NOTES ON RANDOM MATRICES

LESSON 3: POSITIVE MATRICES AND STATISTICS

LUCA GUIDO MOLINARI

---

1. Positive matrices

A complex Hermitian matrix P is positive semi-definite ($P \geq 0$) if

$$\mathbf{z}^\dagger P \mathbf{z} \geq 0, \qquad \forall \mathbf{z} \in \mathbb{C}^n_0 \tag{1}$$

If the inequality is strict, the matrix is positive-definite, $P > 0$.
• Positive semi-definite matrices form a convex set:

$$P \geq 0,\ Q \geq 0 \implies tP + (1-t)Q \geq 0 \qquad \forall t \in [0,1]$$

• For $P \geq 0$ or $P > 0$, suitable choices of vectors $\mathbf{z}$ give the properties:
- the diagonal matrix elements of P are non-negative or positive,
- the eigenvalues of P are non-negative, or positive.
- the principal submatrices of P are positive semi-definite or definite,
- if $\mathbf{z}$ is zero except for $z_i = 1$ and $|z_j| = 1$, then for a suitable z_j:

$$|P_{ij}| \leq \tfrac{1}{2}(P_{ii} + P_{jj}) \tag{2}$$

• Sylvester's Inertia theorem states that the numbers of positive, zero and negative eigenvalues of a Hermitian matrix H and of $K^\dagger H K \geq 0$ are the same for any choice of invertible square matrix K. Then, if $P \geq 0$ it is $K^\dagger P K \geq 0$, and if $P > 0$ it is $K^\dagger P K > 0$.
• **Theorem** (see [21]): If $P \geq 0$ and $Q \geq 0$, there exists an invertible matrix K such that both $K^\dagger P K$ and $K^\dagger Q K$ are diagonal.
If $P > 0$, then K can be chosen so that $K^\dagger P K = 1$ and $K^\dagger Q K$ is diagonal.
• **Theorem** (I. Schur 1911): If $P \geq 0$ and $Q \geq 0$ then $P \circ Q \geq 0$ and $\det(P \circ Q) \geq (\det P)(\det Q)$, where the Hadamard product is $(P \circ Q)_{ij} = P_{ij} Q_{ij}$ (see [18]).
• **Proposition** If $X \in \mathbb{C}^{p \times n}$, then $XX^\dagger \geq 0$ and $X^\dagger X \geq 0$.
If $P \geq 0$ then $P = U \Lambda U^\dagger$ with $\Lambda \geq 0$ diagonal and U unitary. Then $P = XX^\dagger$ where $X = U\Lambda^{1/2}$. $P > 0$ iff $P = XX^\dagger$ for some invertible matrix.

Date: 18 jan 2018 – revised nov. 2022.

1.1. **Inequalities.** For positive definite matrices the famous Hadamard inequality holds (1893, [8]):

$$\det P \le P_{11} \cdots P_{nn} \tag{3}$$

The inequality was improved by Fischer (1907, [5]): if P is partitioned in 4 blocks, then $\det P \le \det P' \det P''$ where P' and P'' are the diagonal blocks. The iteration eventually terminates with Hadamard's statement. Thompson's inequality (1961, [20]) states that if P is partitioned in square blocks P_{ab}, then $\det P \le \det[\det P_{ab}]$. This inequality is due to Hermann Minkowski [4]:

$$\det(P+Q)^{\frac{1}{n}} \ge (\det P)^{\frac{1}{n}} + (\det Q)^{\frac{1}{n}} \tag{4}$$

with equality only if $Q = cP$, $c > 0$.

1.2. **Metric properties.** The set $\mathsf{P}(n,\mathbb{R})$ of positive-definite matrices is convex, and is an open subset in the space of Hermitian matrices with inner product $(A|B)_2 = \mathrm{tr}(AB)$ (see [3]). In other words, for every $P > 0$ there is a disk of matrices that are all positive:

$$\text{if } P > 0 \ \exists\, r > 0 \quad \text{s.t.} \quad \text{if } Q = Q^T \text{ and } \|Q - P\|_2 < r \ \Rightarrow\ Q > 0$$

$\mathsf{P}(n,\mathbb{R})$ is a symmetric space (a Riemannian manifold with a geodesic-reversing isometry at each point). It is a Riemannian manifold with local distance [3]

$$ds^2 = \mathrm{tr}(P^{-1}dPP^{-1}dP)$$

The line element is invariant under the action of $\mathsf{GL}(n,\mathbb{R})$, $P' = G^T PG$.
The distance of two elements is the length of the shortest path joining two matrices. This unique geodesic joining matrices P and Q is

$$\gamma(t) = P^{1/2}(P^{-1/2}QP^{-1/2})^T P^{1/2}, \qquad t \in [0,1]$$

The distance has the explicit expression:

$$d(P,Q) = \|\log(P^{-1/2}QP^{-1/2})\|_2 \tag{5}$$

The exponential map is continuous from the Hilbert space of symmetric matrices with the trace-norm?(check) to the metric space of positive definite matrices with distance d.

2. Parametrizations and invariant measures

2.1. **The invariant measure.** A real positive matrix P is specified by the independent matrix elements P_{ij} $i \le j$, that form a vector $\mathbf{P} \in \mathbb{R}^{\frac{1}{2}n(n+1)}$. A transformation $P' = K^T PK$ induces a linear transformation $\mathbf{P'} = \Omega_K \mathbf{P}$ that is a representation: $\Omega_{KH} = \Omega_K \Omega_H$. In particular, $\Omega_{K^{-1}} = \Omega_K^{-1}$. Let $K = S^{-1}\Lambda S$, where Λ is diagonal, then $\Omega_K = \Omega_S^{-1}\Omega_\Lambda \Omega_S$ and $\det \Omega_K = \det \Omega_\Lambda = (\det \Lambda)^{n+1} = (\det K)^{n+1}$. The Jacobian of the linear transformation $\mathbf{P'} = \Omega_K \mathbf{P}$ is: $\left|\partial \mathbf{P'}/\partial \mathbf{P}\right| = |\det K|^{n+1}$.

If $dP = \prod_{i\le j} dP_{ij}$, the invariant measure is

$$\boxed{d\mu_n(P) = \frac{dP}{(\det P)^{\frac{1}{2}(n+1)}}.} \tag{6}$$

It is $d\mu_n(G^T PG) = d\mu_n(P)$ for all $G \in \mathsf{GL}(n,\mathbb{R})$.

Example: if $Q > 0$, the following integral results with the change $P' = Q^{1/2}PQ^{1/2}$:

$$\int d\mu_n(P)(\det P)^\alpha e^{-\mathrm{tr}(PQ)} = \Gamma_n(\alpha)(\det Q)^{-\alpha} \tag{7}$$

The constant is evaluated in eq.(9).

2.2. **Triangular coordinates.** (Muirhead [13], theorems A9.7 and 2.1.9)
If $P \in \mathsf{P}(n, \mathbb{R})$ then there exists a unique upper triangular matrix T with positive diagonal elements such that $P = T^T T$ (proof by induction). Moreover,

$$d\mu_n(P) = 2^n \prod_{j=1}^n T_{jj}^{-j} \prod_{i\le k} dT_{jk} \tag{8}$$

Proof. We reproduce the proof in [13]. For $i \le j$ write $P_{ij} = \sum_{k\le j} T_{ki}T_{kj}$.
First row: it is $dP_{11} = 2T_{11}dT_{11}$; being $P_{1k} = T_{11}T_{1k}$, in taking the exterior product of differentials, the factor dT_{11} necessarily comes from dP_{11} and should not be repeated. Then, for the purpose of exterior product: $dP_{1k} = T_{11}dT_{1k}$. Then it is $\prod_k dP_{1k} = 2T_{11}^n \prod_k dT_{1k}$
Second row: $P_{22} = T_{12}^2 + T_{22}^2$ gives $dP_{22} = 2T_{22}dT_{22}$. Then $P_{2k} = T_{12}T_{1k} + T_{22}T_{2k}$ only provides the factor $dP_{2k} = T_{22}dT_{2k}$, $k = 2...n$; then $\prod_k dP_{2k} = 2T_{22}^{n-1}\prod_k dT_{2k}$.
And so on. Then:

$$dP = 2^n \prod_{j=1}^n T_{jj}^{n+1-j} \prod_{i\le k} dT_{jk}$$

Since $\det P = (\det T)^2 = \prod_j T_{jj}^2$, the factors T_{jj}^{n+1} simplify to give the invariant measure. $\square$

Example. With $\det P = (\det T)^2 = \prod_j T_{jj}^2$, $\mathrm{tr}P = \sum_{k\le j} T_{kj}^2$ let's evaluate

$$\begin{aligned}\Gamma_n(\alpha) = \int d\mu_n(P)(\det P)^\alpha e^{-\mathrm{tr}P} &= 2^n \prod_{j=1}^n \int_0^\infty dT_{jj} T_{jj}^{-j+2\alpha} e^{-T_{jj}^2} \prod_{i<j}\int_{-\infty}^{+\infty} dT_{ij} e^{-T_{ij}^2} \\ &= \pi^{\frac14 n(n-1)} \prod_{j=1}^n \Gamma(\alpha - \tfrac12 (j-1)). \end{aligned} \tag{9}$$

2.3. **Iwasawa parametrizations.** (from Terras [19]).
For $p + q = n$ the partial Iwasawa[1] block-factorizations of real positive matrices provide useful coordinates. The first one is

$$P = \begin{bmatrix} I_p & X^T \\ 0 & I_q \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & W \end{bmatrix} \begin{bmatrix} I_p & 0 \\ X & I_q \end{bmatrix} = \begin{bmatrix} V + X^T W X & X^T W \\ WX & W \end{bmatrix} \tag{10}$$

where $V \in \mathsf{P}(p, \mathbb{R})$, $W \in \mathsf{P}(q, \mathbb{R})$ and $X \in \mathbb{R}^{q\times p}$. The other factorization is

$$P = \begin{bmatrix} I_p & 0 \\ Y^T & I_q \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & W \end{bmatrix} \begin{bmatrix} I_p & Y \\ 0 & I_q \end{bmatrix} = \begin{bmatrix} V & VY \\ Y^T V & Y^T V Y + W \end{bmatrix} \tag{11}$$

where $Y \in \mathbb{R}^{p\times q}$. The correspondences are one-to-one with P. The invariant length takes the forms:

$$ds^2 = ds_V^2 + ds_W^2 + 2\mathrm{tr}[V^{-1}dX^T W dX] = ds_V^2 + ds_W^2 + 2\mathrm{tr}[dY W^{-1} dY^T V]$$

[1]Kenkichi Iwasawa, 1917-1998

The metric tensors are block-diagonal. In the first case the Haar measure is:

$$d\mu_n(P) = d_p\mu(V) d\mu_q(W)\, (\det V)^{-q/2} (\det W)^{p/2} dX \tag{12}$$

The blocks V, W may be further decomposed. The **full Iwasawa decomposition** is $P = n^T A^2 n$ where A is diagonal positive and $n \in \mathsf{N}$.
N is the group of upper triangular matrices with unit diagonal. Its invariant measure for left or right multiplication by elements in N is $dn = \prod_{i<j} dx_{ij}$.
Example:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ x_{12} & 1 & 0 \\ x_{13} & x_{23} & 1 \end{bmatrix} \begin{bmatrix} a_1^2 & 0 & 0 \\ 0 & a_2^2 & 0 \\ 0 & 0 & a_3^2 \end{bmatrix} \begin{bmatrix} 1 & x_{12} & x_{13} \\ 0 & 1 & x_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

In the full Iwasawa parametrization, the invariant measure is found to be:

$$\int d\mu_n(P) f(P) = 2^n \int_0^\infty \prod_{j=1}^n \frac{da_j}{a_j} a_j^{n-2j+1} \int_{-\infty}^{+\infty} \prod_{i<j} dx_{ij} f(n^T A^2 n) \tag{13}$$

2.4. **Spectral coordinates.** A real positive matrix factors as $P = R^T \Lambda R$, where $R \in \mathsf{SO}(n)$ and $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, where $\lambda_i > 0$ are the eigenvalues of P.
The invariant measure is

$$d\mu_n(P) = C_n\, dR \prod_{k=1}^n d\lambda_k \lambda_k^{-\frac{1}{2}(n+1)} \prod_{j<k} |\lambda_j - \lambda_k| \tag{14}$$

where $\int dR = 1$. The constant C_n is now obtained:

$$C_n = \pi^{\frac{1}{4}n(n+1)} \frac{1}{n!} \prod_{j=1}^n \frac{1}{\Gamma(j/2)}$$

Proof. Consider the integral

$$\int dP\, e^{-\mathrm{tr}\frac{1}{2}P} = C_n \prod_{k=1}^n \int_0^\infty d\lambda_k e^{-\frac{1}{2}\lambda_k} \prod_{j<k} |\lambda_j - \lambda_k|$$

The left-hand side is Gaussian: $2^{\frac{1}{2}n(n+1)} \pi^{\frac{1}{4}n(n-1)} \prod_{j=1}^n \Gamma(\frac{j+1}{2})$. The right-hand side is a form of Selberg's integral (Cor. 8.2.2 in [1] with $k = 0$):

$$\prod_{k=1}^n \int_0^\infty d\lambda_k e^{-\frac{1}{2}\lambda_k} \lambda_k^{\alpha-1} \prod_{j<k} |\lambda_j - \lambda_k| = 2^{\frac{1}{2}n(n+1)} \prod_{j=1}^n \frac{\Gamma(\alpha + \frac{j-1}{2}) \Gamma(1 + \frac{j}{2})}{\Gamma(\frac{3}{2})}$$

The comparison of the integrals gives C_n. □

- If $P > 0$ and $P' = P^{-1}$ then $P' > 0$ and $d\mu_n(P') = d\mu_n(P)$.

3. Integrals on positive matrices

Various integrals of real positive variable extend to integrals on positive matrices [14, 15, 19]. We mention a few.
- Multivariate Gamma function (Albert E. Ingham [11])

$$\Gamma_n(\alpha) = \int_{P>0} dP (\det P)^{\alpha - \frac{1}{2}(n+1)} e^{-\mathrm{tr} P} = \pi^{\frac{1}{4}n(n-1)} \prod_{k=1}^n \Gamma(\alpha - \tfrac{k-1}{2}) \tag{15}$$

We already evaluated the extension with source $Q>0$:

$$\boxed{\int_{P>0} d\mu_n(P)(\det P)^\alpha e^{-\mathrm{tr}(PQ)} = \Gamma_n(\alpha)(\det Q)^{-\alpha}} \tag{16}$$

An application to the evaluation of moments of negative powers of GUE determinants is found in [7] (with a connection to Riemann's zeta function).

• Multivariate Beta function.

Multiply (16) by $(\det Q)^{\alpha+\beta} e^{-\mathrm{tr}(QH)}\, d\mu_n(Q)$ and integrate both sides:

$$\Gamma_n(\alpha+\beta)\int_{P>0} d\mu_n(P)(\det P)^\alpha [\det(P+H)]^{-\alpha-\beta} = \Gamma_n(\alpha)\Gamma_n(\beta)(\det H)^{-\beta} \tag{17}$$

The choice $H=1$ gives an identity by Carl Ludwig Siegel [17]:

$$\boxed{\int_{P>0} d\mu_n(P)\,(\det P)^\alpha \det(1+P)^{-(\alpha+\beta)} = \frac{\Gamma_n(\alpha)\Gamma_n(\beta)}{\Gamma_n(\alpha+\beta)}} \tag{18}$$

The result is the multivariate Beta function, which is defined by the integral

$$B_n(\alpha,\beta) = \int_{0<P<I} d\mu_n(P)\,(\det P)^\alpha \det(1-P)^{\beta-\frac{1}{2}(n+1)} = \frac{\Gamma_n(\alpha)\Gamma_n(\beta)}{\Gamma_n(\alpha+\beta)} \tag{19}$$

that, basically, is Selberg's integral $S_n(\alpha-\frac{1}{2}(n+1), \beta-\frac{1}{2}(n+1), \frac{1}{2})$ (1944) [1]:

$$\begin{aligned} S_N(\mu,\nu,\lambda) &= \prod_{j=1}^N \int_0^1 dx_j x_j^\mu (1-x_j)^\nu |\Delta(\mathbf{x})|^{2\lambda} \\ &= \prod_{j=1}^N \frac{\Gamma(\mu+1+j\lambda)\Gamma(\nu+1+j\lambda)\Gamma(\lambda+1+j\lambda)}{\Gamma(\mu+\nu+2+(N+j-1)\lambda)\Gamma(1+\lambda)} \end{aligned} \tag{20}$$

Eq.(18) may be obtained from (19) with the change $P' = P(1+P)^{-1}$. Then $I > P' > 0$ and $dP = dP'(\det P')^{-n-1}$.

• The choice $H = 1+\epsilon K$ in (17) and expansion in ϵ gives:

$$\int_{P>0} d\mu_n(P)(\det P)^\alpha [\det(P+1)]^{-\alpha-\beta} \mathrm{tr}[(P+1)^{-1}K] = B_n(\alpha,\beta)\frac{\beta}{\alpha+\beta}\mathrm{tr}K \tag{21}$$

• Matrix confluent hypergeometric functions of the I and II kind ([19], p.68)

$$\Phi_n(a,c,Q) = \frac{\Gamma_n(c)}{\Gamma_n(a)\Gamma_n(c-a)}\int_{0<P<I} d\mu_n(P)\, e^{\mathrm{tr}(PQ)}(\det P)^a [\det(1-P)]^{c-a-\frac{1}{2}(n+1)}$$

The function of II kind was introduced by Muirhead ([13], p.472)

$$\Psi_n(a,c,Q) = \frac{1}{\Gamma_n(a)}\int_{P>0} d\mu_n(P)\, e^{-\mathrm{tr}(PQ)}(\det P)^a [\det(1+P)]^{c-a-\frac{1}{2}(n+1)} \tag{22}$$

• Multivariate Gamma function The matrix Gamma function $\Gamma_n(\alpha)$ has a generalization, by Selberg. Define the (complex) power of $P \in \mathsf{P}(n,\mathbb{R})$

$$P_{\mathbf{z}}(P) = \prod_{k=1}^n (\det P_k)^{z_i}, \quad \mathbf{z}\in\mathbb{C}^n \tag{23}$$

where P_k is the matrix $k\times k$ of the first k rows and columns of P.
The multivariate Gamma function is

$$\Gamma_n(\mathbf{z}) = \int_{P>0} d\mu_n(P) P_{\mathbf{z}}(P) e^{-\mathrm{tr} P} = \pi^{\frac{1}{4}n(n-1)} \prod_{k=1}^{n} \Gamma(z_{n-k+1} + \cdots + z_n - \tfrac{k-1}{2}) \tag{24}$$

The function $\Gamma_n(\alpha)$ in eq.(9) corresponds to the choice $\mathbf{z} = (0, ..., 0, \alpha)$.

• Matrix K-Bessel function (Bengtson [2])

If $A > 0$ and $B > 0$ then

$$K_n(\mathbf{z}|A, B) = \int_{P>0} d\mu_n(P) P_{\mathbf{z}}(P)\, e^{-\mathrm{tr}(AP+BP^{-1})} \tag{25}$$

Bengtson proved the correspondence:

$$\int_{\mathbb{R}^{m\times n}} dX P_{\mathbf{z}}((A + X^T X)^{-1})\, e^{2i\mathrm{tr}(RX)} = \pi^{\frac{mn}{2}} K_n(\mathbf{z}'|A, R^T R) \tag{26}$$

where $A > 0$, $R \in \mathbb{R}^{m\times n}$, $\mathbf{z}' = \mathbf{z} + (0, 0, ..., 0, -\frac{m}{2})$.
In particular,

$$\int_{\mathbb{R}^{m\times n}} dX \frac{e^{2i\mathrm{tr}(RX)}}{\det(A + X^T X)^\alpha} = \pi^{\frac{mn}{2}} \int_{P>0} d\mu_n(P)\, \frac{e^{-\mathrm{tr}(AP+R^T RP^{-1})}}{(\det P)^{\frac{m}{2}-\alpha}} \tag{27}$$

• Matrix Laplace transform If $f : \mathsf{P}(n, \mathbb{R}) \to \mathbb{C}$, the Laplace transform of f at a symmetric complex matrix $X = U + iV$ is

$$(\mathscr{L} f)(X) = \int_{P>0} dP\, e^{-\mathrm{tr}(PX)} f(P) \tag{28}$$

that converges for $U > U_0 \in \mathsf{P}(n, \mathbb{R})$. The inversion formula is:

$$f(P) = (2\pi i)^{-\frac{1}{2}n(n+1)} \int_{X=X^T, U>U_0} dV\, e^{\mathrm{tr}(P(U+iV))} (\mathscr{L} f)(U + iV) \tag{29}$$

---

4. Real rectangular random matrices

$\mathbb{R}^{n\times p}$, $n \geq p$, is the set of rectangular real matrices $n \times p$ of rank p. The set is invariant for left and right multiplication by invertible square matrices in $\mathbb{R}^{n\times n}$ and $\mathbb{R}^{p\times p}$. If $M' = MB$ then $M'_{ia} = \delta_{ki} B_{ba} M_{kb}$, $\partial M'/\partial M = I \otimes B$. The Jacobian of the linear transformation is $|J| = \det(I \otimes B) = (\det B)^n$. With $dM = \prod_{ia} dM_{ia}$, the right-invariant measure is

$$d\mu^R_{n,p}(M) = \frac{dM}{(\det M^T M)^{\frac{n}{2}}}$$

It is $d\mu^R_{n,p}(MB) = \frac{(\det B)^n dM}{(\det B^T M^T MB)^{n/2}} = d\mu^R_{n,p}(M)$.

4.1. **Polar decomposition.** Let $M = [\mathbf{m}_1, \dots, \mathbf{m}_p]$, where the columns $\mathbf{m}_k$ are linearly independent vectors in $\mathbb{R}^n$. Right multiplication by a $p \times p$ matrix G^{-1} is MG^{-1} whose columns are linear combinations of the columns of M. Choose G^{-1} such that the columns in $MG^{-1} = V$ are orthonormal.
• $\mathbb{V}_{n,p} = \{V \in \mathbb{R}^{n\times p},\ V^T V = I_p\}$ is the Stiefel[2] manifold of orthonormal p-frames in $\mathbb{R}^n$, i.e. the p columns of V are orthonormal vectors in $\mathbb{R}^n$. The set is invariant for left multiplication by $\mathsf{O}(n)$ and right multiplication by $\mathsf{O}(p)$. $\mathbb{V}_{n,n}$ is the orthogonal group $\mathsf{O}(n)$.
• Now it is $M = VG$ with $M^T M = G^T G = P^2$, with $P \in \mathsf{P}(p, \mathbb{R})$. Then G has polar decomposition $G = RP$ for some $R \in \mathsf{O}(p)$. Since $VR \in \mathbb{V}_{n,p}$, we obtain the polar representation of rectangular matrices:

$$\boxed{M = VP\ , \qquad V \in \mathbb{V}_{n,p}\ ,\ P > 0} \tag{30}$$

• If $M \in \mathbb{R}^{n\times p}$, then M has the unique representation $M = VT$ where $V \in \mathbb{V}_{n,p}$ and T is upper triangular with positive diagonal ([13] Thrm A9.8).
The measure is ([13] thrm 2.1.13):

$$dM = \prod_{j=1}^{p} T_{jj}^{n-j} dT dV$$

4.2. **Joint probability measure.** The integration measure that corresponds to $M = VP^{1/2}$ is (Lemma 1.4 in [9]):

$$\boxed{d\mu^R_{n,p}(M) = 2^{-p} dV d\mu_p(P)}$$

The evaluation $\int dM e^{-\mathrm{tr}(M^T M)} = 2^{-p} \int dV \int_{P>0} d\mu_p(P) (\det P)^{n/2} e^{-\mathrm{tr} P}$ gives the normalization

$$\int_{\mathbb{V}_{n,p}} dV = \sigma_{n,p} = \frac{2^p \pi^{np/2}}{\Gamma_p(n/2)}$$

4.3. **Rectangular integrals.**
• $\int dM e^{-\mathrm{tr}(M^T M Q)} = \pi^{\frac{1}{2}np} (\det Q)^{-n/2}$
• $\int dM \delta(Q - M^T M) = 2^{-p} \sigma_{n,p} (\det Q)^{\frac{1}{2}(n-p-1)}$ where $\delta(P) = \prod_{i\ge j} \delta(P_{ij})$

5. Wishart random matrices

Multivariate Normal. For an historical introduction read [16].
The extension of the normal distribution $N(\mu, \sigma^2)$ to p variables is the multivariate normal distribution.
A real random vector $\mathbf{x} = \{x_1, \dots, x_p\}$ has multivariate normal distribution with mean $\boldsymbol{\mu} = \{\mu_1, ..., \mu_p\}$ and covariance matrix $\Sigma \in \mathbb{R}^{p\times p}$, i.e. $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \Sigma)$, if the joint probability density function of $x_1, ..., x_p$ is

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}\sqrt{\det \Sigma}} \exp\left[-\tfrac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right] \tag{31}$$

It is $\mathbb{E}[x_k] = \mu_k$ and $\mathbb{E}[(x_j - \mu_j)(x_k - \mu_k)] = \Sigma_{jk}$ [16]. The distribution (marginal) of a subset of $\mathbf{x}$, say $\{x_1, ..., x_k\}$, is multivariate normal with covariance matrix given by the rows and columns $1, ..., k$ of Σ.

[2]Eduard Stiefel, 1909–1978

Given n independent random variables with identical distribution $N[0,\sigma^2]$, the sum of their squares has a $\chi^2_n[\sigma]$ distribution. In analogy, consider n independent random vectors $\mathbf{x}_1 \dots \mathbf{x}_n$ in $\mathbb{R}^p$, with same distribution $N_p(\mathbf{0},\Sigma)$. The distribution of the sum $\sum_{j=1}^n \mathbf{x}_j\mathbf{x}_j^T$ is named Wishart distribution[3]. The sum can be rewritten as XX^T, where X is the $p\times n$ matrix $[\mathbf{x}_1,...,\mathbf{x}_n]$. The $p\times p$ matrix $P=XX^T$ is positive.

5.1. **Wishart distribution (1928).** The Wishart distribution is a multivariate generalization of the χ^2 distribution.
A real $p\times p$ random matrix $P>0$ has the (central) Wishart distribution with n degrees of freedom ($n\ge p$) and covariance matrix Σ, i.e. $P\sim W_p(n,\Sigma)$, if its probability density function is

$$p(P) = \frac{1}{\Gamma_p(n/2)(2^p\det\Sigma)^{n/2}}(\det P)^{\frac12(n-p-1)}\exp[-\tfrac12\mathrm{tr}(\Sigma^{-1}P)] \tag{32}$$

with respect to Lebesque measure on the cone of symmetric positive definite matrices. $\Gamma_p(a)$ is the multivariate Gamma function.
If $P\in W_1(n;\sigma^2)$, then $P\in\chi^2_n[\sigma]$.
The normalization in (32) is Ingham's integral (16) with $Q=\frac12\Sigma^{-1}\in\mathsf{P}(p,\mathbb{R})$.

5.2. **Sample statistics.** Let $\mathbf{x}_1,\dots,\mathbf{x}_n$ be a sample of n random vectors in $\mathbb{R}^p$ taken from a distribution $N_p[\mu,\sigma^2]$. The sample mean and sample variance are:

$$\bar{\mathbf{x}} = \frac1n\sum_{k=1}^n \mathbf{x}_k, \qquad \bar S = \frac{1}{n-1}\sum_{k=1}^n(\mathbf{x_k}-\bar{\mathbf{x}})(\mathbf{x_k}-\bar{\mathbf{x}})^T$$

Proposition: the random variables $\bar{\mathbf{x}}$ and $\bar S$ are independent and

$$\sqrt n(\bar{\mathbf{x}}-\mu)\sim N_p[0,\Sigma], \qquad (n-1)\bar S\sim W_p(n-1,\Sigma)$$

Thus a random draw from the Wishart distribution is some matrix that, upon rescaling, is a covariance matrix for multivariate normal data.

6. The Laguerre ensemble

If $\Sigma_{ij}=\sigma^2\delta_{ij}$, the matrix elements of $X\in\mathbb{R}^{p\times n}$ are independent identically distributed (i.i.d.) random variables with distribution $N(0,\sigma^2)$. The Wishart distribution of the $p\times p$ matrix $P=XX^T$ simplifies,

$$p(P) = \frac{1}{(2^{p/2}\sigma)^n\Gamma_p(n/2)}(\det P)^{\frac12(n-p-1)}\exp[-\tfrac{1}{2\sigma^2}\mathrm{tr}P]$$

It only depends on the eigenvalues of P, that have joint distribution

$$p(\lambda_1,\dots,\lambda_p) = C_{n,p,\sigma}|\Delta(\lambda)|\prod_{j=1}^p\lambda_j^{\frac12(n-p-1)}\exp(-\tfrac{1}{2\sigma^2}\lambda_j) \tag{33}$$

The positive-definite matrices with this distribution form the Laguerre ensemble.

[3]John Wishart (Perth 1898, Acapulco 1956) was introduced by the mathematician Edmund Whittaker to the influential statistician Karl Pearson, at University College in London. As PhD student, he was engaged in the preparation of the numerical tables for the incomplete Gamma function. Years later, he became director of the Statistical Laboratory in Cambridge. He was editor of the important journal Biometrika, created in 1901 by K. Pearson, W.F.D. Welton and C. Davenport.

6.1. **Marcenko-Pastur distribution (1967).** [12]
In the large n and p limits, with finite ratio $r = p/n$, the eigenvalues $\tau_1 \dots \tau_p$ of the real positive square matrix $P = \frac{1}{n} X X^T$ have the distribution with support $[a, b]$

$$\boxed{\rho(\tau) = \frac{1}{2\pi r \sigma^2} \frac{\sqrt{(b-\tau)(\tau-a)}}{\tau}} \tag{34}$$

where $a = \sigma^2(1-\sqrt{r})^2$, $b = \sigma^2(1+\sqrt{r})^2$.

Proof. The partition function for the eigenvalues of positive matrices described by the probability (33) is $Z = \int d\lambda_1 ... d\lambda_p \exp[-p^2 S]$ with

$$S = \frac{1}{p}\sum_{k=1}^{p}\left[\frac{1}{2\sigma^2}\frac{\lambda_k}{p} - \frac{1}{2}\left(\frac{n}{p} - 1 - \frac{1}{p}\right)\log\lambda_k\right] - \frac{1}{2p^2}\sum_{j\neq k}\log|\lambda_j - \lambda_k|$$

With the rescaling $\tau_k = \lambda_k/n$, in the limit of large n, p with fixed $r = p/n$, the eigenvalues are described by a density $\rho(\tau)$. Up to a constant, it is:

$$S[\rho] = \int d\tau \rho(\tau)\left[\frac{\tau}{2r\sigma^2} - \left(\frac{1}{2r} - \frac{1}{2}\right)\log\tau\right] - \frac{1}{2}\iint d\tau d\tau' \rho(\tau)\rho(\tau')\log|\tau - \tau'|$$

In the saddle point approximation the spectral density is an extremum. The constraint of normalization is implemented by the additional term $-\mu[\int d\tau\rho(\tau) - 1]$. The extremum solves: $\tau/2r\sigma^2 - (1/2r - 1/2)\log\tau - \mu = \int d\tau' \rho(\tau')\log|\tau - \tau'|$. A derivative in τ, and a multiplication by τ give:

$$\fint_a^b d\tau' \frac{\tau'\rho(\tau')}{\tau - \tau'} = \frac{\tau}{2r\sigma^2} - \frac{1}{2r} - \frac{1}{2}, \quad a \leq \tau \leq b.$$

The solution is $\rho(\tau) = C\sqrt{(b-\tau)(\tau-a)}/\tau$ with normalization constant

$$C = \frac{2}{\pi}\frac{1}{(\sqrt{b} - \sqrt{a})^2}$$

The equations for a and b are:

$$C\int_a^b d\lambda \sqrt{\frac{\lambda - a}{b - \lambda}} = \frac{b}{2r\sigma^2} - \frac{1}{2r} - \frac{1}{2}, \quad C\int_a^b d\lambda\sqrt{\frac{b-\lambda}{\lambda - a}} = -\frac{a}{2\sigma^2} + \frac{1}{2r} + \frac{1}{2}$$

The two integrals are equal, and give: $C\frac{\pi}{2}(b-a) = \frac{b}{2r\sigma^2} - \frac{1}{2r} - \frac{1}{2} = -\frac{a}{2r\sigma^2} + \frac{1}{2r} + \frac{1}{2}$ i.e. $2rC\pi\sigma^2 = 1$ and $b + a = 2\sigma^2(1+r)$. $\square$

6.2. **The complex case.** Let $P = X^\dagger X$ where $X \in \mathbb{C}^{n\times m}$, $n \geq m$ with independent Gaussian elements. The joint distribution of eigenvalues of P is:

$$p(\lambda_1, \dots, \lambda_m) = \frac{1}{C_{m,n}}\Delta(\lambda)^2 \prod_{j=1}^m e^{-\lambda_j/a}\lambda_j^{n-m} \tag{35}$$

$$C_{m,n} = a^{mn}\prod_{j=1}^m (n-j)!(m-j)! \tag{36}$$

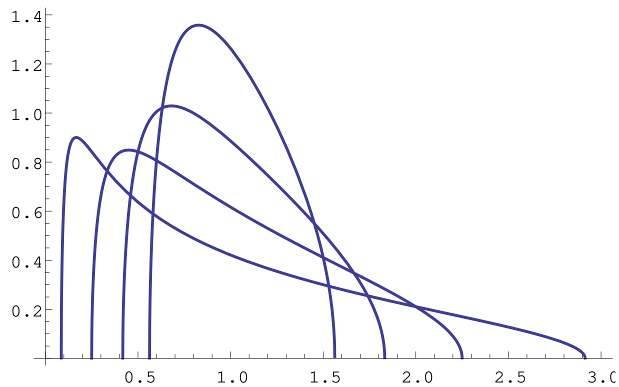

FIGURE 1. The Marcenko-Pastur distributions (34) for ratios $r = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$ (inner curve for smaller r) and $\sigma^2 = 1$.

The orthogonal polynomials associated to this probability are the associated Laguerre polynomials:

$$\int_0^\infty dx\, e^{-x} x^\ell L_j^\ell(x) L_k^\ell(x) = \frac{(\ell+k)!}{k!}\delta_{jk} \tag{37}$$

$$(k+1)L_{k+1}^\ell(x) + (x-1-2k-\ell)L_k^\ell(x) + (k+\ell)L_{k-1}^\ell(x) = 0 \tag{38}$$

$$\begin{aligned} K_n(x,y) &= \sum_{j=0}^n \frac{j!}{(j+\ell)!} L_j^\ell(x) L_j^\ell(y) \\ &= \frac{(n+1)!}{(n+\ell)!} \frac{L_n^\ell(x)L_{n+1}^\ell(y) - L_n^\ell(y)L_{n+1}^\ell(x)}{x-y} \end{aligned} \tag{39}$$

7. Spherical distributions

The simple assumption of normality is not always appropriate. A generalization is provided by spherical distributions of random vectors, studied by Isaac Jacob Schoenberg (1938). In general, a *spherical distribution* of a vector $\mathbf{x}$ is characterized by $p(\mathbf{x}) = p(O\mathbf{x})$, for the orthogonal group. The multivariate distribution with $\Sigma_{ij} = \sigma^2\delta_{ij}$ is spherically symmetric.

7.1. **Spherical coordinates in $\mathbb{R}^d$.** Let $\mathbf{x} = r\mathbf{n}$, where $r = \|\mathbf{x}\|$ and $\mathbf{n}$ belongs to the unit surface S^{d-1} in $\mathbb{R}^d$. In polar angles the components are:

$$\begin{aligned} n_1 &= \sin\theta_1 \sin\theta_2 \dots \sin\theta_{d-2} \sin\theta_{d-1} \\ n_2 &= \sin\theta_1 \sin\theta_2 \dots \sin\theta_{d-2} \cos\theta_{d-1} \\ n_3 &= \sin\theta_1 \sin\theta_2 \dots \cos\theta_{d-2} \\ \dots & \quad \dots \\ n_{d-1} &= \sin\theta_1 \cos\theta_2 \\ n_d &= \cos\theta_1 \end{aligned}$$

where $\theta_1 \in [0, 2\pi]$ and $\theta_2, ..., \theta_{d-1} \in [0, \pi]$. The volume element is $d\mathbf{x} = r^{d-1} dr d\mathbf{n}$ with $d\mathbf{n} = (\sin^{d-2}\theta_1\, d\theta_1)(\sin^{d-3}\theta_2\, d\theta_2)\dots(\sin\theta_{d-2}\, d\theta_{d-2})d\theta_{d-1}$.

The area of the surface S^{d-1} is

$$\omega_d = 2\pi \prod_{k=1}^{d-2} \int_0^\pi \sin\theta^k d\theta = 2\pi \prod_{k=1}^{d-2} \frac{\sqrt{\pi}\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k+2}{2})} = \frac{2\pi^{d/2}}{\Gamma(d/2)}$$

Let $\mathbf{n}$ be uniformly distributed in S^{d-1}. Because of the rotational invariance, the characteristic function $\mathbb{E}(e^{i\mathbf{t}\cdot\mathbf{n}})$ is a function $\Omega_d(t)$. To evaluate the function, we may choose $\mathbf{t}$ in the direction d. Then:

$$\Omega_d(t) = \frac{1}{\omega_d}\int d\mathbf{n}\, e^{itn_d} = \frac{\int_0^{2\pi} d\theta_1 e^{it\cos\theta_1}(\sin\theta_1)^{d-2}}{\int_0^{2\pi} d\theta_1 (\sin\theta_1)^{d-2}} = {}_0F_1(\tfrac{1}{2}d, -\tfrac{1}{4}t^2)$$

Note the large d limit: $\Omega_d(t\sqrt{d}) \to \exp(-t^2/2)$, uniformly in x.

7.2. **Borel's lemma.** Let $\mathbf{n}$ be a random point of the unit surface in $\mathbb{R}^d$. Then,

$$P(\sqrt{d}\, n_1 \le t) \to \int_{-\infty}^T \frac{dx}{\sqrt{2\pi}} e^{-x^2/2} \quad (d\to\infty) \tag{40}$$

With the same hypothesis:

$$\mathbb{E}[|n_1|^{\alpha_1}\dots|n_n|^{\alpha_n}] = \Gamma(\frac{n}{2}) n^{\frac{1}{2}\sum\alpha_k} \frac{\prod_k \Gamma(\frac{\alpha_k+1}{2})}{\Gamma(\sum_k \frac{\alpha_k+1}{2})} \tag{41}$$

A random vector has spherical distribution with characteristic function $\Phi(t)$ if and only if

$$\Phi(t) = \int d\mathbf{x} p(\mathbf{x}) \Omega_p(xt)$$

Theorem 7.1 (Schoenberg, 1938)**.**

The following integrals were proven in [7]. Let Q, H be Hermitian matrices of size n, $Q>0$, $\mathrm{Im}\mu > 0$, $N > n$ then;

$$\int dH \frac{\exp i\mathrm{tr}(HQ)}{\det(H-\mu)^N} = C(\det Q)^{N-n} e^{i\mu \mathrm{tr} Q} \tag{42}$$

$$C = \frac{i^{n^2+N-n}(2\pi)^{\frac{1}{2}n(n+1)}}{\Gamma(N)\Gamma(N-1)\dots\Gamma(N-n+1)}$$

For real symmetric matrices, $Q>0$ and $N \ge \frac{1}{2}(n+1)$:

$$\int dS \frac{\exp i\mathrm{tr}(SQ)}{\det(S-\mu)^N} = C(\det Q)^{N-n} e^{i\mu \mathrm{tr} Q} \tag{43}$$

$$C = \frac{2^n i^{N+\frac{1}{2}(n^2-1)} \pi^{\frac{1}{4}n(n+3)}}{\Gamma(N)\Gamma(N-1)\dots\Gamma(N-\frac{n-1}{2})}$$

The integrals are zero if we allow for a negative eigenvalue of Q.

7.3. **Bernstein identity.** Let $\Delta = \det(\partial^T\partial)$, where $\partial = (\partial^2_{ij})_{n\times m}$, then:

$$\Delta(\det X^T X)^{\lambda/2} = B(\lambda)(\det X^T X)^{\frac{1}{2}\lambda - 1} \tag{44}$$

$$B(\lambda) = (-1)^m \prod_{i=0}^{m-1} (\lambda+i)(2-n-\lambda+i) \tag{45}$$

References

[1] G. E. Andrews, R. Askey and R. Roy, Special functions, Encyclopedia of Mathematics and its Applications 71, Cambridge University Press (1999).
[2] Thomas E. Bengtson, *Bessel functions on P_n*, Pacific Journal of Mathematics **108** (1)(1983) 19–29.
[3] Rajendra Bhatia, Positive definite matrices, Princeton University Press, 2007.
[4] Rajendra Bathia, Matrix analysis, GTM 169, Springer, 1997.
[5] Ernst Fischer, *Über den Hadamardschen Determinentsatz*, Arch. Math. U. Phys. **13** n.3 (1907) 32–40.
[6] Peter J. Forrester, Log-gases and Random Matrices, London Mathematical Society Monograph, 2010 Princeton University Press.
[7] Yan V. Fyodorov, *Negative moments of characteristic polynomials of random matrices: Ingham-Siegel integral as an alternative to Hubbard-Stratonovich transformation*, Nucl. Phys. B **621** [PM] (2002) 643–674.
[8] Jacques Hadamard, *Résolution d'une question relative aux déterminant*, Bull. Sci. Math. **17** n.2 (1893) 240–246.
[9] Carl S. Herz, *Bessel functions of matrix argument*, Annals of Mathematics **61** (3) (1955) 474–523. https://www.jstor.org/stable/1969810.
[10] Plamen Koev and Alan Edelman, *The efficient evaluation of the hypergeometric function of a matrix argument*, Mathematics of Computation **75** (254) (2006) 833–846. https://math.mit.edu/~plamen/files/hyper.pdf
[11] Albert Edward Ingham, *An integral which occurs in statistics*, Proc. Cambridge Philos. Soc. **29** (1933) 271–276. https://doi.org/10.1017/S0305004100011075.
[12] V. A. Marcenko and L. A. Pastur, *Distribution of eigenvalues for some sets of random matrices*, Mathematics of the USSR-Sbornik **1** n.4 (1967) 457–483. https://doi.org/10.1070/SM1967v001n04ABEH001994
[13] Robb J. Muirhead, Aspects of Multivariate Statistics, John Wiley 1982.
[14] D. K. Nagara, A. Roldán-Correa, A. K. Gupta, *Extended matrix variate gamma and beta functions*, J. Multivariate Analysis **122** (2013) 53–69. https://doi.org/10.1016/j.jmva.2013.07.001
[15] D. St. P. Richards, *Functions of Matrix Argument*, Ch.35 in NIST Handbook of Mathematical Functions, 2010. https://dlmf.nist.gov/35
[16] C. Radhakrishna Rao, *Multivariate Analysis: Some Reminiscences on Its Origin and Development*, The Indian Journal of Statistics, Series B (1960-2002) **45** n.2 (1983) pp. 284–299. https://www.jstor.org/stable/25052296.
[17] Carl Ludwig Siegel, Über die analytische theorie der quadratischen formen, Ann. of Math. **36** (3) (1935) 527–606.
[18] George P. H. Styan, *Hadamard products and Multivariate statistical analysis*, Linear Algebra and its Applications **6** (1973) 217–240.
[19] Audrey Terras, Harmonic analysis on symmetric spaces - higher rank spaces, positive definite matrix space and generalizations, 2nd Ed. Springer, 2016.
[20] R. C. Thompson, *A determinantal inequality for positive definite matrices*, Canadian Math. Bull. **4** n.1 (1961) 57–62.
[21] Fuzhen Zhang, Matrix theory, basic results and techniques, Springer (1991).