Some solutions to 542 HW#2.

MDP Exercise 2.2. Because $\left.\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}\right| X_3$ is multivariate normal, it suffices to show that $\mathrm{Cov}(X_1, X_2)|X_3 = 0 \iff \Sigma^{12} = 0$. But

$$\mathrm{Cov}(X_1, X_2)|X_3 = [\Sigma_{(12)\cdot 3}]_{12} = \{[\Sigma^{(12)}]^{-1}\}_{12},$$

which $= 0$ iff $\Sigma^{12} = 0$. □

MDP Exercise 3.19. Let $\Delta \equiv \{\delta_{ij}\}$ be a real symmetric $p \times p$ matrix such that $\Gamma\Delta\Gamma' = \Delta\ \forall$ orthogonal matrices $\Gamma \equiv \{\gamma_{ij}\}$. Set $\Gamma = \mathrm{Diag}(-1, 1, \ldots, 1)$ to see that $\delta_{12} = \cdots = \delta_{1p} = 0$. Similarly all off-diagonal $\delta_{ij} = 0$, so Δ is a diagonal matrix. Because $\Gamma\Delta = \Delta\Gamma$, $\gamma_{ij}\delta_{jj} = \delta_{ii}\gamma_{ij}$ for all $i \neq j$. Now choose Γ such that $\gamma_{ij} \neq 0$ to conclude that $\delta_{ii} = \delta_{jj} \equiv \delta$, QED. □.

MDP Exercise 3.22. From (A.7) in Appendix A, the conditional distribution $F_{p-1,n-p+1}(\zeta Z)$ of $U \mid Z$ can be represented as a Poisson mixture of central F distributions:

$$(3.69)\quad F_{p-1,\,n-p+1}(\zeta Z)\,\big|\,Z, K \sim F_{p-1+2K,\,n-p+1}, \quad K|Z \sim \mathrm{Poisson}\,(\zeta Z/2)\,.$$

Thus the conditional distribution of $U|K$ also can be represented as a mixture of central F rvs:

$$(3.70)\qquad U \,\big|\, K \;\sim\; F_{p-1+2K,\,n-p+1}\,,$$

where, since $Z \sim \chi^2_n$, the mixing probabilities are now given by

$$\begin{aligned}
\Pr[K = k] &= \mathrm{E}\{\Pr[K = k|Z]\} \\
&= \frac{1}{k!}\int_0^\infty e^{-\frac{\zeta z}{2}}(\tfrac{\zeta z}{2})^k \tfrac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} z^{\frac{n}{2}-1} e^{-\frac{z}{2}}\,dz \\
\ldots &= \frac{\Gamma(\frac{n}{2}+k)\zeta^k}{\Gamma(\frac{n}{2})\,k!(\zeta+1)^{\frac{n}{2}+k}} \\
&= \frac{\Gamma\left(\frac{n}{2}+k\right)}{\Gamma\left(\frac{n}{2}\right)\,k!}\left(1-\rho^2\right)^{\frac{n}{2}}\left(\rho^2\right)^k,
\end{aligned}$$

that is,

$$(3.72)\qquad K \sim \text{Negative binomial}\,(\rho^2). \quad □$$

MDP Exercise 4.1. From (4.11),

$$\int_{S>0} |S|^{\frac{n-p+1}{2}} e^{-\frac{1}{2}\operatorname{tr}\Sigma^{-1}S} dS = c_{p,n}^{-1} |\Sigma|^{\frac{n}{2}},$$

so

$$\begin{aligned}
\mathrm{E}(|S|^k) &= \frac{c_{p,n}}{|\Sigma|^{\frac{n}{2}}} \int_{S>0} |S|^{\frac{n+2k-p+1}{2}} e^{-\frac{1}{2}\operatorname{tr}\Sigma^{-1}S} dS \\
&= \frac{c_{p,n}}{|\Sigma|^{\frac{n}{2}}} c_{p,n+2k}^{-1} |\Sigma|^{\frac{n+2k}{2}} \\
&= |\Sigma|^k \cdot \frac{2^{pk}\,\Gamma_p\left(\frac{n}{2}+k\right)}{\Gamma_p\left(\frac{n}{2}\right)}.
\end{aligned}$$

MDP Exercise 4.2. Clearly the range of (U, V) is contained in the Cartesian product $\{0 < U < I\} \times \{V > 0\}$. Conversely, for any (U, V) in this product, (S, T) is given uniquely by $S = V^{\frac{1}{2}} U V^{\frac{1}{2}\prime}$, $T = V^{\frac{1}{2}}(I - U)V^{\frac{1}{2}\prime}$, thus the mapping $(S, T) \mapsto (U, V)$ is 1-1 and onto the Cartesian product. The remainder of the solution is straightforward except perhaps for the Jacobian. For this, apply the chain rule to the sequence of mappings

$$(S, T) \mapsto (S, V) \mapsto (U, V)$$

and apply the extended combination rule to each of the intermediate mappings:

$$\begin{aligned}
\left|\frac{\partial(S, V)}{\partial(S, T)}\right| &= \left|\frac{\partial V}{\partial T}\right| = 1, \\
\left|\frac{\partial(U, V)}{\partial(S, V)}\right| &= \left|\frac{\partial U}{\partial S}\right| = |V|^{-\frac{p+1}{2}},
\end{aligned}$$

since $V = S + T$ and $U = V^{-\frac{1}{2}} S V^{-\frac{1}{2}\prime}$. □

MDP Exercise 4.3. Use the relation $s_{ij} = s_{ii}^{1/2} r_{ij} s_{jj}^{1/2}$ to show that the Jacobian is

$$\left|\frac{\partial(s_{12}, \ldots, s_{p-1,p})}{\partial R}\right| = \prod_{i=1}^{p} s_{ii}^{\frac{p-1}{2}}.$$

MDP Exercise 4.4. Since $d(S^{-1}) = -S^{-1}(dS)S^{-1}$, the Jacobian of the mappine $W = S^{-1}$ (both symmetric matrices) is $|S^{-1}|^{p+1}$, from which the result follow. □

Solution to Exercise 5.5. For general dimension p, Takemura's estimator can be written as

$$\hat{\Sigma}(S) = \int_{\mathcal{O}} \Psi'\, T(\Psi S \Psi') \Delta T(\Psi S \Psi')'\, \Psi\, d\nu(\Psi). \tag{1}$$

Here $T(\Psi S \Psi')$ is the lower triangular square root of $\Psi S \Psi'$ so

$$T(\Psi S \Psi')[T(\Psi S \Psi')]' = \Psi S \Psi',$$

and

$$\Delta = \mathrm{Diag}(\delta_1, \ldots, \delta_p) = \sum_{j=1}^{p} \delta_j e_j e_j'$$

where $\delta_j = \frac{1}{n+p+1-2l}$ and $e_j = (0, \ldots, 0, 1, 0, \ldots 0)'$ is the j-th unit coordinate vector. Let $S = \Gamma\, D_l\, \Gamma'$ be the spectral decomposition of S, where $l \equiv (l_1 \geq \cdots \geq l_p > 0)$ are the ordered eigenvalues of S and the columns of Γ are the corresponding unit eigenvectors. Because $\hat{\Sigma}_O(S)$ is orthogonally equivariant, $\hat{\Sigma}(S) = \Gamma\, \hat{\Sigma}(D_l)\, \Gamma'$, and $\hat{\Sigma}(D_l)$ is diagonal:

$$\hat{\Sigma}(D_l) = D_{\phi(l)}. \tag{2}$$

where $\phi(l) = (\phi_1(l) \geq \cdots \geq \phi_p(l))$. Thus from (1) and (2),

$$D_{\phi(l)} = \int_{\mathcal{O}} \Psi'\, T(\Psi D_l \Psi') \Delta T(\Psi D_l \Psi')'\, \Psi\, d\nu(\Psi),$$

where now $$T(\Psi D_l \Psi')[T(\Psi D_l \Psi')]' = \Psi D_l \Psi',$$

so $$\Psi' T(\Psi D_l \Psi')\, [\Psi' T(\Psi D_l \Psi')]' \overset{(*)}{=} D_l.$$

and $$D_l^{-1/2} \Psi' T(\Psi D_l \Psi')\, [D_l^{-1/2} \Psi' T(\Psi D_l \Psi')]' \overset{(**)}{=} I.$$

Therefore

$$\begin{aligned}
\phi_i(l) &= e_i' D_{\phi(l)} e_i \\
&= e_i' \Big[\int_{\mathcal{O}} \Psi'\, T(\Psi D_l \Psi') \Big(\sum_j \delta_j e_j e_j' \Big) T(\Psi D_l \Psi')'\, \Psi\, d\nu(\Psi) \Big] e_i \\
&= \sum_{j=1}^{p} \delta_j \int_{\mathcal{O}} [e_i' \Psi'\, T(\Psi D_l \Psi') e_j]^2 d\nu(\Psi) \\
&= \sum_{j=1}^{p} \delta_j \int_{\mathcal{O}} ([\Psi'\, T(\Psi D_l \Psi')]_{ij})^2 d\nu(\Psi) \quad \equiv \sum_{j=1}^{p} \delta_j c_{ij}. \qquad (3)
\end{aligned}$$

Lemma. Define $w_{ij} = c_{ij}/l_i$. The matrix $W \equiv \{w_{ij}\}$ is nonnegative and doubly stochastic, i.e., $\sum_i w_{ij} = \sum_j w_{ij} = 1$. (Note that c_{ij}, w_{ij}, and W depend on l and that $w_{ij}(l)$ is scale-invariant: $w_{ij}(a\,l) = a\,w_{ij}(l)$ for any real $a > 0$.)

Proof. Nonnegativity is clear. Next,

$$\begin{aligned} l_i \sum_j w_{ij} = \sum_j c_{ij} &= \sum_j \int_{\mathcal{O}} ([\Psi' \, T(\Psi D_l \Psi')]_{ij})^2 d\nu(\Psi) \\ &= \int_{\mathcal{O}} \sum_j ([\Psi' \, T(\Psi D_l \Psi')]_{ij})^2 d\nu(\Psi) \\ &\overset{(***)}{=} \int_{\mathcal{O}} l_i \, d\nu(\Psi) \quad = l_i, \end{aligned}$$

where (***) follows from (*); thus $\sum_j w_{ij} = 1$. Finally,

$$\begin{aligned} \sum_i w_{ij} = \sum_i \frac{c_{ij}}{l_i} &= \sum_i \frac{1}{l_i} \int_{\mathcal{O}} ([\Psi' \, T(\Psi D_l \Psi')]_{ij})^2 d\nu(\Psi) \\ &= \int_{\mathcal{O}} \sum_i ([D_l^{-1/2} \Psi' \, T(\Psi D_l \Psi')]_{ij})^2 d\nu(\Psi) \\ &\overset{(****)}{=} \int_{\mathcal{O}} 1 \, d\nu(\Psi) \quad = 1, \end{aligned}$$

where (****) follows from (**); thus $\sum_j w_{ij} = 1$. $\square$

Now consider the case $p = 2$. From (3),

$$\begin{aligned} \phi_1 &= (\delta_1 w_{11} + \delta_2 w_{12}) l_1 = [\delta_1 w_{22} + \delta_2 (1 - w_{22})] l_1, \\ \phi_2 &= (\delta_1 w_{21} + \delta_2 w_{22}) l_2 = [\delta_1 (1 - w_{22}) + \delta_2 w_{22}] l_2, \end{aligned}$$

since W is doubly stochastic. Thus it suffices to show that

$$w_{22} \equiv w_{22}(l_1, l_2) = \frac{\sqrt{l_1}}{\sqrt{l_1} + \sqrt{l_2}}.$$

Because $w_{22}(l_1, l_2)$ is scale-invariant, it suffices to show that

$$(4) \qquad w_{22}(a, 1) = \frac{\sqrt{a}}{\sqrt{a} + 1}.$$

The group of 2×2 orthogonal matrices is the disjoint union $\mathcal{O} = \mathcal{O}^+ \cup \mathcal{O}^-$, where

$$\mathcal{O}^+ = \left\{ \Psi_+(\theta) \equiv \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \,\middle|\, 0 \le \theta < 2\pi \right\},$$
$$\mathcal{O}^- = \left\{ \Psi_-(\theta) \equiv \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \,\middle|\, 0 \le \theta < 2\pi \right\}.$$

By symmetry,

$$\begin{aligned} w_{22}(a,1) &= \frac{1}{2\pi}\int_0^{2\pi} ([\Psi_+(\theta)' \, T(\Psi_+(\theta)\,\mathrm{Diag}(a,1)\Psi_+(\theta)')]_{22})^2 d\theta \\ &\overset{(\dagger)}{=} \frac{1}{2\pi}\int_0^{2\pi} \frac{a\cos^2\theta}{a\cos^2\theta + \sin^2\theta}\, d\theta \\ &\overset{(\ddagger)}{=} \frac{\sqrt{a}}{\sqrt{a}+1}. \end{aligned}$$

Notes: 1. For (†), first show that

$$T(\Psi_+(\theta)\,\mathrm{Diag}(a,1)\Psi_+(\theta)') = \begin{pmatrix} t_{11} & 0 \\ t_{21} & t_{22} \end{pmatrix}$$

with

$$t_{11} = a\cos^2\theta, \quad t_{21} = \frac{(a-1)\sin\theta\cos\theta}{\sqrt{a\cos^2\theta + \sin^2\theta}}, \quad t_{22}^2 = \frac{a}{a\cos^2\theta + \sin^2\theta}.$$

2. A nice derivation of (‡) is given by our student Zhen Miao, as follows:

$$\begin{aligned} \tfrac{1}{2\pi}\int_0^{2\pi} \tfrac{a\cos^2\theta}{a\cos^2\theta+\sin^2\theta}\, d\theta &= \tfrac{1}{2\pi}\int_0^{2\pi} \tfrac{a}{a+\tan^2\theta}\, d\theta = \tfrac{2}{\pi}\int_0^{\frac{\pi}{2}} \tfrac{a}{a+\tan^2\theta}\, d\theta \\ &= \tfrac{2}{\pi}\int_0^{\infty} \tfrac{a}{a+t^2}\tfrac{dt}{1+t^2} = \tfrac{2}{\pi}\tfrac{a}{1-a}\Big[\int_0^\infty \tfrac{dt}{a+t^2} - \int_0^\infty \tfrac{dt}{1+t^2}\Big] \\ &= \tfrac{2}{\pi}\tfrac{a}{1-a}\Big[\tfrac{1}{\sqrt{a}} - 1\Big]\int_0^\infty \tfrac{dt}{1+t^2} = \tfrac{2}{\pi}\tfrac{\sqrt{a}}{1+\sqrt{a}}\tfrac{\pi}{2} = \tfrac{\sqrt{a}}{1+\sqrt{a}}. \end{aligned}$$

Exercise A.19. Prove Fact A.4: *If* $f(x,y) > 0$ *and* $(*)\frac{\partial^2 \log f}{\partial x \partial y} \ge 0$ *on* $A \times B$ *then* f *is TP2.*

Solution. $(*) \Rightarrow \frac{\partial \log f(x_2,y)}{\partial y} \ge \frac{\partial \log f(x_1,y)}{\partial y}$ when $x_2 > x_1$. This implies that $\log f(x_2,y) - \log f(x_1,y)$ is nondecreasing in y, which is equivalent to $(A.2)$.

Exercise 6.37c. Let $U(p,\, r,\, n)$ denote the null ($\mu = 0$) distribution of $|U|$. ($U(p,\, r,\, n)$ is called *Wilks' distribution.*) Show that this distribution can be represented as the product of independent Beta distributions:

$$U(p,r,n) \sim \prod_{i=1}^{r} B\big(\tfrac{n-p+i}{2},\, \tfrac{p}{2}\big), \tag{6.33}$$

where the Beta variates are mutually independent.

Solution. Here $|U| \equiv \frac{|T|}{|XX'+T|}$. Write $X = (X_1, X_2)$ with $X_1 : p \times 1$ and $X_2 : p \times (r-1)$, so

$$|U| = \frac{|T|}{|X_2X_2' + T|} \cdot \frac{|X_2X_2' + T|}{|X_1X_1' + X_2X_2' + T|} \equiv A \cdot B. \tag{1}$$

Because $\frac{|T|}{|X_2X_2'+T|} = \frac{1}{|X_2'T^{-1}X_2+I_r|}$, it follows from Exercise 6.37a that

$$\frac{|T|}{|X_2X_2' + T|} \perp\!\!\!\perp (X_2X_2' + T) \perp\!\!\!\perp X_1, \tag{2}$$

hence $A \perp\!\!\!\perp B$. Because $A \sim U(p, r-1, n)$ and

$$\begin{aligned} B = \frac{1}{X_1'(X_2X_2' + T)^{-1}X_2 + 1} &\sim B\left(\frac{n + (r-1) - p + 1}{2}, \frac{p}{2}\right) \\ &= B\left(\frac{n + r - p}{2}, \frac{p}{2}\right), \end{aligned}$$

the result (6.33) follows by induction.