## L06 Extended definition of normal distributions

1. Extended definition of normal distributions

   (1) Defintions
   Random vector $\mathbf{x} \in R^p$ has normal distribution with mean $\mu$ and variance-covariance matrix $\Sigma$ denoted as $\mathbf{x} \sim N(\mu, \Sigma)$ if $\mathbf{x} = A\mathbf{z} + \mu$ where $\mathbf{z} \sim N(0, I_r)$ by its pdf and $AA' = \Sigma$.
   By the definition with $\mathbf{z} \sim N(0, I_r)$ all $A\mathbf{z} + \mu$ are normal.

   (2) Extended definition
   Suppose $\mathbf{x} \sim N(\mu, \Sigma)$ by its pdf. Then $\Sigma > 0$. So $\Sigma^{-1/2}$ exists. Let $\mathbf{z} = \Sigma^{-1/2}(\mathbf{x} - \mu)$. Then $\mathbf{z} \sim N(0, I)$ and $\mathbf{x} = \Sigma^{1/2}\mathbf{z} + \mu$ with $\Sigma^{1/2}\left(\Sigma^{1/2}\right)' = \Sigma$. Hence by the new definition, $\mathbf{x} \sim N(\mu, \Sigma)$.

   (3) Support
   The support of $\mathbf{z} \sim N(0, I_r)$ is $R^r$, i.e., the values of $\mathbf{z}$ occupy whole $R^r$. Thus the support of $\mathbf{x} = A\mathbf{z} + \mu$ is $AR^r + \mu$. Here $AR^r = L(A)$ is the column space of $A$. But $L(A) = L(AA') = L(\Sigma)$. So $\mathbf{x}$ has support $\mu + L(\Sigma)$.
   **Ex1:** $\mathbf{x} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right)$ has support $\begin{pmatrix} 1 \\ 2 \end{pmatrix} + L\left[\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + L\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right]$.
   **Comment:** When $\Sigma \in R^{p\times p}$ is singular, $\mathbf{x}$ does not have a pdf, and the support is not $R^p$.

   (4) A transformation
   $\mathbf{x} \sim N(\mu, \Sigma) \Longrightarrow \mathbf{y} = Bx + b \sim N(B\mu + b, B\Sigma B')$.
   **Proof** $\mathbf{x} \sim N(\mu, \Sigma) \Longleftrightarrow \mathbf{x} = A\mathbf{z} + \mu$, $\mathbf{z} \sim (0, I_r)$ and $AA' = \Sigma$.
   So $\mathbf{y} = B\mathbf{x} + b = BA\mathbf{z} + B\mu + b$ with $BA(BA)' = B\Sigma B'$. Thus $\mathbf{y} \sim N(B\mu + b, B\Sigma B')$.

2. Probability and parameters

   (1) Probability
   $\mathbf{x} \sim N(\mu, \Sigma)$ and $\mathcal{D} \subset R^p$. Find $P(\mathbf{x} \in \mathcal{D})$.
   If $\Sigma > 0$, then pdf $f(\mathbf{x})$ exists. So $P(\mathbf{x} \in \mathcal{D}) = \iint_{\mathcal{D}} f(\mathbf{x})dx_1, .., dx_p$.
   If $\Sigma$ is singular, then $\Sigma = AA'$, $A$ has full column rank $r$. So $\mathbf{x} = A\mathbf{z} + \mu$ with $\mathbf{z} \sim N(0, I_r)$.
   Let $\mathcal{D}_r = \{\mathbf{z} \in R^r : A\mathbf{x} + \mu \in \mathcal{D}\}$. Then

   $$P(\mathbf{x} \in \mathcal{D}) = P(A\mathbf{z} + \mu \in \mathcal{D}) = P(\mathbf{z} \in \mathcal{D}_r) = \iint_{R^r} f_z(\mathbf{z})dz_1, .., dz_r.$$

   **Ex2:** For $\mathbf{x} \sim N\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right)$ let $\mathcal{D} = \{-1 \le X_1 \le 1, 0 \le X_2 \le 10\}$. Find $P(\mathbf{x} \in \mathcal{D})$.

   $$\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}'. \text{ So } \mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} Z + \begin{pmatrix} 1 \\ 2 \end{pmatrix} \text{ where } Z \sim (0, 1^2).$$

   $$\begin{aligned} P(\mathbf{x} \in \mathcal{D}) &= P\left(\begin{pmatrix} Z+1 \\ Z+2 \end{pmatrix} \in \mathcal{D}\right) \\ &= P(-1 \le Z+1 \le 1, 0 \le Z+2 \le 10) = P(-2 \le Z \le 0, -2 \le Z \le 8) \\ &= P(-2 \le Z \le 0) = P(0 \le Z \le 2) = 0.4772. \end{aligned}$$

   (2) Parameters
   $\mathbf{x} \sim N(\mu, \Sigma) \Longrightarrow \mathbf{x} \sim (\mu, \Sigma)$.
   The above is true when $\Sigma > 0$.
   When $\Sigma$ is singular, $\mathbf{x} = A\mathbf{z} + \mu$ and $\mathbf{z} \sim N(0, I) \Longrightarrow \mathbf{x} = A\mathbf{z} + \mu$ and $\mathbf{z} \sim (0, I)$.
   Thus $E(\mathbf{x}) = E(A\mathbf{z} + \mu) = A\mathbf{0} + \mu = \mu$ and $\text{Cov}(\mathbf{x}) = \text{Cov}(A\mathbf{z} + \mu) = AIA' = \Sigma$.
   **Comment:** To have $E(\mathbf{x})$, the joint pdf for $\mathbf{x}$ is not a necessary condition. It is only required to have marginal pdfs for each components of $\mathbf{x}$.
   To have $\text{Cov}(\mathbf{x})$, the joint pdf is not a necessary condition. It is only requred to have marginal pdfs for all $\begin{pmatrix} X_i \\ X_j \end{pmatrix}$.

3. Extended definition for independence

(1) Extended definition for independence
$\mathbf{x} \in R^p$ and $\mathbf{y} \in R^q$ are independent if $\mathbf{x} = g_1(\mathbf{u})$, $\mathbf{y} = g_1(\mathbf{v})$ and $\mathbf{u} \in R^{r_1}$ and $\mathbf{v} \in R^{r_2}$ are independent by $f(\mathbf{u}\,\mathbf{v}) = f_1(\mathbf{u})\, f_2(\mathbf{v})$.
By this definition, if $\mathbf{u}$ and $\mathbf{v}$ are independent, then all functions of $\mathbf{u}$ are independent to all functions of $\mathbf{v}$.

(2) Extended definiton
If $\mathbf{x}$ and $\mathbf{y}$ are independent by the definition using the pdfs, then they are still independent by the extended definition.

(3) Relation to uncorrelation
If $\mathbf{x}$ and $\mathbf{y}$ are independent, then $\mathbf{x}$ and $\mathbf{y}$ are uncorrelated.

**Proof** $X_i = g_{1i}(\mathbf{u})$ and $Y_j = g_{2j}(\mathbf{v})$. So

$$\begin{aligned} E(X_iY_j) &= \iint_{R^{r_1+r_2}} g_{1i}(\mathbf{u})g_{2j}(\mathbf{v})f(\mathbf{u},\,\mathbf{v})\,du_1,..,du_{r_1}dv_1,..,dv_{r_2} \\ &= \iint_{R^{r_1}} g_{1i}(\mathbf{u})f_1(\mathbf{u})\,du_1,..,du_{r_1} \iint_{R^{r_2}} g_{2j}(\mathbf{v})f_2(\mathbf{v})\,dv_1,..,dv_{r_2} \\ &= E(X_i)E(Y_j). \end{aligned}$$

Thus $\text{cov}(X_i,\, Y_j) = E(X_iY_j) - E(X_i)E(Y_j) = 0$. Hence $\text{Cov}(\mathbf{x},\,\mathbf{y}) = \mathbf{0}$.

(4) Suppose $\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix} \sim N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\right)$. Then

$$\mathbf{x} \text{ and } \mathbf{y} \text{ are independent} \Longleftrightarrow \Sigma_{12} = 0 \Longleftrightarrow \Sigma_{21} = 0.$$

**Proof** Only show $\Leftarrow$: When $\Sigma_{12} = 0$ and $\Sigma_{21} = 0$, $\Sigma = \Sigma^{1/2}\left(\Sigma^{1/2}\right)'$ where $\Sigma^{1/2} = \begin{pmatrix}\Sigma_{11}^{1/2} & 0\\ 0 & \Sigma_{22}^{1/2}\end{pmatrix}$.
So $\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix} = \begin{pmatrix}\Sigma_{11}^{1/2} & 0\\ 0 & \Sigma_{22}^{1/2}\end{pmatrix}\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2\end{pmatrix} + \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$ where $\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2\end{pmatrix} \sim N\left(\begin{pmatrix}0\\ 0\end{pmatrix}, \begin{pmatrix}I & 0\\ 0 & I\end{pmatrix}\right)$. So $\mathbf{z}_1$ and $\mathbf{z}_2$ are independent. But $\mathbf{x} = \Sigma_{11}^{1/2}\mathbf{z}_1 + \mu_1$ is a function of $\mathbf{z}_1$ and $\mathbf{y} = \Sigma_{22}^{1/2}\mathbf{z}_2 + \mu_2$ is a function of $\mathbf{z}_2$. Hence $\mathbf{x}$ and $\mathbf{y}$ are independent.

**Ex3:** Random vector $\mathbf{x} \in R^p$ has distribution $\mathbf{x} \sim N(\mu,\,\Sigma)$. Let $A \in R^{m\times p}$ and $B \in R^{n\times p}$.

$$A\mathbf{x} + \alpha \text{ and } B\mathbf{x} + \beta \text{ are independent} \Longleftrightarrow A\Sigma B' = 0.$$

**Proof** $\begin{pmatrix}A\mathbf{x}+\alpha\\ B\mathbf{x}+\beta\end{pmatrix} = \begin{pmatrix}A\\ B\end{pmatrix}\mathbf{x} + \begin{pmatrix}\alpha\\ \beta\end{pmatrix} \sim N\left(\begin{pmatrix}A\mu+\alpha\\ B\mu+\beta\end{pmatrix}, \begin{pmatrix}A\Sigma A' & A\Sigma B'\\ B\Sigma A' & B\Sigma B'\end{pmatrix}\right)$.
So $A\mathbf{x} + \alpha$ and $B\mathbf{x} + \beta$ are independent if and only if $A\Sigma B' = 0$.