Multivariate Normal Random Variables.

$(X_1, X_2, \dots, X_n)$

(I) Bivariate Normal r.v.'s $(n=2)$.

Def 1 [R, Sec 6.3, p.233]

(X,Y) has bi-variate normal distribution if their joint pdf. is given by:

$$p(x,y) = \frac{e^{-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_X}{\sigma_Y}\right)\right\}}}{2\pi\sigma_X\sigma_X\sqrt{1-\rho^2}}$$

where $\boxed{\sigma_X, \sigma_Y > 0, \ -1<\rho<1.}$

$\boxed{\text{Def 2 [R, Sec 7.8]}}$

(X, Y) has bi-variate normal distribution if there (X, Y) can be written as:

$$\begin{cases} X = aZ_1 + bZ_2 + \cancel{\mu_X} \to 0 \\ Y = cZ_1 + dZ_2 + \cancel{\mu_Y} \to 0 \end{cases}$$

where Z_1, Z_2 are iid $N(0,1)$, a, b, c, d, μ_X, μ_Y are some constants.

[Assume $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible, (A^{-1} exists)

i.e. $\det A = (ad - bc) \neq 0$.]

---

(1) Def 1 $\Longleftrightarrow$ Def 2

How to relate, σ_X, σ_Y, ρ to a, b, c, d?

(2) Assume $\mu_X = 0, \mu_Y = 0$. (mean zero condition)

Start from Def 2

$$p(x,y) = q(z_1,z_2)\left|\frac{dz_1\,dz_2}{dx\,dy}\right|$$

$$= \frac{q(z_1,z_2)}{\left|\frac{dx\,dy}{dz_1\,dz_2}\right|}$$

$$= \frac{e^{-\frac{1}{2}z_1^2-\frac{1}{2}z_2^2}}{(\sqrt{2\pi})^2 \det\begin{vmatrix} \frac{\partial x}{\partial z_1} & \frac{\partial x}{\partial z_2} \\ \frac{\partial y}{\partial z_1} & \frac{\partial y}{\partial z_2}\end{vmatrix}}$$

$$= \frac{e^{-\frac{1}{2}(z_1^2+z_2^2)}}{2\pi\begin{vmatrix} a & b \\ c & d\end{vmatrix}}$$

$$= \frac{e^{-\frac{1}{2}(z_1^2+z_2^2)}}{2\pi|ad-bc|}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}}{ad-bc}$$

$$= \left(\left(\frac{dx-by}{ad-bc} \right), \left(\frac{-cx+ay}{ad-bc} \right) \right)^T$$

$$\boxed{\begin{gathered} \text{Hence} \\ p(x,y) = \frac{1}{2\pi |ad-bc|} e^{-\frac{1}{2}\left[\left(\frac{dx-by}{ad-bc} \right)^2 + \left(\frac{-cx+ay}{ad-bc} \right)^2 \right]} \\ = \frac{e^{-\frac{1}{2(ad-bc)^2}\left\{ (c^2+d^2)x^2 + (a^2+b^2)y^2 - 2(ac+bd)xy \right\}}}{2\pi |ad-bc|} \end{gathered}}$$

$$
p(x,y) = \frac{e^{-\dfrac{\left\{ \dfrac{x^2}{a^2+b^2} + \dfrac{y^2}{c^2+d^2} - \dfrac{2(ac+bd)\,xy}{(a^2+b^2)(c^2+d^2)} \right\}}{2\left\{ \dfrac{(ad-bc)^2}{(a^2+b^2)(c^2+d^2)} \right\}}}}{2\pi \sqrt{a^2+b^2}\sqrt{c^2+d^2} \left| \dfrac{ad-bc}{\sqrt{a^2+b^2}\sqrt{c^2+d^2}} \right|}
$$

2 dim case

(Bivariate normal.)

$$
A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}
$$

$\Downarrow$

Link to Def 1 (find σ_X, σ_Y, ρ)

(II) Multivariate normal r.v.

Def 2 $(X_1, X_2, \cdots, X_m)$ are called multi-variate, if they can be written as:

$$X_1 = a_{11}Z_1 + a_{12}Z_2 + \cdots + a_{1n}Z_n + \mu_{X_1}$$

$$X_2 = a_{21}Z_1 + a_{22}Z_2 + \cdots + a_{2n}Z_n + \mu_{X_2}$$

$$\vdots \qquad \vdots$$

$$X_m = a_{m1}Z_1 + a_{m2}Z_2 + \cdots + a_{mn}Z_n + \mu_{X_m}$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & \\ & a_{ij} & & \\ & & & a_{mn} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix} + \begin{bmatrix} \mu_{X_1} \\ \mu_{X_2} \\ \vdots \\ \mu_{X_m} \end{bmatrix}$$

$$\boxed{\underset{m\times 1}{\vec{X}} = \underset{m\times n}{A}\ \underset{n\times 1}{\vec{Z}} + \underset{m\times 1}{\vec{\mu}}}$$

Assume $m=n$, A is invertible, $\mu = 0$

Def 2 $\longrightarrow$ Def 1

$$p(\vec{X}) = q(Z)\left|\frac{dZ}{dX}\right|$$

$$= \frac{q(Z)}{\left|\frac{dX}{dZ}\right|} = \frac{e^{-\frac{1}{2}(Z_1^2+\cdots+Z_n^2)}}{(\sqrt{2\pi})^n\left|\det\left(\frac{dX}{dZ}\right)\right|}$$

$$= \frac{e^{-\frac{1}{2}\|Z\|^2}}{(\sqrt{2\pi})^n|\det A|}$$

$\|Z\|^2 = \langle Z, Z\rangle$

$= Z^T Z$

$$= \frac{e^{-\frac{1}{2}\langle Z, Z\rangle}}{(2\pi)^{n/2}|\det A|}$$

$Z = A^{-1}X$

$$= \frac{e^{-\frac{1}{2}\langle A^{-1}X, A^{-1}X\rangle}}{(2\pi)^{n/2}|\det A|}$$

$$= \frac{e^{-\frac{1}{2}\langle (A^{-1})^T A^{-1}X, X\rangle}}{(2\pi)^{n/2}|\det A|}$$

$$= \frac{e^{-\frac{1}{2}\langle (AA^T)^{-1}X, X\rangle}}{(2\pi)^{n/2}|\det A|}$$

---

(1) $\langle A^{-1}X, A^{-1}X\rangle = \langle (AA^T)^{-1}X, X\rangle$

$$\overset{\shortparallel}{\langle (A^{-1})^T A^{-1}X, X\rangle}$$

$(A^{-1})^T = (A^T)^{-1}$ $\quad = \langle (A^T)^{-1}A^{-1}X, X\rangle$ $\quad \left((AB)^{-1} = B^{-1}A^{-1}\right)$

(2) Let $\Gamma = AA^T$

(3) $\det\Gamma = \det(AA^T) = (\det A)(\det A^T)$

$$= (\det A)^2$$

Hence $|\det A| = \sqrt{\det\Gamma}$.

Hence

$$\boxed{p(X) = \frac{e^{-\frac{1}{2}\langle \Gamma^{-1}X, X\rangle}}{(2\pi)^{n/2}(\det\Gamma)^{1/2}}}$$

where $\Gamma = AA^T$

(1) Interpretation of Γ = Covariant matrix $= (\Gamma_{ij})$

$$\Gamma = AA^T$$

$$\boxed{\Gamma_{ij}} = (AA^T)_{ij} = \sum_k A_{ik}(A^T)_{kj}$$

$$= \sum_k A_{ik}A_{jk}$$

$$\boxed{= \sum_k a_{ik}a_{jk}}$$

$$X_i = a_{i1}Z_1 + a_{i2}Z_2 + \cdots + a_{in}Z_n$$

$$X_j = a_{j1}Z_1 + a_{j2}Z_2 + \cdots + a_{jn}Z_n$$

$$\boxed{E(X_iX_j)} = E\left(\sum_{k,l} a_{ik}a_{jl}Z_kZ_l\right)$$

$$= \sum_{k,l} a_{ik}a_{jl}\underbrace{E[Z_kZ_l]}_{=\begin{cases}1 & k=l\\ 0 & k\neq l\end{cases}}$$

$$= \sum_k a_{ik}a_{jk} = \boxed{\Gamma_{ij}}$$

(2) $\boxed{\langle \Gamma^{-1}X, X\rangle \geq 0}$ $\langle (AA^T)^{-1}X, X\rangle = \|A^{-1}X\|^2$

i.e. Γ^{-1} is a positive matrix.

(3) $\Gamma = \Gamma^T$, $\Gamma_{ij} = \Gamma_{ji}$

i.e. $\boxed{\Gamma \text{ is a symmetric, positive matrix.}}$

(4) Given A, $\Gamma = AA^T$ is symmetric, positive matrix.

Given any symmetric, positive, there is a (symmetric) A s.t. $\Gamma = AA^T$, $A = "\sqrt{\Gamma}"$

Back to 2x2

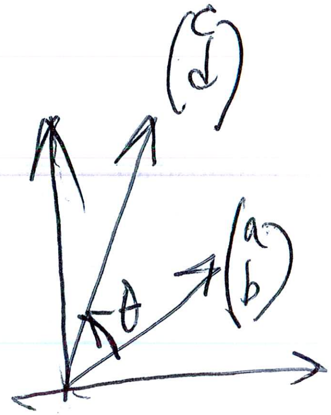
$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$

$$\begin{cases} X = aZ_1 + bZ_2 \\ Y = cZ_1 + dZ_2 \end{cases}$$

(1) $\sigma_X^2 = Var(X) = Var(aZ_1 + bZ_2)$

$= a^2 + b^2.$

$\sigma_Y^2 = Var(Y) = c^2 + d^2$



(2) $E(XY) = ac + bd$

$\begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} c \\ d \end{pmatrix}$

$= \sqrt{a^2+b^2}\sqrt{c^2+d^2} \cos\theta$

$= \sigma_X \sigma_Y \cos\theta$

$\left(\text{Let } \rho = \cos\theta \right) \quad -1 < \rho < 1$

$= \sigma_X \sigma_Y \rho$

(3) $|ad-bc| = |\det(ad-bc)|$

$= \text{Area of parallelogram}$

$= \sqrt{a^2+b^2}\sqrt{c^2+d^2}\,|\sin\theta|$

$= \sigma_X\sigma_Y\sqrt{1-\cos^2\theta}$

$= \sigma_X\sigma_Y\sqrt{1-\rho^2}$

$$\boxed{\frac{|ad-bc|}{\sigma_X\sigma_Y} = \sqrt{1-\rho^2}}$$

<u>Hence</u>

$$\boxed{p(x,y) = \frac{e^{-\left\{\frac{\frac{x^2}{\sigma_X^2}+\frac{y^2}{\sigma_Y^2}-\frac{2\rho xy}{\sigma_X\sigma_Y}}{2\sqrt{1-\rho^2}}\right\}}}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}}$$

Remark

(1) (2×2) If $\rho = 0$, i.e. $E(XY)=0$, uncorrelated,

then
$$p(x,y) = \frac{e^{-\left(\frac{x^2}{\sigma_X^2}+\frac{y^2}{\sigma_Y^2}\right)/2}}{2\pi\sigma_X\sigma_Y}$$

$$= \underbrace{\left(\frac{e^{-\frac{x^2}{2\sigma_X^2}}}{\sqrt{2\pi}\,\sigma_X}\right)}_{p_X(x)}\underbrace{\left(\frac{e^{-\frac{y^2}{2\sigma_Y^2}}}{\sqrt{2\pi}\,\sigma_Y}\right)}_{p_Y(y)}.$$

Hence X, Y are independent.

(2) $\underline{n\times n}$, If $\Gamma_{ij}=0$, then

$X_1, X_2, \cdots, X_n$ are <u>independent</u>.

Multivariate Normal r.v's and their MGF

Let $X_1, X_2, \dots, X_m$ be given by

$$\begin{cases} X_1 = a_{11} Z_1 + \dots + a_{1n} Z_n + \mu_1 \\ X_2 = a_{21} Z_1 + \dots + a_{2n} Z_n + \mu_2 \\ \vdots \quad \vdots \quad \vdots \\ X_m = a_{m1} Z_1 + \dots + a_{mn} Z_n + \mu_m \end{cases}$$

$$\vec{X} = A\vec{Z} + \vec{\mu}$$

Consider the multi-dim MGF

$$M_{X_1,\dots,X_m}(t_1, t_2, \dots, t_m) = E\, e^{(t_1 X_1 + \dots + t_m X_m)}$$

$$= E\, e^{\sum_i (t_i X_i)} = E\, e^{\sum_{i=1}^{m} t_i \left(\sum_{j=1}^{n} a_{ij} Z_j \right) + \sum_{i=1}^{m} t_i \mu_i}$$

$$= E\, e^{\sum_{j=1}^{n} \left(\sum_{i=1}^{m} t_i a_{ij} \right) Z_j + \sum_{i=1}^{m} t_i \mu_i}$$

$$= e^{\sum_{i=1}^{m} t_i \mu_i} \prod_{j=1}^{n} e^{\frac{1}{2}\left(\sum_{i=1}^{m} t_i a_{ij}\right)^2}$$

$$= e^{\sum_{i=1}^{m} t_i \mu_i + \frac{1}{2}\sum_{j=1}^{n}\left(\sum_{i=1}^{m} t_i a_{ij}\right)^2}$$

$$= e^{\sum_{i=1}^{m} t_i \mu_i + \frac{1}{2}\sum_{j=1}^{n} \sum_{k=1}^{m}\sum_{l=1}^{n} t_k t_l a_{kj} a_{lj}}$$

$$= e^{\sum_{i=1}^{m} t_i \mu_i + \frac{1}{2}\sum_{k}^{m}\sum_{l=1}^{m} t_k t_l \underbrace{\left(\sum_{j} a_{kj} a_{lj}\right)}_{\Gamma_{kl}}}$$

$$\boxed{= e^{\sum_{i=1}^{m} t_i \mu_i + \frac{1}{2}\sum_{ij} t_i t_j \Gamma_{ij}}}$$

$$\Gamma_{kl} = (AA^T)_{kl}$$

Correlation coefficients

$$\boxed{Cov(X,Y) = E[(X-EX)(Y-EY)]}$$

$$\underset{\boxed{\text{C.-S. Ineq}}}{\leq} \sqrt{E(X-EX)^2}\sqrt{E(Y-EY)^2}$$

$$= \sqrt{Var(X)}\sqrt{Var(Y)}$$

Hence

$$-1 \leq \underbrace{\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}} \leq 1$$

<u>Def</u> $\rho_{XY} = \dfrac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$

<u>Remark</u>:

$X /\!/ Y$

$X \perp Y$

(1) $\rho = 1 \Rightarrow X = \alpha Y, \quad \alpha > 0$

(2) $\rho = -1 \Rightarrow X = \alpha Y, \quad \alpha < 0$

(3) $\rho = 0 \Rightarrow X$ & Y are <u>uncorrelated</u>.

(4) $\rho > 0 \Rightarrow X$ & Y are <u>positively</u> correlated

(5) $\rho < 0 \Rightarrow X$ & Y are <u>negatively</u> correlated

Computation of Condition Probability for Bi-Variate Normal r.v.'s [R, Sec 6.5, p. 253]

$$p(x,y) = \frac{e^{-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right\}}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$$

$$p_{X|Y}(x/y) = \frac{p(x,y)}{p_Y(y)}$$

$$p_Y(y) = \frac{e^{-\frac{1}{2\sigma_y^2}(y-\mu_y)^2}}{\sqrt{2\pi}\,\sigma_y}$$

$$
= \frac{e^{-\frac{(y-\mu_x)^2}{2(1-\rho^2)}}\, e^{-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_x}{\sigma_y}\right)\right\}}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}\;\frac{e^{-\frac{1}{2\sigma_y^2}(y-\mu_y)^2}}{\sqrt{2\pi}\,\sigma_y}}
$$

$= \text{Const}(y)$

$$
= C_1(y)\, e^{-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \rho^2\left(\frac{y-\mu_y}{\sigma_y}\right)^2 - \rho^2\left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}}
$$

$$
= C_2(y)\, e^{-\frac{1}{2(1-\rho^2)}\left[\frac{x-\mu_x}{\sigma_x} - \frac{\rho}{\sigma_y}(y-\mu_y)\right]^2}
$$

$$
= C_2(y)\, e^{-\frac{1}{2(1-\rho^2)}\left[\frac{x-\mu_x-\frac{\rho\sigma_x}{\sigma_y}(y-\mu_y)}{\sigma_x}\right]^2}
$$

$$= C_2(y)\, e^{-\frac{\left[x - \underbrace{\left(\mu_x + \frac{\rho\sigma_x}{\sigma_y}(y-\mu_y)\right)}_{\tilde{\mu}_x}\right]^2}{2\underbrace{\sigma_x^2(1-\rho^2)}_{\tilde{\sigma}_x^2}}}$$

$$= \boxed{\frac{e^{-\frac{1}{2}\left(\frac{x-\tilde{\mu}_x}{\tilde{\sigma}_x}\right)^2}}{\sqrt{2\pi}\,\tilde{\sigma}_x}}$$

where $\tilde{\mu}_x = \mu_x + \frac{\rho\sigma_x}{\sigma_y}(y-\mu_y)$

$$\tilde{\sigma}_x = \sigma_x\sqrt{1-\rho^2}$$