

第七讲 Wishart分布II

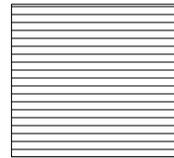
2025.3.24

$$Z = \begin{pmatrix} * \\ \vdots \\ * \end{pmatrix}$$

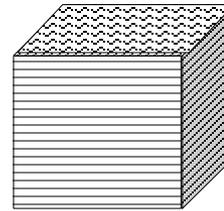
各“行”独立
同分布



vector



matrix



tensor

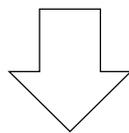
Recap

上节课：从多元正态样本的密度出发，应用矩阵外积微分形式等工具，给出了求解Wishart分布概率密度的大致过程：

$\mathbf{z}_1, \dots, \mathbf{z}_m$ iid $\sim N_p(\mathbf{0}, \Sigma)$, $Z = (\mathbf{z}_1, \dots, \mathbf{z}_m)^\top$ 的概率密度

$$p_Z(Z) = p(\mathbf{z}_1, \dots, \mathbf{z}_m) = C \exp\left(-\frac{1}{2} \text{tr}(\Sigma^{-1} Z^\top Z)\right) \triangleq h(W)$$

球对称

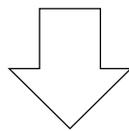


$$Z = UW^{1/2}, W = Z^\top Z, U = ZW^{-1/2}, U^\top U = I_p$$

$$J(Z \rightarrow (W, U)) = J(W) = 2^{-p} |W|^{(m-p-1)/2}$$

$$V_{m,p} = \{U: U^\top U = I_p\}$$

$$p_Z(Z) dZ = h(W) J(W) (dW) (dU) = p(W) (dW) \times \frac{1}{|V_{m,p}|} (dU), U \in V_{m,p}$$



$$p(W) = |V_{m,p}| h(W) J(W), |V_{m,p}| = \frac{2^p \pi^{mp/2}}{\Gamma_p(\frac{m}{2})}$$

$U \sim U(V_{m,p})$, $U \perp W$, W 的概率密度 (Wishart)

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

其中球对称性有基本的重要性。

球对称（正交不变）随机矩阵

球对称 函数

记号： $\mathcal{O}(m) = \{H \in R^{m \times m}: H^T H = I_m\}$

假设 $m \geq p$ ， $Z \in R^{m \times p}$ ($m \times p$ 矩阵)，假设 $Z^T Z > 0$ 。矩阵函数 $f(Z)$ 称为是(左)球对称的或正交不变的，如果对任何 $\forall H \in \mathcal{O}(m)$ ， $f(HZ) = f(Z)$ 。当 $p = 1$ 时，这是第2-3讲的随机向量情形。

命题1: $f(Z)$ 球对称, $Z \in R^{m \times p} \Leftrightarrow f(Z) = h(Z^T Z)$, 某个 $R^{p \times p}$ 上的函数 h .

证: (\Rightarrow) 对任何给定的 $Z \in R^{m \times p}$, 令 $H_1 = (Z^T Z)^{-1/2} Z^T$, 它是行正交的, 即 $H_1 H_1^T = I_p$ 。补全 H_1 成正交矩阵

$$H = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} = \begin{pmatrix} (Z^T Z)^{-1/2} Z^T \\ H_2 \end{pmatrix},$$

由 $0 = H_2 H_1^T = H_2 Z (Z^T Z)^{-1/2} \Rightarrow H_2 Z = 0$ 。则 $HZ = \begin{pmatrix} (Z^T Z)^{1/2} \\ 0 \end{pmatrix}$, 则

$$f(Z) = f(HZ) = f\left(\begin{pmatrix} (Z^T Z)^{1/2} \\ 0 \end{pmatrix}\right) \triangleq h(Z^T Z)。$$

$N_p(\mathbf{0}, \Sigma)$

样本的球
对称性

$\mathbf{z}_1, \dots, \mathbf{z}_m \text{ iid } \sim N_p(\mathbf{0}, \Sigma)$, \mathbf{z} 's 转置后罗列成“列向量”

$$Z = \begin{pmatrix} \mathbf{z}_1^\top \\ \vdots \\ \mathbf{z}_m^\top \end{pmatrix} \quad \begin{array}{|c|} \hline \text{---} \\ \hline \text{---} \\ \hline \text{---} \\ \hline \text{---} \\ \hline \end{array}$$

Z 的概率密度即 $\mathbf{z}_1, \dots, \mathbf{z}_m$ 的联合概率密度

$$p_Z(Z) = p(\mathbf{z}_1, \dots, \mathbf{z}_m) = C \exp\left(-\frac{1}{2} \text{tr}(\Sigma^{-1} Z^\top Z)\right)$$

仅依赖于 $Z^\top Z$, 是球对称的, 即对 $\forall H \in \mathcal{O}(m)$,

$$HZ \stackrel{d}{=} Z, p_Z(HZ) = p_Z(Z),$$

即 Z 的分布不依赖于正交坐标系的特定选取。

虽然 Z 行内分量不独立, 但按列来看

$$Z = (\mathbf{z}_{(1)}, \dots, \mathbf{z}_{(p)})$$

每个列向量在 R^m 中都是球对称的:

$$\mathbf{z}_{(j)} \sim N_m(\mathbf{0}, \sigma_{jj} I_m),$$

$$\Sigma = (\sigma_{ij})$$

列向量之间的相关结构也具有球对称性:

$$\text{cov}(\mathbf{z}_{(j)}, \mathbf{z}_{(k)}) = \sigma_{jk} I_m$$

模仿第2-3讲，我们可以考虑一般的球对称分布 $m \times p$ 矩阵 Z ：

$$HZ \stackrel{d}{=} Z, \quad H \in \mathcal{O}(m) \text{ 正交}$$

Z 的概率密度仅依赖于 $Z^\top Z$ (未必正态)：

$$p_Z(Z) = h(W), \quad W = Z^\top Z$$

与Wishart分布的求解完全相同 (P2)：

$$Z = UW^{1/2}, W = Z^\top Z, U = ZW^{-1/2}, U^\top U = I_p \\ J(Z \rightarrow (W, U)) = J(W) = 2^{-p} |W|^{(m-p-1)/2}$$

$$p_Z(Z) dZ = h(W) J(W) (dW) (dU) = p(W) (dW) \times \frac{1}{|V_{m,p}|} (dU)$$

则我们得到 $U \sim U(V_{m,p})$, $U \perp W$, $W = Z^\top Z$ 的概率密度

$$p(W) = |V_{m,p}| J(W) h(W) = \frac{\pi^{mp/2}}{\Gamma_p(\frac{m}{2})} h(W) |W|^{(m-p-1)/2}$$

反之， $W = Z^\top Z$ 的密度 $p(W)$ 唯一决定了球对称分布：由 $h(W) = p(W) / |V_{m,p}| J(W)$ ，可得 $m \times p$ 随机矩阵 Z 的分布

$$p_Z(Z) = \frac{p(Z^\top Z)}{|V_{m,p}| J(Z^\top Z)}$$

如果 Z 球对称，且各行独立，则各行服从 $N_p(\mathbf{0}, \Sigma)$ ，这是多元版本的 Maxwell-Herschel (MH) 定理

多元MH 定理

一元Herschel-Maxwell Theorem (HM) :

假设随机变量 z_1, \dots, z_m 独立, 若 $m \times 1$ 随机向量 $\mathbf{z} = (z_1, \dots, z_m)^\top$ 服从球对称分布, 则 z_1, \dots, z_m iid $\sim N(0, \sigma^2)$ 。

Mukherjee(2017)证明大概:

取 $\xi \sim N_m(\mathbf{0}, I_m)$ 是球对称的。 \mathbf{z} 服从球对称分布 \Rightarrow

$$\frac{\mathbf{z}}{\|\mathbf{z}\|} \stackrel{d}{=} \frac{\xi}{\|\xi\|} \sim U(S^{m-1}) \Rightarrow \frac{z_1}{\|\mathbf{z}\|/\sqrt{m}} \stackrel{d}{=} \frac{\xi_1}{\|\xi\|/\sqrt{m}},$$

由大数定律, $m \rightarrow \infty$ 时, $\frac{\|\xi\|}{\sqrt{m}} = \sqrt{(\xi_1^2 + \dots + \xi_m^2)/m} \rightarrow E(\xi_1^2) = 1,$

$$\|\mathbf{z}\|/\sqrt{m} \rightarrow \sqrt{E(z_1^2)} \triangleq \sigma \Rightarrow z_1/\sigma \stackrel{d}{=} \xi_1 \sim N(0,1).$$

S. Mukherjee (2017) A Proof of the Herschel-Maxwell Theorem Using the Strong Law of Large Numbers. *Pi Mu Epsilon Journal*. Vol 14(6).

该文花了较大篇幅处理上述证明的漏洞。

命题2(多元MH). 假设 $\mathbf{z}_1, \dots, \mathbf{z}_m \in R^p$ 独立, $Z = (\mathbf{z}_1, \dots, \mathbf{z}_m)^\top$, 若 $m \times p$ 矩阵 Z 服从球对称分布, 则 $\mathbf{z}_1, \dots, \mathbf{z}_m$ iid $\sim N_p(\mathbf{0}, \Sigma)$ 。

证明: 对任何常向量 $\mathbf{t} \in R^p$, $Z\mathbf{t} \in R^m$ 球对称分布且分量独立, 由MH定理, $\mathbf{z}_i^\top \mathbf{t}$ 服从一元正态分布, 则 \mathbf{z}_m 服从多元正态 $N_p(\mathbf{0}, \Sigma)$ 。

课程中有些问题可以进一步思考PME:

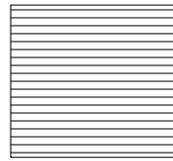
- 多元MH定理（我们自己的命名）在物理学中的应用？
- 矩阵球对称分布在文献中是否有过足够讨论？
- 简单（直观）地证明矩阵的极分解变换 $Z = UW^{1/2}$, $W = Z^T Z$, $U = ZW^{-1/2}$ 的雅可比 $J(Z \rightarrow U, W)$ 仅与 $|W|$ 有关。
- 推广的阿基米德定理（我们自己的命名）的几何意义或应用？
- $V_{m,p} = \{U \in R^{m \times p} : U^T U = I_p\}$ 上的均匀分布 $U(V_{m,p})$ 的边际分布？比如第一列、第一行的分布？藉此考察Stiefel流形的几何性质。
- $Z = \begin{pmatrix} * \\ \vdots \\ * \end{pmatrix}$ 的第 i 个 * 分量/位置代表第 i 个样本。* 位置如果放置矩阵，即矩阵样本数据，所有的样本构成一个tensor张量 Z 。张量的球对称性？

$$Z = \begin{pmatrix} * \\ \vdots \\ * \end{pmatrix}$$

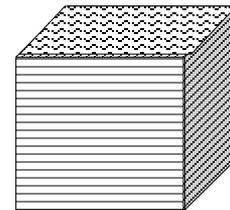
各“行”独立同分布



vector



matrix



tensor

几个引理

首先介绍几个有关判定独立、条件独立、等分布的引理。

引理1(参见Bilodeau & Brenner,2009, P27, Proposition 2.13)

(1) 假设 x, y 是随机元(随机变量、随机向量、随机矩阵甚至其它复杂的随机数学对象)，我们有

$$x, y \text{ 独立} \Leftrightarrow \text{对任何可测(可积)函数 } f, g, \quad Ef(x)g(y) = Ef(x)Eg(y).$$

(2) $x \stackrel{d}{=} y \Leftrightarrow$ 对任何可测函数 $f, Ef(x) = Ef(y)$

对于通常的随机变量， f, g 取任何示性函数就足够了。

引理2. u, v 是两个随机变量，若 $u|v$ 的条件分布仅与 v 的某个函数 $\varphi(v)$ 有关，则 $u|\varphi(v) \stackrel{d}{=} u|v$.

将条件期望看作投影，条件分布也是条件期望，条件分布 $P(u|v)$ 可看作是 u 在 v 空间上的投影，若 $P(u|v)$ 恰好落在 v 的子空间 $\varphi(v)$ 上，则 $P(u|v)$ 可看成是 u 直接在 $\varphi(v)$ 上的投影，即 $P(u|v) = P(u|\varphi(v))$.

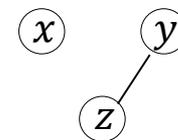
引理2的证明: 对任何给定的 u 的函数 f , 因为给定 v 时, u 的条件分布仅与 $\varphi(v)$ 有关, 所以 $E(f(u)|v)$ 仅与 $\varphi(v)$ 有关, 记作 $E(f(u)|v) = h(\varphi(v))$, 由条件期望的平滑性质/tower性质,

$$E[E(f(u)|v)|\varphi(v)] = E(f(u)|\varphi(v)) \quad (*)$$

左端 $E[E(f(u)|v)|\varphi(v)] = E[h(\varphi(v))|\varphi(v)] = h(\varphi(v)) = E(f(u)|v)$,

所以 $E(f(u)|\varphi(v)) = E(f(u)|v)$, 这说明 $u|\varphi(v) \stackrel{d}{=} u|v$.

引理3. 若 $x \perp\!\!\!\perp y | z$, 且 $x \perp\!\!\!\perp z$, 则 $x \perp\!\!\!\perp y$



形式化证明 (P 表示概率或密度):

$$\begin{aligned} P(x, y) &= \int P(x, y|z) P(z) dz = \int P(x|z) P(y|z) P(z) dz \\ &= \int P(x) P(y|z) P(z) dz = P(x) \int P(y|z) P(z) dz = P(x) P(y) \end{aligned}$$

引理3的证明: 任取 f, g 。由 $x \perp\!\!\!\perp z \Rightarrow E(f(x)|z) = Ef(x)$

由 $x \perp\!\!\!\perp y | z \Rightarrow E(f(x)g(y)|z) = E(f(x)|z) \times E(g(y)|z)$ 。

所以 $Ef(x)g(y) = E(Ef(x)g(y)|z) = E\{[Ef(x)|z][Eg(y)|z]\}$

$= E\{Ef(x) [Eg(y)|z]\} = Ef(x) E\{Eg(y)|z\} = Ef(x)Eg(y) \Rightarrow x \perp\!\!\!\perp y$ 。

引理4. 假设 $\mathbf{x} \sim N_m(\mathbf{0}, I_m)$, P 是秩为 r 的 $m \times m$ 投影矩阵(对称幂等矩阵), 则 $\mathbf{x}^\top P \mathbf{x} \sim \chi_r^2$, $\mathbf{x}^\top (I_m - P) \mathbf{x} \sim \chi_{m-r}^2$, 两者独立

该结果在线性回归分析中已有介绍和广泛应用, 它是Cochran定理的特殊情况 (见后面定理2)

证明: 因为 $P_{m \times m}$ 是秩为 r 的对称幂等阵, 存在正交矩阵 $H_{m \times m}$ 使得

$$P = H^\top \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} H, \text{ 故 } \mathbf{y} = H\mathbf{x} \sim N_m(\mathbf{0}, I_m), \mathbf{y} \text{的前 } r \text{ 个分量组成的 } \mathbf{y}_1 \sim N_r(\mathbf{0}, I_r)$$

$$\mathbf{x}^\top P \mathbf{x} = \mathbf{x}^\top H^\top \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} H \mathbf{x} = \mathbf{y}^\top \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \mathbf{y} = (\mathbf{y}_1^\top, \mathbf{y}_2^\top) \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \mathbf{y}_1^\top \mathbf{y}_1 \sim \chi_r^2.$$

$$\mathbf{x}^\top (I_m - P) \mathbf{x} = \mathbf{y}_2^\top \mathbf{y}_2 \sim \chi_{m-r}^2$$

例子：求解 $W_2(m, I_2)$ 的概率密度

从多元正态 $N_p(\mathbf{0}, \Sigma)$ 样本出发，基于矩阵微分形式，我们求得了Wishart分布的概率密度，但推导过程不完整(也不要求掌握)。下面使用基本概率论方法求解最简单的Wishart分布 $W_2(m, I_2)$ 的概率密度。

例1. 证明 $W = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \sim W_2(m, I_2)$ 的概率密度为

$$p(W) = p(w_{12}, w_{11}, w_{22}) = c_m e^{-(w_{11}+w_{22})/2} (w_{11}w_{22} - w_{12}^2)^{(m-3)/2},$$

即 $p(W) = c_m |W|^{(m-3)/2} e^{-\text{tr}(W)/2}$ ，其中 $c_m = \left(2^m \sqrt{\pi} \Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{m-1}{2}\right)\right)^{-1}$

假设 $W = Z^\top Z$ ，其中 $Z = (\mathbf{z}_1, \dots, \mathbf{z}_m)^\top$ ， $\mathbf{z}_1, \dots, \mathbf{z}_m \text{ iid} \sim N_2(0, I_2)$ ，记 Z 的两列 $Z = (\mathbf{x}, \mathbf{y})$ ， $\mathbf{x}, \mathbf{y} \text{ iid} \sim N_m(0, I_m)$ ，

$$W = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} = Z^\top Z = \begin{pmatrix} \mathbf{x}^\top \mathbf{x} & \mathbf{x}^\top \mathbf{y} \\ \mathbf{x}^\top \mathbf{y} & \mathbf{y}^\top \mathbf{y} \end{pmatrix},$$

我们已知 $w_{11}, w_{22} \text{ iid} \sim \chi_m^2$ ，但 w_{12} 与它们不独立。为了求解它们的联合分布，我们使用（1）变量变换（2）Schmidt正交化。

证明1(变量变换)

令 $r = w_{12}/\sqrt{w_{11}w_{22}} = \mathbf{x}^\top \mathbf{y}/\|\mathbf{x}\|\|\mathbf{y}\| = \mathbf{u}^\top \mathbf{v} \stackrel{d}{=} u_1$, $\mathbf{u}, \mathbf{v} \text{ iid } \sim U(S^{m-1})$
 $r \perp \{w_{11}, w_{22}\}$, 所以 r, w_{11}, w_{22} 的联合概率密度容易得到:

$$p(r, w_{11}, w_{22}) = p_0(r)p_{11}(w_{11})p_{22}(w_{22})$$

其中

$$p_{11}(w_{11}) = \frac{1}{2^{m/2}\Gamma(m/2)} w_{11}^{m/2-1} e^{-w_{11}/2}$$

$$p_{22}(w_{22}) = \frac{1}{2^{m/2}\Gamma(m/2)} w_{22}^{m/2-1} e^{-w_{22}/2}$$

$$p_0(r) = \frac{\Gamma(m/2)}{\pi^{1/2}\Gamma((m-1)/2)} (1-r^2)^{(m-3)/2}$$

χ_m^2

χ_m^2

$U(S^{n-1})$ 的一元边际
 第一讲推论1

然后转换到 w_{12}, w_{11}, w_{22} 的分布:

$$p(w_{12}, w_{11}, w_{22})$$

$$= p_0(r)p_{11}(w_{11})p_{22}(w_{22}) \frac{1}{\sqrt{w_{11}w_{22}}}$$

$$J((r, w_{11}, w_{22}) \rightarrow (w_{12}, w_{11}, w_{22}))$$

$$= 1/\sqrt{w_{11}w_{22}}$$

$$= c_m e^{-(w_{11}+w_{22})/2} w_{11}^{m/2-1} w_{22}^{m/2-1} (1 - (w_{12}/\sqrt{w_{11}w_{22}})^2)^{(m-3)/2} \frac{1}{\sqrt{w_{11}w_{22}}}$$

$$= c_m e^{-(w_{11}+w_{22})/2} (w_{11}w_{22} - w_{12}^2)^{(m-3)/2}$$

注解：证明1的变量变换不太显然，我们从联合分布更常规的条件概率求解方法角度解释一下证明1：

- $p(w_{12}, w_{11}, w_{22}) = p(w_{12}|w_{11}, w_{22})p(w_{11}, w_{22})$,

只需求 $p(w_{12}|w_{11}, w_{22})$. 注意

$$w_{12} = \mathbf{x}^\top \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \left(\frac{\mathbf{x}}{\|\mathbf{x}\|} \right)^\top \frac{\mathbf{y}}{\|\mathbf{y}\|} \triangleq \sqrt{w_{11} w_{22}} \mathbf{u}^\top \mathbf{v} \stackrel{d}{=} \sqrt{w_{11} w_{22}} u_1$$

其中 $\mathbf{u}, \mathbf{v} \text{ iid} \sim U(S^{m-1}) \perp w_{11}, w_{22}$, 故给定 w_{11}, w_{22} 时,

$$p(w_{12}|w_{11}, w_{22}) = p_0(w_{12}/\sqrt{w_{11} w_{22}}) \times \frac{1}{\sqrt{w_{11} w_{22}}}$$

p_0 如上页。这实际上等同于证明1。

- 我们也可以尝试求 $p(w_{12}|w_{22})$ 或 $p(w_{12}|w_{11})$ 。

首先考虑 $w_{12} | \mathbf{y} = \mathbf{x}^\top \mathbf{y} | \mathbf{y} \sim N(0, \mathbf{y}^\top \mathbf{y}) = N(0, w_{22})$, 由引理2,

$w_{12} | w_{22} \sim N(0, w_{22})$, 同样, $w_{12} | w_{11} \sim N(0, w_{11})$ 。但这些分布不足

以完全确定 $p(w_{12}, w_{11}, w_{22})$ 。 w_{12} 与 w_{11} 、 w_{22} 都有关, 我们下面变换 (投影) $w_{11} \rightarrow w_{11 \cdot 2}$, 使得变换后的 $w_{11 \cdot 2}$ 与 w_{12} , w_{22} 都独立, 从而

$$p(w_{12}, w_{11}, w_{22}) = p(w_{11 \cdot 2}, w_{12}, w_{22}) = p(w_{11 \cdot 2})p(w_{12}, w_{22})$$

证明2
(正交化)

令 $w_{11\cdot 2} \triangleq w_{11} - w_{12}w_{12}^{-1}w_{21}$, 则

$$w_{11\cdot 2} = \mathbf{x}^\top \mathbf{x} - \mathbf{x}^\top \mathbf{y}(\mathbf{y}^\top \mathbf{y})^{-1} \mathbf{y}^\top \mathbf{x} = \mathbf{x}^\top (I_m - P_{\mathbf{y}}) \mathbf{x}$$

所以 $w_{11\cdot 2} | \mathbf{y} = \mathbf{x}^\top (I_m - P_{\mathbf{y}}) \mathbf{x} | \mathbf{y} \sim \chi_{m-1}^2$ 该分布与 \mathbf{y} 无关, 故

$$w_{11\cdot 2} \perp \mathbf{y} \Rightarrow w_{11\cdot 2} \perp w_{22}, \text{ 且 } w_{11\cdot 2} \sim \chi_{m-1}^2$$

因为 $(I_m - P_{\mathbf{y}}) \mathbf{y} = \mathbf{0}$, $\mathbf{x} \sim N_m(\mathbf{0}, I_m)$, 以及 $\mathbf{x} \perp \mathbf{y}$, 所以在 \mathbf{y} 给定时, $w_{12} = \mathbf{y}^\top \mathbf{x}$ 与 $(I_m - P_{\mathbf{y}}) \mathbf{x}$ 条件独立, 进而与 $w_{11\cdot 2} = \mathbf{x}^\top (I_m - P_{\mathbf{y}}) \mathbf{x}$ 条件独立: $w_{11\cdot 2} \perp w_{12} | \mathbf{y}$, 但因 $w_{11\cdot 2} \perp \mathbf{y}$, 由引理2, $w_{11\cdot 2} \perp w_{12}$ 。至此, 我们证明了 $w_{11\cdot 2} \perp \{w_{22}, w_{12}\}$ (参见下页直观解释)

另外, $w_{12} | \mathbf{y} = \mathbf{y}^\top \mathbf{x} | \mathbf{y} \sim N_m(\mathbf{0}, \mathbf{y}^\top \mathbf{y}) = N_m(\mathbf{0}, w_{22})$, 该条件分布仅与 w_{22} 有关, 故 $w_{12} | w_{22} \sim N_m(\mathbf{0}, w_{22})$. 所以

$$\begin{aligned} p(w_{12}, w_{11}, w_{22}) &= p(w_{12}, w_{11\cdot 2}, w_{22}) = p(w_{11\cdot 2})p(w_{12}, w_{22}) \\ &= p(w_{11\cdot 2})p(w_{12} | w_{22})p(w_{22}) = \dots \end{aligned}$$

直观解释 $w_{11 \cdot 2} \perp \{w_{12}, w_{22}\}$:

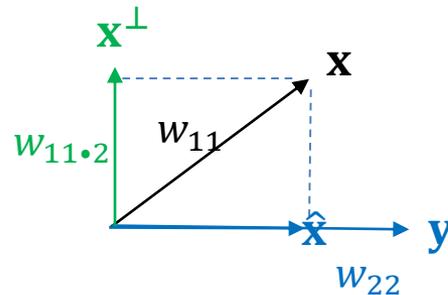
记 $\hat{\mathbf{x}} = P_{\mathbf{y}}\mathbf{x}$ 为 \mathbf{x} 在 \mathbf{y} 方向的投影, $\mathbf{x}^\perp = \mathbf{x} - \hat{\mathbf{x}} \perp \mathbf{y}$, 则

□ $w_{11 \cdot 2} = \|\mathbf{x}^\perp\|^2$ 仅与图中竖轴方向 (\mathbf{x}^\perp 方向) 有关;

□ 而 $w_{12} = \mathbf{x}^\top \mathbf{y} = \hat{\mathbf{x}}^\top \mathbf{y}$, $w_{22} = \mathbf{y}^\top \mathbf{y}$ 仅与图中横轴方向 (\mathbf{y} 方向) 有关,

因为 $\mathbf{x}^\perp \perp \{\hat{\mathbf{x}}, \mathbf{y}\}$, 且都是正态分布 $\Rightarrow \mathbf{x}^\perp \perp \{\hat{\mathbf{x}}, \mathbf{y}\}$

$$\Rightarrow w_{11 \cdot 2} \perp \{w_{12}, w_{22}\}$$



$$\mathbf{x} = \mathbf{x}^\perp + \hat{\mathbf{x}}, \quad \hat{\mathbf{x}} = P_{\mathbf{y}}\mathbf{x}$$

一般阶数的 $W_p(m, I_p)$ 的概率密度求解与 $W_2(m, I_2)$ 类似, 但需要应用 Wishart 分布的一些性质, 主要是 Cochran 定理。