TEORIA

$X = (X_1 \dots X_p)^T$ – wektor losowy

$\mu^X = (EX_1 \dots EX_p)^T$ – wektor wartości oczekiwanych

$\Sigma^X(i,j) = cov(X_i, X_j) = E(X_i X_j) - EX_i EX_j$

Właściwości cov:

1. $\Sigma^X(i,j) = \Sigma^X(j,i)$
   / $cov(X_i, X_j) = cov(X_j, X_i)$ /
   (symetryczność)
2. $cov(\alpha X_i + \beta X_j, X_k) = \alpha\, cov(X_i, X_k) + \beta\, cov(X_j, X_k)$
   (liniowość na każdym argumencie)
3. $\Sigma^X(i,i) = cov(X_i, X_i) = EX_i^2 - (EX_i)^2 =$
   $= var\ X_i \geq 0.$
   (nieujemność)

$\Sigma^X_{p\times p} = (\Sigma^X(i,j))_{i,j=1}^{p}$; $\Sigma^X = \Sigma^X_{p\times p}$

$$\Sigma^X_{p\times p} = \begin{pmatrix} Var X_1 & \cdots & cov(X_1, X_p) \\ \vdots & \ddots & \vdots \\ cov(X_p, X_1) & \cdots & Var X_p \end{pmatrix}$$ macierz kowariancji

Właściwości $\Sigma^X_{p\times p}$:

1. symetryczność $\left((\Sigma^X_{p\times p})^T = \Sigma^X_{p\times p} \right)$
2. wszystkie elementy na głównej przekątnej ≥ 0.

3. $\Sigma^{X}_{p\times p}$ jest diagonalna

$\Leftrightarrow$ $X_1 \ldots X_p$ są nieskorelowane ($X_1 .. X_p$ - norm $\Rightarrow$ niezależne)

4. $\Sigma^{X}_{p\times p}$ - nieujemne określona

(FORMA kwadratowa: $\sum_{i,j=1}^{p} \Sigma(i,j) x_i x_j \geq 0$
$\Leftrightarrow$ wartości własne ≥ 0
kryterium Sylwestera +)

<u>Przykład:</u>

ξ_1, ξ_2, ξ_3 - zmienne losowe, niezależne
$var\, \xi_i = \sigma_i^2$

$X_1 = a_1\xi_1 + a_2\xi_2 + a_3\xi_3$
$X_2 = b_1\xi_1 + b_2\xi_2 + b_3\xi_3$ $\qquad a_i, b_i, c_i \in \mathbb{R}$
$X_3 = c_1\xi_1 + c_2\xi_2 + c_3\xi_3$

$cov(X_1, X_1) = var\, X_1 \overset{\xi_i - niez.}{=} a_1^2 var\, \xi_1 + a_2^2 var\, \xi_2 + a_3^2 var\, \xi_3 = \sum_{i=1}^{3} a_i^2 \sigma_i^2$

$cov(X_1, X_2) = cov(a_1\xi_1 + a_2\xi_2 + a_3\xi_3, b_1\xi_1 + b_2\xi_2 + b_3\xi_3) =$ / liniowość cov + niezal. / $=$

$= a_1 b_1 cov(\xi_1, \xi_1) + a_1 b_2 \underbrace{cov(\xi_1, \xi_2)}_{0} + \ldots + a_3 b_3 \underbrace{cov(\xi_3, \xi_3)}_{var\,\xi_3 = \sigma_3^2} =$

$= a_1 b_1 \sigma_1^2 + a_2 b_2 \sigma_2^2 + a_3 b_3 \sigma_3^2 = \sum_{i=1}^{3} a_i b_i \sigma_i^2$

$$\Sigma^{X}_{3\times 3} = \begin{pmatrix} \sum a_i^2 \sigma_i^2 & \sum a_i b_i \sigma_i^2 & \sum a_i c_i \sigma_i^2 \\ \cdot\ \cdot & \sum b_i^2 \sigma_i^2 & \sum b_i c_i \sigma_i^2 \\ \cdot\ \cdot & \cdot\ \cdot & \sum c_i^2 \sigma_i^2 \end{pmatrix}$$

$X = (X_1 \dots X_p)$ w. l.

μ^X – wektor wartości oczekiwanych

$\Sigma^X_{p\times p}$ – macierz kowariancji

Dla dowolnej ustalonej macierzy $A = A_{k\times p}$ i wektora $B \in \mathbb{R}^k$: definiujemy przekształcenie liniowe (afiniczne)

$Y = AX + B$ $\qquad Y = (Y_1 \dots Y_k)$

$$\Rightarrow \quad \mu^Y = A\mu^X + B$$

$$\Sigma^Y = \underset{k\times p}{A}\,\underset{p\times p}{\Sigma^X}\,\underset{p\times k}{A^T} \qquad \Sigma^Y = \Sigma^Y_{k\times k}$$

/dowód: liniowość przekształcenia E, $\operatorname{cov}(\dots,\dots)$ /

<u>Przykład:</u>

$\xi = (\xi_1, \xi_2, \xi_3)^T \qquad \mu^\xi = (E\xi_1, E\xi_2, E\xi_3)^T$

$$\Sigma^\xi = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}$$

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

$X = (X_1, X_2, X_3)^T \quad \Rightarrow \quad X = A\xi$

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}$$

$$\Rightarrow \quad \mu^X = A\mu^\xi;$$

$$\Sigma^X = A \Sigma^\xi A^T =$$

$$= \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} =$$

$$= \dots$$

Wielowymiarowy rozkład normalny: $N(\mu, \Sigma)$.

$X \sim N(\mu, \Sigma) \iff$ gęstość wyraża się wzorem:

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

$x = (x_1 \dots x_p)^T$; $\mu = (\mu_1 \dots \mu_p)$

$\Sigma^{-1} = (\tilde{\sigma}_{ij})_{i,j=1}^{p}$

$(x-\mu)^T \Sigma^{-1} (x-\mu) = \sum \tilde{\sigma}_{ij} (x_i - \mu_i)(x_j - \mu_j)$
(forma kwadratowa)

<u>Przykład</u>

$X = (X_1 \dots X_p)$ X_i - niezależne $\sim N(0,1)$

$\iff \Sigma = I$ - (jednostkowa)

$\Rightarrow \Sigma^{-1} = I$

$\mu^X = (0 \dots 0)$

$\Rightarrow (x-\mu)^T \Sigma^{-1} (x-\mu) = \sum x_i^2$

$X \sim N(\vec{0}, I) \iff f(x) = (2\pi)^{-\frac{p}{2}} e^{-\frac{1}{2}(x_1^2 + \dots + x_p^2)}$

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

$Y = AX + B$, $\quad A = A_{k \times p}$; $B \in \mathbb{R}^k$

$Y = (Y_1 \ldots Y_k)$

$\Rightarrow Y \sim N(A\mu + B, \; A \Sigma A^T)$

Pytanie

$X \sim N(\mu, \Sigma)$ jak wygenerować?

$\mu = (\mu_1 \ldots \mu_p)^T$; $\Sigma = \Sigma_{p \times p}$

I krok

$\xi = (\xi_1 \ldots \xi_p)^T$ $\quad \xi_i$ - niezależne $N(0,1)$

/znorm(1) p razy/
znorm(p)

$\xi \sim N(\vec{0}, I)$.

II krok

wybieramy: A, B: $X = A\xi + B$:

$X \sim N(B, AA^T) = N(\mu, \Sigma)$

$\mu = B$

$\Sigma = AA^T$. (A nie jest jednoznacznie zdefiniowane)

A - dolna macierz trójkątna

Rozkład Choleskiego.

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix} = \begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & 0 \\ \vdots & \vdots & \ddots & \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{p1} \\ & a_{22} & \cdots & a_{p2} \\ & & \ddots & \vdots \\ 0 & & & a_{pp} \end{pmatrix}$$

$$\sigma_{11} = a_{11}^2 \quad \Rightarrow \quad a_{11} = \sqrt{\sigma_{11}}$$

$$\sigma_{12} = \sigma_{21} = a_{21} a_{11} \Rightarrow \quad a_{21} = \frac{\sigma_{21}}{a_{11}}$$

$$\sigma_{22} = a_{21}^2 + a_{22}^2 \quad \Rightarrow \quad a_{22} = \sqrt{\sigma_{22} - a_{21}^2}$$

...

$$a_{ii} = \sqrt{\sigma_{ii} - \sum_{k=1}^{i-1} a_{ik}^2}$$

$$a_{ij} = \frac{\sigma_{ij} - \sum_{k=1}^{i-1} a_{jk} a_{ik}}{a_{ii}}$$

Przykład:

$$\mu = (1, 1, 1)^T, \quad \Sigma = \begin{pmatrix} 4 & 1 & 2 \\ 1 & 9 & 3 \\ 2 & 3 & 4 \end{pmatrix}$$

wygenerować $X \sim N(\mu, \Sigma)$

I: $\xi = (\xi_1\ \xi_2\ \xi_3)$ $\quad \xi_i$ - niezal. $\sim N(0,1)$

$X = A\xi + B$; $\quad B = (1, 1, 1)^T$

A: $\Sigma = AA^T$

$$\begin{pmatrix} 4 & 1 & 2 \\ 1 & 9 & 3 \\ 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ 0 & a_{22} & a_{32} \\ 0 & 0 & a_{33} \end{pmatrix}$$

$a_{11} = 2$

$\sigma_{21} = 1 = a_{21} \cdot a_{11} \Rightarrow a_{21} = \frac{1}{2}$

$\sigma_{22} = 9 = a_{21}^2 + a_{22}^2 \Rightarrow a_{22}^2 = 9 - \frac{1}{4} = \frac{35}{4}; a_{22} = \frac{\sqrt{35}}{2}$

$\sigma_{31} = 2 = a_{31} a_{11} \Rightarrow a_{31} = 1$

$\sigma_{32} = 3 = a_{31} a_{21} + a_{32} a_{22} = \frac{1}{2} + a_{32} \frac{\sqrt{35}}{2}$

$\Rightarrow a_{32} = \frac{5}{\sqrt{35}} = \frac{\sqrt{5}}{\sqrt{7}}$

$\sigma_{33} = 4 = a_{31}^2 + a_{32}^2 + a_{33}^2 = 1 + \frac{5}{7} + a_{33}^2 \Rightarrow a_{33} = \frac{4}{\sqrt{7}}$

$$A = \begin{pmatrix} 2 & 0 & 0 \\ \frac{1}{2} & \frac{\sqrt{35}}{2} & 0 \\ 1 & \frac{\sqrt{5}}{\sqrt{7}} & \frac{4}{\sqrt{7}} \end{pmatrix}$$

$$X = A\xi + B$$