Homework 1

April 1, 2014

1. Assume the expectations of X and Y exist. Prove that, for $g(x, y)$,

$$\mathbb{E}(\mathbb{E}(g(X,Y) \mid X)) = \mathbb{E}(g(X,Y)).$$

Solution

$$\begin{aligned}
\mathbb{E}[\mathbb{E}(g(X,Y)|X)] &= \int_{-\infty}^{+\infty} \mathbb{E}(g(X,Y)|X = x) dF_X(x) \\
&= \int_{-\infty}^{+\infty} (\int_{-\infty}^{+\infty} g(x,y) dF_{Y|X}(y|x)) dF_X(x) \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) dF(x,y) \\
&= \mathbb{E}[g(X,Y)]
\end{aligned}$$

2. Assume X and Y are random variables. Prove that

$$\mathrm{Var}(Y) = \mathbb{E}(\mathrm{Var}(Y \mid X)) + \mathrm{Var}(\mathbb{E}(Y \mid X)).$$

Solution

$$\mathrm{Var}(Y|X) = \mathbb{E}(Y^2|X) - [\mathbb{E}(Y|X)]^2$$

$$\mathbb{E}(\mathrm{Var}(Y|X)) = \mathbb{E}(Y^2) - \mathbb{E}[\mathbb{E}(Y|X)]^2$$

$$\begin{aligned}
\mathrm{Var}(\mathbb{E}(Y|X)) &= \mathbb{E}[\mathbb{E}(Y|X)]^2 - \{\mathbb{E}[\mathbb{E}(Y|X)]\}^2 \\
&= \mathbb{E}[\mathbb{E}(Y|X)]^2 - (\mathbb{E}(Y))^2
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 \\
&= \mathbb{E}(\mathrm{Var}(Y|X)) + \mathrm{Var}(\mathbb{E}(Y|X))
\end{aligned}$$

3. Assume $X = (X_1, \ldots, X_m)^T \sim \mathcal{N}_m(\mathbf{0}, \mathbf{\Sigma})$. Let

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \begin{matrix} p \\ q \end{matrix}, \quad \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix} \begin{matrix} p \\ q \end{matrix}, \quad \mathbf{\Theta} = \mathbf{\Sigma}^{-1} = \begin{pmatrix} \mathbf{\Theta}_{11} & \mathbf{\Theta}_{12} \\ \mathbf{\Theta}_{21} & \mathbf{\Theta}_{22} \end{pmatrix} \begin{matrix} p \\ q \end{matrix},$$

and $\mathbf{\Sigma}_{11\cdot 2} = \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$. Prove that $\mathbf{\Sigma}_{11\cdot 2} = \mathbf{\Theta}_{11}^{-1}$ and $\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12} = -\mathbf{\Theta}_{12}\mathbf{\Theta}_{22}^{-1}$.

Solution

$$\mathbf{\Sigma\Theta} = \begin{pmatrix} \mathbf{\Sigma}_{11}\mathbf{\Theta}_{11} + \mathbf{\Sigma}_{12}\mathbf{\Theta}_{21} & \mathbf{\Sigma}_{11}\mathbf{\Theta}_{12} + \mathbf{\Sigma}_{12}\mathbf{\Theta}_{22} \\ \mathbf{\Sigma}_{21}\mathbf{\Theta}_{11} + \mathbf{\Sigma}_{22}\mathbf{\Theta}_{21} & \mathbf{\Sigma}_{21}\mathbf{\Theta}_{12} + \mathbf{\Sigma}_{22}\mathbf{\Theta}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{pmatrix}$$

Thus we have

$$\begin{cases} \mathbf{\Sigma}_{11}\mathbf{\Theta}_{11} + \mathbf{\Sigma}_{12}\mathbf{\Theta}_{21} = \mathbf{I} & (1) \\ \mathbf{\Sigma}_{11}\mathbf{\Theta}_{12} + \mathbf{\Sigma}_{12}\mathbf{\Theta}_{22} = 0 & (2) \\ \mathbf{\Sigma}_{21}\mathbf{\Theta}_{11} + \mathbf{\Sigma}_{22}\mathbf{\Theta}_{21} = 0 & (3) \\ \mathbf{\Sigma}_{21}\mathbf{\Theta}_{12} + \mathbf{\Sigma}_{22}\mathbf{\Theta}_{22} = \mathbf{I} & (4) \end{cases}$$

Solve the equations. Then we can get

$$(1),(3) \Rightarrow \mathbf{\Theta}_{11}^{-1} = \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$$
$$(2) \Rightarrow \mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12} = -\mathbf{\Theta}_{12}\mathbf{\Theta}_{22}^{-1}$$

4. Suppose the multivariate t distribution has density

$$st_m(\mathbf{x} \mid \boldsymbol{\mu}, \mathbf{\Sigma}, \nu) = \frac{\Gamma[(\nu+m)/2]}{\Gamma(\nu/2)(\nu\pi)^{m/2}} |\mathbf{\Sigma}|^{-1/2} \left[1 + \frac{1}{\nu}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{-(\nu+m)/2},$$

with $\nu > 0$. Let

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \begin{matrix} p \\ q \end{matrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \begin{matrix} p \\ q \end{matrix}, \quad \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix} \begin{matrix} p \\ q \end{matrix}.$$

Compute $p(\mathbf{x}_1)$ and $p(\mathbf{x}_2 \mid \mathbf{x}_1)$.

Solution

Since $X \sim st_m(\mathbf{x}|\boldsymbol{\mu}, \mathbf{\Sigma}, \nu)$, then

$$X = \boldsymbol{\mu} + \frac{1}{\sqrt{V}}Y,$$

where $Y \sim \mathcal{N}_m(0, \mathbf{\Sigma})$ and $\nu \sim Gamma(\nu/2, \nu/2)$. Applying the linear transformation we have

$$\mathbf{A}X = \mathbf{A}\boldsymbol{\mu} + \frac{1}{\sqrt{V}}\mathbf{A}Y,$$

where $\mathbf{A} \in \mathbb{R}^{p\times m}$. Then we have

$$\begin{aligned}
\mathbf{A}Y &\sim \mathcal{N}_p(0, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T), \\
\mathbf{A}X &\sim st_p(\mathbf{x}|\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T, \nu).
\end{aligned}$$

Let $\mathbf{A} = [\mathbf{I}_p, 0]$, we obtain

$$X_1 = \mathbf{A}X \sim st_p(\mathbf{x}_1|\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, \nu).$$

Then

$$p(\mathbf{x}_1) \quad = \quad \frac{\Gamma[(\nu+p)/2]}{\Gamma(\nu/2)(\nu\pi)^{p/2}}|\boldsymbol{\Sigma}_{11}|^{-1/2}\left[1+\frac{1}{\nu}(\mathbf{x}_1-\boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1)\right]^{-(\nu+p)/2},$$

and

$$\begin{aligned}
& p(\mathbf{x}_2|\mathbf{x}_1) \\
= \;& p(\mathbf{x})/p(\mathbf{x}_1) \\
= \;& \frac{\frac{\Gamma[(\nu+m)/2]}{\Gamma(\nu/2)(\nu\pi)^{m/2}}|\boldsymbol{\Sigma}|^{-1/2}\left[1+\frac{1}{\nu}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{-(\nu+m)/2}}{\frac{\Gamma[(\nu+p)/2]}{\Gamma(\nu/2)(\nu\pi)^{p/2}}|\boldsymbol{\Sigma}_{11}|^{-1/2}\left[1+\frac{1}{\nu}(\mathbf{x}_1-\boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1)\right]^{-(\nu+p)/2}} \\
= \;& \frac{\Gamma[(\nu+m)/2]}{\Gamma[(\nu+p)/2](\nu\pi)^{q/2}}|\boldsymbol{\Sigma}_{22.1}|^{-1/2}\frac{\left[1+\frac{1}{\nu}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{-(\nu+m)/2}}{\left[1+\frac{1}{\nu}(\mathbf{x}_1-\boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1)\right]^{-(\nu+m)/2+q/2}} \\
= \;& \frac{\Gamma[(\nu+m)/2]}{\Gamma[(\nu+p)/2](\nu\pi)^{q/2}\left[1+\frac{1}{\nu}(\mathbf{x}_1-\boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1)\right]^{q/2}}|\boldsymbol{\Sigma}_{22.1}|^{-1/2} \\
& \frac{\left[1+\frac{1}{\nu}(\mathbf{x}_1-\boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1)+\frac{1}{v}(\mathbf{x}_{2.1}-\boldsymbol{\mu}_{2.1})^T\boldsymbol{\Sigma}_{22.1}^{-1}(\mathbf{x}_{2.1}-\boldsymbol{\mu}_{2.1})\right]^{-(v+m)/2}}{\left[1+\frac{1}{\nu}(\mathbf{x}_1-\boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1)\right]^{-(\nu+m)/2}} \\
= \;& \frac{\Gamma[(\nu_{2.1}+q)/2]}{\Gamma(\nu_{2.1}/2)(\nu_{2.1}\pi)^{q/2}}|\boldsymbol{\Sigma}_{2.1}|^{-1/2}\left[1+\frac{1}{v_{2.1}}(\mathbf{x}_{2.1}-\boldsymbol{\mu}_{2.1})^T\boldsymbol{\Sigma}_{2.1}^{-1}(\mathbf{x}_{2.1}-\boldsymbol{\mu}_{2.1})\right]^{-(v_{2.1}+q)/2} \\
= \;& st_m(\mathbf{x}_{2.1}|\boldsymbol{\mu}_{2.1}, \boldsymbol{\Sigma}_{2.1}, \nu_{2.1}),
\end{aligned}$$

where

$$
\begin{cases}
\nu_{2.1} = \nu + p, \\
\mathbf{x}_{2.1} = \mathbf{x}_2 + \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1), \\
\boldsymbol{\mu}_{2.1} = \boldsymbol{\mu}_2 + \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1), \\
\mathbf{\Sigma}_{22.1} = \mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}, \\
\mathbf{\Sigma}_{2.1} = \dfrac{\nu + (\mathbf{x}_1 - \boldsymbol{\mu}_1)^T \mathbf{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \mu_1)}{\nu + p}\mathbf{\Sigma}_{22.1}.
\end{cases}
$$