1.3 HW 3

作业 3 链接

✍ **练习 1.1** 假设随机向量 $\mathbf{x},\mathbf{y},\mathbf{z}$ 的联合分布具有如下的形式:

$$p(\mathbf{x},\mathbf{y},\mathbf{z}) = C\exp\left(-||\mathbf{x}-\mathbf{z}||_2^2 - ||\mathbf{y}-\mathbf{z}||_2^2\right)$$

其中 C 是常数。判断(不用推导)给定 $\mathbf{z}$ 时,$\mathbf{x}$ 和 $\mathbf{y}$ 是否独立,以及 $\mathbf{x}$ 的条件分布的形式 $N(\_\_,\_\_)$ 。

证明 给定 $\mathbf{z}$ 时,$\mathbf{x}$ 和 $\mathbf{y}$ 独立. $\mathbf{x}$ 的条件分布的形式 $N(\mathbf{z},\frac{1}{2}\mathbf{I})$. 其中 $\mathbf{I}$ 是一个单位矩阵,边际维数和随机变量的维数相同.

✍ **练习 1.2** $\mathbf{x}\sim\mathcal{N}_p(\mu,\Omega^{-1})$,试求给定 $\mathbf{x}_{-(ij)}=(x_k,k\neq i,j)$ 条件下,$(x_i,x_j)^\top$ 的条件分布。

解 不妨设 $i=1,j=2$, 记 $\mathbf{x}=\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix}$ 其中 $\mathbf{x}_1^\top=(x_1,x_2)^\top$, $\mathbf{x}_2=\mathbf{x}_{-(12)}$. 则我们可以知道

$$\mathbf{x}_1|\mathbf{x}_2\sim\mathcal{N}\left(\mu_1-\Sigma_{12}\Sigma_{22}^{-1}(\mu_2-\mathbf{x}_2),\Sigma_{11\cdot 2}\right)$$

由于 $\Sigma_{12}=-\Omega_{11}^{-1}\Omega_{12}\Omega_{22\cdot 1}^{-1}$, $\Sigma_{22}^{-1}=\Omega_{22\cdot 1}$ 且 $\Sigma_{11\cdot 2}^{-1}=\Omega_{11}$. 我们有

$$\mathbf{x}_1|\mathbf{x}_2\sim\mathcal{N}\left(\mu_1+\Omega_{11}^{-1}\Omega_{12}(\mu_2-\mathbf{x}_2),\Omega_{11}^{-1}\right)$$

✍ **练习 1.3** 假设随机向量 $\mathbf{z}_1,...,\mathbf{z}_m\in\mathbb{R}^p$ 独立同分布,记 $Z_{m\times p}=(\mathbf{z}_1,...,\mathbf{z}_m)^\top$,假设对任何 $m\times m$ 的正交矩阵 H ,都有 $HZ\overset{d}{=}Z$,证明 $\mathbf{z}_1\sim\mathcal{N}_p(\mathbf{0},\Sigma)$

证明 对于任意一个 $\mathbf{t}\in\mathbb{R}^p$,我们来考察 $Z\mathbf{t}\in\mathbb{R}^m$ 这个新随机向量。可以发现首先这个随机向量各个维度仍然是相互独立的,这是因为 $(\mathbf{z}_1^\top,...,\mathbf{z}_m^\top)$ 是独立的,所以 $(\mathbf{z}_1^\top\mathbf{t},...,\mathbf{z}_m^\top\mathbf{t})$ 也是相互独立的。又因为 $HZ\overset{d}{=}Z$,所以我们有 $HZ\mathbf{t}\overset{d}{=}Z\mathbf{t}$,这意味着 $Z\mathbf{t}$ 服从球对称分布。因此根据 lecture02 里的 Maxwell-Hershell 定理,我们可以发现 $Z\mathbf{t}\sim\mathcal{N}_m(0,\sigma_t^2I_m)$,从而可以得到 $\mathbf{z}_1^\top\mathbf{t}\sim\mathcal{N}(0,\sigma_t^2)$.

在这里值得注意的一件事情是 σ_t^2 是一个和 t 有关的量,并且可以显式算出来,这是因为 $\sigma_t^2=\mathrm{Var}(\mathbf{z}_1^\top\mathbf{t})=\mathbf{E}\left[(\mathbf{z}_1^\top\mathbf{t})^2\right]=\mathbf{E}\left[\mathbf{t}^\top\mathbf{z}_1\mathbf{z}_1^\top\mathbf{t}\right]=\mathbf{t}^\top\Sigma\mathbf{t}$. 因此我们有 $\mathbf{z}_1^\top\mathbf{t}\sim\mathcal{N}(0,\mathbf{t}^\top\Sigma\mathbf{t})$.

现在这个问题就变成了,如果对于任意一个给定的 $\mathbf{t}\in\mathbb{R}^p$,都有 $\mathbf{z}_1^\top\mathbf{t}\sim\mathcal{N}(0,\mathbf{t}^\top\Sigma\mathbf{t})$,那么是不是可以说明 $\mathbf{z}_1\sim\mathcal{N}_p(\mathbf{0},\Sigma)$ 。答案是肯定的! 我们可以用特征函数说明一下这个事情。因为对于任意给定的实数 s ,$\mathbf{E}\left[e^{is\mathbf{z}_1^\top\mathbf{t}}\right]=e^{-\frac{1}{2}\mathbf{t}^\top\Sigma\mathbf{t}s^2}$,所以我们有 $\mathbf{E}\left[e^{i(s\mathbf{t})^\top\mathbf{z}_1}\right]=e^{-\frac{1}{2}(s\mathbf{t})^\top\Sigma(s\mathbf{t})}$. 由于 $\mathbf{t}$ 是任意给定的 p 维度的向量,s 是任意给定的标量,我们可以让 $\mathbf{t}$ 是一个任意给定的 p 维的单位向量,以上的论述也会全部成立。因此我们有:

$$\forall\mathbf{t}\in\mathbb{R}^p,||\mathbf{t}||=1,\forall s,\mathbf{E}\left[e^{i(s\mathbf{t})^\top\mathbf{z}_1}\right]=e^{-\frac{1}{2}(s\mathbf{t})^\top\Sigma(s\mathbf{t})}$$

注意到上述命题等价于,

$$\forall\mathbf{t}\in\mathbb{R}^p,\mathbf{E}\left[e^{i\mathbf{t}^\top\mathbf{z}_1}\right]=e^{-\frac{1}{2}\mathbf{t}^\top\Sigma\mathbf{t}}$$

这表明 $\mathbf{z}_1\in\mathcal{N}_p(\mathbf{0},\Sigma)$.

注 上述证明依赖于 $\mathbf{E}\left[\mathbf{z}_1\mathbf{z}_1^\top\right]$ 存在,如果它不存在呢? 答曰,它是一定存在的。这是因为 $\mathbf{t}\in\mathbb{R}^p$,因为我们已经知道了 $\mathbf{z}_1^\top\mathbf{t}\sim\mathcal{N}(0,\sigma_t^2)$,从而必须有:

$$\mathbf{E}(\mathbf{z}_1^\top\mathbf{t})^2<\infty$$

即

$$\mathbf{t}^\top\mathbf{E}\left[\mathbf{z}_1\mathbf{z}_1^\top\right]\mathbf{t}<\infty,\forall\mathbf{t}\in\mathbb{R}^p$$

取 $\mathbf{t}=\mathbf{e}_i,\forall i=1,...,p$ 可知各个维度的方差存在,再取 $\mathbf{t}=\mathbf{e}_i+\mathbf{e}_j$ 可知任意两个维度的协方差存在,协方差矩阵的各个元素都存在,因此 $\mathbf{E}\left[\mathbf{z}_1\mathbf{z}_1^\top\right]$ 存在,我们记 $\mathbf{E}\left[\mathbf{z}_1\mathbf{z}_1^\top\right]=\Sigma$ 时,该 Σ 是定义良好的.

✍ **练习 1.4** 计算二元 Gamma 函数

$$\Gamma_2(x)=\int_D e^{-a_{11}-a_{22}}|a_{11}a_{22}-a_{12}^2|^{x-\frac{3}{2}}da_{11}da_{12}da_{22},\quad x>\frac{1}{2},$$

其中积分区域

$$D = \{(a_{11}, a_{12}, a_{22}) \in \mathbb{R}^3 : a_{11}a_{22} - a_{12}^2 > 0, a_{11} > 0, a_{22} > 0\}.$$

提示：可以做变量代换：

$$a_{11} = t_{11}^2, \quad a_{12} = t_{11}t_{12}, \quad a_{22} = t_{12}^2 + t_{22}^2$$

（参见第 6 讲定理 A5），或其他任何形式的合理代换。

证明 换元后的积分区域为

$$D' = \{(t_{11}, a_{12}, t_{22}) \in \mathbb{R}^3 : t_{11} > 0, t_{22} > 0\}.$$

所以我们有：

$$\begin{aligned}
\Gamma_2(x) &= \int_D e^{-a_{11}-a_{22}} |a_{11}a_{22} - a_{12}^2|^{x-\frac{3}{2}} da_{11}da_{12}da_{22} \\
&= \int_D e^{-t_{11}^2-t_{12}^2-t_{22}^2} |t_{11}^2 t_{22}^2|^{x-\frac{3}{2}} |4t_{11}^2 t_{22}| dt_{11}dt_{12}dt_{22} \\
&=4 \int_0^{+\infty} e^{-t_{11}^2} t_{11}^{2x-1} dt_{11} \int_{-\infty}^{+\infty} e^{-t_{12}^2} t_{12}^{2\cdot\frac{1}{2}-1} dt_{12} \int_0^{+\infty} e^{-t_{22}^2} t_{22}^{2(x-\frac{1}{2})-11} dt_{22} \\
&=\left(2\int_0^{+\infty} e^{-t_{11}^2} t_{11}^{2x-1} dt_{11}\right)\left(2\int_0^{+\infty} e^{-t_{12}^2} t_{12}^{2\cdot\frac{1}{2}-1} dt_{12}\right)\left(2\int_0^{+\infty} e^{-t_{22}^2} t_{22}^{2(x-\frac{1}{2})-11} dt_{22}\right) \\
&=\sqrt{\pi}\Gamma(x)\Gamma(x-\frac{1}{2}), \forall x > \frac{1}{2}
\end{aligned}$$

注 在这里面为了要保证 $a_{11} > 0, a_{22} > 0$, t_{11} 和 t_{22} 的积分区域就不可以经过原点，但是 t_{12} 的可以允许经过原点。