1. 假设随机向量 $\mathbf{x}_1,...,\mathbf{x}_m$ iid $\sim N_p(\mathbf{0},\Sigma)$, $X=(\mathbf{x}_1,...,\mathbf{x}_m)^\top$。假设 H 是任一 $m\times m$ 常数正交矩阵，$Y=HX\stackrel{\Delta}{=}(\mathbf{y}_1,...,\mathbf{y}_m)^\top$, 证明 $\mathbf{y}_1,...,\mathbf{y}_m$ iid $\sim N_p(\mathbf{0},\Sigma)$。

   证明 1: X 的概率密度函数 (即 $\mathbf{x}_1,...,\mathbf{x}_m$ 的联合概率密度函数)

   $$f(X)=c\exp(-\frac{1}{2}tr(\Sigma^{-1}X^\top X)$$

   所以 Y 的密度函数

   $$g(Y)=J(X\to Y)\times c\exp(-\frac{1}{2}tr\Sigma^{-1}(Y^\top HH^\top Y))=c\exp(-\frac{1}{2}tr(\Sigma^{-1}Y^\top Y)),$$

   其中 Jacobian $J(X\to Y)=|det(H)^p|=1$（参见第 10 讲 P22 定理 A2)。所以 Y 的各行 iid $\sim N_p(0,\Sigma)$。

   证明 2: 令 $\mathbf{z}_i=\Sigma^{-1/2}\mathbf{x}_i\sim N_p(0,I_p)$, $Z=(\mathbf{z}_1,...,\mathbf{z}_m)^\top=X\Sigma^{-1/2}$ 的所有元素 iid 服从 $N(0,1)$ 分布。对于正交矩阵 H, 故 $\tilde{Z}=HZ=(\tilde{\mathbf{z}}_1,...,\tilde{\mathbf{z}}_m)^\top$ 的所有元素 iid 服从 $N(0,1)$ 分布，从而 $\tilde{Z}\Sigma^{1/2}=(\Sigma^{1/2}\tilde{\mathbf{z}}_1,...,\Sigma^{1/2}\tilde{\mathbf{z}}_m)^\top$, 其中 $\Sigma^{1/2}\tilde{\mathbf{z}}_i\sim N_p(0,\Sigma), i=1,...,m$, 且各行独立。

   证明 3: 利用拉直运算和 Kronecker 乘积。已知

   $$vec(X^\top)=\begin{pmatrix}\mathbf{x}_1\\ \vdots\\ \mathbf{x}_m\end{pmatrix}\sim N_{mp}(\mathbf{0},I_m\otimes\Sigma)$$

   由第 7 讲引理 4,$vec(AXB)=(B^\top\otimes Avec(X)))$

   $$vec(Y^\top)=vec(X^\top H)=(H^\top\otimes I_p)vec(X^\top)\sim N_{mp}(\mathbf{0},(H^\top\otimes I_p)(I_m\otimes\Sigma)(H\otimes I_p))=N_{mp}(0,I_m\otimes\Sigma).$$

   这表明 Y 的各个行向量独立同分布服从 $N_p(\mathbf{0},\Sigma)$。

2. 若 $W\sim W_p(m,\Sigma)$, $\Sigma>0, m\ge p$, 则 $|W|\stackrel{d}{=}|\Sigma|\prod_{i=1}^p\chi^2_{m-p+i}$.

   证明: 划分 $W=\begin{pmatrix}w_{11} & \mathbf{w}_{12}\\ \mathbf{w}_{21} & W_{22}\end{pmatrix}$, 其中左上角为标量，则

   $$|W|=w_{11\bullet2}|W_{22}|,\quad w_{11\bullet2}\sim\sigma_{11\bullet2}\chi^2_{m-p+1}$$

   进一步划分 W_{22}, 递归得证。

3. 假设 $W\sim W_p(m),\mathbf{z}\sim N_p(0,I_p)$, 两者独立，证明 $\mathbf{z}^\top\mathbf{z}\left(1+\frac{1}{\mathbf{z}^\top W^{-1}\mathbf{z}}\right)\sim\chi^2_{m+1}$.

   证明: 由第 7 讲引理 1，给定 $\mathbf{z}$ 时，

   $$\mathbf{z}^\top\mathbf{z}/\mathbf{z}^\top W^{-1}\mathbf{z}\sim\chi^2_{m-p+1}$$

   该分布与 $\mathbf{z}$ 无关，故与 $\mathbf{z}$ 独立，进而与 $\mathbf{z}^\top\mathbf{z}\sim\chi^2_p$ 独立, 所以

   $$\mathbf{z}^\top\mathbf{z}+\mathbf{z}^\top\mathbf{z}/\mathbf{z}^\top W^{-1}\mathbf{z}\sim\chi^2_{m+1}.$$

4. 假设 $W \sim W_p(m, \Sigma)$，$\Sigma > 0$.

   (a) 证明 $E(W) = m\Sigma$.

   (b) 假设 $m > p+1$，对任何常数向量 $\mathbf{t} \in S^{p-1}$ (即 $\|\mathbf{t}\| = 1$), 求 $E(\mathbf{t}^\top W^{-1}\mathbf{t})$。

   (c) 假设 $m \geq p+2$，证明 $E(W^{-1}) = \frac{1}{m-p-1}\Sigma^{-1}$。

证明: (a) 按定义 $W \stackrel{d}{=} \sum_{i=1}^m \mathbf{z}_i\mathbf{z}_i^\top$, 其中 $\mathbf{z}_1, ..., \mathbf{z}_m$ iid $\sim N_p(0, \Sigma)$, 因为 $E(\mathbf{z}_i\mathbf{z}_i^\top) = \Sigma$, 所以 $E(W) = m\Sigma$。

(b) 容易验证 $E(1/\chi_k^2) = 1/(k-2), k > 2$ (即, 若 r.v. $x \sim \chi_k^2$, 则 $E(1/x) = 1/(k-2)$)。

首先证明 $\Sigma = I_p$ 的情形。由第 7 讲定理 1，对于 $\mathbf{t} \in R^p$,

$$\mathbf{t}^\top\mathbf{t}/\mathbf{t}^\top W^{-1}\mathbf{t} \sim \chi^2_{m-p+1} \Rightarrow E(\mathbf{t}^\top W^{-1}\mathbf{t}) = \mathbf{t}^\top\mathbf{t}/(m-p-1).$$

对于一般的 Σ, 令 $U = \Sigma^{-1/2}W\Sigma^{-1/2} \sim W_p(m, I_p)$，记 $\mathbf{s} = \Sigma^{1/2}\mathbf{t}$, 则

$$\mathbf{t}^\top W^{-1}\mathbf{t} = \mathbf{t}^\top\Sigma^{-1/2-1}\Sigma^{-1/2}\mathbf{t} = \mathbf{s}^\top U^{-1}\mathbf{s}$$

其中 $\|\mathbf{s}\|^2 = \mathbf{s}^\top\mathbf{s} = \mathbf{t}^\top\Sigma^{-1}\mathbf{t}$. 由单位阵情形下的结果，

$$E(\mathbf{s}^\top U^{-1}\mathbf{s}) = \mathbf{s}^\top\mathbf{s}/(m-p-1) \Rightarrow E(\mathbf{t}^\top W^{-1}\mathbf{t}) = \mathbf{t}^\top\Sigma^{-1}\mathbf{t}/(m-p-1).$$

(c) 记 $A = \Sigma^{-1}$。(b) 中最后一式说明对任何 $\mathbf{t} \in R^p, \|\mathbf{t}\| = 1$，$\mathbf{t}^\top A\mathbf{t} = \mathbf{t}^\top\Sigma^{-1}\mathbf{t}/(m-p-1)$ ，所以必有 $EW^{-1} = \Sigma^{-1}/(m-p-1)$。

5. （选）假设 $W \sim W_p(m, \Sigma)$, $A_{k\times p}$ 是行满秩常数矩阵，证明 $(AW^{-1}A^\top)^{-1} \sim W_k(m-p+k, (A\Sigma^{-1}A^\top)^{-1})$. (提示: 先证明 $\Sigma = I_p$ 情形)。

证明: 先证明 $\Sigma = I_p$ 情形。对于 $W \sim W_p(m)$，我们需要证

$$\begin{aligned}
&(AW^{-1}A^\top)^{-1} \sim W_k(m-p+k, (AA^\top)^{-1}) \\
\Leftrightarrow &\left((AA^\top)^{-1/2}AW^{-1}A^\top(AA^\top)^{-1/2}\right)^{-1} \sim W_k(m-p+k, I_k) \\
\Leftrightarrow &\left(BW^{-1}B^\top\right)^{-1} \sim W_k(m-p+k, I_k), \quad B = (AA^\top)^{-1/2}A
\end{aligned}$$

因为 $B^\top B = A^\top(AA^\top)^{-1}A = P_{A^\top}$ 为 A 行空间的投影矩阵，存在 p 阶正交矩阵 H 使得

$$H^\top B^\top BH = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$$

划分 $C_{k\times p} = BH = (C_1, C_2)$, 其中 C_1 是 $k \times k$ 方阵，则上式为

$$C^\top C = \begin{pmatrix} C_1^\top C_1 & C_1^\top C_2 \\ C_2^\top C_1 & C_2^\top C_2 \end{pmatrix} = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$$

故 $C_2 = 0, C_1^\top C_1 = I_k$, C_1 是 $k \times k$ 正交矩阵。令 $V = H^\top WH \sim W_p(m)$, 所以

$$BW^{-1}B^\top = BH(H^\top WH)^{-1}H^\top B^\top = CV^{-1}C^\top = (C_1, 0)V^{-1}\begin{pmatrix} C_1^\top \\ 0 \end{pmatrix} = C_1V_{11\bullet 2}^{-1}C_1^\top$$

其中我们划分了 $V=\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}, V^{-1}=\begin{pmatrix} V_{11\bullet 2}^{-1} & * \\ * & * \end{pmatrix}$, 由 Wishart 矩阵划分定理 3（第 5 讲），$V_{11\bullet 2}\sim W_k(m-(p-k))$, 而 C_1 是正交矩阵，所以

$$(BW^{-1}B^\top)^{-1}=C_1V_{11\bullet 2}C_1^\top\sim W_k(m-(p-k)).$$

对 $\Sigma\neq I_p$ 情形，令 $V=\Sigma^{-1/2}W\Sigma^{-1}\sim W_p(m,I_p)$, 则 $(AW^{-1}A^\top)^{-1}=(A\Sigma^{-1/2}V^{-1}\Sigma^{-1/2}A^\top)^{-1}\triangleq(BV^{-1}B^\top)^{-1}$, 其中 $B=A\Sigma^{-1/2}$, 利用前述已证结论知 $(BV^{-1}B^\top)^{-1}\sim W_k(m-p+k,(BB^\top)^{-1})=W_k(m-p+k,(A\Sigma^{-1}A^\top)^{-1})$.

6.（选）假设 $W\sim W_p(m)$, 考虑 (Galton) 相关系数矩阵 $R=(r_{ij}), r_{ij}=\frac{w_{ij}}{\sqrt{w_{ii}w_{jj}}}$, 试证明 R 的概率密度函数 (上三角 $p(p-1)/2$ 个元素的联合分布)

$$f(R)=\frac{(\Gamma(m/2))^p}{\Gamma_p(m/2)}|R|^{(m-p-1)/2}.$$

其中 $\Gamma_p(x)=\pi^{p(p-1)/4}\prod_{i=1}^p\Gamma(x-(i-1)/2)$.

证明: 按行排列 W 上三角元素成 $p(p-1)/2\times p(p-1)/2$ 向量 $\mathbf{w}$, 对角元排列为 $p\times p$ 向量 $\mathbf{u}$:

$$\mathbf{w}=(w_{12},...,w_{13},....,w_{1p},w_{23},w_{24},...,w_{2p},....,w_{p-1,p})^\top,\quad \mathbf{u}=(w_{11},...,w_{pp})^\top,\quad D=\mathrm{diag}(\mathbf{u})$$

记标准 Wishart $W_p(m)$ 的密度

$$p(W)=\frac{1}{2^{mp/2}\Gamma_p(m/2)}|W|^{(m-p-1)/2}\exp\left(-\frac{1}{2}tr(W)\right),$$

则 (R,D) 的联合概率密度为 $f(R,D)=p(D^{1/2}RD^{1/2})\times|J|$, 其中 Jacobian

$$J=J(W\to R,D)=J(\mathbf{w},\mathbf{u}\to\mathbf{r},\mathbf{u})=\frac{\partial(\mathbf{w},\mathbf{u})}{\partial(\mathbf{r},\mathbf{u})}=\begin{pmatrix}\partial\mathbf{w}/\partial\mathbf{r} & \partial\mathbf{w}/\partial\mathbf{u}\\ \partial\mathbf{u}/\partial\mathbf{r} & \partial\mathbf{u}/\partial\mathbf{u}\end{pmatrix}=\begin{pmatrix}\partial\mathbf{w}/\partial\mathbf{r} & 0\\ 0 & I_p\end{pmatrix}$$

而

$$\partial\mathbf{w}/\partial\mathbf{r}=\mathrm{diag}(\sqrt{u_1u_2},\sqrt{u_1u_3},...,\sqrt{u_1u_p},\sqrt{u_2u_3},...,\sqrt{u_{p-1}u_p})^\top,$$

所以

$$|J|=|\partial\mathbf{w}/\partial\mathbf{r}|=(u_1u_2\cdots u_p)^{(p-1)/2}=|D|^{(p-1)/2},$$

所以 (R,D) 的联合概率密度函数

$$\begin{aligned}
f(R,D) &= p(D^{1/2}RD^{1/2})\times|J|=\frac{1}{2^{mp/2}\Gamma_p(m/2)}|R|^{(m-p-1)/2}|D|^{(m-p-1)/2}\exp\left(-\frac{1}{2}tr(D)\right)\times|D|^{(p-1)/2}\\
&= \frac{1}{2^{mp/2}\Gamma_p(m/2)}|R|^{(m-p-1)/2}|D|^{m/2-1}\exp\left(-\frac{1}{2}tr(D)\right)\\
&= \frac{1}{2^{mp/2}\Gamma_p(m/2)}|R|^{(m-p-1)/2}\prod_{i=1}^p u_i^{m/2-1}\exp(-u_i/2)
\end{aligned}$$

因此 R 与 D 独立，且 R 的边际概率密度函数为

$$f(R)=\int f(R,D)dD=\frac{(\Gamma(m/2))^p}{\Gamma_p(m/2)}|R|^{(m-p-1)/2}.$$

7.（选）假设 $\mathbf{z}_1,\dots,\mathbf{z}_m$ iid $\sim N_p(\mathbf{0},\Sigma)$, $\Sigma>0$, $p\geq 1$, $Z=(\mathbf{z}_1,\dots,\mathbf{z}_m)^\top$。我们已知 Z 的（即 $\mathbf{z}_1,\dots,\mathbf{z}_m$ 的联合）概率密度函数为

$$p(Z)=\frac{1}{(2\pi)^{mp/2}|\Sigma|^{m/2}}\exp\left(-\frac{1}{2}tr(\Sigma^{-1}Z^\top Z)\right)\triangleq h(W) \tag{1}$$

其中 $W=Z^\top Z$。第 7 讲定理 3、4 利用分布 (1) 和 W 的矩阵分块和归纳法求得 Wishart 分布的概率密度函数，但归纳法一般不是求解答案未知的问题的最佳方式（除非我们事先能猜出答案）。我们的问题是：**能否通过应用随机变量变换密度公式的常规方式简单地求出 $W\sim \mathrm{Wishart}_p(m,\Sigma)$ 的概率密度?**

讨论：第 10 讲附录定理 A5-A9 实际上正是采用了这种常规求解密度的方法，但其中涉及大量外微分和上三角变换等内容，过程过于复杂以至于其求解过程不甚清晰。我们将第 10 讲附录中的求解步骤整理如下（未必完全正确，仅供参考）：

- 定义 $U=Z(Z^\top Z)^{-1/2}=ZW^{-1/2}$, $\quad U^\top U=I_p$;
- 基于 Z 的密度 (1) 以及 Jacobian $J(Z\to U,W)$ 得到 U,W 的联合概率密度；
- 联合密度 $p(U,W)$ 对 U 积分求解得到 W 的边际概率密度。

具体过程如下。给定 $Z^\top Z=W$ 时，$p(Z)$ 是常数，因此 Z 在 $\{Z_{m\times p}:Z^\top Z=W\}$ 上服从均匀分布，即 $Z|_{Z^\top Z=W}\sim \mathrm{Uniform}\{Z_{m\times p}:Z^\top Z=W\}$。令 $U=Z(Z^\top Z)^{-1/2}$，则

$$U|_{Z^\top Z=W}\sim \mathrm{Uniform}(V_{m,p})$$

其中 $V_{m,p}=\{V_{m\times p}:V^\top V=I_p\}$ 是 R^{mp} 空间中的一个几何形体，而 $\mathrm{Uniform}(V_{m,p})$ 表示集合 $V_{m,p}$ 上的均匀分布。条件分布 $U|_W\sim \mathrm{Uniform}(V_{m,p})$ 与条件 W 无关，所以 $W\perp\!\!\!\perp U$，U 的概率密度

$$p(U)=1/|V_{m,p}|,$$

其中 $|V_{m,p}|$ 代表 $V_{m,p}$ 的体积，所以

$$p(U,W)=p(U)p(W)=h(W)J(Z\to U,W)$$

$\Rightarrow$

$$p(W)=h(W)J/p(U)=|V_{m,p}|\times h(W)J(Z\to U,W)$$

第 10 讲附录定理 A8 利用上三角正交坐标表示（Z 的 Gram-Schmidt 变换）

$$Z=\text{列正交矩阵}\times\text{上三角矩阵},$$

完成了求解 $J(Z\to U,W)$ 和 $|V_{m,p}|$。最终的 Jacobian 和体积并不依赖于这些具体的坐标表示，因此上三角变换并不是计算过程中必须的手段（甚至这些工具的引入可能导致计算过于复杂）, 但遗憾的是，我们尚不知道是否存在更简单和直接的方法求体积 $|V_{m,p}|$ 以及 Jacobian。