1. 假设 2×2 随机矩阵 $W=\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22}\end{pmatrix}\sim W_2(m)$, $m\geq 2$。（注：(a),(b),(c) 没有相承关系）

   (a) 证明 w_{11}, w_{22} iid $\sim \chi^2_m$.

   (b) 令 $\xi=(w_{11}+w_{22})/2+w_{12}, \eta=(w_{11}+w_{22})/2-w_{12}$, 证明 $\xi,\eta\ iid\sim\chi^2_m$.

   (c) 证明 $\frac{2(w_{11}w_{22}-w_{12}^2)}{w_{11}+w_{22}-2w_{12}}\sim\chi^2_{m-1}$.

   本题可应用第 8 讲定理 6：

$$W\sim W_p(m)\Rightarrow \mathbf{t}^\top W\mathbf{t}\sim\chi^2_m,\quad 1/\mathbf{t}^\top W^{-1}\mathbf{t}\sim\chi^2_{m-p+1},\quad \forall\mathbf{t}\in S^{m-1}.$$

   也可以直接按照 Wishart 分布定义证明。

证明 1:

(a) 取 $\mathbf{t}=(1,0)^\top$。

(b) 取 $A=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1\end{pmatrix}$, $AA^\top=I_2$, 则由第 8 讲定理 1，

$$AWA^\top=\frac{1}{2}\begin{pmatrix} w_{11}+w_{22}-2w_{12} & w_{11}-w_{22} \\ w_{11}-w_{22} & w_{11}+w_{22}+2w_{12}\end{pmatrix}=\begin{pmatrix} \xi & (w_{11}-w_{22})/2 \\ (w_{11}-w_{22})/2 & \eta\end{pmatrix}\sim W_2(m)$$

因此 ξ,η iid $\sim\chi^2_m$。

(c) 因为

$$W^{-1}=\begin{pmatrix} w_{11} & w_{12} \\ w_{12} & w_{22}\end{pmatrix}^{-1}=\frac{1}{w_{11}w_{22}-w_{12}^2}\begin{pmatrix} w_{22} & -w_{12} \\ -w_{12} & w_{11}\end{pmatrix}$$

取 $\mathbf{t}=(1,1)^\top$, 则有 $1/\mathbf{t}^\top W^{-1}\mathbf{t}=2(w_{11}w_{22}-w_{12}^2)/(w_{11}+w_{22}-2w_{12})\sim\chi^2_{m-2+1}$.

证明 2 (Wishart **分布定义**):

(b) 设

$$W=\begin{pmatrix} w_{11} & w_{12} \\ w_{12} & w_{22}\end{pmatrix}=\begin{pmatrix} \mathbf{x}^\top\mathbf{x} & \mathbf{x}^\top\mathbf{y} \\ \mathbf{x}^\top\mathbf{y} & \mathbf{y}^\top\mathbf{y}\end{pmatrix},$$

其中 $\mathbf{x},\mathbf{y}$ iid $\sim N_m(0,I_m)$, 则 $w_{11}=\mathbf{x}^\top\mathbf{x}\sim\chi^2_m$,

$$\xi=\frac{1}{2}(\mathbf{x}^\top\mathbf{x}+\mathbf{y}^\top\mathbf{y}+2\mathbf{x}^\top\mathbf{y})=\frac{1}{2}(\mathbf{x}+\mathbf{y})^\top(\mathbf{x}+\mathbf{y})\stackrel{\Delta}{=}\mathbf{z}^\top\mathbf{z},$$

其中 $\mathbf{u}=(\mathbf{x}+\mathbf{y})/\sqrt{2}\sim N_m(0,I_m)$, 所以 $\xi\sim\chi^2_m$. 同理 $\eta=\frac{1}{2}(\mathbf{x}-\mathbf{y})^\top(\mathbf{x}-\mathbf{y})\stackrel{\Delta}{=}\mathbf{v}^\top\mathbf{v}$, 其中 $\mathbf{v}=(\mathbf{x}-\mathbf{y})/\sqrt{2}\sim N_m(0,I_m)$, 且 $\mathbf{u}\perp\!\!\!\perp\mathbf{v}$, 所以 $\eta\sim\chi^2_m$ 且 $\eta\perp\!\!\!\perp\xi$。

(c) 定义 $\mathbf{u},\mathbf{v}$ 同 (b)，$\mathbf{u},\mathbf{v}$ iid $\sim N_m(0,I_m)$. 可以验证

$$\frac{2(w_{11}w_{22}-w_{12}^2)}{w_{11}+w_{22}-2w_{12}} = \mathbf{u}^\top\mathbf{u}-(\mathbf{v}^\top\mathbf{u})^2/\mathbf{v}^\top\mathbf{v}$$

给定 $\mathbf{v}$, 令正交矩阵 $H=\begin{pmatrix}\mathbf{v}^\top/\|\mathbf{v}\| \\ * \end{pmatrix}$ 即 H 的第一行为 $\mathbf{v}^\top/\|\mathbf{v}\|$, 以及 $\mathbf{z}=H\mathbf{u}$, 则 $H\mathbf{v}/\|\mathbf{v}\|=(1,0,...,0)^\top$, 所以给定 $\mathbf{v}$ 条件下

$$\mathbf{u}^\top\mathbf{u}-(\mathbf{v}^\top\mathbf{u})^2/\mathbf{v}^\top\mathbf{v}=\mathbf{z}^\top\mathbf{z}-\left\{[H\mathbf{v}/\|\mathbf{v}\|]^\top\mathbf{z}\right\}^2=\mathbf{z}^\top\mathbf{z}-z_1^2=z_2^2+...+z_m^2\sim\chi^2_{m-1}$$

该分布与 $\mathbf{v}$ 无关，所以上式无条件成立。

2. 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: 我们已知，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

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

3. 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 $W\overset{d}{=}\sum_{i=1}^m\mathbf{z}_i\mathbf{z}_i^\top$, 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 $E(1/\chi^2_k)=1/(k-2),k>2$ (即, 若 r.v. $x\sim\chi^2_k$, 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 $EW^{-1}=\Sigma^{-1}/(m-p-1)$。

4. 假设 $W\sim W_p(m,\Sigma)$, $A_{k\times p}$ 是行满秩常数矩阵，证明 $(AW^{-1}A^\top)^{-1}\sim W_k(m-p+k,(A\Sigma^{-1}A^\top)^{-1})$.

思路：考虑 $\Sigma = I_p$ 情形。划分 $W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}$, 其中 W_{11} 是 $k \times k$ 矩阵，则

$$W^{-1} = \begin{pmatrix} W_{11\bullet 2}^{-1} & * \\ * & * \end{pmatrix},$$

- 若 A 具有（最）简单的形式 $A = E_1 = (I_k, 0_{k\times(p-k)})$, 则由第 8 讲定理 4，

$$AW^{-1}A^\top = W_{11\bullet 2}^{-1} \ \Rightarrow\ (AW^{-1}A^\top)^{-1} = W_{11\bullet 2} \sim W_k(m-(p-k), I_k)$$

- 若 A 具有简单形式 $A = (C, 0)$, C 是 $k \times k$ 可逆矩阵，同样地有 $(AW^{-1}A^\top)^{-1} = C^{-1}W_{11\bullet 2}(C^{-1})^\top \sim W_k(m-p+k, C^{-1}(C^{-1})^\top = (AA^\top)^{-1})$;
- 对于一般的 A，注意对任何正交矩阵 H,

$$AW^{-1}A^\top = AH^\top (HWH^\top)^{-1} HA^\top$$

其中 $HWH^\top \overset{d}{=} W \sim W_p(m)$, 我们可选取适当的 H 使得 $AH^\top$ 具有形式 $(*, 0)$。

证明 1: 先证明 $\Sigma = I_p$ 情形，即需要证明：$W \sim W_p(m) \Rightarrow (AW^{-1}A^\top)^{-1} \sim W_k(m-p+k, (AA^\top)^{-1})$. 令 $B = (AA^\top)^{-1/2}A$ 为 A 的行“单位化”, 即 $BB^\top = I_k$。构造 $p \times p$ 正交矩阵

$$H = \begin{pmatrix} (AA^\top)^{-1/2}A \\ H_2 \end{pmatrix},$$

其中 $(AA^\top)^{-1/2}A$ 的行与 H_2 的行正交：$(AA^\top)^{-1/2}AH_2^\top = 0 \Rightarrow AH_2 = 0$, 故 $HA^\top = \begin{pmatrix} (AA^\top)^{1/2} \\ 0 \end{pmatrix}$. 记 $V = HWH^\top = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \sim W_p(m)$, 由第 8 讲定理 4，$V_{11\bullet 2} \sim W_k(m-p+k)$, 所以

$$AW^{-1}A^\top = AH^\top(HWH^\top)^{-1}HA^\top = ((AA^\top)^{1/2}, 0)\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}^{-1}\begin{pmatrix} (AA^\top)^{1/2} \\ 0 \end{pmatrix} = (AA^\top)^{1/2}V_{11\bullet 2}^{-1}(AA^\top)^{1/2}$$

进而

$$(AW^{-1}A^\top) = (AA^\top)^{-1/2}V_{11\bullet 2}(AA^\top)^{-1/2} \sim W_k(m-p+k, (AA^\top)^{-1}).$$

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 $(BV^{-1}B^\top)^{-1} \sim W_k(m-p+k, (BB^\top)^{-1}) = W_k(m-p+k, (A\Sigma^{-1}A^\top)^{-1})$.

证明 2: 先证明 $\Sigma = I_p$ 情形。对于 $W \sim W_p(m)$，我们需要证

$$(AW^{-1}A^\top)^{-1} \sim W_k(m-p+k, (AA^\top)^{-1})$$

$$\Leftrightarrow \left((AA^\top)^{-1/2}AW^{-1}A^\top(AA^\top)^{-1/2}\right)^{-1} \sim W_k(m-p+k, I_k)$$

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

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

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

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

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

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 $(BV^{-1}B^\top)^{-1} \sim W_k(m-p+k, (BB^\top)^{-1}) = W_k(m-p+k, (A\Sigma^{-1}A^\top)^{-1})$.

5.（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 $\mathbf{u}$:

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

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

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

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

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

评注：

- $p=2$ 时，$r=r_{12}$，
$$
f(r) = \frac{\Gamma(m/2)}{\sqrt{\pi}\Gamma((m-1)/2)}(1-r^2)^{(m-3)/2}, |r|\le 1.
$$
- $p=3$，(r_{12},r_{13},r_{23}) 的联合概率密度为
$$
f(R) = f(r_{12},r_{13},r_{23}) = \frac{\Gamma(m/2)^2}{\pi^{3/2}\Gamma((m-1)/2)\Gamma((m-2)/2)}(1-r_{12}^2-r_{13}^2-r_{23}^2+2r_{12}r_{13}r_{23})^{(m-4)/2},
$$
$$
1-r_{12}^2-r_{13}^2-r_{23}^2+2r_{12}r_{13}r_{23}\ge 0.
$$