Question 1 [4 marks]

Suppose that A is a 2×2 Wishart distributed random matrix (i.e., $A \sim W_p(n, \Sigma)$ with $p = 2$ and $n \geq 2$). Prove using elementary methods that $\det(A)$ is distributed like $\det(\Sigma)$ times two independent chi-squared random variables with degrees of freedom n and $n-1$.

As $A \sim W_p(n, \Sigma)$, we have

$$A = \sum_{i=1}^{n} z_i z_i^T, \quad z_i \sim N_p(0, \Sigma) \quad (\text{As } \Delta = 0 \text{ in our case})$$

Let's define the unique square-root of Σ:

$$CC^T = \Sigma$$

We define y_k such that $z_k := C y_k$, such that

$$y_1, y_2, \dots, y_N \sim N_p(0, I_p)$$

$$\Rightarrow y_k = C^{-1} z_k$$

Define $B := \sum_{k=1}^{N} y_k y_k^T = \sum_{k=1}^{N} C^{-1} z_k z_k^T (C^{-1})^T$

$$= C^{-1} A (C^{-1})^T$$

then; $|A| = |C| \cdot |B| \cdot |C^T| = |B| \cdot |\Sigma|$

$$B \sim W_p(n, I_p), \quad p = 2$$

Define $B = \sum_{i=1}^{n} y_i y_i^T = Y^T Y$, $Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{R}^{2 \times n}$.

Define $y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \end{bmatrix} \in \mathbb{R}^{2 \times 1}$, $y_{i1}, y_{i2} \overset{iid}{\sim} N(0, 1)$

Then $B = \begin{bmatrix} \sum_{i=1}^n y_{i1}^2 & \sum_{i=1}^n y_{i1} y_{i2} \\ \sum_{i=1}^n y_{i2} y_{i1} & \sum_{i=1}^n y_{i2}^2 \end{bmatrix}$

Let's perform cholesky decomposition here.

Denote the Cholesky L as $\begin{bmatrix} a & 0 \\ b & c \end{bmatrix}$

Then, for a 2×2 matrix:

$$B = LL^T = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \cdot \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} = \begin{bmatrix} a^2 & ab \\ ba & b^2+c^2 \end{bmatrix}$$

$$\Rightarrow \begin{cases} a^2 = \sum_{i=1}^n y_{i1}^2 \\ ab = \sum_{i=1}^n y_{i1} y_{i2} \\ b^2 + c^2 = \sum_{i=1}^n y_{i2}^2 \end{cases} \Rightarrow \begin{cases} a = \sqrt{\sum_{i=1}^n y_{i1}^2} \\ b = \sum_{i=1}^n y_{i1} y_{i2} \Big/ \sqrt{\sum_{i=1}^n y_{i1}^2} \\ c = \sqrt{\sum_{i=1}^n y_{i2}^2 - \left(\sum_{i=1}^n y_{i1} y_{i2}\right)^2 \Big/ \sum_{i=1}^n y_{i1}^2} \end{cases}$$

$$|B| = |L|\,|L^T| = |a \cdot c| \cdot |a \cdot c|$$

$$= a^2 \cdot c^2$$

$$= \left(\sum_{i=1}^n y_{i1}^2\right) \cdot \left(\sum_{i=1}^n y_{i2}^2 - \left(\sum_{i=1}^n y_{i1} y_{i2}\right)^2 \Big/ \sum_{i=1}^n y_{i1}^2 \right)$$

Observe that $\sum_{i=1}^n y_{i1}^2 \sim \chi(n)$

$Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{R}^{2\times n}$, $y_1, y_2 \overset{iid}{\sim} N_n(\vec{0}, I_n)$

Let's do a Gram-Schmidt factorization

Define $w_1 = y_1$

$$w_2 = y_2 - w_1 \frac{w_1^T y_2}{w_1^T w_1}$$

Define $t_{11} = \frac{y_1^T y_1}{\|y_1\|}$, $t_{11}^2 = \frac{y_1^T y_1 \cdot y_1 y_1}{\cancel{y_1^T y_1}} = y_1^T y_1$

$$t_{22} = \frac{y_2^T w_2}{\|w_2\|}, \quad t_{22}^2 = \frac{(y_2^T w_2)^2}{w_2^T w_2}$$

$$y_2^T w_2 = y_2^T \left(y_2 - w_1 \frac{w_1^T y_2}{w_1^T w_1}\right)$$

$$= y_2^T y_2 - y_2^T w_1 \frac{w_1^T y_2}{w_1^T w_1}$$

$$w_2^T w_2 = \left(y_2 - w_1 \frac{w_1^T y_2}{w_1^T w_1}\right)^T \left(y_2 - w_1 \frac{w_1^T y_2}{w_1^T w_1}\right)$$

$$= \left(y_2^T y_2 - y_2^T w_1 \frac{w_1^T y_2}{w_1^T w_1}\right) - w_1^T \frac{y_2^T w_1}{w_1^T w_1} \left(y_2 - w_1 \frac{w_1^T y_2}{w_1^T w_1}\right)$$

$$= \left(y_2^T y_2 - y_2^T w_1 \frac{w_1^T y_2}{w_1^T w_1}\right) - w_1^T w_2 \frac{y_2^T w_1}{w_1^T w_1}$$

From the Gram-Schmidt factorization: $w_1 \cdot w_2 = 0$, thus

$$= \left(y_2^T y_2 - y_2^T w_1 \frac{w_1^T y_2}{w_1^T w_1}\right) = y_2^T w_2$$

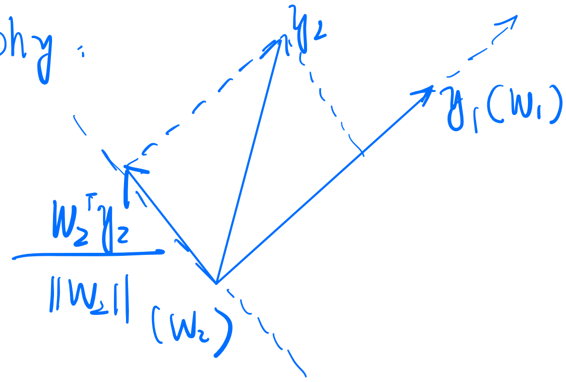
Thus, $t_{22}^2 = \frac{(y_2^T W_2)^2}{W_2^T W_2} = \frac{(y_2^T W_2)^2}{y_2^T W_2} = y_2^T W_2$

$$= y_2^T y_2 - y_2^T W_1 \frac{W_1^T y_2}{W_1^T W_1}$$

$$= \sum_{i=1}^{n} y_{i2}^2 - \left(\sum_{i=1}^{n} y_{i1} y_{i2}\right)^2 \Big/ \sum_{i=1}^{n} y_{i1}^2$$

$$t_{11}^2 = y_1^T y_1 = \sum_{i=1}^{n} y_{i1}^2$$

From geography:



We know $t_{11} \perp t_{22}$ are independently distributed.

t_{22}^2 is the $y_2^T W_2$ which is the projection of V_2 to W_2 which is the sum of y_{2i} projected to W_2 direction omitting the $y_1(W_1)$, thus:

$$t_{22}^2 \sim \chi(n-1)$$

$$t_{11}^2 \sim \chi(n)$$

Thus

$$|A| = |B| \cdot |\Sigma| = |\Sigma| \cdot \chi(n) \cdot \chi(n-1)$$

Question 2 [4 marks]

Consider a sequence of two-dimensional random vectors $\mathbb{x}_1, \mathbb{x}_2, \ldots, \mathbb{x}_n$ where the coordinates of $\mathbb{x}_n := (x_{n1}, x_{n2})'$. Suppose that

$$\sqrt{n}\,\mathbb{x}_n \to \mathbf{N}_2\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}3 & 1\\1 & 2\end{bmatrix}\right),$$

then what is the asymptotic distribution of $\xi := \sqrt{n}\,(x_{n1} + x_{n2}^2)$ as $n \to \infty$?

$$\sqrt{n}\,(x_n - \theta) \xrightarrow{D} N_2\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}3 & 1\\1 & 2\end{bmatrix}\right)$$

Here θ is a consistent estimater of μ of x_n:

$$\theta \xrightarrow{P} \mu = \begin{bmatrix}0\\0\end{bmatrix}$$

From the delta method:

$$\sqrt{n}\left[g(x_n) - g(\theta)\right] \xrightarrow{D} N_2\left(\begin{pmatrix}0\\0\end{pmatrix}, \nabla g(\theta)^T \begin{bmatrix}3 & 1\\1 & 2\end{bmatrix} \cdot \nabla g(\theta)\right)$$

Define $g: R^2 \to R$ as $g(x_{n1}, x_{n2}) = x_{n1} + x_{n2}^2$

$$g(\theta) = g(\mu) = g\left(\begin{bmatrix}0\\0\end{bmatrix}\right) = 0$$

$$\nabla g(x_n) = \begin{bmatrix}\frac{\partial g(x_n)}{\partial x_{n1}}\\[4pt] \frac{\partial g(x_n)}{\partial x_{n2}}\end{bmatrix} = \begin{bmatrix}1\\2x_{n2}\end{bmatrix}$$

$$\nabla g(x_n)\Big|_{\theta=0} = \begin{bmatrix}1\\0\end{bmatrix}$$

$$\nabla g(\theta)^T \begin{bmatrix}3 & 1\\1 & 2\end{bmatrix} \nabla g(\theta) = \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}3 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}$$

$$= \begin{bmatrix}3 & 1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = 3$$

Thus

$$\sqrt{n}\left[g(X_n) - g(\theta)\right] \xrightarrow{D} N(0,3)$$

$$\Rightarrow \sqrt{n}\left(X_{n1} + X_{n2}^2\right) \xrightarrow{D} N(0,3)$$

(b) The paper defines (see Theorem 1) that

$$\eta := \mathbb{E}[f(\sigma_w \sigma_x \xi)^2], \quad \zeta := \mathbb{E}[\sigma_w \sigma_x f'(\sigma_w \sigma_x \xi)]^2,$$

where $\xi \sim N(0,1)$. The paper claims that when $f = f_\alpha$, it is straightforward to check that (see near equation (18) in paper),

$$\eta = 1, \quad \zeta = \frac{(1-\alpha)^2}{2(1+\alpha)^2 - \frac{2}{\pi}(1+\alpha)^2}.$$

Proceed to check this is true. Note that $[x]_+ = \max(0,x) = x\mathbf{1}_{\{x>0\}}$.

$$\eta = \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} f(\sigma_w \sigma_x z)^2$$

$$= \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \cdot \frac{\left([\sigma_w\sigma_x z]_+ + \alpha[-\sigma_w\sigma_x z]_+ - \frac{1+\alpha}{\sqrt{2\pi}}\right)^2}{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2}$$

$$= \frac{1}{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2} \left\{ \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \left([\sigma_w\sigma_x z]_+ + \alpha[-\sigma_w\sigma_x z]_+\right)^2 \right.$$

$$-2\int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \left([\sigma_w\sigma_x z]_+ + \alpha[-\sigma_w\sigma_x z]_+\right)\left(\frac{(1+\alpha)}{\sqrt{2\pi}}\right)$$

$$\left. + \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \left(\frac{1+\alpha}{\sqrt{2\pi}}\right)^2 \right\}$$

$$\int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} [z]_+ = \int_0^{+\infty} dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \cdot z = \frac{1}{\sqrt{2\pi}}\left(-e^{-z^2/2}\right)\Big|_0^{+\infty}$$

$$= 0 - \left(-\frac{1}{\sqrt{2\pi}}\right) = \frac{1}{\sqrt{2\pi}}$$

$$\int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} [-z]_+ = \int_{-\infty}^0 dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \cdot (-z) = \frac{1}{\sqrt{2\pi}}\left(e^{-z^2/2}\right)\Big|_{-\infty}^0$$

$$= \frac{1}{\sqrt{2\pi}}$$

$$\int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} [z]_+^2 = \int_0^{+\infty} dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \cdot z^2 = \int_0^{+\infty} dz \left(z\frac{e^{-z^2/2}}{\sqrt{2\pi}}\right) \cdot z$$

$$= \frac{1}{\sqrt{2\pi}}\left(-z e^{-z^2/2}\right)\Big|_0^{+\infty} + \int_0^{+\infty} \left(e^{-z^2/2}/\sqrt{2\pi}\right)$$

$= 0 + \frac{1}{2}$

$= \frac{1}{2}$

$$\int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} [-z]_+^2 = \int_{-\infty}^{0} dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \cdot z^2 = \frac{1}{2} \quad \text{By Symmetryic.}$$

Thus: $\int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \left([\bar{\sigma}_w \bar{\sigma}_x z]_+ + \alpha [-\bar{\sigma}_w \bar{\sigma}_x z] \right)^2$

$$= \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sigma_w^2 \sigma_x^2 [z]_+^2 + 2 \cdot \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sigma_w^2 \sigma_x^2 [z]_+ [-z]_+ \alpha$$

$$+ \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sigma_w^2 \sigma_x^2 \alpha^2 [-z]_+^2$$

$\int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} [z]_+ [-z]_+ = 0$, as if $z \geq 0$, $[-z]_+ = 0$

and if $z < 0$, $[z] = 0$, thus.

$$= \sigma_w^2 \sigma_x^2 \cdot \frac{1}{2} + \sigma_w^2 \sigma_x^2 \cdot \alpha^2 \cdot \frac{1}{2}$$

$$-2 \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \left([\bar{\sigma}_w \bar{\sigma}_x z]_+ + \alpha [-\bar{\sigma}_w \bar{\sigma}_x z] \right) \left(\frac{(1+\alpha)}{\sqrt{2\pi}} \right)$$

$$= \frac{-2(1+\alpha)}{\sqrt{2\pi}} \left\{ \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} [\sigma_w \sigma_x z] + \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \alpha [-\sigma_w \sigma_x z] \right\}$$

$$= \frac{-2(1+\alpha)}{\sqrt{2\pi}} \left\{ \frac{\sigma_w \sigma_x}{\sqrt{2\pi}} + \frac{\alpha \sigma_w \sigma_x}{\sqrt{2\pi}} \right\}$$

$$\int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \left(\frac{1+\alpha}{\sqrt{2\pi}} \right)^2 = \frac{(1+\alpha)^2}{2\pi} \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} = \frac{(1+\alpha)^2}{2\pi}.$$

$$\eta = \frac{1}{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2} \Big\{ \sigma_w^2 \sigma_x^2 \cdot \frac{1}{2} + \sigma_w^2 \sigma_x^2 \cdot \alpha^2 \cdot \frac{1}{2}$$

$$\frac{-2(1+\alpha)}{\sqrt{2\pi}} \Big\{ \frac{\sigma_w \sigma_x}{\sqrt{2\pi}} + \frac{\alpha\, \sigma_w \sigma_x}{\sqrt{2\pi}} \Big\} + \frac{(1+\alpha)^2}{2\pi} . \Big\}$$

$$= \frac{1}{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2} \Big\{ \frac{1}{2}(1+\alpha^2)\sigma_w^2 \sigma_x^2 - \frac{2(1+\alpha)(1+\alpha)\sigma_w \sigma_x}{2\pi} + \frac{(1+\alpha)^2}{2\pi} \Big\}$$

Let's take $\sigma_w = \sigma_x = 1$

Then,

$$= \frac{1}{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2} \Big\{ \frac{1}{2}(1+\alpha^2) - \frac{2(1+\alpha)^2}{2\pi} + \frac{(1+\alpha)^2}{2\pi} \Big\}$$

$$= 1$$

$$\varsigma = \Big[\sigma_w \sigma_x \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} f'(\sigma_w \sigma_x z) \Big]^2$$

$$f_\alpha(x) = \frac{[x]_+ + \alpha[-x]_+ - \frac{1+\alpha}{\sqrt{2\pi}}}{\sqrt{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2}}$$

Let's define a subgradient for ReLU function:

$$([x]_+)' = \begin{cases} 1, & x > 0 \\ [0,1], & x = 0 \\ 0, & x < 0 \end{cases}, \text{ define subdivision } x^+ \to 0 = 1 \text{ and } x^- \to 0 = 0$$

Thus

$$([x]_+)' = \begin{cases} 1, & x \geq 0 \\ 0, & x \leq 0 \end{cases}$$

If $x>0$: $f_\alpha'(\alpha) = \dfrac{1}{\sqrt{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2}}$

If $x<0$: $f_\alpha'(x) = \dfrac{-\alpha}{\sqrt{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2}}$

If $x=0$: $f_\alpha'(\alpha) = \dfrac{c-\alpha c}{\sqrt{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2}}$, $c \in [0,1]$

In all the cases, the $f_\alpha'(\alpha)$ is not a function of x.

Thus

$$\varsigma = \left[\sigma_w\sigma_x \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} f'(\sigma_w\sigma_x z)\right]^2 , \text{ take } \sigma_w = \sigma_x = 1$$

$$= \left\{\sigma_w\sigma_x\left[\int_0^{+\infty} dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \frac{1}{\sqrt{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2}}\right.\right.$$

$$\left.\left. + \int_{-\infty}^{0} dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \frac{-\alpha}{\sqrt{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2}}\right]\right\}^2$$

$$= \left\{\frac{1}{2}\frac{1}{\sqrt{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2}} + \frac{1}{2}\frac{-\alpha}{\sqrt{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2}}\right\}^2$$

$$= \frac{(1-\alpha)^2}{\frac{1}{2}(1+\alpha^2) - \frac{1}{2\pi}(1+\alpha)^2} \cdot \frac{1}{4}$$

$$= \frac{(1-\alpha)^2}{2(1+\alpha^2) - 2/\pi(1+\alpha)^2}$$

$\Rightarrow$ proofed.