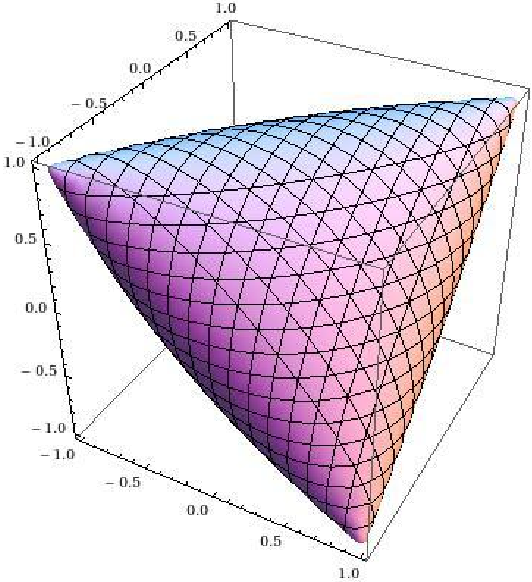
$\Theta = \mathbb{R}^p \times \mathbb{R}^p_+ \times \mathcal{E}_{\binom{p}{2}}$. The 3-dimensional elliptope:



$\hat{\mathbf{C}} = \frac{1}{n}\mathbf{S}$ indeed produces the maximum of the log-likelihood function

$$
\begin{aligned}
&\ell(\hat{\mathbf{C}}^{-1}) - \ell(\mathbf{C}^{-1}) \\
&= c + \frac{n}{2}\ln|\hat{\mathbf{C}}^{-1}| - \frac{1}{2}\mathtt{tr}\,(\hat{\mathbf{C}}^{-1}\mathbf{S}) \\
&- c - \frac{n}{2}\ln|\mathbf{C}^{-1}| + \frac{1}{2}\mathtt{tr}\,(\mathbf{C}^{-1}\mathbf{S}) \\
&= \frac{n}{2}\ln|(\frac{\mathbf{S}}{n})^{-1}| - \frac{1}{2}\mathtt{tr}\,((\frac{\mathbf{S}}{n})^{-1}\mathbf{S}) \\
&- \frac{n}{2}\ln|\mathbf{C}^{-1}| + \frac{1}{2}\mathtt{tr}\,(\mathbf{C}^{-1}\mathbf{S}) \\
&= \frac{n}{2}\left[-\ln|\mathbf{C}^{-1}\frac{\mathbf{S}}{n}| - \frac{p}{2} + \mathtt{tr}\,(\mathbf{C}^{-1}\frac{\mathbf{S}}{n})\right]
\end{aligned}
$$

$$\ell(\hat{\mathbf{C}}^{-1}) - \ell(\mathbf{C}^{-1}) = \frac{n}{2}\sum_{i=1}^{p}[-\ln\lambda_i - 1 + \lambda_i] \geq 0,$$

where λ_i ($i = 1, \ldots, p$) are positive real eigenvalues of the matrix $\mathbf{C}^{-1}\frac{\mathbf{S}}{n}$. Equality is attained if $\mathbf{C}^{-1}\frac{\mathbf{S}}{n} = \mathbf{I}_p$, i.e., at $\hat{\mathbf{C}}$.

Lemma: If $\mathbf{A}$ and $\mathbf{B}$ are symmetric, positive definite matrices, then $\mathbf{AB}$ (not singular) has real positive eigenvalues.

Proof:

$$\mathbf{ABx} = \lambda\mathbf{x}$$

$$(\mathbf{B}^{1/2}\mathbf{AB}^{1/2})(\mathbf{B}^{1/2}\mathbf{x}) = \lambda(\mathbf{B}^{1/2}\mathbf{x}),$$

where $\mathbf{B}^{1/2}\mathbf{AB}^{1/2} = (\mathbf{B}^{1/2}\mathbf{A}^{1/2})(\mathbf{B}^{1/2}\mathbf{A}^{1/2})^T$ is positive definite. (For pos. def. $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^T \Longrightarrow \mathbf{A}^{1/2} = \mathbf{U}\Lambda^{1/2}\mathbf{U}^T$.)

Converse statement of Theorem 2: Eugene Lukács

A characterization of the normal distribution, *Ann Math. Statistics* **13** (1942), 91–93.
"A necessary and sufficient condition for the normality of the parent distribution is that the sampling distributions of the mean and of the variance be independent."
This was first proved in
R. C. Geary, Distribution of Student's ratio for nonnormal samples, Roy. Stat. Soc. Jour., Supp. Vol. 3, no. 2 (1936).
Here it is proved (for absolutely continuous distributions), by using characteristic functions.
"This reasoning applies also to the multivariate case."

Eugene Lukács

The stochastic independence of symmetric and homogeneous linear and quadratic statistics. *Ann Math. Statistics* **23** (1952), 442–449.

"If a univariate distribution has moments of first and second order and admits a homogeneous and symmetric quadratic statistic Q which is independently distributed of the mean of a sample of n drawn from this distribution, then it is either the normal distribution (Q is then proportional to the variance) or the degenerate distribution (in this case no restriction is imposed on Q) or a step function with two symmetrically located steps (in this case Q is the sum of the squared observations).
The converse of this statement is also true."

The standard Wishart density

The density of the $p \times p$ standard Wishart-matrix $\mathbf{W}$:

$$c_{np}|\mathbf{W}|^{\frac{n-p-1}{2}} e^{-\frac{1}{2}\mathrm{tr}\,\mathbf{W}},$$

where $n > p$ and

$$c_{np} = \left[\sqrt{2}^{np}\sqrt{\pi}^{\binom{p}{2}} \prod_{i=1}^{p} \Gamma\left(\frac{n+1-i}{2}\right)\right]^{-1}.$$

J. Wishart, **The generalized product moment distribution in samples from a normal multivariate population**, *Biometrika* **20** (1928), 32–52.

I. Olkin, **The 70th anniversary of the distribution of random matrices: A survey**, *Lin. Alg. Appl.* **354** (2002), 231–243.

Olkin's method

Old: $\mathbf{X} \longrightarrow$ New: $\mathbf{Y} = t(\mathbf{X})$, t is bijection.

$pdf_{\mathbf{X}}(\mathbf{x}) \longrightarrow pdf_{\mathbf{Y}}(\mathbf{y}) = pdf_{\mathbf{X}}(t^{-1}(\mathbf{y})) \cdot \left|\mathtt{det}\left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}}\right)\right|$

$\mathbf{X}_{p\times n} \longrightarrow \mathbf{W}_{p\times p} = \mathbf{X}\mathbf{X}^T$ $(p < n)$ not a bijection, but

pdf of $\mathbf{X}$: $\prod_{i=1}^{n}\prod_{j=1}^{p}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}X_{ij}^2} = \frac{1}{\sqrt{2\pi}^{np}}e^{-\frac{1}{2}\mathtt{tr}\,(\mathbf{W})} =: f(\mathbf{W})$

pdf of $\mathbf{W}$: $f(\mathbf{W}) \cdot h(\mathbf{W})$, $\quad h = ?$

Find h!

1. $\mathbf{X}_{p\times n} \longrightarrow \mathbf{Y}_{p\times n}$: $\mathbf{X} = \mathbf{AY}$ with $\mathbf{A}_{p\times p}$ non-singular.

$$pdf_{\mathbf{Y}}(\mathbf{Y}) = pdf_{\mathbf{X}}(\mathbf{AY}) \cdot \left| \det\left(\frac{\partial \mathbf{X}}{\partial \mathbf{Y}}\right) \right| = f(\mathbf{AYY}^T\mathbf{A}^T) \cdot |\det(\mathbf{A})|^n,$$

since

$$\det\left(\frac{\partial \mathbf{X}}{\partial \mathbf{Y}}\right) = \det(\mathbf{I}_n \otimes \mathbf{A}) = (\det(\mathbf{I}_n))^p \cdot (\det(\mathbf{A}))^n = (\det(\mathbf{A}))^n.$$

2. $\mathbf{Y}_{p\times n} \longrightarrow \mathbf{V}_{p\times p}$: $\mathbf{V} = \mathbf{Y}\mathbf{Y}^T$. Find *pdf* of $\mathbf{V}$!

On the one hand, since $\mathbf{W} = \mathbf{X}\mathbf{X}^T = \mathbf{A}\mathbf{Y}\mathbf{Y}^T\mathbf{A}^T = \mathbf{A}\mathbf{V}\mathbf{A}^T$:

$$\begin{aligned} pdf_{\mathbf{V}}(\mathbf{V}) &= pdf_{\mathbf{W}}(\mathbf{A}\mathbf{V}\mathbf{A}^T) \cdot \left| \det\left(\frac{\partial \mathbf{W}}{\partial \mathbf{V}}\right) \right| \\ &= f(\mathbf{A}\mathbf{V}\mathbf{A}^T) \cdot h(\mathbf{A}\mathbf{V}\mathbf{A}^T) \cdot |\det(\mathbf{A})|^{p+1}. \end{aligned}$$

On the other hand, since $\mathbf{V}$ depends on $\mathbf{Y}$ in the same way as $\mathbf{W}$ depends on $\mathbf{X}$:

$$pdf_{\mathbf{V}}(\mathbf{V}) = pdf_{\mathbf{Y}}(\mathbf{V}) \cdot h(\mathbf{V}) = f(\mathbf{A}\mathbf{V}\mathbf{A}^T) \cdot |\det(\mathbf{A})|^n \cdot h(\mathbf{V}).$$

We make the two expressions for the *pdf* of V equal:

$$f(\mathbf{AVA}^T) \cdot h(\mathbf{AVA}^T) \cdot |\mathtt{det}\,(\mathbf{A})|^{p+1} = f(\mathbf{AVA}^T) \cdot |\mathtt{det}\,(\mathbf{A})|^{n} \cdot h(\mathbf{V})$$

$$h(\mathbf{AVA}^T) = h(\mathbf{V}) \cdot |\mathtt{det}\,(\mathbf{A})|^{n-p-1}$$

Choosing $\mathbf{A}$ such that $\mathbf{V} = \mathbf{I}_p$, yields

$$h(\mathbf{W}) = h(\mathbf{I}_p) \cdot (\mathtt{det}\,(\mathbf{W}))^{\frac{n-p-1}{2}} = \kappa_p \cdot (\mathtt{det}\,(\mathbf{W}))^{\frac{n-p-1}{2}}$$

and

$$c_{np} = \kappa_p \frac{1}{\sqrt{2\pi}^{np}}.$$

(The non-standard Wishart density is obtained by an easy transformation.)

Distribution of the standard Wishart eigenvalues

M. L. Mehta: **Random matrices**. Academic Press, London, 1991.
pdf of $\mathbf{W}$:

$$c_{np}|\mathbf{W}|^{\frac{n-p-1}{2}} e^{-\frac{1}{2}\mathrm{tr}\,\mathbf{W}} = c_{np} \prod_{j=1}^{p} \left(\lambda_j^{\frac{n-p-1}{2}} e^{-\frac{1}{2}\lambda_j} \right)$$

where $\lambda_1, \dots, \lambda_p > 0$ are the eigenvalues of $\mathbf{W}$ $(n > p)$.

- *pdf* of $\mathbf{W} = \mathbf{U}\Lambda\mathbf{U}^T \longrightarrow$ *pdf* of $(\Lambda, \mathbf{U})$
- integrating with respect to $\mathbf{U}$, the joint *pdf* of $\lambda_1, \dots, \lambda_p$:

$$\kappa_{np} \left(\prod_{j=1}^{p} \lambda_j\right)^{\frac{n-p-1}{2}} e^{-\frac{1}{2}\sum_{j=1}^{p}\lambda_j} \prod_{j<k} |\lambda_j - \lambda_k|$$

Remark

Without absolute values, this is the van der Monde determinant:

$$\prod_{j<k}(\lambda_j - \lambda_k) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_p \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_p^2 \\ \dots & \dots & \dots & \dots \\ \lambda_1^{p-1} & \lambda_2^{p-1} & \dots & \lambda_p^{p-1} \end{vmatrix}$$

If $n > p$, then the eigenvalues are (w. p. 1) positive and distinct. The *pdf* of their ordered sample is $p!$ times the above.

$\mathbf{W} \to (\Lambda, \mathbf{U})$

$\#$ of parameters: $p(p+1)/2 = p + p(p-1)/2$:

$\underline{\lambda} = (\lambda_1, \ldots, \lambda_p), \quad \mathbf{q} = (q_1, \ldots, q_{p(p-1)/2})$

pdf of $(\underline{\lambda}, \mathbf{q})$:

$$g(\underline{\lambda}, \mathbf{q}) = c_{np} \prod_{j=1}^{p} \left(\lambda_j^{\frac{n-p-1}{2}} e^{-\frac{1}{2}\lambda_j} \right) \cdot \left| \frac{\partial \mathbf{W}}{\partial(\underline{\lambda}, \mathbf{q})} \right|$$

$$= c_{np} \prod_{j=1}^{p} \left(\lambda_j^{\frac{n-p-1}{2}} e^{-\frac{1}{2}\lambda_j} \right) \cdot |\mathbf{J}(\underline{\lambda}, \mathbf{q})|$$

Some technical steps

We differentiate both sides of $\mathbf{U}^T\mathbf{U} = \mathbf{I}_p$ with respect to q_z:

$$\frac{\partial \mathbf{U}^T}{\partial q_z}\mathbf{U} + \mathbf{U}^T\frac{\partial \mathbf{U}}{\partial q_z} = \mathbf{O}_{p\times p},$$

from where

$$-\frac{\partial \mathbf{U}^T}{\partial q_z}\mathbf{U} = \mathbf{U}^T\frac{\partial \mathbf{U}}{\partial q_z} = \mathbf{S}^{(z)}$$

is skew symmetric:

for $1 \leq x \leq y \leq p$: $s_{xy}^{(z)} = -s_{yx}^{(z)}, \quad z = 1, \ldots, \binom{p}{2}$.

(Consequently, $s_{xx}^{(z)} = 0$, $x = 1, \ldots, p$.)

Some technical steps

We differentiate both sides of $\mathbf{W} = \mathbf{U}\Lambda\mathbf{U}^T$ with respect to q_z:

$$\frac{\partial \mathbf{W}}{\partial q_z} = \frac{\partial \mathbf{U}}{\partial q_z}\Lambda\mathbf{U}^T + \mathbf{U}\Lambda\frac{\partial \mathbf{U}^T}{\partial q_z},$$

from where, multiplying by $\mathbf{U}^T$ from the left, and by $\mathbf{U}$ from the right:

$$\mathbf{U}^T\frac{\partial \mathbf{W}}{\partial q_z}\mathbf{U} = \mathbf{S}^{(z)}\Lambda - \Lambda\mathbf{S}^{(z)}.$$

Equivalently, for the entries:

$$\sum_{j,k}\frac{\partial w_{jk}}{\partial q_z}u_{jx}u_{ky} = s_{xy}^{(z)}(\lambda_y - \lambda_x),$$

$1 \leq x \leq y \leq p, \quad z = 1, \ldots, \binom{p}{2}.$
(For $x > y$ we get the same because of the skew symmetry of $\mathbf{S}^{(z)}$.)

Some technical steps

We differentiate both sides of $\mathbf{U}^T\mathbf{W}\mathbf{U} = \Lambda$ with respect to λ_t:

$$\sum_{j,k} \frac{\partial w_{jk}}{\partial \lambda_t} u_{jx} u_{ky} = \frac{\partial \Lambda_{xy}}{\lambda_t} = \delta_{xy}\delta_{xt},$$

$1 \leq x \leq y \leq p, \quad t = 1, \ldots p.$

Now we collect our knowledge together

The $p(p+1)/2 \times p(p+1)/2$ Jacobian $\mathbf{J} = \mathbf{J}(\underline{\lambda}, \mathbf{q})$:

$$\mathbf{J} = \begin{pmatrix} \frac{\partial w_{jj}}{\partial \lambda_t} & \frac{\partial w_{jk}}{\partial \lambda_t} \\ \frac{\partial w_{jj}}{\partial q_z} & \frac{\partial w_{jk}}{\partial q_z} \end{pmatrix}$$

Rows: $t = 1, \ldots p, \quad z = 1, \ldots, p(p-1)/2.$
Columns: $j = 1, \ldots, p, \quad 1 \le j < k \le p \quad (\#: \ p + p(p-1)/2)$

$$\text{Auxiliary matrix} \quad \mathbf{V} := \begin{pmatrix} u_{jx} u_{jy} \\ 2u_{jx} u_{ky} \end{pmatrix}$$

Rows: $j = 1, \ldots, p, \quad 1 \le j < k \le p$
Columns: $1 \le x \le y \le p \quad (\#: \ p + p(p-1)/2)$

$$\mathbf{J}\cdot\mathbf{V} = \begin{pmatrix} \delta_{xy}\delta_{xt} \\ s_{xy}^{(z)}(\lambda_y - \lambda_x) \end{pmatrix}$$

Rows: $t = 1, \ldots p, \quad z = 1, \ldots, p(p-1)/2$
Columns: $1 \leq x \leq y \leq p$.

$$|\mathbf{J}|\cdot|\mathbf{V}| = |\mathbf{J}\cdot\mathbf{V}| = \prod_{x<y} |\lambda_y - \lambda_x| \cdot \begin{vmatrix} \delta_{xy}\delta_{xt} \\ s_{xy}^{(z)} \end{vmatrix}.$$

Since the right hand side determinant and $|\mathbf{V}|$ depends merely on $\mathbf{q}$:

$$|\mathbf{J}(\underline{\lambda}, \mathbf{q})| = \prod_{x<y} |\lambda_y - \lambda_x| \cdot f(\mathbf{q}).$$

Last step: we integrate with respect to the eigenvectors

The joint *pdf* of the eigenvalues and eigenvectors of **W**:

$$g(\underline{\lambda}, \mathbf{q}) = c_{np} (\prod_{j=1}^{p} \lambda_j)^{\frac{n-p-1}{2}} e^{-\frac{1}{2}\sum_{j=1}^{p} \lambda_j} \prod_{j<k} |\lambda_j - \lambda_k| \cdot f(\mathbf{q})$$

$\Longrightarrow$ The eigenvalues and eigenvectors of **W** are independent.
The *pdf* of $\mathbf{q}$: $cst \cdot f(\mathbf{q})$ (Haar).
The joint *pdf* of the eigenvalues of **W**:

$$\int g(\underline{\lambda}, \mathbf{q})\, d\mathbf{q} = \kappa_{np} (\prod_{j=1}^{p} \lambda_j)^{\frac{n-p-1}{2}} e^{-\frac{1}{2}\sum_{j=1}^{p} \lambda_j} \prod_{j<k} |\lambda_j - \lambda_k|$$

where $\kappa_{np} = c_{np}/cst$.