1. 假设 $\mathbf{x}_{11},...,\mathbf{x}_{1n_1}$ iid $\sim N_p(\boldsymbol{\mu}_1,\Sigma)$, $\mathbf{x}_{21},...,\mathbf{x}_{2n_2}$ iid $\sim N_p(\boldsymbol{\mu}_2,\Sigma)$, 两组样本独立, 两组的样本均值分别是 $\bar{\mathbf{x}}_1,\bar{\mathbf{x}}_2$, 样本方差矩阵分别是 S_1,S_2。定义 $S_{pooled}=[(n_1-1)S_1+(n_2-1)S_2]/(n-2)$, $n=n_1+n_2$。考虑两样本检验问题 $H_0:\boldsymbol{\mu}_1-\boldsymbol{\mu}_2=\mathbf{0}_{p\times 1}$，两样本 Hotelling T^2 检验统计量定义为

$$T^2=\frac{n_1n_2}{n_1+n_2}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)^\top S_{pooled}^{-1}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2).$$

另一方面，两样本问题作为多组 MANOVA 的最简单的情况，其检验也可使用 MANOVA 的 Wilks 检验 $\Lambda^*=|W|/|W+B|$ （W,B 的定义，第 9 讲 P3），试验证

$$-n\log\Lambda^*=n\log(1+T^2/(n-2)).$$

2. 假设 $\mathbf{x}_1,...,\mathbf{x}_n$ iid $\sim N_p(\boldsymbol{\mu},\Sigma)$, 我们考虑如下球对称假设

$$H_0:\Sigma=\gamma I_p,\gamma>0(\text{未知}),$$

记号 S 为样本协方差矩阵，似然比统计量 $\Lambda=\frac{\max L(\boldsymbol{\mu},\Sigma)}{\max L(\boldsymbol{\mu},\gamma I_p)}$。

(a) 证明 Wilks 统计量

$$\Lambda^*=\Lambda^{2/n}=\frac{det(S)}{(tr(S)/p)^p}$$

(b) 已知 $-2\log(\Lambda)=-n\log(\Lambda^*)\xrightarrow{d}\chi_f^2$, 求出自由度 f。

3. 假设 $\mathbf{x}_1,...,\mathbf{x}_n$ iid $\sim N_p(\boldsymbol{\mu},\Sigma)$, $p>1$, 样本均值和方差分别是 $\bar{\mathbf{x}},S$。考虑原假设 $H_0:\boldsymbol{\mu}=\boldsymbol{\mu}_0,\quad \boldsymbol{\mu}_0$已知。如果将数据和参数向一维投影方向投影，利用投影坐标构建一元检验统计量，效果如何？

对 $\forall\mathbf{a}\in S^{p-1}$, 考虑

$$H_0(\mathbf{a}):\ \mathbf{a}^\top\boldsymbol{\mu}=\mathbf{a}^\top\boldsymbol{\mu}_0,$$

因为 $y_1=\mathbf{a}^\top\mathbf{x}_1,...,y_n=\mathbf{a}^\top\mathbf{x}_n$ iid $\sim N_1(\mathbf{a}^\top\boldsymbol{\mu},\mathbf{a}^\top\Sigma\mathbf{a})$, $\bar{y}=\mathbf{a}^\top\bar{\mathbf{x}},s_y^2=\mathbf{a}^\top S\mathbf{a}$, 因此 $H_0(\mathbf{a})$ 的 t 检验统计量

$$t(\mathbf{a})=\frac{\sqrt{n}(\bar{y}-\mathbf{a}^\top\boldsymbol{\mu}_0)}{s_y}=\frac{\sqrt{n}(\mathbf{a}^\top\bar{\mathbf{x}}-\mathbf{a}^\top\boldsymbol{\mu}_0)}{\sqrt{\mathbf{a}^\top S\mathbf{a}}}\overset{H_0(\mathbf{a})}{\sim}t_{n-1}.$$

对所有 $\mathbf{a}$, 我们取最显著的一个结果，即取 $t_{\max}^2=\max_{\mathbf{a}\in S^{p-1}}|t(\mathbf{a})|^2$ 作为 H_0 的检验统计量。证明

$$t_{\max}^2=n(\bar{\mathbf{x}}-\boldsymbol{\mu}_\mathbf{0})^\top S^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0).$$

注：（1）该结果说明, 试图在所有方向上使用一元边际检验方法进行检验是徒劳的，最终还是归结于一个多元检验，即 Hotelling T^2 检验。（2）但是，如果只取少数有意义的方向 $\mathbf{a}$ 进行检验或许是有益的。比如，如果我们相信数据几乎分布在 $\mathbf{v}$ 方向上，那么我们只需做一个一元检验：

$$t(\mathbf{v})=\frac{\sqrt{n}(\mathbf{v}^\top\bar{\mathbf{x}}-\mathbf{v}^\top\boldsymbol{\mu}_0)}{\sqrt{\mathbf{v}^\top S\mathbf{v}}}\overset{H_0(\mathbf{v})}{\sim}t_{n-1}.$$

其功效高于 Hotelling T^2 检验。