Chapter 2 Equations and Algebra

Key-words

- Linear model – method of least squares, normal equations
- Inverse matrix
- Rank
- Non-singularity
- Determinant

2.1 Linear model, least squares method and normal equations

■ Definition of a linear model

Y : dependent (response) variables
$X_1, \cdots, X_p$: independent (explanatory) variables

$$\Rightarrow \; Y = b_1 X_1 + b_2 X_2 + \cdots + b_p X_p + e \;;\; e \sim (0, \sigma^2)$$

Note: In general, $X_1 = 1$

· Regression analysis (a case of n observations)

$$y_i = b_1 X_{i1} + b_2 X_{i2} + \cdots + b_p X_{ip} + e_i \;;\; i = 1, 2, \cdots, n$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} + \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

$\Leftrightarrow \; \underline{y} = X\underline{b} + \underline{e}$ $\quad$ (X : incidence matrix or design matrix): Matrix equation

Example: Aerobic Fitness

	age	weight	runtime	rst pulse	run pulse	max pulse	oxy
1	44	89.47	11.37	62	178	182	44.61
2	40	75.07	10.07	62	185	185	45.31
3	44	85.84	8.65	45	156	168	54.30
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
29	49	76.32	9.4	56	186	188	48.67
30	48	61.24	11.5	52	170	176	47.92
31	52	82.78	10.5	53	170	172	47.47

* The number of observations : $n=31$
* The number of variables : $p=7$
* Dependent variable(1) : oxy
* Independent variables(6) : age, weight, runtime, rstpulse, runpulse, maxpulse

* Regression analysis : $y_i = b_1 X_{i1} + b_2 X_{i2} + \cdots + b_p X_{ip} + e_i$; $i=1,2,\cdots,n$

$$\Rightarrow \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} + \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \Leftrightarrow \quad \underline{y} = X\underline{b} + \underline{e}$$

$$\Rightarrow \begin{pmatrix} 44.61 \\ \vdots \\ 47.47 \end{pmatrix} = \begin{pmatrix} 1 & 44 & \cdots & 182 \\ \vdots & \vdots & & \vdots \\ 1 & 52 & \cdots & 172 \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} + \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

$\therefore \quad \hat{\underline{b}} = (X^t X)^{-1} X^t \underline{y}$ if $\exists\ (X^t X)^{-1}$

$\Rightarrow y_i = 102.9153\, X_{i1} - 0.2267\, X_{i2} + \cdots + 0.3037\, X_{i7} + e_i$; $i=1,2,\cdots,31$

- Analysis of variance (ANOVA)

$$y_{ij} = \mu + \alpha_i + e_{ij}$$

Example: weights of 6 plants $(i = 1,2,3 \quad j = 1,\cdots,n_i)$

Normal	Off-type	Aberrant
$y_{11} = 101$	$y_{21} = 84$	$y_{31} = 32$
$y_{12} = 105$	$y_{22} = 88$	
$y_{13} = 94$		

$$\Rightarrow \quad \begin{array}{l} 101 = y_{11} = \mu + \alpha_1 \qquad\qquad + e_{11} \\ 105 = y_{12} = \mu + \alpha_1 \qquad\qquad + e_{12} \\ 94 = y_{13} = \mu \qquad + \alpha_1 \qquad + e_{13} \\ 84 = y_{21} = \mu \qquad + \alpha_2 \qquad + e_{21} \\ 88 = y_{22} = \mu \qquad + \alpha_2 \qquad + e_{22} \\ 32 = y_{31} = \mu \qquad\qquad + \alpha_3 + e_{31} \end{array} \quad \Leftrightarrow \quad \begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{31} \end{pmatrix} = \begin{pmatrix} 1\,1\,0\,0 \\ 1\,1\,0\,0 \\ 1\,1\,0\,0 \\ 1\,0\,1\,0 \\ 1\,0\,1\,0 \\ 1\,0\,0\,1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} + \begin{pmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \\ e_{31} \end{pmatrix} \quad \Leftrightarrow \quad \underline{y} = X\underline{b} + \underline{e}$$

$\therefore \quad \hat{\underline{b}} = \left(X^t X\right)^{-1} X^t \underline{y} \quad$ if $\exists \ (X^t X)^{-1}$ ········· (1)

■ Least squares method(LSM)

For

$$y_i = b_1 X_{i1} + b_2 X_{i2} + \cdots + b_p X_{ip} + e_i \ ; \ i = 1, 2, \cdots, n \ (\ n \geq p \ ) \quad \Leftrightarrow \quad \underline{y} = X\underline{b} + \underline{e}$$

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (\ y_i - f(x_i, b) \ )^2 = \underline{e}^t \underline{e} = (\underline{y} - X\underline{b})^t (\underline{y} - X\underline{b})$$: Sum of squared residuals

To find parameter values $b_1, b_2, \cdots, b_p$ for which the model "best" fits data,
the least squares method defines "best" as when the sum S of squared residuals is a minimum

I.e.

$$Min_{b_1, \ldots, b_p} \sum_{i=1}^{n} (y_i - b_1 x_{i1} - \cdots - b_p x_{ip})^2 \quad \Leftrightarrow \quad Min_{\underline{b}} \ (\underline{y} - X\underline{b})^t (\underline{y} - X\underline{b})$$

: The least squares method

History of Least Squares Method

Adrien Marie Legendre (1752–1833)

Adrien–Marie Legendre (born 18 September 1752 – died 10 January 1833)

In 1805, the idea of least squares analysis was formulated by the Legendre.

・ In 1805, he published the work by which he is chiefly known in the history of statistics:

Nouvelles methodes pour la determination des orbites des cometes.

– Sur la methods des moindres quarres

<혜성궤도를 결정하기 위한 새로운 방법> –부록 (최소제곱부분에 대하여)

・ Ten years after Legendre's 1805 appendix, the method of LS was a standard tool in astronomy and geodesy in France, Italy, and Prussia. By 1825 the same was true in England.

・ LSM was closely associated with three of the major scientific problems of 18c:

i) to determine and represent mathematically the motion of the moon

ii) Motions of the planets Jupiter and Saturn

iii) to determine the shape of figure of the earth.

Note : Matrix Derivatives(Diffentiation)

$$\forall\ \underline{a}, \underline{v}, \underline{w} \in R^p$$

i) Scalar

$$- \ \frac{\partial}{\partial\ \underline{v}} \underline{a}^t \underline{v} = \frac{\partial}{\partial\ \underline{v}} \underline{v}^t \underline{a} = \begin{pmatrix} \frac{\partial}{\partial\ v_1} \underline{a}^t \underline{v} \\ \vdots \\ \frac{\partial}{\partial\ v_p} \underline{a}^t \underline{v} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix} = \underline{a}$$: round $\underline{a}^t \underline{v}$ over round $\underline{v}$(derivative of $\underline{a}^t \underline{v}$ wrt $\underline{v}$)

$$- \ \frac{\partial \underline{a}^t \underline{v}}{\partial \underline{v}^t} = \left(\frac{\partial \underline{a}^t \underline{v}}{\partial v_1}, \cdots, \frac{\partial \underline{a}^t \underline{v}}{\partial v_p} \right) = (a_1, \cdots, a_p) = \underline{a}^t$$

ii) Vectors

$$A_{p \times p} = \left(\underline{a_1}, \underline{a_2}, \cdots, \underline{a_p} \right) \quad \text{and} \quad \underline{v}^t A = \left(\underline{v}^t \underline{a}_1, \underline{v}^t \underline{a}_2, \cdots, \underline{v}^t \underline{a}_p \right)$$

$$- \ \frac{\partial \underline{v}^t A}{\partial \underline{v}} = \left(\frac{\partial \underline{v}^t \underline{a}_1}{\partial \underline{v}}, \frac{\partial \underline{v}^t \underline{a}_2}{\partial \underline{v}}, \cdots, \frac{\partial \underline{v}^t \underline{a}_p}{\partial \underline{v}} \right) = (\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_p) = A$$

$$- \ \left(\frac{\partial}{\partial\ \underline{v}^t} (A \underline{v}) \right)^t = \frac{\partial}{\partial\ \underline{v}} (A \underline{v})^t = \frac{\partial}{\partial\ \underline{v}} (\underline{v}^t A^t) = A^t$$

iii) Quadratic forms

$\forall\ \underline{v}, \underline{w} \in R^p,\ \ A_{p\times p} = \left(\underline{a_1}\ , \underline{a_2}\ , \cdots\ , \underline{a_p}\right)$

- $\dfrac{\partial \underline{v}^t A \underline{v}}{\partial \underline{v}} = (A + A^t)\underline{v}$

- $\dfrac{\partial\ \underline{v}^t A \underline{v}}{\partial\ \underline{v}} = 2A\underline{v}$ If $A = A^t$

- $\dfrac{\partial \underline{v}^t A \underline{w}}{\partial \underline{v}} = A\underline{w}$

- $\dfrac{\partial \underline{v}^t \underline{v}}{\partial \underline{v}} = 2\underline{v}$

Proof of equation (1): $\hat{\underline{b}} = (X^t X)^{-1} X^t \underline{y}$

• Regression analysis and analysis of variance (ANOVA)

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} + \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \Rightarrow \underline{y} = X\underline{b} + \underline{e} \Leftarrow \begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{31} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} + \begin{pmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \\ e_{31} \end{pmatrix}$$

ⅰ) Matrix proof

$$S = \underline{e}^t \underline{e} = (\underline{y} - X\underline{b})^t (\underline{y} - X\underline{b})$$

$$= \underline{y}^t \underline{y} - \underline{b}^t X^t \underline{y} - \underline{y}^t X \underline{b} + \underline{b}^t X^t X \underline{b}$$

$$= \underline{y}^t \underline{y} - 2\underline{b}^t X^t \underline{y} + \underline{b}^t X^t X \underline{b}$$

$$\frac{\partial S}{\partial \underline{b}} = -2X^t \underline{y} + 2X^t X \underline{b} = 0 \quad \Leftrightarrow \quad X^t X \underline{b} = X^t \underline{y} \ : \text{normal equation}$$

$$\Leftrightarrow \hat{\underline{b}} = (X^t X)^{-1} X^t \underline{y}$$

(∵ rank(X)=full column rank =p $\Rightarrow$ $\exists\ (X^t X)^{-1}$)

: least square solution for $\underline{b}$ = the unique least squares estimator ii) Elementwise proof

$$y_i = b_1 x_{i1} + b_2 x_{i2} + \cdots + b_p x_{ip} + e_i \ , \ i = 1, 2, \cdots, n$$

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - b_1 x_{i1} - \cdots - b_p x_{ip})^2$$: sum of squared errors

$$\Rightarrow \ \min_{b_1, \cdots, b_p} S$$

$$f(b_{1,} b_{2,} \cdots, b_p) = \sum_{i=1}^{n} (y_i - b_1 x_{i1} - \cdots - b_p x_{ip})^2$$: objective function

$$\begin{cases} \dfrac{\partial f}{\partial b_1} = \displaystyle\sum_{i=1}^{n} 2(y_i - b_1 x_{i1} - \cdots - b_p x_{ip})(-x_{i1}) = 0 \\ \dfrac{\partial f}{\partial b_2} = \displaystyle\sum_{i=1}^{n} 2(y_i - b_1 x_{i1} - \cdots - b_p x_{ip})(-x_{i2}) = 0 \\ \vdots \\ \dfrac{\partial f}{\partial b_p} = \displaystyle\sum_{i=1}^{n} 2(y_i - b_1 x_{i1} - \cdots - b_p x_{ip})(-x_{ip}) = 0 \end{cases}$$

: partial derivative of $f(\bullet)$ w.r.t b_j

$$\Leftrightarrow \begin{cases} \sum_{i=1}^{n}(b_1x_{i1}x_{i1}+\ \cdots\ +b_px_{i1}x_{ip})=\sum_{i=1}^{n}x_{i1}y_i \\ \sum_{i=1}^{n}(b_1x_{i2}x_{i1}+\ \cdots\ +b_px_{i2}x_{ip})=\sum_{i=1}^{n}x_{i2}y_i \\ \vdots \\ \sum_{i=1}^{n}(b_1x_{ip}x_{i1}+\ \cdots\ +b_px_{ip}x_{ip})=\sum_{i=1}^{n}x_{ip}y_i \end{cases} \Leftrightarrow \begin{cases} b_1\sum_{i=1}^{n}x_{i1}x_{i1}+\ \cdots\ +b_p\sum_{i=1}^{n}x_{i1}x_{ip}=\sum_{i=1}^{n}x_{i1}y_i \\ b_1\sum_{i=1}^{n}x_{i2}x_{i1}+\ \cdots\ +b_p\sum_{i=1}^{n}x_{i2}x_{ip}=\sum_{i=1}^{n}x_{i2}y_i \\ \vdots \\ b_1\sum_{i=1}^{n}x_{ip}x_{i1}+\ \cdots\ +b_p\sum_{i=1}^{n}x_{ip}x_{ip}=\sum_{i=1}^{n}x_{ip}y_i \end{cases}$$

$$\Leftrightarrow \begin{cases} b_1a_{11}+\ \cdots\ +b_pa_{1p}=c_1 \\ b_1a_{21}+\ \cdots\ +b_pa_{2p}=c_2 \\ \vdots \\ b_1a_{p1}+\ \cdots\ +b_pa_{pp}=c_p \end{cases} \qquad \left(\because\ \sum_{i=1}^{n}x_{ij}x_{ik}=a_{jk},\ j,k=1,\ldots,p\ \right)$$

$$\Leftrightarrow \begin{pmatrix} a_{11}\cdots a_{1p} \\ \vdots \qquad \vdots \\ a_{p1}\cdots a_{pp} \end{pmatrix}\begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}=\begin{pmatrix} c_1 \\ \vdots \\ c_p \end{pmatrix}$$

$\Leftrightarrow$ $A\underline{b}=\underline{c}$: normal equation

In fact, $\begin{cases} A=X^tX \\ \underline{c}=X^t\underline{y} \end{cases} \quad \Rightarrow \quad X^tX\underline{b}=X^t\underline{y}$: normal equation

2.2 Solution of (simultaneous) equations and inverse matrices

$$\begin{cases} 2b_1 - 5b_2 + 4b_3 = -3 \\ b_1 - 2b_2 + b_3 = 5 \\ b_1 - 4b_2 + 6b_3 = 10 \end{cases} \Rightarrow \begin{array}{l} b_1 = 124 \\ b_2 = 75 \\ b_3 = 31 \end{array}$$ (by Gaussian elimination method)

$$\Leftrightarrow \begin{pmatrix} 2 & -5 & 4 \\ 1 & -2 & 1 \\ 1 & -4 & 6 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} -3 \\ 5 \\ 10 \end{pmatrix}$$: Matrix equation

$$\Leftrightarrow \quad A\underline{b} = \underline{c}$$

$$\exists A^{-1} \Rightarrow A^{-1}A\underline{b} = A^{-1}\underline{c} \Rightarrow \hat{\underline{b}} = A^{-1}\underline{c} = \begin{pmatrix} 124 \\ 75 \\ 31 \end{pmatrix}$$

■ Definition of identity matrix

$I = diag(1, 1, \cdots, 1)$
: An identity matrix
having all diagonal elements equal to 1

For any matrices A and B

$$AI = A, \quad IB = B$$

■ Definition of inverse matrix

For square matrix A, $\exists$ square matrix B s.t.

$$AB = BA = I$$

$\Rightarrow \quad B = A^{-1}$: The inverse of A

Remark: simple properties of inverse matrix

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

since $AB = I$,

$$AB = I \rightarrow A^{-1}AB = A^{-1}I \rightarrow IB = A^{-1}$$

$\boxed{\text{Note : } A \rightarrow I \text{ , } I \rightarrow A^{-1}}$

$$\Rightarrow \quad (A|I) \rightarrow (I|A^{-1})$$

$$\left(\begin{array}{ccc|ccc} 2 & -5 & 4 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 & 0 \\ 1 & -4 & 6 & 0 & 0 & 1 \end{array}\right) \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -8 & 14 & 3 \\ 0 & 1 & 0 & -5 & 8 & 2 \\ 0 & 0 & 1 & -2 & 3 & 1 \end{array}\right)$$: Augmented matrix

- Gauss-Jordan elimination method

M1) $A\underline{b} = \underline{c}$

$$\exists A^{-1} \Rightarrow A^{-1}A\underline{b} = A^{-1}\underline{c} \Leftrightarrow I\underline{b} = A^{-1}\underline{c}$$

$\boxed{\text{Note : } A \rightarrow I \text{ , } \underline{c} \rightarrow A^{-1}\underline{c}}$

$$\Rightarrow \quad (A|\underline{c}) \rightarrow (I|A^{-1}\underline{c})$$

$$\left(\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 124 \\ 0 & 1 & 0 & 75 \\ 0 & 0 & 1 & 31 \end{array}\right)$$

M2) $A\underline{b} = \underline{c}$

$$\exists A^{-1} \;\Rightarrow\; A\underline{b} = I\underline{c} \;\Leftrightarrow\; I\underline{b} = A^{-1}\underline{c}$$

$\boxed{\text{Note : } A \to I \;,\; I \to A^{-1} \;,\; \underline{c} \to A^{-1}\underline{c}}$

$$\Rightarrow\; (A|\underline{c}|I) \to (I|A^{-1}\underline{c}|A^{-1})$$

$$\left(\begin{array}{ccc|c|ccc} 2 & -5 & 4 & -3 & 1 & 0 & 0 \\ 1 & -2 & 1 & 5 & 0 & 1 & 0 \\ 1 & -4 & 6 & 10 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c|ccc} 1 & 0 & 0 & 124 & -8 & 14 & 3 \\ 0 & 1 & 0 & 75 & -5 & 8 & 2 \\ 0 & 0 & 1 & 31 & -2 & 3 & 1 \end{array}\right)$$

Summary : Elimination Methods

- $(A|I) \to (I|A^{-1})$
- $(A|\underline{c}) \to (I|A^{-1}\underline{c})$
- $(A|\underline{c}|I) \to (I|A^{-1}\underline{c}|A^{-1})$

Home work : Do it !

2.3 Conditions for existence of the inverse

■ Conditions for the existence of A^{-1}

i) $A_{n\times n}$: non-singular

ii) $|A| \neq 0$

iii) $rank(A) = n$: full-rank

iv) A has n linearly independent rows

v) A has n linearly independent columns

vi) $A\underline{x} = \underline{0}$ has the sole solution $\underline{x} = \underline{0}$

vii) Number of eigenvalues of A = n

■ Properties of A^{-1}

- $A^{-1}A = AA^{-1} = I$ (commutative law)
- $\exists\,!A^{-1}$ (uniqueness)
- $|A^{-1}| = \dfrac{1}{|A|}$
- $|A^{-1}| \neq 0$
- $(A^{-1})^{-1} = A$
- $(A^t)^{-1} = (A^{-1})^t$
- $A^t = A \;\Rightarrow\; (A^{-1})^t = A^{-1}$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A+B)^{-1} = A^{-1} - A^{-1}(B^{-1} + A^{-1})A^{-1} \quad (\Leftarrow A, B \text{ and } A+B : nonsingular)$

2.4 Calculation of the inverse in R and SAS/IML

$A\underline{b} = \underline{c} \quad \Rightarrow \underline{b} = A^{-1}\underline{c}$	
R	SAS/IML
• solve(A) b<-solve(A)%*%c (b<- solve(A, c))	• inv(A) b=inv(A)*c;

Refer to the text p. 40

2.5 Linear dependence and independence

$$A_1 = \begin{pmatrix} 2 & -5 & 4 \\ 1 & -2 & 1 \\ 1 & -4 & 5 \end{pmatrix} \quad \Rightarrow \quad |A_1| = \begin{vmatrix} 2 & -5 & 4 \\ 1 & -2 & 1 \\ 1 & -4 & 5 \end{vmatrix}$$

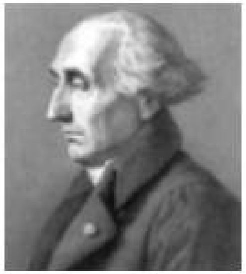
$$= [2(-2)5 + 1(-4)4 + 1(-5)1] - [4(-2)1 + 1(-5)5 + 2(-4)1]$$

$$= [-20 - 16 - 5] - [-8 - 25 - 8] = -25 - 16 + 16 + 25 = 0$$ ($\because$ Rule of Sarrus)

$\Rightarrow \quad \not\exists A^{-1} \Leftrightarrow A$: singular

Note

Rule of Sarrus or Sarrus' Rule is a method and a memorization scheme to compute the determinant of a 3×3 matrix. It is named after the French mathematician Pierre Frédéric Sarrus.



Sarrus(1798-1861)

- Rows

$$2\begin{pmatrix} 2 \\ -5 \\ 4 \end{pmatrix} - 3\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \\ 5 \end{pmatrix}$$: linear dependence

- Columns

$$-3\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - 2\begin{pmatrix} -5 \\ -2 \\ -4 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 5 \end{pmatrix}$$: linear dependence

■ LIN (linear independence or linearly independent)

$$X_{n\times n}\underline{k} = (\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = \sum_{i=1}^{n} k_i \underline{x}_i \quad : \text{linear combination}$$

$X\underline{k} = \underline{0} \quad \Rightarrow \quad \underline{k} = \underline{0}$

$\Leftrightarrow \quad K\underline{k} = \underline{0}$ only for $\underline{k} = \underline{0}$

: The columns of X are <u>LIN</u> vectors

■ LDN (linear dependence or linearly dependent)

$$X_{n\times n}\underline{k} = (\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = \sum_{i=1}^{n} k_i \underline{x}_i \quad : \text{linear combination}$$

$X\underline{k} = \underline{0} \quad \Rightarrow \quad \underline{k} \neq \underline{0}$

$\Leftrightarrow \quad X\underline{k} = \underline{0}$ for some non-null $\underline{k}$ $(\underline{k} \neq \underline{0})$

: The columns of X are LDN vectors

Example

i) $2\begin{pmatrix}2\\-5\\4\end{pmatrix}-3\begin{pmatrix}1\\-2\\1\end{pmatrix}-1\begin{pmatrix}1\\-4\\5\end{pmatrix}=\underline{0} \iff \begin{pmatrix}2&1&1\\-5&-2&-4\\4&1&5\end{pmatrix}\begin{pmatrix}2\\-3\\1\end{pmatrix}=\underline{0}$

$\iff A_1\underline{k}=\underline{0}$

$\Rightarrow \underline{k}=\begin{pmatrix}2\\-3\\-1\end{pmatrix}\neq\underline{0}$: LDN,

Note: $|A_1|=0$

ii) $A_2=\begin{pmatrix}2&-5&4\\1&-2&1\\1&-4&6\end{pmatrix}$

$k_1\begin{pmatrix}2\\1\\1\end{pmatrix}+k_2\begin{pmatrix}-5\\-2\\-4\end{pmatrix}+k_3\begin{pmatrix}4\\1\\6\end{pmatrix}=\underline{0} \iff \begin{pmatrix}2&-5&4\\1&-2&1\\1&-4&6\end{pmatrix}\begin{pmatrix}k_1\\k_2\\k_3\end{pmatrix}=\underline{0} \iff A_2\underline{k}=\underline{0}$

Note: $|A_2|=17\neq0$

$\exists A_2^{-1} \Rightarrow A_2^{-1}A_2\underline{k}=A_2^{-1}\underline{0}$

$\iff \underline{k}=\underline{0}$ $\therefore$ The columns of A_2 are LIN

■ Definition of rank

Rank of $A \equiv rank(A) \equiv r(A)$

: The number of linearly independent rows (and columns) in A

Example

$$A_1 = \begin{pmatrix} 2 & -5 & 4 \\ 1 & -2 & 1 \\ 1 & -4 & 5 \end{pmatrix} , \quad A_2 = \begin{pmatrix} 2 & -5 & 4 \\ 1 & -2 & 1 \\ 1 & -4 & 6 \end{pmatrix}$$

$$rank(A_1) = 2 \Leftrightarrow \det(A_1) = 0$$

$$rank(A_2) = 3 \Leftrightarrow \det(A_2) \neq 0$$

■ Properties of $rank(A_{m \times n})$

i) $rank(A_{m\times n}) = \min(m,n) \Rightarrow A$: full rank

ii) $rank(A_{m\times n}) = m \Rightarrow A$: full row rank

iii) $rank(A_{m\times n}) = n \Rightarrow A$: full column rank

2.6 Determinants

■ Application to statistics

i) Variance

$$X_{n\times p} = (x_{ij}) \ , \ i=1, \cdots, n \ , \ j=1, \cdots, p$$

: Data matrix

$$\rightarrow \ Y = (x_{ij} - \overline{x_j})$$

: Centering matrix

$$Y^t Y = \begin{pmatrix} \sum_{i=1}^{n}(x_{i1}-\overline{x_1})^2 & \sum_{i=1}^{n}(x_{i1}-\overline{x_1})(x_{i2}-\overline{x_2}) & \cdots & \sum_{i=1}^{n}(x_{i1}-\overline{x_1})(x_{ip}-\overline{x_p}) \\ & \sum_{i=1}^{n}(x_{i2}-\overline{x_2})^2 & \cdots & \sum_{i=1}^{n}(x_{i2}-\overline{x_2})(x_{ip}-\overline{x_p}) \\ & & \ddots & \vdots \\ sym & & & \sum_{i=1}^{n}(x_{ip}-\overline{x_p})^2 \end{pmatrix}$$

$$= (n-1)S = (n-1)\begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ & s_{22} & \cdots & s_{2p} \\ & & \ddots & \vdots \\ sym & & & \\ & & & s_{pp} \end{pmatrix}$$

(S : The variance-covariance matrix)

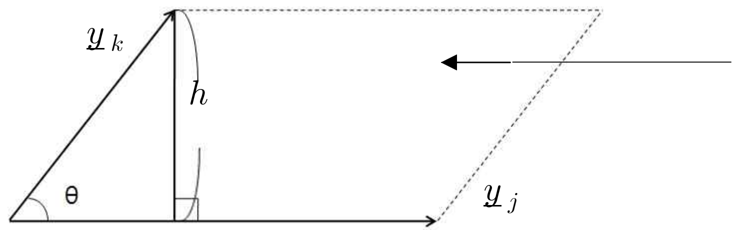
- Variance

  - Generalized : $|S|$
  - Total : $tr(S) = s_{11} + \cdots + s_{pp}$

- Geometric view of $|S|$, $with\ S = \begin{pmatrix} S_{kk} & S_{kj} \\ S_{jk} & S_{jj} \end{pmatrix}\ and\ p = 2$

$$\underline{y}_k = (x_{1k} - \overline{x_k}\ ,\ \cdots, x_{nk} - \overline{x_k})^t$$

$$\underline{y}_j = (x_{1j} - \overline{x_j}\ ,\ \cdots, x_{nj} - \overline{x_j})^t$$



$$Area = \|\underline{y}_k\| \times \|\underline{y}_j\| \times \sin\theta$$

$$= \sqrt{\sum_{i=1}^{n}(x_{ik} - \overline{x_k})^2}\sqrt{\sum_{i=1}^{n}(x_{ij} - \overline{x_j})^2} \times \sqrt{1-\cos^2\theta}$$

Note

- $\sum_{i=1}^{n} (x_{ik} - \overline{x_k})^2 = (n-1) s_{kk}$
- $\sum_{i=1}^{n} (x_{ij} - \overline{x_j})^2 = (n-1) s_{jj}$
- $\cos\theta = \dfrac{\underline{y}_k^t \underline{y}_j}{\| \underline{y}_k \| \ \| \underline{y}_j \|} = \dfrac{\sum_{i=1}^{n} (x_{ik} - \overline{x_k})(x_{ij} - \overline{x_j})}{\sqrt{\sum_{i=1}^{n} (x_{ik} - \overline{x_k})^2} \sqrt{\sum_{i=1}^{n} (x_{ij} - \overline{x_j})^2}}$

$$= \frac{(n-1) s_{kj}}{\sqrt{(n-1) s_{kk}} \sqrt{(n-1) s_{jj}}}$$

$$= \frac{s_{kj}}{\sqrt{s_{kk}} \sqrt{s_{jj}}} = r_{kj}$$

$\therefore \ Area = (n-1) \sqrt{s_{kk}} \sqrt{s_{jj}} \sqrt{1 - r_{kj}^2}$

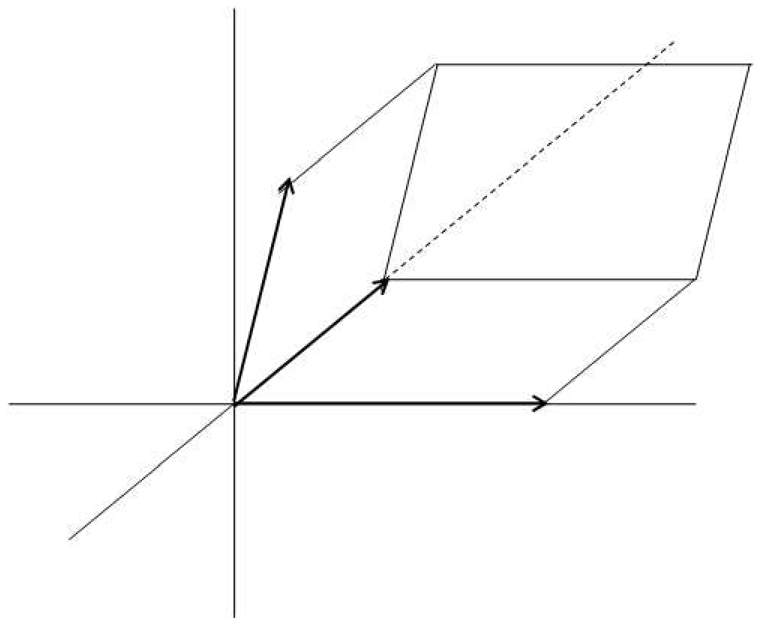
$$|S| = \begin{vmatrix} s_{kk} & s_{kj} \\ s_{jk} & s_{jj} \end{vmatrix} = s_{kk}s_{jj} - s_{kj}s_{jk}$$

$$= s_{kk}s_{jj} - (r_{jk}\sqrt{s_{kk}}\sqrt{s_{jj}})(r_{jk}\sqrt{s_{kk}}\sqrt{s_{jj}}) \quad (\because r_{kj} = \frac{s_{kj}}{\sqrt{s_{kk}}\sqrt{s_{jj}}} = r_{jk} \ )$$

$$= s_{kk}s_{jj}(1 - r_{jk}^2)$$

$$\therefore \ \ If \ \ p = 2, \quad |S| = \frac{(Area)^2}{(n-1)^2}$$

– $If \ \ p > 2, \ \ |S| = \dfrac{(Volume)^2}{(n-1)^p}$



ii) Multivariate normal distribution

$$x \sim N(\mu, \sigma^2), \quad \sigma > 0$$

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

$$= (2\pi)^{-1/2}(\sigma^2)^{-1/2}\exp\left(-\frac{1}{2}(x-\mu)(\sigma^2)^{-1}(x-\mu)\right)$$

$$\Rightarrow \ \underline{x} = (x_{1,} \cdots, x_p)^t \sim N_p(\underline{\mu}, \Sigma) \ , \ \Sigma > 0$$

$$\boxed{f(\underline{x}) = (2\pi)^{-p/2}|\Sigma|^{-1/2}\exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu})^t\Sigma^{-1}(\underline{x}-\underline{\mu})\right)}$$

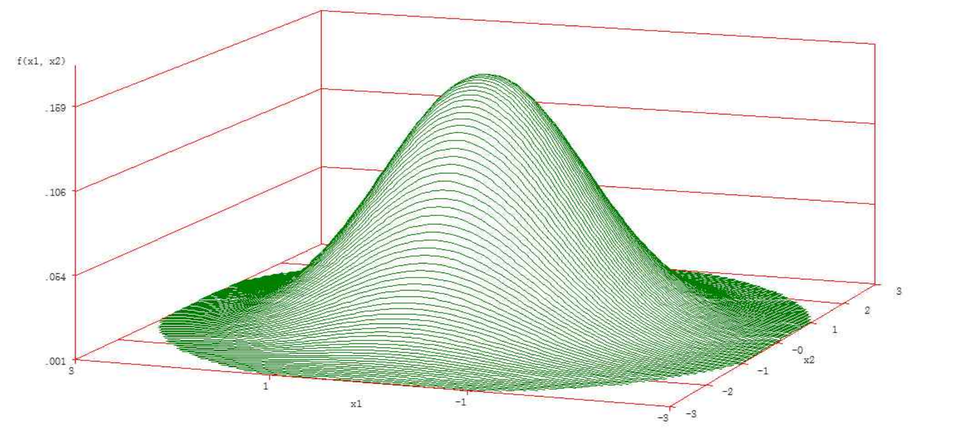
↳ Ellipsoid

- $p = 2$, Bivariate normal

$$\underline{x} = (x_1, x_2) \sim N_2(\underline{\mu}, \Sigma) \quad \text{where } \underline{\mu} = (\mu_1, \mu_2)^t \ , \ \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

$$|\Sigma| = \sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2 = (1-\rho^2)\sigma_1^2\sigma_2^2$$

$$\Sigma^{-1} = \frac{1}{|\Sigma|}\begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}$$



$$\begin{pmatrix} x_1-\mu_1 & x_2-\mu_2 \end{pmatrix} \Sigma^{-1} \begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix} = \frac{1}{1-\rho^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 \right]$$

$$\therefore \ f(\underline{x}) = f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp(-q/2)$$

$$\text{where } q = \frac{1}{1-\rho^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 \right]$$

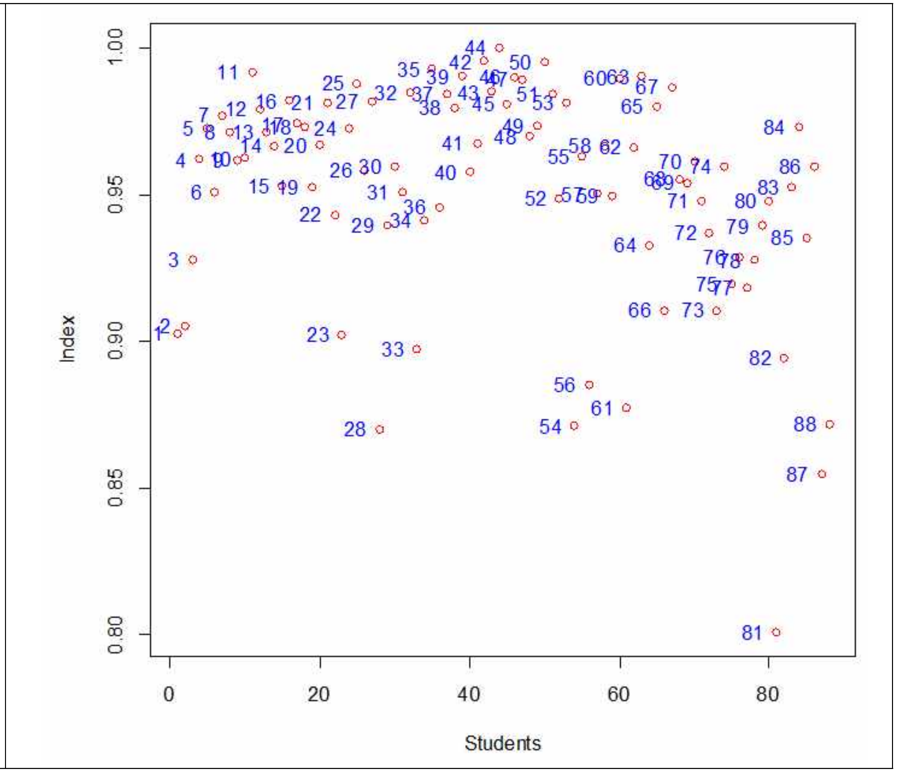
- $\rho = 0 \quad \rightarrow \quad f(x_1, x_2) = f(x_1)f(x_2)$

iii) Outlier

- $r_i^2 = \dfrac{|S_{(i)}|}{|S|}$: for single outlier where $S_{(i)}$ covariance matrix with x_i omitted.
- $\dfrac{|S_{(\underline{i})}|}{|S|}$: for multiple outlier where $\underline{i} = (i_1, i_2, \cdots, i_t)$, t =(number of observations)

Criterion: A relatively small value of r_i^2 indice x_i is potential outlier.

	single outlier				
[1,]	0.9472527	[26,]	1.0055308	[56,]	0.9289727
[2,]	0.9500749	[27,]	1.0300230	[57,]	0.9970349
[3,]	0.9736236	[28,]	0.9129172	[58,]	1.0140230
[4,]	1.0093201	[29,]	0.9858263	[59,]	0.9962962
[5,]	1.0204481	[30,]	1.0069715	[60,]	1.0380680
[6,]	0.9976272	[31,]	0.9978288	[61,]	0.9208771
[7,]	1.0250582	[32,]	1.0333239	[62,]	1.0137424
[8,]	1.0192968	[33,]	0.9416272	[63,]	1.0389501
[9,]	1.0092410	[34,]	0.9877500	[64,]	0.9787027
[10,]	1.0097776	[35,]	1.0416638	[65,]	1.0281443
[11,]	1.0404495	[36,]	0.9924048	[66,]	0.9554624
[12,]	1.0273055	[37,]	1.0326007	[67,]	1.0350832
[13,]	1.0189797	[38,]	1.0278448	[68,]	1.0021611
[14,]	1.0142306	[39,]	1.0389238	[69,]	1.0010168
[15,]	0.9998236	[40,]	1.0047766	[70,]	1.0087386
[16,]	1.0302959	[41,]	1.0150475	[71,]	0.9946434
[17,]	1.0221134	[42,]	1.0444162	[72,]	0.9832537
[18,]	1.0208078	[43,]	1.0335075	[73,]	0.9554405
[19,]	0.9996248	[44,]	1.0493268	[74,]	1.0068274
[20,]	1.0143645	[45,]	1.0290247	[75,]	0.9647899
[21,]	1.0294478	[46,]	1.0385267	[76,]	0.9746207
[22,]	0.9894279	[47,]	1.0379448	[77,]	0.9633846
[23,]	0.9466385	[48,]	1.0175129	[78,]	0.9733431
[24,]	1.0203104	[49,]	1.0214440	[79,]	0.9856481
[25,]	1.0365367	[50,]	1.0440324	[80,]	0.9946146
		[51,]	1.0327465	[81,]	0.8404648
		[52,]	0.9954448	[82,]	0.9384063
		[53,]	1.0297299	[83,]	0.9996551
		[54,]	0.9141544	[84,]	1.0210044
		[55,]	1.0104093	[85,]	0.9811699
				[86,]	1.0066482
				[87,]	0.8970167
				[88,]	0.9146625



■ Axiomatic definition of determinant :

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{bmatrix} \underline{a}_1^t \\ \underline{a}_2^t \end{bmatrix}$$

Axiom i) Factorization

$$\begin{vmatrix} ka_{11} & ka_{12} \\ a_{21} & a_{22} \end{vmatrix} = ka_{11}a_{22} - ka_{12}a_{21} = k(a_{11}a_{22} - a_{12}a_{21}) = k\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$\Leftrightarrow \begin{vmatrix} k\underline{a}_1^t \\ \underline{a}_2^t \end{vmatrix} = \det\begin{bmatrix} k\underline{a}_1^t \\ \underline{a}_2^t \end{bmatrix} = k\det\begin{bmatrix} \underline{a}_1^t \\ \underline{a}_2^t \end{bmatrix} \qquad \Rightarrow \det\begin{bmatrix} \underline{a}_1^t \\ \vdots \\ k\underline{a}_i^t \\ \vdots \\ \underline{a}_p^t \end{bmatrix} = k\det\begin{bmatrix} \underline{a}_1 \\ \vdots \\ \underline{a}_i \\ \vdots \\ \underline{a}_p \end{bmatrix}$$

Axiom ii) Adding multiples of a row(column) to a row(column)

$$\begin{vmatrix} a_{11}+k & a_{12}+k \\ a_{21} & a_{22} \end{vmatrix} = (a_{11}+k)a_{22} - a_{21}(a_{12}+k) = a_{11}a_{22} - a_{21}a_{12} + ka_{22} - ka_{21} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} k & k \\ a_{21} & a_{22} \end{vmatrix}$$

$$\Leftrightarrow \det\begin{bmatrix} \underline{a}_1^t + k1_2^t \\ \underline{a}_2^t \end{bmatrix} = \det\begin{bmatrix} \underline{a}_1^t \\ \underline{a}_2^t \end{bmatrix} + \det\begin{bmatrix} k1_2^t \\ \underline{a}_2^t \end{bmatrix} \Rightarrow \det\begin{bmatrix} \underline{a}_1^t \\ \vdots \\ \underline{a}_i^t + k\underline{1}_p^t \\ \vdots \\ \underline{a}_p^t \end{bmatrix} = \det\begin{bmatrix} \underline{a}_1^t \\ \vdots \\ \underline{a}_i^t \\ \vdots \\ \underline{a}_p^t \end{bmatrix} + \det\begin{bmatrix} \underline{a}_1^t \\ \vdots \\ k\underline{1}_p^t \\ \vdots \\ \underline{a}_p^t \end{bmatrix}$$

Axiom iii) If two rows(columns) of A are the same, $|A| = 0$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{11} & a_{12} \end{vmatrix} = 0$$

Axiom iv) The determinant of the Identity matrix is 1

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

$$\Rightarrow \ \det(I_p) = \det(\underline{e}_1, \underline{e}_1, \cdots, \underline{e}_p) = 1$$

Axiom v) For $kA_{p\times p}$, $\det(kA) = k^p \det(A)$

■ Zero determinants

$\det(A) \neq 0 \quad \Leftrightarrow$ The vectors are linearly independent (LIN)

$\Leftrightarrow$ A is non-singular $\quad \Leftrightarrow \ \exists A^{-1}$

2.7 Determinants and the theory of the inverse

■ Calculate the determinants

- $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$

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

$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - (a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{11}a_{23}a_{32})$: Sarrus' Rule

$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$: Expansion by minors

$= a_{11}(-1)^{1+1}|M_{11}| + a_{12}(-1)^{1+2}|M_{12}| + a_{13}(-1)^{1+3}|M_{13}|$

- $|A_{n\times n}| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{i1}(-1)^{i+1}|M_{i1}| + a_{i2}(-1)^{i+2}|M_{i2}| + \cdots + a_{in}(-1)^{i+n}|M_{in}|$

$$= \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|M_{ij}|$$

$$|A_{n\times n}| = \sum_{j=1}^{n} a_{ij}c_{ij}, \text{ for any } i$$

$$= \sum_{i=1}^{n} a_{ij}c_{ij}, \text{ for any } j$$

where $c_{ij} = (-1)^{i+j}|M_{ij}|$: Cofactor

: Cofactor Expansion

- $0 = \sum_{j=1}^{n} a_{ij}c_{kj}$, for $i \neq k$

$$\because \sum_{j=1}^{3} a_{1j}c_{2j} = a_{11}(-1)^{2+1}|M_{21}| + a_{12}(-1)^{2+2}|M_{22}| + a_{13}(-1)^{2+3}|M_{23}|$$

$$= a_{11}(-1)^{2+1}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)^{2+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}(-1)^{2+3}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0$$

■ Derivation of the inverse by cofactors

Note

$$\left.\begin{array}{l} A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\[2em] |A| = a_{11}a_{22} - a_{12}a_{21} \neq 0 \\[2em] A^{-1} = \dfrac{1}{|A|}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \end{array}\right\} \quad \Rightarrow \quad A^{-1}A = I$$

$$\Leftrightarrow \quad \frac{1}{|A|}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

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

$$= \frac{1}{|A|}\begin{pmatrix} |A| & 0 \\ 0 & |A| \end{pmatrix} = \frac{|A|}{|A|}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

- Adjugate of A

$$c_{ij} = (-1)^{i+j} |M_{ij}|$$

$$\Rightarrow \{c_{ij}\} = C$$

$\Rightarrow adjA = C^t$: Adjugate (or adjoint) matrix of A

<u>Theorem</u>

- $A_{n \times n}$ with $|A| \neq 0$

$$\Rightarrow A^{-1} = \frac{1}{|A|} C^t = \frac{1}{|A|} adjA$$

<u>Theorem</u> $AC^t = |A| \times I$

$$A^{-1} = \frac{1}{|A|} C^t$$

Proof) $C^t = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix}^t = \begin{pmatrix} c_{11} & \cdots & c_{n1} \\ \vdots & & \vdots \\ c_{1n} & \cdots & c_{nn} \end{pmatrix}$

$$AC^t = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} c_{11} & \cdots & c_{n1} \\ \vdots & & \vdots \\ c_{1n} & \cdots & c_{nn} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}c_{1j} & 0 & \cdots & 0 \\ & \ddots & & \vdots \\ & & \ddots & 0 \\ symm. & & & \sum_{j=1}^{n} a_{nj}c_{nj} \end{pmatrix} \quad (\because \sum_{j=1}^{n} a_{ij}c_{kj} = 0,\ i \neq k)$$

$$= \begin{pmatrix} |A| & 0 & \cdots & 0 \\ & |A| & & \vdots \\ & & \ddots & 0 \\ sy\ mm. & & & |A| \end{pmatrix} = |A|\, I$$

<u>Lemma</u>

$\therefore\ |A| \neq 0 \quad \Rightarrow \quad A\dfrac{C^t}{|A|} = I \quad \Leftrightarrow \quad A^{-1} = \dfrac{C^t}{|A|}$ (of course $C^t A = |A| I \quad \Leftrightarrow \quad \dfrac{C^t}{|A|} A = I$)

■ Cramer's rule

$$A\underline{b} = \underline{c}$$

$$|A| \neq 0 \quad \Rightarrow \quad \exists\, A^{-1} = \frac{1}{|A|} C^t$$

$$\underline{b} = A^{-1}\underline{c} = \frac{1}{|A|} C^t \underline{c} \quad \Rightarrow \quad b_k = \frac{1}{|A|}|A_k|,\ k = 1, \ldots, n$$

$$\therefore\ \underline{b} = \frac{1}{|A|}\begin{pmatrix} c_{11} & \cdots & c_{n1} \\ \vdots & & \vdots \\ c_{1k} & \cdots & c_{nk} \\ \vdots & & \vdots \\ c_{1n} & \cdots & c_{nn} \end{pmatrix}\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

$$= \frac{1}{|A|}\begin{pmatrix} c_{11}c_1 + \cdots + c_{n1}c_n \\ \vdots \\ c_{1k}c_1 + \cdots + c_{nk}c_n \\ \vdots \\ c_{1n}c_1 + \cdots + c_{nn}c_n \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} \sum_{j=1}^{n} c_{j1}c_j \\ \vdots \\ \sum_{j=1}^{n} c_{jk}c_j \\ \vdots \\ \sum_{j=1}^{n} c_{jn}c_j \end{pmatrix}$$

$$\therefore\ b_k = \frac{1}{|A|}\sum_{j=1}^{n} c_{jk}c_j$$

$$\sum_{j=1}^{n} c_{jk}c_j = c_{1k}c_1 + \ \cdots\ + c_{nk}c_n$$

$$= c_1 \times (-1)^{1+k}|M_{1k}| + \ \cdots\ + c_n \times (-1)^{n+k}|M_{nk}|$$

$$= |A_k|$$

with A_k is a matrix replacing the k^{th} column of A by c

Note

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

$$\Rightarrow \sum_{j=1}^{3} c_{j1}c_j = c_{11}c_1 + c_{21}c_2 + c_{31}c_3$$

$$= c_1 \times (-1)^{1+1}|M_{11}| + c_2 \times (-1)^{2+1}|M_{21}| + c_3 \times (-1)^{3+1}|M_{31}|$$

$$= \begin{vmatrix} c_1 & a_{12} & a_{13} \\ c_2 & a_{22} & a_{23} \\ c_3 & a_{32} & a_{33} \end{vmatrix} = |A_1|$$

$$\Rightarrow b_1 = \frac{1}{|A|}|A_1|$$

■ Properties of $|A|$

i) $A = \begin{pmatrix} a_{11} \cdots & a_{1n} \\ \ddots & \vdots \\ \underline{0} & a_{nn} \end{pmatrix}$: upper triangular matrix , $A = \begin{pmatrix} a_{11} & \underline{0} \\ \vdots & \ddots \\ a_{n1} \cdots & a_{nn} \end{pmatrix}$: lower triangular matrix

$\Rightarrow \quad |A| = a_{11} \times \cdots \times a_{nn}$

ii) $|AB| = |A||B|$

iii) $|A| = |A^t|$

iv) $\begin{vmatrix} A_{11} & \underline{0} \\ \underline{0} & A_{22} \end{vmatrix} = |A_{11}||A_{22}|$

v) $A_{n \times n}$: The orthogonal matrix $\Leftrightarrow$ $A^t A = I_n$ with $\underline{a}_j{}^t \underline{a}_k = 0$, $j \neq k$ and $\underline{a}_j{}^t \underline{a}_j = 1$

- $|A| = \pm 1$ ($\because$ $1 = |I_n| = |A^t A| = |A^t||A| = |A|^2$)

- $B_{n \times n}$: Any matrix $\Rightarrow$ $|A^t BA| = |B|$

vi) $|cA| = c^n |A|$ if $A_{n \times n} \wedge$ c is constant

vii) $A_{n \times n} = \begin{pmatrix} \underset{n_1 \times n_1}{A_{11}} & \underset{n_1 \times (n-n_1)}{A_{12}} \\ \underset{(n-n_1) \times n_1}{A_{21}} & \underset{(n-n_1) \times (n-n_1)}{A_{22}} \end{pmatrix}$: Partitioned matrix

$$|A| = |A_{11}||A_{22} - A_{21}A_{11}^{-1}A_{12}| \quad \text{if } |A_{11}| \neq 0$$

$$|A| = |A_{22}||A_{11} - A_{12}A_{22}^{-1}A_{21}| \quad \text{if } |A_{22}| \neq 0$$

viii)* $\lambda_j,\ j = 1, \ldots, n$: eigenvalues of $A_{n \times n}$

- $|A| = \prod_{j=1}^{n} \lambda_j$

- $tr(A) = \sum_{j=1}^{n} \lambda_j$

Application 1

$$\underline{x} = (x_1, x_2, \cdots, x_p)^t \sim N_p(\underline{\mu}, \Sigma)$$

$$\underline{x}_{p\times 1} = \begin{array}{c} \begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \end{pmatrix} \end{array} \begin{array}{c} q \\ p-q=r \end{array} \sim \underline{\mu} = \begin{pmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{pmatrix}, \ \Sigma = \begin{pmatrix} \underset{q\times q}{\Sigma_{11}} & \underset{q\times r}{\Sigma_{12}} \\ \underset{r\times q}{\Sigma_{21}} & \underset{r\times r}{\Sigma_{22}} \end{pmatrix}$$

$$\underline{x}_1 \perp \underline{x}_2 \quad \Leftrightarrow \quad \Sigma_{12} = \underline{0}$$

- $f(\underline{x}) = (2\pi)^{-p/2} |\Sigma|^{-1/2} \exp[-\frac{1}{2}(\underline{x} - \underline{\mu})^t \Sigma^{-1} (\underline{x} - \underline{\mu})]$

i) $|\Sigma| = |\Sigma_{11}||\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}| = |\Sigma_{11}||\Sigma_{22}|$ ($\because \ \Sigma_{12} = \underline{0}$)

ii) $(\underline{x} - \underline{\mu})^t \Sigma^{-1} (\underline{x} - \underline{\mu}) = \left((\underline{x}_1 - \underline{\mu}_1)^t \ \ (\underline{x}_2 - \underline{\mu}_2)^t \right) \begin{pmatrix} \Sigma_{11}^{-1} & \underline{0} \\ \underline{0} & \Sigma_{22}^{-1} \end{pmatrix} \begin{pmatrix} \underline{x}_1 - \underline{\mu}_1 \\ \underline{x}_2 - \underline{\mu}_2 \end{pmatrix}$

$$= (\underline{x}_1 - \underline{\mu}_1)^t \Sigma_{11}^{-1} (\underline{x}_1 - \underline{\mu}_1) + (\underline{x}_2 - \underline{\mu}_2)^t \Sigma_{22}^{-1} (\underline{x}_2 - \underline{\mu}_2)$$

iii) $f(\underline{x}) = (2\pi)^{-q/2}|\Sigma_{11}|^{-1/2}\exp[-\frac{1}{2}(\underline{x}_1 - \underline{\mu}_1)^t\Sigma_{11}^{-1}(\underline{x}_1 - \underline{\mu}_1)]$

$$\times (2\pi)^{-r/2}|\Sigma_{22}|^{-1/2}\exp[-\frac{1}{2}(\underline{x}_2 - \underline{\mu}_2)^t\Sigma_{22}^{-1}(\underline{x}_2 - \underline{\mu}_2)]$$

$$= f(\underline{x}_1) \times f(\underline{x}_2)$$

$$\underline{x}_1 \sim N_q(\underline{\mu}_1, \Sigma_{11}) \quad \perp \quad \underline{x}_2 \sim N_r(\underline{\mu}_2, \Sigma_{22})$$

Application 2

LRT for
$$\begin{cases} H_0 : \underline{\mu} = \underline{\mu}_0 \quad \Rightarrow \quad \Lambda^{2/n} = \dfrac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \dfrac{|\sum(\underline{x}_i - \overline{\underline{x}})(\underline{x}_i - \overline{\underline{x}})^t|}{|\sum(\underline{x}_i - \underline{\mu}_0)(\underline{x}_i - \underline{\mu}_0)^t|} \\ H_0 : \Sigma = \Sigma_0 \quad \Rightarrow \quad \Lambda = (\dfrac{e}{n})^{np/2} etr(-\dfrac{1}{2}\Sigma_0^{-1}A)\left|\Sigma_0^{-1}A\right| \qquad where\ A = \sum(\underline{x}_i - \overline{\underline{x}})(\underline{x}_i - \overline{\underline{x}})^t \end{cases}$$

- Multivariate normal distribution $\sim f(\underline{x}) = (2\pi)^{-p/2}|\Sigma|^{-1/2}\exp[-\frac{1}{2}(\underline{x} - \underline{\mu})^t\Sigma^{-1}(\underline{x} - \underline{\mu})]$
- Likelihood function : $L(\underline{\mu}, \Sigma) = \prod_{i=1}^{n} f(\underline{x}_i) = (2\pi)^{-np/2}|\Sigma|^{-n/2}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}(\underline{x}_i - \underline{\mu})^t\Sigma^{-1}(\underline{x}_i - \underline{\mu})\right)$
- $\Lambda = \max_{\Theta_0} L(\underline{\mu}, \Sigma) \,/\, \max_{\Theta_1} L(\underline{\mu}, \Sigma) \quad \Rightarrow \quad -2\ln\Lambda \sim \chi^2 \ \deg = \dim(\Theta) - \dim(\Theta_0) \quad where\ \Theta = \Theta_0 \bigcup \Theta_1$