    -4    10
  [2,]     5    -1
```



```
#---------------------------------
# Multiplication by a scalar
#---------------------------------


> c <-matrix(c(2, -1, 3, 0,
        4, -2), 2, 3, byrow=T)
> c
        [,1]  [,2]  [,3]
  [1,]     2    -1     3
  [2,]     0     4    -2

> d <-2*c
> d
        [,1]  [,2]  [,3]
  [1,]     4    -2     6
  [2,]     0     8    -4
```

#——————————————
Transpose of a matrix
#——————————————

```
> ct <-t(c)
> ct
      [,1] [,2]
[1,]     2    0
[2,]    -1    4
[3,]     3   -2
```


#——————————————-
Matrix multiplication
#——————————————-

```
> a <-matrix(c(3, 0, -2, 1, -1, 4),
      2,3,byrow=T)
> a
      [,1] [,2] [,3]
[1,]     3    0   -2
[2,]     1   -1    4

> b <-matrix(c(1,1,1,2,1,3), 3,2,byrow=T)
> b
      [,1] [,2]
[1,]     1    1
[2,]     1    2
[3,]     1    3

> c <-a %*% b
> c
      [,1] [,2]
[1,]     1   -3
[2,]     4   11
```


#———————————
Inner product
#———————————

```
> x <-c(1,7,-6,4)

> y <-c(2,-2,1,5)

> x
[1]  1 7 -6 4

> y
[1] 2 -2 1 5

> t(x)%*%y
     [,1]
[1,] 2

> x%*%y
     [,1]
[1,] 2
```



```
> crossprod(x,y)
     [,1]
[1,] 2
```

#———————————
Length of a vector
#———————————

```
> ynorm<-sqrt(crossprod(y,y))
> ynorm
     [,1]
[1,] 5.830952
```

#————————————————-
Number of elements in a vector
#————————————————-

```
> length(y)
[1] 4
```

```
#————————————————————
# Elementwise multiplication
#————————————————————

> a <-matrix(c(3, 6, 2, 1),2,2,byrow=T)
> a
      [,1] [,2]
[1,]     3    6
[2,]     2    1

> b <-matrix(c(7, -4, -3, 2),2,2,byrow=T)
> b
      [,1] [,2]
[1,]     7   -4
[2,]    -3    2

> a*b
      [,1] [,2]
[1,]    21  -24
[2,]    -6    2
```



```
# ———————————————
# Kronecker Product)
# ———————————————

> a<-matrix(c(2,4,0,-2,3,-1),
      ncol=2,byrow=T)
> a
      [,1] [,2]
[1,]     2    4
[2,]     0   -2
[3,]     3   -1

> b<-matrix(c(5,3,2,1),2,2,byrow=T)
> b
      [,1] [,2]
[1,]     5    3
[2,]     2    1

> kronecker(a,b)
      [,1] [,2] [,3] [,4]
[1,]    10    6   20   12
[2,]     4    2    8    4
[3,]     0    0  -10   -6
[4,]     0    0   -4   -2
[5,]    15    9   -5   -3
[6,]     6    3   -2   -1
```



```
#————————————————————————
# What happens when the dimensions
# of the matrices or vectors are
# not appropriate for the operation
#————————————————————————

> a<-matrix(c(1, 1, 1, 2), 2, 2, byrow=T)
> b<-matrix(c(3, 0, -2, 1, -1, 4), 2, 3,
        byrow=T)
> a
      [,1] [,2]
[1,]     1    1
[2,]     1    2
> b
      [,1] [,2] [,3]
[1,]     3    0   -2
[2,]     1   -1    4

> a+b
Error in a + b: Dimension attributes do not
match

> b+a
Error in b + a: Dimension attributes do not
match
```



```
> a%*%b
      [,1] [,2] [,3]
[1,]     4   -1    2
[2,]     5   -2    6

> b%*%a
Error in "%*%.default"(b, a): Number of
columns of x should be the same as number
of rows of y

> a*b
Error in a * b: Dimension attributes do not
match
```

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.

Defn 1.16: 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, nonsingular matrix A is the matrix, denoted by A^{-1}, such that

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

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

(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

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

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

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

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


```
# This script is also stored in
#
#       matrix.ssc


# ------------------------------------
# Create an nxn identity matrix
# ------------------------------------


> diag(rep(1,4))
      [,1] [,2] [,3] [,4]
 [1,]    1    0    0    0
 [2,]    0    1    0    0
 [3,]    0    0    1    0
 [4,]    0    0    0    1
```



```
#--------------------------
# Trace of a matrix
#--------------------------


> w<-matrix(c(1,2,3,4,5,6,7,8,10),
     3,3,byrow=T)


> w
      [,1] [,2] [,3]
 [1,]    1    2    3
 [2,]    4    5    6
 [3,]    7    8   10

> tr<-sum(diag(w))
> tr
[1] 16
```



```
#---------------------------------
#      Inverse of a matrix
#---------------------------------


> w<-matrix(c(1,2,3,4,5,6,7,8,10),
     3,3,byrow=T)
> w
      [,1] [,2] [,3]
 [1,]    1    2    3
 [2,]    4    5    6
 [3,]    7    8   10

> winv<-solve(w)
> winv
            [,1]       [,2] [,3]
 [1,] -0.6666667  -1.333333    1
 [2,] -0.6666667   3.666667   -2
 [3,]  1.0000000  -2.000000    1

> w%*%winv
              [,1]          [,2]          [,3]
 [1,] 1.000000e+00  4.440892e-15 -2.664535e-15
 [2,] 8.881784e-16  1.000000e+00  8.881784e-16
 [3,] 0.000000e+00  0.000000e+00  1.000000e+00
```

```
# ————————————————
# Determinant of a matrix
# ————————————————

# Build your own function

> determ<-function(M)  Re(prod(
    eigen(M, only.values=T)$values))

> determ(w)
[1] -3

# Another function (V&R, page 101)

> absdet <- function(M)
      abs(prod(diag(qr(M)$qr)))

> absdet(w)
    [1] 3
```



```
# Another example

> x1 <- matrix(c(1,2,3,4,5,6,7,8,9),
      ncol=3,byrow=T)
> x1
      [,1] [,2] [,3]
 [1,]    1    2    3
 [2,]    4    5    6
 [3,]    7    8    9

> determ(x1)
[1] 3.154999e-15

> absdet(x1)
[1] 1.631688e-15
```



```
#————————————————————
# Rank of a matrix: use the "qr"
# function (V&R on p.101)
#————————————————————

> A <- matrix(c(1, 1, 1,
2, 5, -1,
0, 1, -1),3,3,byrow=T)

> A
      [,1] [,2] [,3]
 [1,]    1    1    1
 [2,]    2    5   −1
 [3,]    0    1   −1

> qr(A)$rank
[1] 2
```



```
# Another example

> A <- matrix(c(1,1, 1,
+                2,5,-1,
+                0,1, 1),3,3,byrow=T)

> A
      [,1] [,2] [,3]
 [1,]    1    1    1
 [2,]    2    5   −1
 [3,]    0    1    1

> qr(A)$rank
[1] 3
```

```
# Another example

> X <- matrix(c(1,1,0,0,
+              1,1,0,0,
+              1,0,1,0,
+              1,0,1,0,
+              1,0,0,1,
+              1,0,0,1),ncol=4,byrow=T)


> X
     [,1] [,2] [,3] [,4]
[1,]    1    1    0    0
[2,]    1    1    0    0
[3,]    1    0    1    0
[4,]    1    0    1    0
[5,]    1    0    0    1
[6,]    1    0    0    1

> qr(X)$rank
[1] 3
```



```
# Note that the sum of squares
# and crossproducts matrix has
# the same rank as X


> XtX ¡- t(X)%*%X
> XtX
     [,1] [,2] [,3] [,4]
[1,]    6    2    2    2
[2,]    2    2    0    0
[3,]    2    0    2    0
[4,]    2    0    0    2

> qr(XtX)$rank  [1] 3

# This is a square symmetric matrix
# but the inverse does not exist


> solve(XtX)
Problem in solve.qr(a): apparently
singular matrix
Use traceback() to see the call stack
```



```
# Note that the function "rank" in Splus
# is related to sorting. It computes the
# ranks of the elements of a vector.
# (V&R on page 45)


> rank(c(1.2, 5.1, 3.5, 9.8))
[1] 1 3 2 4
```



```
# ————————————————————-
# Create an nxn identity matrix
# ————————————————————-


> I4<-diag(rep(1,4))
> I4
     [,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]    0    1    0    0
[3,]    0    0    1    0
[4,]    0    0    0    1
```

```
#----------------------------
# Trace of a matrix
#----------------------------

> w<-matrix(c(1,2,3,4,5,6,7,8,10),
            3,3,byrow=T)
> w
      [,1]  [,2]  [,3]
[1,]     1     2     3
[2,]     4     5     6
[3,]     7     8    10

> tr<-sum(diag(w))
> tr
[1] 16
```



```
#-----------------------------------------
# Compute row sums or column sums
#-----------------------------------------

> sum(w)
[1] 46

> apply(w,1,sum)
[1] 6 15 25

> apply(w,2,sum)
[1] 12 15 19

> apply(w,1,prod)
[1] 6 120 560

> apply(w,1,mean)
[1] 2.000000 5.000000 8.333333

> apply(w,1,var)
[1] 1.000000 1.000000 2.333333
```


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:

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

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

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

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

Example

$$P = \begin{bmatrix} \frac{5}{6} & \frac{2}{6} & -\frac{1}{6} \\ \frac{2}{6} & \frac{2}{6} & \frac{2}{6} \\ -\frac{1}{6} & \frac{2}{6} & \frac{5}{6} \end{bmatrix}$$

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 idempodent 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

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

$\uparrow$ 1 $\quad\uparrow$ -3.5 $\quad\uparrow$ 2.5

Solutions to a quadratic equation:

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

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

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

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

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

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

then

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

To satisfy our convention we must have

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

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 $\lambda_2 = 1$

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

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

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

Then

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

To satisfy our convention, we must have

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

Consequently,

$c = -0.6$ 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 ($AA = A$), 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

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

where

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

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

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

<u>Result 1.14:</u> 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.


```
# -----------------------
# Eigenvalues & Eigenvectors
# -----------------------

> A <- matrix(c(1.96,.72,.72,1.54),
             2,2,byrow=T)

> A
      [,1]  [,2]
 [1,] 1.96  0.72
 [2,] 0.72  1.54

> EA <- eigen(A)
> EA
$values:
[1] 2.5 1.0

$vectors:
      [,1]  [,2]
 [1,] -0.8   0.6
 [2,] -0.6  -0.8
```



```
# --------------------------
# Singular Value Decomposition
# --------------------------

> A<-matrix(c(2,0,1,1,0,2,1,1,1,1,1,1),
             ncol=4,byrow=T)

> A
      [,1] [,2] [,3] [,4]
 [1,]    2    0    1    1
 [2,]    0    2    1    1
 [3,]    1    1    1    1

> svdA <- svd(A)
> svdA
$d:
[1] 3.464102 2.000000 0.000000
```

```
$v:
        [,1]           [,2]   [,3]
  [1,]  -0.5   7.071068e-001   0.5
  [2,]  -0.5  -7.071068e-001   0.5
  [3,]  -0.5  -1.226910e-016  -0.5
  [4,]  -0.5  -9.065285e-017  -0.5

$u:
              [,1]            [,2]        [,3]
  [1,]  -0.5773503   7.071068e-001  -0.4082483
  [2,]  -0.5773503  -7.071068e-001  -0.4082483
  [3,]  -0.5773503  -7.597547e-017   0.8164966
```

```
> svdA$u %*% t(svdA$u)

                 [,1]            [,2]            [,3]
  [1,]  1.000000e+000   9.310586e-018  -3.089976e-018
  [2,]  9.310586e-018   1.000000e+000   3.244475e-017
  [3,] -3.089976e-018   3.244475e-017   1.000000e+000
```



```
> t(svdA$v) %*% svdA$v

                 [,1]            [,2]            [,3]
  [1,]  1.000000e+000   5.116079e-017  -2.775558e-017
  [2,]  5.116079e-017   1.000000e+000  -3.405750e-017
  [3,] -2.775558e-017  -3.405750e-017   1.000000e+000
```

```
> svdA$u%*%diag(svdA$d)%*%t(svdA$v)

                 [,1]            [,2]  [,3]  [,4]
  [1,]  2.000000e+000  -1.557456e-016     1     1
  [2,] -9.074772e-017   2.000000e+000     1     1
  [3,]  1.000000e+000   2.000000e+000     1     1
```



```
> diag(svdA$d) %*% diag(svdA$d)
       [,1]  [,2]  [,3]
  [1,]   12     0     0
  [2,]    0     4     0
  [3,]    0     0     0

> eigen(A %*% t(A))$values
[1] 1.200000e+001 4.000000e+000
  4.440892e-016

> eigen.(t(A) %*% A)$values
[1] 1.200000e+001 4.000000e+000
  -2.167786e-016 -3.238078e-015
```



```
# An example where the singular values
# are the eigenvalues

> A <- matrix(c(1.96,.72,.72,1.54),
        2,2,byrow=T)
> A
       [,1]  [,2]
  [1,] 1.96  0.72
  [2,] 0.72  1.54

> svdA <- svd(A)
> svdA
$d:
[1] 2.5 1.0

$v:
       [,1]  [,2]
  [1,] -0.8  -0.6
  [2,] -0.6   0.8

$u:
       [,1]  [,2]
  [1,] -0.8  -0.6
  [2,] -0.6   0.8
```

```
#———————————————————
# Trace and determinant of a matrix
#———————————————————

> A <- matrix(c(1,1, 1,
2,5,-1,
0,1, 1),3,3,byrow=T)

> A
     [,1] [,2] [,3]
[1,]    1    1    1
[2,]    2    5   -1
[3,]    0    1    1

> traceA <- sum(diag(A))
> traceA
[1] 7
```



```
> eigenA <- eigen(A)
> eigenA
$values:
[1]   5.336912+0.0000000i    0.831544-0.6578603i
0.831544+0.6578603i

$vectors:
                    [,1]                    [,2]
[1,] -0.2852888 + 0i     1.7077352 + 0.3055786i
[2,] -1.0054394 + 0i    -0.7100770 - 0.2551929i
[3,] -0.2318330 + 0i     0.6234268 - 0.9197348i
                    [,3]
[1,]  1.7077352 - 0.3055786i
[2,] -0.7100770 + 0.2551929i
[3,]  0.6234268 + 0.9197348i
```



```
> traceA <- sum(eigenA$values)
> traceA
[1] 7+0i

> Re(traceA)
[1] 7

> detA <- Re(prod(eigenA$values))
> detA
[1] 6
```



```
# An example where the eigenvalues
# are real numbers.

> A <- matrix(c(1,1, 1,
+               2,5,-1,
+               0, 1, -1),3,3,byrow=T)

> A
     [,1] [,2] [,3]
[1,]    1    1    1
[2,]    2    5   -1
[3,]    0    1   -1

> eigenA <- eigen(A)
> eigenA

$values:
[1] 5.372281e+00 -3.722813e-01
  -1.405092e-16
```

```
$vectors:
           [,1]      [,2]       [,3]
[1,] -0.2608539 -3.269275 -0.8164966
[2,] -0.9858217  1.730136  0.4082483
[3,] -0.1547047  2.756228  0.4082483

> traceA <- sum(eigenA$values)
> traceA
[1] 5

> detA <- Re(prod(eigenA$values))
> detA
[1] 2.810183e-16
```



```
#-----------------------------------------
# Eigenvalues of a square symmetric matrix
#-----------------------------------------

> A<-matrix(c(4,2,-1,2,6,-4,-1,-4,9),
     3,3,byrow=T)
> A
     [,1] [,2] [,3]
[1,]    4    2   -1
[2,]    2    6   -4
[3,]   -1   -4    9

> EA <- eigen(A)
> EA
$values:
[1] 12.245772 4.433349 2.320879

$vectors:
           [,1]       [,2]       [,3]
[1,] -0.2347350 -0.7321107  0.6394634
[2,] -0.5764345 -0.4248579 -0.6980108
[3,]  0.7827022 -0.5324563 -0.3222848
```



```
> SVDA <- svd(A)
> SVDA
$d:
[1] 12.245772 4.433349 2.320879

$v:
           [,1]      [,2]       [,3]
[1,] -0.2347350 0.7321107 -0.6394634
[2,] -0.5764345 0.4248579  0.6980108
[3,]  0.7827022 0.5324563  0.3222848

$u:
           [,1]      [,2]       [,3]
[1,] -0.2347350 0.7321107 -0.6394634
[2,] -0.5764345 0.4248579  0.6980108
[3,]  0.7827022 0.5324563  0.3222848
```



```
#------------------------------------
# An example of a square symmetric
# matrix that is not positive definite
#------------------------------------

> W<-matrix(c(4,2,-1,2,6,-4,-1,-4,-9),
     3,3,byrow=T)

> W
     [,1] [,2] [,3]
[1,]    4    2   -1
[2,]    2    6   -4
[3,]   -1   -4   -9

> EW <- eigen(W)
> EW
$values:
[1] 8.151345 2.865783 -10.017128

$vectors:
           [,1]        [,2]      [,3]
[1,]  0.4665008 -0.88381658 0.0352886
[2,]  0.8550024  0.46079428 0.2379907
[3,] -0.2266009 -0.08085101 0.9706262
```

```
> t(EW$vectors)%*%EW$vectors
              [,1]          [,2]          [,3]
[1,]  1.000000e+00 -1.353084e-16 -1.665335e-16
[2,] -1.353084e-16  1.000000e+00 -6.938894e-17
[3,] -1.665335e-16 -6.938894e-17  1.000000e+00

> SVDW <- svd(W)

> SVDW
$d:
[1] 10.017128 8.151345 2.865783

$v:
           [,1]       [,2]        [,3]
[1,] -0.0352886 -0.4665008 -0.88381658
[2,] -0.2379907 -0.8550024  0.46079428
[3,] -0.9706262  0.2266009 -0.08085101

$u:
           [,1]       [,2]        [,3]
[1,] -0.0352886 -0.4665008 -0.88381658
[2,] -0.2379907 -0.8550024  0.46079428
[3,] -0.9706262  0.2266009 -0.08085101
```



```
#-----------------------------------
# Inverse of a matrix
#-----------------------------------

> A<-matrix(c(1.96,.72,.72,1.54),
      2,2,byrow=T)

> Ainv <- solve(A)
> Ainv
       [,1]   [,2]
[1,]  0.616 -0.288
[2,] -0.288  0.784

> A%*%Ainv
              [,1]           [,2]
[1,] 1.000000e+000 -1.638772e-016
[2,] 7.548758e-017  1.000000e+000
```



```
# Use the spectral decomposition
# to compute the inverse of a matrix

> Aev<-eigen(A)$vectors
> Aeval<-eigen(A)$values
> Ainv2<-Aev%*%diag(1/Aeval)%*%t(Aev)
> Ainv2
       [,1]   [,2]
[1,]  0.616 -0.288
[2,] -0.288  0.784
```



```
#---------------------------------------
# Solutions to linear equations
#---------------------------------------

> x<-c(1,1)
> x
[1] 1 1

> b<-solve(A,x)
> b
[1] 0.328 0.496
```

2. Vector Spaces

Euclidean space:

A vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ of order 2 represents

a point in a plane

Note that any point in the plane can be represented as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

↖ ↗
basis vectors

The entire plane is denoted by R^2.

A vector of order 3, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

represents a point in 3-dimensional Euclidean space (denoted by R^3).

Note that any $\mathbf{x} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \epsilon\ R^3$ can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{x}_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \mathbf{x}_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \mathbf{x}_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

↗ ↗ ↗
basis vectors for R^3

A vector of order n, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ represents

a point in n-dimensional Euclidean space (denoted by R^n).

R^n is a special case of a more general concept of a **vector space**.

Defn 2.1: A set of vectors, denoted by S, is a **vector space** if for every pair of vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ in S we have

(i) $\mathbf{x}_i + \mathbf{x}_j$ is a vector in S

(ii) $a\,\mathbf{x}_i$ is in S for any real scalar.

Defn 2.2: If every vector in some vector space S can be expressed as a linear combination

$$a_1\,\mathbf{x}_1 + a_2\,\mathbf{x}_2 + \cdots + a_k\,\mathbf{x}_k$$

of a set of k vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$, this set of vectors is said to **span** the vector space S.

Defn 2.3: If a set of vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ span S and are linearly independent, then the set is called a **basis** for S.

Comments:

(i) The number of vectors in a basis for a vector space S is called the dimension of S (dim(S)).

(ii) $\mathbf{0}$ belongs to every vector space in R^n.

(iii) A vector space can have many bases.

Example:

$$\mathrm{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \mathrm{x}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \mathrm{x}_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \mathrm{x}_4 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

span R^3, but are not a basis for R^3.

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

are a basis for R^3.

Note that

$$\frac{1}{3}\mathbf{x}_1 + \frac{1}{3}\mathbf{x}_2 + \frac{1}{3}\mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\frac{1}{3}\mathbf{x}_1 - \frac{2}{3}\mathbf{x}_2 + \frac{1}{3}\mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$frac13\mathbf{x}_1 + \frac{1}{3}\mathbf{x}_2 - \frac{2}{3}\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

then

$$\begin{aligned} \begin{bmatrix} a \\ b \\ c \end{bmatrix} &= a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \left(\frac{a+b+c}{3}\right)\mathbf{x}_1 \\ &\quad + \left(\frac{a-2b+c}{3}\right)\mathbf{x}_2 \\ &\quad + \left(\frac{a+b-2c}{3}\right)\mathbf{x}_3 \end{aligned}$$

Example

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

do not span R^3. Any two of these vectors provides a basis for a 2-dimensional subspace of R^3.

Note that $\mathbf{x}_3 = \mathbf{x}_1 + 2\,\mathbf{x}_2$, which implies that $\mathbf{x}_1 = \mathbf{x}_3 - 2\,\mathbf{x}_2$ and $\mathbf{x}_2 = 0.5\,(\mathbf{x}_3 - \mathbf{x}_1)$.

Then, for any $\mathbf{z} = a\,\mathbf{x}_1 + b\,\mathbf{x}_2$

we have

$$\begin{aligned} \mathbf{z} &= a(\mathbf{x}_3 - 2\,\mathbf{x}_2) + b\,\mathbf{x}_2 \\ &= (b - 2a)\mathbf{x}_2 + a\,\mathbf{x}_3 \end{aligned}$$

and

$$\begin{aligned} \mathbf{z} &= a\,\mathbf{x}_1 + \frac{b}{2}(\mathbf{x}_3 - \mathbf{x}_1) \\ &= (a - \frac{b}{2})\mathbf{x}_1 + \frac{b}{2}\mathbf{x}_3 \end{aligned}$$

This 2-dimensional subspace of R^3 is the vector space consisting of all vectors of the form

$$\begin{aligned} \mathbf{z} = a\,\mathbf{x}_1 + b\,\mathbf{x}_2 &= a\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + b\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} a+b \\ 2a \\ b-a \end{bmatrix} \end{aligned}$$

<u>Random vectors:</u>

<u>Defn 2.4:</u> A random vector $\mathbf{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$ is a vector whose elements are random variables.

<u>Mean vectors:</u>

$$E(\mathbf{Y}) = \begin{bmatrix} E(Y_1) \\ \vdots \\ E(Y_n) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = \boldsymbol{\mu}$$

where

$$\mu_i = E(Y_i) = \begin{cases} \int_{-\infty}^{\infty} y\, f_i(y)dy & \text{if } Y_i \text{ is a continuous random variable with density function } f_i(y) \\ \sum_{\substack{all\ possible \\ y\ values}} y\, p_i(y) & \text{if } Y_i \text{ is a discrete random variable with probability function } p_i(y). \end{cases}$$

<u>Covariance matrix:</u>

$$\Sigma = Var(\mathbf{Y}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \cdots & \sigma_n^2 \end{bmatrix}$$

with variances

$$Var(Y_i) = \sigma_i^2 = E(Y_i - \mu_i)^2$$

$$= \begin{cases} \int_{-\infty}^{\infty} (y - \mu_i)^2 f_i(y)dy & \text{if } y_i \text{ is a continuous random variable} \\ \sum_{\substack{all \\ y}} (y - \mu_i)^2\, p_i(y) & \text{if } y_i \text{ is a discrete random variable} \end{cases}$$

and covariances:

$$\sigma_{ij} = Cov(Y_i, Y_j) = E\left[(Y_i - \mu_i)(Y_j - \mu_j)\right]$$

where

$$\sigma_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \mu_i)(v - \mu_j) f_{ij}(y, v)\, dy\, dv$$

if Y_i and Y_j are continuous random variables with joint density function $f_{ij}(y, v)$

and

$$\sigma_{ij} = \sum_{\substack{\text{all} \\ y}} \sum_{\substack{\text{all} \\ v}} (y - \mu_i)(v - \mu_j) P_{ij}(y, v)$$

if Y_i and Y_j are discrete random variables with joint probability function $P_{ij}(y, v) = Pr(Y_i = y, V_j = v)$

Result 2.1:

Let $\mathbf{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$ be a random vector with

$$\boldsymbol{\mu} = E(\mathbf{Y}) \quad \text{and} \quad \sum = Var(\mathbf{Y}),$$

and let

$$A_{p \times n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{p1} & \cdots & a_{pn} \end{bmatrix}$$

be a matrix of non-random elements, and let

$$\mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \quad \text{and} \quad \mathbf{d} = \begin{bmatrix} d_1 \\ \vdots \\ d_p \end{bmatrix}$$

be vectors of non-random elements, then

(i) $E(A\mathbf{Y} + \mathbf{d}) = A\boldsymbol{\mu} + \mathbf{d}$

(ii) $Var(A\mathbf{Y} + \mathbf{d}) = A\Sigma A^T$

(iii) $E(\mathbf{c}^T\mathbf{Y}) = \mathbf{c}^T\boldsymbol{\mu}$

(iv) $Var(\mathbf{c}^T\mathbf{Y}) = \mathbf{c}^T\Sigma\mathbf{c}$