Properties of the multivariate normal distribution (H: 2.6, 4.4, 5.1)

TMA4267 L5
24.01.2017

$$\underset{p\times 1}{X} \sim N_p(\mu, \Sigma)$$

$$f(x) = (2\pi)^{-\frac{p}{2}} \det(\Sigma)^{-\frac{1}{2}} \exp\{-\tfrac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\}$$

$$M_X(t) = E(e^{t^T X}) = e^{t^T\mu + \frac{1}{2}t^T \Sigma t}$$

1 Contours of the mvN

For some constant a $(a>0)$ the solution to $f(x) = a$ are called the contours of X.

$$f(x) = \text{const.} \exp\{-\tfrac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\}$$

or work with

$$(X-\mu)^T \Sigma^{-1}(X-\mu) = b \qquad (b>0)$$

What is this graphical object?

Let $\Sigma = P \Lambda P^T$

$$P = [e_1 \cdots e_p] \qquad \Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_p \end{bmatrix} \qquad \begin{bmatrix} \frac{1}{\lambda_1} & & & 0 \\ & \frac{1}{\lambda_2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\lambda_p} \end{bmatrix}$$

Work with $\Sigma^{-1} = (P\Lambda P^T)^{-1} = P\Lambda^{-1}P^T$

$$= \frac{1}{\lambda_1} e_1 e_1^T + \frac{1}{\lambda_2} e_2 e_2^T + \dots + \frac{1}{\lambda_p} e_p e_p^T$$

Study $p=2$:

$P = [e_1 \; e_2]$, $\Lambda^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & 0 \\ 0 & \frac{1}{\lambda_2} \end{pmatrix}$

$\Sigma^{-1} = \frac{1}{\lambda_1} e_1 e_1^T + \frac{1}{\lambda_2} e_2 e_2^T$

$e_1 = 2\times1$
$e_2 = 2\times1$

$(\underline{X}-\mu)^T \Sigma^{-1} (\underline{X}-\mu) = b$

$$\underbrace{(\underline{X}-\mu)^T}_{1\times2} \left(\frac{1}{\lambda_1} \underbrace{e_1 e_1^T}_{2\times2} + \frac{1}{\lambda_2} e_2 e_2^T \right) \underbrace{(\underline{X}-\mu)}_{2\times1} = b$$

$$\frac{1}{\lambda_1} \underbrace{\overbrace{(x-\mu)^T}^{1\times2} \overbrace{e_1}^{2\times1}}_{w_1^T} e_1^T (x-\mu) + \frac{1}{\lambda_2} \underbrace{(x-\mu)^T e_2}_{w_2^T} e_2^T (x-\mu) = b$$

$$\frac{1}{\lambda_1} \underbrace{w_1^T w_1}_{w_1^2} + \frac{1}{\lambda_2} \underbrace{w_2^T w_2}_{w_2^2} = b$$

$$\frac{1}{\lambda_1 b} \cdot w_1^2 + \frac{1}{\lambda_2 b} \cdot w_2^2 = 1$$

What is this?

This is an ellipse ($p=2$)
Halflengths: $\sqrt{\lambda_1 b}$, $\sqrt{\lambda_2 b}$
Axis: are in the direction of the eigenvectors e_1, e_2
Centered: in μ

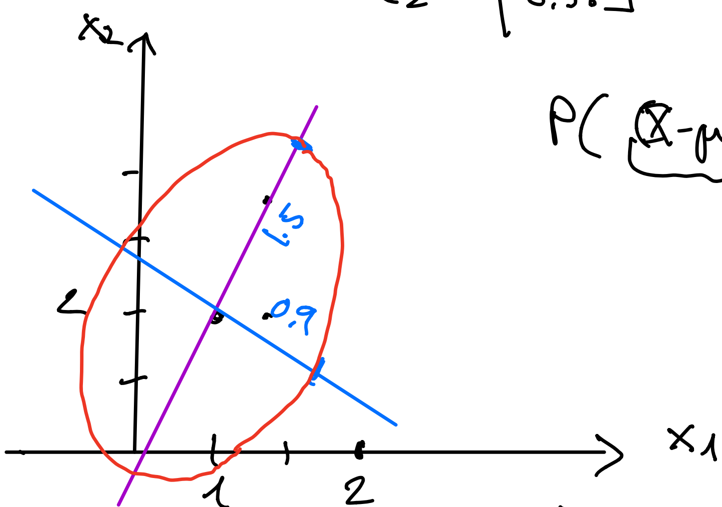
$P > 2$: general: ellipsoide.

Exam K2014 1b modified

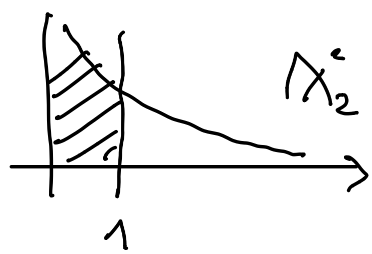
$$\underset{2\times1}{X} \qquad \mu = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 2 \end{bmatrix}, \quad b = 1$$

$\lambda_1 = 2.2 \qquad e_1 = \begin{bmatrix} 0.38 \\ 0.92 \end{bmatrix} \qquad \sqrt{\lambda_1 \cdot b} = 1.5$

$\lambda_2 = 0.8 \qquad e_2 = \begin{bmatrix} -0.92 \\ 0.38 \end{bmatrix} \qquad \sqrt{\lambda_2 \cdot b} = 0.9$



$$P\Big(\underbrace{(X-\mu)^T \Sigma^{-1} (X-\mu)}_{\chi^2_{p},\ p=2} \leq 1\Big) = \underline{\underline{0.39}}$$



pchisq(1,2) = 0.39

② Linear combinations

$\underset{p\times 1}{X} \sim N_p(\mu, \Sigma)$ and $\underset{q\times p}{B}$ and $c \in \mathbb{R}^q$

constants

$$\underset{q\times 1}{Y} = BX + c \sim N_q(B\mu + c, B\Sigma B^T)$$

PROOF: $M_X(t) = E(e^{t^T X}) = \exp\left(t^T\mu + \frac{1}{2} t^T \Sigma t\right)$

$$M_Y(t) = M_{BX+c}(t) = E\left(e^{t^T(BX+c)}\right)$$

$$= e^{t^T c} \underbrace{E\left(e^{t^T B X}\right)}_{M_X(B^T t)} = e^{t^T c}\, e^{t^T B\mu + \frac{1}{2} t^T B \Sigma B^T t}$$

$$= \exp\{ t^T \underbrace{(B\mu + c)}_{E(Y)} + \frac{1}{2} t^T \underbrace{B\Sigma B^T}_{Cov(Y)} t \}$$

This is mvN with $E(Y) = B\mu + c$ and $Cov(Y) = B\Sigma B^T$

Exam K2014 1a

$$Y = \underbrace{\begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix}}_{C} X \sim N_2(C\mu, C\Sigma C^T)$$

μ, Σ in task.

③ Subsets

$\mathbf{X} \sim N_p(\mu, \Sigma)$, partition into

$$p = p_1 + p_2$$

$$\underset{p\times 1}{\mathbf{X}} = \begin{bmatrix} \underset{p_1\times 1}{\mathbf{X}_1} \\ \underset{p_2\times 1}{\mathbf{X}_2} \end{bmatrix} \qquad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

$$\underset{p_1\times p_1}{\Sigma_{11}} = Cov(\mathbf{X}_1)$$

$$\underset{p_1\times p_2}{\Sigma_{12}} = Cov(\mathbf{X}_1, \mathbf{X}_2)$$

$$\Sigma_{21} = \Sigma_{12}^T$$

$$\underset{p_2\times p_2}{\Sigma_{22}} = Cov(\mathbf{X}_2)$$

Let $\underset{p_1\times p}{B} = [\underset{p_1\times p_1}{I} \quad \underset{p_1\times p_2}{0}]$

Then $\mathbf{X}_1 = B\mathbf{X} \sim N_{p_1}(B\mu = \mu_1,\ B\Sigma B^T = \Sigma_{11})$

$$\uparrow$$

$$[I \quad 0] \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix}$$

④ "Covariance = 0" implies independence for mvN

---

a) If $\underset{p_1\times 1}{X_1}$ and $\underset{p_2\times 1}{X_2}$ are independent.

$$\Downarrow$$

$$Cov(X_1, X_2) = \underset{p_1\times p_2}{0}$$

for general X's.

But: $Cov(X_1, X_2) = 0$ does not in general imply that X_1 and X_2 are independent, e.g.

$X_1 = Y$, $X_2 = Y^2$, $Cov(Y, Y^2) = 0$

How does $M_X(t)$ look like when X_1 and X_2 are independent when $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N_p(\mu, \Sigma)$.

$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{bmatrix}$

$$M_X(\underset{p}{t}) = M_{X_1}(\underset{p_1}{t_1}) \cdot M_{X_2}(\underset{p_2}{t_2})$$

$$= \exp\left(t_1^T \mu_1 + \tfrac{1}{2} t_1^T \Sigma_{11} t_1\right) \cdot \exp\left(t_2^T \mu_2 + \tfrac{1}{2} t_2^T \Sigma_{22} t_2\right)$$

$t = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$, $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_{11} & \overset{0}{\Sigma_{12}} \\ \underset{0}{\Sigma_{21}} & \Sigma_{22} \end{bmatrix}$

$$M_X(t) = \exp\left(t^T \mu + \tfrac{1}{2} t^T \Sigma t\right)$$

b) Let $\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$

$\mathbf{X}_1$ and $\mathbf{X}_2$ are independent iff
$\Sigma_{12} = \mathrm{Cov}(\mathbf{X}_1, \mathbf{X}_2) = \underset{p_1 \times p_2}{0}$

Why: $\Sigma_{12} = 0 \Leftrightarrow \Sigma_{21} = 0$ and then

$M_X(t) = M_{X_1}(t_1) \cdot M_{X_2}(t_2)$.

NB only if N_p. ← multivariate normal.

Ex: Independent variables:
$(1,3)$, $(1,4)$, $(2,3)$

$\begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}$ $\begin{bmatrix} X_3 \\ X_4 \end{bmatrix}$ not independent

Homework: K2014 1a on slide 17.

(5) Independence of AX and BX

$X \sim N_p(\mu, \Sigma)$

AX and BX are independent iff $A\Sigma B^T = 0$.

Why? $Y = \begin{bmatrix} A \\ B \end{bmatrix} X$ $\qquad A \qquad B$

$q \times p \quad p \times 1 \qquad q_1 \times p \quad q_2 \times p$

$q_1 + q_2 = q$

$$M_Y(t) = E\left[\exp\left\{t^T \begin{bmatrix} A \\ B \end{bmatrix} X\right\}\right]$$

know: $M_X(t) = \exp\left(t^T\mu + \frac{1}{2} t^T \Sigma t\right)$

$$M_Y(t) = \exp\left\{t^T \underline{\begin{bmatrix} A \\ B \end{bmatrix}\mu} + \frac{1}{2} t^T \underline{\begin{bmatrix} A \\ B \end{bmatrix} \Sigma \begin{bmatrix} A^T & B^T \end{bmatrix}} t\right\}$$

which:

$$\begin{bmatrix} A \\ B \end{bmatrix} X \sim N_q\left(\begin{bmatrix} A \\ B \end{bmatrix}\mu,\ \underbrace{\begin{bmatrix} A \\ B \end{bmatrix} \Sigma \begin{bmatrix} A^T & B^T \end{bmatrix}}_{\begin{bmatrix} A\Sigma A^T & A\Sigma B^T \\ B\Sigma A^T & B\Sigma B^T \end{bmatrix}}\right)$$

so if $A\Sigma B^T = 0 \iff AX$ and BX are independent.

(This version did not rely on result (2) and (4).)
→ as pointed out by student after class.

Added after the lecture: shorter version of "why" that builds on both ② and ④.

$\underline{X} \sim N_p(\mu, \Sigma)$ and $\underset{q_1 \times p}{A}$ $\underset{q_2 \times p}{B}$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} AX \\ BX \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix} \underline{X} = \underset{q \times p}{C}\underline{X} .$$

From ② we know that $Y \sim N_q(C\mu, C\Sigma C^T)$.

From ④ we know that Y_1 and Y_2 are independent iff $\mathrm{Cov}(Y_1, Y_2) = 0$,

and $\mathrm{Cov}(Y_1, Y_2) = \mathrm{Cov}(A\underline{X}, B\underline{X})$

$$= A \underbrace{\underbrace{\mathrm{Cov}(\underline{X}, \underline{X})}_{\underbrace{\mathrm{Cov}(\underline{X})}_{\Sigma}}} B^T = A \Sigma B^T, \text{ so}$$

$A\Sigma B^T = 0 \iff A\underline{X}$ and $B\underline{X}$ independent.

⑥ Conditional distribution of mvN

$$\underline{X} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \end{bmatrix} \sim N_p\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$$

Q: What is the distribution of $\underline{X}_2 \mid \underline{X}_1 = x_1$.

A: $\underline{X}_2 \mid \underline{X}_1 = x_1 \sim N_{p_2}\left(\mu_2 + \Sigma_{12} \Sigma_{11}^{-1} (x_1 - \mu_1), \right.$
$\left. \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \right)$

Observe $E(\underline{X}_2 \mid \underline{X}_1 = x_1)$ is linear in x_1

$Cov(\underline{X}_2 \mid \underline{X}_1 = x_1)$ not dependent on x_1.