Next: Appendix B: Kernels and **Up:** Appendix A: Conditional and **Previous:** Inverse and determinant of

Marginal and conditional distributions of multivariate normal distribution

Assume an n-dimensional random vector

$$\mathbf{x}=\left[\begin{array}{c}\mathbf{x}_1\\ \mathbf{x}_2\end{array}\right]$$

has a normal distribution $N(\mathbf{x},\mu,\Sigma)$ with

$$\mu=\left[\begin{array}{c}\mu_1\\ \mu_2\end{array}\right]\quad\text{and}\quad\Sigma=\left[\begin{array}{cc}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{array}\right]$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are two subvectors of respective dimensions p and q with $p+q=n$. Note that $\Sigma=\Sigma^T$, and $\Sigma_{21}=\Sigma_{21}^T$.

Theorem 4:

Part a The marginal distributions of $\mathbf{x}_1$ and $\mathbf{x}_2$ are also normal with mean vector μ_i and covariance matrix Σ_{ii} ($i=1,2$), respectively.

Part b The conditional distribution of $\mathbf{x}_i$ given $\mathbf{x}_j$ is also normal with mean vector

$$\mu_{i|j}=\mu_i+\Sigma_{ij}\Sigma_{jj}^{-1}(\mathbf{x}_j-\mu_j)$$

and covariance matrix

$$\Sigma_{i|j}=\Sigma_{jj}-\Sigma_{ij}^T\Sigma_{ii}^{-1}\Sigma_{ij}$$

Proof: The joint density of $\mathbf{x}$ is:

$$f(\mathbf{x})=f(\mathbf{x}_1,\mathbf{x}_2)=\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}exp[-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)]=\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}exp[-\frac{1}{2}Q(\mathbf{x}_1,\mathbf{x}_2)]$$

where Q is defined as

$$Q(\mathbf{x}_1,\mathbf{x}_2)=(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)$$

$$=$$

$$[(\mathbf{x}_1-\mu_1)^T,(\mathbf{x}-\mu_2)^T]\left[\begin{array}{cc}\Sigma^{11} & \Sigma^{12}\\ \Sigma^{21} & \Sigma^{22}\end{array}\right]\left[\begin{array}{c}\mathbf{x}_1-\mu_1\\ \mathbf{x}_2-\mu_2\end{array}\right]$$

$$=(\mathbf{x}_1-\mu_1)^T\Sigma^{11}(\mathbf{x}_1-\mu_1)+2(\mathbf{x}_1-\mu_1)^T\Sigma^{12}(\mathbf{x}_2-\mu_2)+(\mathbf{x}_2-\mu_2)^T\Sigma^{22}(\mathbf{x}_2-\mu_2)$$

Here we have assumed

$$\Sigma^{-1}=\left[\begin{array}{cc}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{array}\right]^{-1}=\left[\begin{array}{cc}\Sigma^{11} & \Sigma^{12}\\ \Sigma^{21} & \Sigma^{22}\end{array}\right]$$

According to theorem 2, we have

$$\Sigma^{11}=(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^T)^{-1}=\Sigma_{11}^{-1}+\Sigma_{11}^{-1}\Sigma_{12}(\Sigma_{22}-A_{12}^T\Sigma_{11}^{-1}\Sigma_{12})^{-1}\Sigma_{12}^T\Sigma_{11}^{-1}$$

$$\Sigma^{22}=(\Sigma_{22}-\Sigma_{12}^T\Sigma_{11}^{-1}\Sigma_{12})^{-1}=\Sigma_{22}^{-1}+\Sigma_{22}^{-1}\Sigma_{12}^T(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^T)^{-1}\Sigma_{12}\Sigma_{22}^{-1}$$

$$\Sigma^{12}=-\Sigma_{11}^{-1}\Sigma_{12}(\Sigma_{22}-\Sigma_{12}^T\Sigma_{11}^{-1}\Sigma_{12})^{-1}=(\Sigma^{21})^T$$

Substituting the second expression for Σ^{11}, first expression for Σ^{22}, and Σ^{12} into $Q(\mathbf{x}_1,\mathbf{x}_2)$ to get:

$$\begin{aligned}
Q(\mathbf{x}_1,\mathbf{x}_2) &= (\mathbf{x}_1-\mu_1)^T[\Sigma_{11}^{-1}+\Sigma_{11}^{-1}\Sigma_{12}(\Sigma_{22}-A_{12}^T\Sigma_{11}^{-1}\Sigma_{12})^{-1}\Sigma_{12}^T\Sigma_{11}^{-1}](\mathbf{x}_1-\mu_1)\\
&\quad -2(\mathbf{x}_1-\mu_1)^T[\Sigma_{11}^{-1}\Sigma_{12}(\Sigma_{22}-\Sigma_{12}^T\Sigma_{11}^{-1}\Sigma_{12})^{-1}](\mathbf{x}_2-\mu_2)\\
&\quad +(\mathbf{x}_2-\mu_2)^T[(\Sigma_{22}-\Sigma_{12}^T\Sigma_{11}^{-1}\Sigma_{12})^{-1}](\mathbf{x}_2-\mu_2)\\
&= (\mathbf{x}_1-\mu_1)^T\Sigma_{11}^{-1}(\mathbf{x}_1-\mu_1)\\
&\quad +(\mathbf{x}_1-\mu_1)^T\Sigma_{11}^{-1}\Sigma_{12}(\Sigma_{22}-A_{12}^T\Sigma_{11}^{-1}\Sigma_{12})^{-1}\Sigma_{12}^T\Sigma_{11}^{-1}](\mathbf{x}_1-\mu_1)\\
&\quad -2(\mathbf{x}_1-\mu_1)^T[\Sigma_{11}^{-1}\Sigma_{12}(\Sigma_{22}-\Sigma_{12}^T\Sigma_{11}^{-1}\Sigma_{12})^{-1}](\mathbf{x}_2-\mu_2)\\
&\quad +(\mathbf{x}_2-\mu_2)^T[(\Sigma_{22}-\Sigma_{12}^T\Sigma_{11}^{-1}\Sigma_{12})^{-1}](\mathbf{x}_2-\mu_2)\\
&= (\mathbf{x}_1-\mu_1)^T\Sigma_{11}^{-1}(\mathbf{x}_1-\mu_1)\\
&\quad +[(\mathbf{x}_2-\mu_2)-\Sigma_{12}^T\Sigma_{11}^{-1}(\mathbf{x}_1-\mu_1)]^T(\Sigma_{22}-\Sigma_{12}^T\Sigma_{11}^{-1}\Sigma_{12})^{-1}[(\mathbf{x}_2-\mu_2)-\Sigma_{12}^T\Sigma_{11}^{-1}(\mathbf{x}_1-\mu_1)]
\end{aligned}$$

The last equal sign is due to the following equations for any vectors u and v and a symmetric matrix $A=A^T$:

$$\begin{aligned} & u^T A u - 2u^T A v + v^T A v = u^T A u - u^T A v - u^T A v + v^T A v \\ = \; & u^T A(u-v) - (u-v)^T A v = u^T A(u-v) - v^T A(u-v) \\ = \; & (u-v)^T A(u-v) = (v-u)^T A(v-u) \end{aligned}$$

We define

$$b \stackrel{\triangle}{=} \mu_2 + \Sigma_{12}^T \Sigma_{11}^{-1}(\mathbf{x}_1 - \mu_1)$$

$$A \stackrel{\triangle}{=} \Sigma_{22} - \Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12}$$

and

$$\left\{ \begin{array}{ll} Q_1(\mathbf{x}_1) & \stackrel{\triangle}{=} (\mathbf{x}_1 - \mu_1)^T \Sigma_{11}^{-1} (\mathbf{x}_1 - \mu_1) \\ Q_2(\mathbf{x}_1, \mathbf{x}_2) & \stackrel{\triangle}{=} [(\mathbf{x}_2 - \mu_2) - \Sigma_{12}^T \Sigma_{11}^{-1}(\mathbf{x}_1 - \mu_1)]^T (\Sigma_{22} - \Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12})^{-1} [(\mathbf{x}_2 - \mu_2) - \Sigma_{12}^T \Sigma_{11}^{-1}(\mathbf{x}_1 - \mu_1)] \\ & = (\mathbf{x}_2 - b)^T A^{-1} (\mathbf{x}_2 - b) \end{array} \right.$$

and get

$$Q(\mathbf{x}_1, \mathbf{x}_2) = Q_1(\mathbf{x}_1) + Q_2(\mathbf{x}_1, \mathbf{x}_2)$$

Now the joint distribution can be written as:

$$\begin{aligned} f(\mathbf{x}) \; = \; & f(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} exp[-\frac{1}{2} Q(\mathbf{x}_1, \mathbf{x}_2)] \\ = \; & \frac{1}{(2\pi)^{n/2} |\Sigma_{11}|^{1/2} |\Sigma_{22} - \Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12}|^{1/2}} exp[-\frac{1}{2} Q(\mathbf{x}_1, \mathbf{x}_2)] \\ = \; & \frac{1}{(2\pi)^{p/2} |\Sigma_{11}|^{1/2}} exp[-\frac{1}{2}(\mathbf{x}_1 - \mu_1)^T \Sigma_{11}^{-1} (\mathbf{x}_1 - \mu_1)] \frac{1}{(2\pi)^{q/2} |A|^{1/2}} exp[-\frac{1}{2}(\mathbf{x}_2 - b)^T A^{-1} (\mathbf{x}_2 - b)] \\ = \; & N(\mathbf{x}_1, \mu_1, \Sigma_{11}) \; N(\mathbf{x}_2, b, A) \end{aligned}$$

The third equal sign is due to theorem 3:

$$|\Sigma| = |\Sigma_{11}||\Sigma_{22} - \Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12}|$$

The marginal distribution of $\mathbf{x}_1$ is

$$f_1(\mathbf{x}_1)=\int f(\mathbf{x}_1,\mathbf{x}_2)d\mathbf{x_2}=\frac{1}{(2\pi)^{p/2}|\Sigma_{11}|^{1/2}}exp[-\frac{1}{2}(\mathbf{x}_1-\mu_1)^T\Sigma_{11}^{-1}(\mathbf{x}_1-\mu_1)]$$

and the conditional distribution of $\mathbf{x}_2$ given $\mathbf{x}_1$ is

$$f_{2|1}(\mathbf{x}_2|\mathbf{x}_1)=\frac{f(\mathbf{x}_1,\mathbf{x}_2)}{f(\mathbf{x}_1)}=\frac{1}{(2\pi)^{q/2}|A|^{1/2}}exp[-\frac{1}{2}(\mathbf{x}_2-b)^T A^{-1}(\mathbf{x}_2-b)]$$

with

$$b=\mu_2+\Sigma_{12}^T\Sigma_{11}^{-1}(\mathbf{x}_1-\mu_1)$$

$$A=\Sigma_{22}-\Sigma_{12}^T\Sigma_{11}^{-1}\Sigma_{12}$$

Next: Appendix B: Kernels and **Up:** Appendix A: Conditional and **Previous:** Inverse and determinant of
Ruye Wang 2006-11-14