L17: Extended concepts of Normal distribution and independence

1. Normal distributions

    (1) New definition for $N(\mu, \Sigma)$

    $$X \sim N(\mu,\,\Sigma) \;\overset{def}{\Longleftrightarrow}\; X \overset{d}{=} AZ + \mu \text{ where } Z \sim N(0,\, I_r) \text{ by its pdf and } AA' = \Sigma$$
    $$\Longrightarrow\; E(X) = \mu \text{ and } \mathrm{Cov}(X) = \Sigma.$$

    (2) Extended concept of $N(\mu, \Sigma)$

    $X \sim N(\mu,\, \Sigma)$ by its pdf $\Longrightarrow X \sim N(\mu,\, \Sigma)$ by new definitionh

    Proof If $X \sim N(\mu,\, \Sigma)$ by its pdf, then $\Sigma > 0$.
    So $Z = \Sigma^{-1/2}(X - \mu) \sim N(0, I)$ by its pdf.
    But $X = \Sigma^{1/2}[\Sigma^{-1/2}(X - \mu)] + \mu = \Sigma^{1/2}Z + \mu$ where $\Sigma^{1/2}[\Sigma^{1/2}]' = \Sigma$.
    Thus by new definition $Z \sim N(\mu,\, \Sigma)$.

    (3) A transformation

    $X \sim N(\mu,\, \Sigma) \Longrightarrow AX + b \sim N(A\mu + b,\, A\Sigma A')$

    Proof $X \sim N(\mu,\, \Sigma) \Longleftrightarrow X = BZ + \mu,\; Z \sim N(0,\, I)$ and $BB' = \Sigma$.
    So $AX + b = (AB)Z + (A\mu + b)$ where $(AB)(AB)' = A\Sigma A'$.
    Thus $AX + b \sim N(A\mu + b,\, A\Sigma A')$.

    (4) Support of $X \sim N(\mu,\, \Sigma)$.

    The support for $X \sim N(\mu,\, \Sigma)$ is $\mu + L(\Sigma)$. $(L(A) = \mathcal{R}(A) = \mathcal{C}(A) = \mathrm{S}(A))$

    Proof $X \sim N(\mu,\, \Sigma) \Longleftrightarrow X = AZ + \mu$ where $AA' = \Sigma$ and $Z \sim N(0,\, I)$.
    Thus $X = AZ + \mu \in \mu + L(A) = \mu + L(AA') = \mu + L(\Sigma)$.

    Ex1: For $Z \sim N(0,\, 1^2)$ define $X_1 = Z + 1$ and $X_2 = Z + 2$.
    Then $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} Z + \begin{pmatrix} 1 \\ 2 \end{pmatrix} \sim N\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right)$.

    Ex2: The support for $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ in Ex1 is $\begin{pmatrix} 1 \\ 2 \end{pmatrix} + L\left[\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + L\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right]$.

2. Independence

    (1) New definition for independence

    Random vectors $X \in R^m$ and $Y \in R^n$ are independent if X is a vector-valued function of Z_I, Y is a vector-valued function of Z_{II}, and Z_I and Z_{II} are independent by classical definitions, i.e., Z_I has pdf $f_1(z_1)$, Z_{II} has pdf $f_2(z_{II})$, $\begin{pmatrix} Z_I \\ Z_{II} \end{pmatrix}$ has joint pdf $f(z_I, z_{II})$, and $f(z_I,\, z_{II}) = f_1(z_I)\, f_2(z_{II})$.

    (2) Extended concept of independece

    If X and Y are independent by classical definitions, they are independent by the new definition.

    (3) Property I

    If X and Y are independent by new definition, so are functions of X and functions of Y.

    (4) Property II

    If X and Y are independent, then $P(X \in A | Y \in B) = P(X \in A)$ and
    $P(Y \in B | X \in A) = P(Y \in B)$

Proof $X \in A \Longrightarrow Z_I \in A_1$ and $Y \in B \Longrightarrow Z_{II} \in B_1$.

So $P(X \in A) = P(X_I \in A_1)$ and $P(Y \in B) = P(X_{II} \in B_1)$

Thus $P(X \in A|Y \in B) = P(Z_I \in A_1|Z_{II} \in B_1) = P(Z_I \in A_1) = P(X \in A)$.

(5) Property III

If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$.

Pf: X is a vector valued function of Z_I, so X_i is a function of Z_I, $X_i = g_i(Z_I)$.

Y is a vector-valued function of Z_{II}, so Y_j is a function of Z_{II}, $Y_j = h_j(Z_{II})$.

$$\begin{aligned} E(X_iY_j) &= E[g_i(Z_I)h_j(Z_{II})] = \int\int_{z_I,\, z_{II}} g_i(z_I)h_j(z_{II})f(z_I,\, z_{II})\, dz_I\, dz_{II} \\ &= \int\int_{z_I,\, z_{II}} g_i(z_I)h_j(z_{II})f_1(z_I)f_2(z_{II})\, dz_I\, dz_{II} \\ &= \int\int_{z_I)} g_i(z_I)f_1(z_I)dz_I \int\int_{z_{II}} h_j(z_{II})f_2(z_{II})\, dz_{II} \\ &= E[g_i(Z_I)]E[h_j(Z_{II})] = E(X_i)E(Y_j). \end{aligned}$$

So $E(XY^T) = E(X)E(Y^T)$. Hence $\mathrm{Cov}(X, Y) = E(XY^T) - E(X)E(Y^T) = 0$.

3. Independence in normal distributions

Suppose $X \sim N_n(\mu, \Sigma)$.

(1) $AX \in R^p$ and $BX \in R^q$ are independent $\Longleftrightarrow \mathrm{Cov}(AX, BX) = A\Sigma B' = 0$.

Pf: $\Longrightarrow$ has been established by (5) of 2. Now consider $\Longleftarrow$.

$X \sim N(\mu, \Sigma) \Longleftrightarrow X \overset{d}{=} DZ + \mu$ where $Z \in N(0, I_k)$ and $DD' = \Sigma$.

In $AX = ADZ + A\mu$, $AD \in R^{p\times k}$ with rank r_1 has a sub-matrix $T_1 \in R^{r_1\times k}$ with full row rank such that ADZ is a function of T_1Z and hence

$$AX = ADZ + A\mu \text{ is a function of } T_1Z.$$

In $BX = BDZ + B\mu$, $BD \in R^{q\times k}$ with rank r_2 has a sub-matrix $T_2 \in R^{r_2\times k}$ with full row rank such that BDZ is a function of T_2Z and hence

$$BX = BDZ + B\mu \text{ is a function of } T_2Z.$$

Note that $0 = A\Sigma B' = ADD'B' = (AD)(BD)'$ has a sub-matrix T_1T_2'. So $T_1T_2' = 0$. From $\begin{pmatrix} T_1 \\ T_2 \end{pmatrix} Z \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} T_1T_1' & 0 \\ 0 & T_2T_2' \end{pmatrix}\right)$, to claim the independence of T_1Z and T_2Z we need the full row rank of $\begin{pmatrix} T_1 \\ T_2 \end{pmatrix}$ or full column rank of (T_1', T_2'). But

$$0 = (T_1', T_2')\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = T_1'\alpha + T_2'\beta \Longrightarrow \begin{cases} 0 = T_1T_1'\alpha \\ 0 = T_2T_2'\beta \end{cases} \Longrightarrow \begin{cases} \alpha = (T_1T_1')^{-1}0 = 0 \\ \beta = (T_2T_2')^{-1}0 = 0 \end{cases}$$

So (T_1', T_2') has full column rank. Conclusion follows.

(2) $A\Sigma B = 0$ where $B' = B \Longrightarrow AX$ and $X'BX$ are independent.

Pf: $B' = B$ has compact form of EVD $B = P_I\Lambda_rP_I'$.

$0 = A\Sigma B = A\Sigma P_I\Lambda_rP_I' \Longrightarrow A\Sigma P_I = 0 \Longrightarrow AX$ and $P_I'X$ are independent by (1) of 3.

So AX and $X'BX = (P_I'X)'\Lambda_r(P_I'X)$ are independent by (3) of 2.

L18 Matrices with normal distributions

1. Two notations for random matrices

   (1) Settings
   $M = (m_1, ..., m_n) \in R^{p\times n}$, $\Sigma' = \Sigma = (\sigma_{ij})_{p\times p} \geq 0$ where $\sigma_{ii} = \sigma_i^2$ for $i = 1, ..., p$, and $\Psi' = \Psi = (\psi_{ij})_{n\times n} \geq 0$ where $\psi_{ii} = \psi_i^2$ for $i = 1, ..., n$ are non-random matrices. $X \in R^{p\times n}$ is a random matrix.

   (2) Notation $X \sim (M, \Sigma, \Psi)$
   We write $X \sim (M, \Sigma, \Psi)$ if $E(X) = M$, $\mathrm{Cov}(\mathrm{vec}(X), \mathrm{vec}(X)) = \Psi \otimes \Sigma$, i.e.,

   $$X \sim (M, \Sigma, \Psi) \Longleftrightarrow \mathrm{vec}(X) \sim (\mathrm{vec}(M), \Psi \otimes \Sigma).$$

   (3) Notation $X \sim N_{p\times n}(M, \Sigma, \Psi)$

   $$X \sim N_{p\times n}(M, \Sigma, \Psi) \Longleftrightarrow \mathrm{vec}(X) \sim N(\mathrm{vec}(M), \Psi \otimes \Sigma)$$

   (4) Relations

   $$\begin{aligned} X \sim N_{p\times n}(M, \Sigma, \Psi) &\Longrightarrow X \sim (M, \Sigma, \Psi) \\ &\Longleftrightarrow X_i \sim (m_i, \Sigma, \psi_i^2) = (\mu_i, \psi_i^2\Sigma) \text{ for all } i = 1, .., n. \text{ and} \\ & (X_i, X_j) \sim \left((m_i, m_j), \Sigma, \begin{pmatrix} \psi_i^2 & \psi_{ij} \\ \psi_{ji} & \psi_j^2 \end{pmatrix}\right) \text{ for all } i \neq j. \end{aligned}$$

   Ex1: $X_1, ..., X_n$ is a random sample from a $N(\mu, \sigma^2)$, then $\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim N(\mu 1_n, \sigma^2 I_n)$.

   Ex2: $X_1, ..., X_n$ is a random sample from a p-dimensional population.
   If the population has parameters (μ, Σ), then $(X_1, ..., X_n) \sim (\mu 1_n', \Sigma, I_n)$.
   If the population has the distribution $N(\mu, \Sigma)$, then $(X_1, .., X_n) \sim N_{p\times n}(\mu 1_n', \Sigma, I_n)$.

   Ex3: $X \sim (\mu, \Sigma) = (\mu, \Sigma, 1)$ and $X \sim N(\mu, \Sigma) = N(\mu, \Sigma, 1)$. $N(\mu 1_n, \sigma^2 I_n) = N(\mu 1_n, I_n, \sigma^2)$.

2. X and X'

   (1) If $X \sim N_{p\times n}(M, \Sigma, \Psi)$, then $X' \sim N_{n\times p}(M', \Psi, \Sigma)$.

   Pf: If $X \sim N_{p\times n}(M, \Sigma, \Psi)$, then $\mathrm{vec}(X) \sim N(\mathrm{vec}(M), \Psi \otimes \Sigma)$.
   With commutation matrix K_{pn}, $K_{pn}\mathrm{vec}(X) \sim N(K_{pn}\mathrm{vec}(M), K_{pn}(\Psi \otimes \Sigma)K_{np})$.
   Thus $\mathrm{vec}(X') \sim N(\mathrm{vec}(M'), \Sigma \otimes \Psi)$. Hence $X' \sim N_{n\times p}(M', \Psi, \Sigma)$.

   (2) If $X \sim (M, \Sigma, \Psi)$, then $X' \sim (M', \Psi, \Sigma)$.

   Pf: Skipped

   Ex4: Recall: For random vectors X and Y, $E(X'AY) = [E(X)]'A[E(Y)] + \mathrm{tr}(A\mathrm{Cov}(Y, X))$.
   So if $X = (X_1, ..., X_n) \sim (M, \Sigma, \Psi)$ where $M = (m_1, .., m_n)$, then

   $$E(X_i'AX_j) = (m_i'Am_j) + \mathrm{tr}(A\psi_{ji}\Sigma).$$

   Note that $E(X_i'AX_j) = (E(X'AX))_{ij}$, $m_i'Am_j = (M'AM)_{ij}$ and $\mathrm{tr}(\Sigma)\psi_{ij} = \mathrm{tr}(A\Sigma)(\Psi)_{ij}$.
   So

   $$E(X'AX) = M'AM + \mathrm{tr}(A\Sigma)\Psi.$$

Comment: In Ex4 $A \in R^{p\times p}$. What about $E(XBX')$ where $B \in R^{n\times n}$?
Hint: $X \sim (M, \Sigma, \Psi) \Longrightarrow Y = X' \sim (M', \Psi, \Sigma)$.
$E(XBX') = E(Y'BY) = \cdots$.

3. More properties

(1) Distribution of a transformation
If $X \sim N_{p\times n}(M, \Sigma, \Psi)$, then with $A \in R^{q\times p}$, $B \in R^{n\times m}$ and $C \in C^{q\times m}$

$$AXB + C \sim N_{q\times m}(AMB + C,\, A\Sigma A',\, B'\Psi B).$$

Pf: If $X \sim N_{p\times n}(M, \Sigma, \Psi)$, then $\mathrm{vec}(X) \sim N(\mathrm{vec}(M), \Psi \otimes \Sigma)$. So

$$\begin{aligned} \mathrm{vec}(AXB + C) &= (B' \otimes A)\mathrm{vec}(X) + \mathrm{vec}(C) \\ &\sim N((B' \otimes A)\mathrm{vec}(M) + \mathrm{vec}(C),\, (B' \otimes A)(\Psi \otimes \Sigma)(B \otimes A')) \\ &= N(\mathrm{vec}(AMB + C),\, (B'\Psi B) \otimes (A\Sigma A')) \end{aligned}$$

Thus $AXB + c \sim N_{q\times m}(AMB + C,\, A\Sigma A',\, B'\Psi B)$.

(2) Parameters of a transformation
If $X \sim (M, \Sigma, \Psi)$, then $AXB + C \sim (AMB + C,\, A\Sigma A',\, B'\Psi B)$.

Pf: Skipped

(3) Independence
$X \sim N_{p\times n}(M, \Sigma, \Psi)$.

$$AXB \text{ and } CXD \text{ are independent} \Longleftrightarrow A\Sigma C' = 0 \text{ or } B'\Psi D = 0$$

Pf: AXB and CXD are independent if and only if $\mathrm{vec}(AXB) = (B' \otimes A)\mathrm{vec}(X)$ and $\mathrm{vec}(CXD) = (D' \otimes C)\mathrm{vec}(X)$ are independent.
But $\mathrm{vec}(X) \sim N(\mathrm{vec}(M), \Sigma, \Psi)$. So AXB and CXD are independent if and only if $\mathrm{Cov}\,(\mathrm{vec}(AXB), \mathrm{vec}(CXD)) = 0$. But

$$\mathrm{Cov}(\mathrm{vec}(AXB), \mathrm{vec}(CXD)) = (B' \otimes A)(\Psi \otimes \Sigma)(D' \otimes C)' = (B'\Psi D) \otimes (A\Sigma C').$$

So AXB and CXD are independent if and only if $A\Sigma C' = 0$ or $B'\Psi D = 0$.