SAMsi

Johnstone’s introductory talk

Three groups of attendees

Statistics
Applications
Mathematics

I will oscillate between the three groups

Brief comments on some earlier lectures
Zeitouni’s tutorial

\( O(p) \): The group of \( p \times p \) orthogonal matrices

\( H \): A generic member of \( G \)

\( dH \): normalized Haar measure on \( O(p) \)

\( A \) and \( B \): symmetric \( p \times p \) matrices

\[ \int_{O(p)} \exp \left( \text{tr} \, AH BH^* \right) \, dH \]

The integral appears in the density of the eigenvalues of a Wishart matrix
The zonal polynomial expansion

\[
\int_{O(p)} \exp\left( \text{tr} \, AHBH^* \right) \, dH = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{|\tau|=k} \frac{C_{\tau}(A) \, C_{\tau}(B)}{C_{\tau}(I)}
\]

\(C_{\tau}(A)\): No “explicit” formula

WLOG, \( A = \text{diag}(a_1, \ldots, a_p) \) and \( B = \text{diag}(b_1, \ldots, b_p) \)

For \( p = 2 \), an explicit formula:

\[
\int_{O(2)} \exp\left( \text{tr} \, AHBH^* \right) \, dH = c \, \exp(\cdots) \, I_\nu(|a_1 - a_2| \cdot |b_1 - b_2|)
\]

\(I_\nu\) is a modified Bessel function
Eichinger (1980’s): polymer chemistry, Gaussian macromolecules


Open Problem: Generalize Eichinger’s formula to general $p$

The integral is a generalized Bessel function: Opdam, Heckman
$U(p)$: Similar expansion, but with different zonal polynomials

$C^\tau(A)$ is a multiple of the well-known Schur function, $s^\tau(A)$

$\tau = (\tau_1, \ldots, \tau_p)$, $\tau_1 \geq \cdots \geq \tau_p \geq 0$

$A$ is Hermitian, $A = \text{diag}(a_1, \ldots, a_p)$

Vandermonde: $V(A) = \prod_{i<j}(a_i - a_j)$

$$s^\tau(A) = \frac{\det (a_i^{\tau_j+p-j})}{V(A)}$$
\[
\int_{U(p)} \exp(\text{tr } ABH^*) \ dH \\
= \frac{c}{V(A) V(B)} \sum_{\tau_1 \geq \cdots \geq \tau_p \geq 0} \det(a^{\tau_j+p-j}) \det(b^{\tau_j+p-j}) \prod_{j=1}^{p} c_{\tau_j+p-j}
\]

Apply Cauchy-Binet/Andréief/Pólya-Szegö Formula:

\[
\cdots \int \det(f_i(x_j)) \det(g_i(x_j)) \prod_{j=1}^{p} d\mu(x_j) \\
= p! \det(\int f_i(x) g_j(x) d\mu(x))
\]
\[
\int_{U(p)} \exp \left( \text{tr} \ A H B H^* \right) \, dH = c \frac{\det (e^{a_i b_j})}{V(A)V(B)}
\]

Hua Loo-Keng, “Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains,” AMS Transl., 1963.

C. G. Khatri (1970), Sankhyā: A statistician!

Harish-Chandra-Itzykson-Zuber formula

2-D normal p.d.f., $f(x, y) \propto \exp(-x^2 - y^2) \exp(xy)$,

For $x_j, y_j \in \mathbb{R}, j = 1, \ldots, p$, when is $\det(f(x_i, y_j)) \geq 0$? Karlin, Lehmann, Perlman-Olkin, ...

Let $X = \text{diag}(x_i), Y = \text{diag}(y_i)$. Then

$$\det(e^{\rho x_i y_j}) = c V(X)V(Y) \int_{U(p)} \exp(\text{tr } XHYH^*) \, dH$$

Conclude: $\det(f(x_i, y_j)) \geq 0$ if the $x_i$ and $y_i$ are similarly ordered

Gross and D.R., JAT (1995): Many types of total positivity, related to compact Lie groups

Open Area: Develop a theory of variation-diminishing transformations for these theories of total positivity
Ingram Olkin: I’ve heard that he knows \( N \gg 8 \) derivations of the Wishart distribution

I myself know four

The formula for the multivariate gamma function,

\[
\Gamma_p(a) = \pi^{p(p-1)/4} \prod_{j=1}^{p} \Gamma(a - \frac{1}{2}(j - 1)),
\]

is implicit in Wishart’s derivation of his distribution

Ingham (1939?) read Wishart’s paper and noted the importance of Wishart’s formula for $\Gamma_p(a)$ for analytic number theory.

C. L. Siegel (1940?) gave an independent calculation of $\Gamma_p(a)$.

The Ingham-Siegel formula.


A remarkable calculation of $\Gamma_p(a)$ (and more) using integration-by-parts on the space of positive definite matrices, treated as a Riemannian manifold.
Back to applications

In all noncentral multivariate problems we find the “troublesome” factors, the hypergeometric functions of matrix argument

Perlman and Olkin (1980), Bondar (1980’s)

When is
\[
\frac{\partial^2}{\partial a_1 \partial a_2} \log V(A)V(B) \int_{O(p)} \exp(\text{tr } AHBH^*) \, dH \geq 0?
\]

How about
\[
\frac{\partial^2}{\partial a_1 \partial b_1} \log V(A)V(B) \int_{O(p)} \exp(\text{tr } AHBH^*) \, dH \geq 0?
\]

Replace \( O(p) \) by \( U(p) \), then both are true for all Hermitian \( A, B \)

D.R. (J. Statist. Phys., 2004) used the condensation formulas of C. L. Dodgson [Lewis Carroll]
When Hegerl mentioned “missing data,” I had this happy feeling.

Monotone missing data: The matrix-argument hypergeometric functions appear even in the central problems

$$(p + q)$$-dimensional normal random vector, $$\left(\frac{X}{Y}\right) \sim N(\mu, \Sigma)$$

Random sample

$$
\begin{array}{cccc}
X_1 & \cdots & X_n & * \\
Y_1 & \cdots & Y_n & Y_{n+1} \\
\end{array}
\cdots
\begin{array}{ccc}
\cdots & \cdots & \cdots \\
& & \\
\end{array}
\begin{array}{c}
* \\
Y_N \\
\end{array}
$$

Panel survey data, astronomy, early detection of diseases, wildlife research, covert communications, mental-health research, ...
Explicit formulas for the MLE’s of $\mu$ and $\Sigma$ are available, but nothing yet for the EXACT (fixed $n, N$) distributions

In some applications, it could be unethical to hope for large sample sizes.

Calling Alan Edelman! Lots and lots of random matrices ...

Theorem (Wan-Ying Chang and D.R., November 1, 2006):

$$\hat{\mu} \sim \mathcal{N}_{p+q}(\mu, \Omega), \quad \sqrt{1 - \frac{n}{N}} R \left( \begin{pmatrix} V_2 \\ 0 \end{pmatrix} \right)$$

where $V_1 \sim \mathcal{N}_{p+q}(\mu, \Omega)$, $R$ is a positive r.v., and $V_2 \sim \mathcal{N}_p(0, \ldots)$
$\hat{\Sigma}$ has a density function of the form

$$W(\hat{\Sigma}_{11.2}) W(\hat{\Sigma}_{22}) {_1F_1}(\hat{\Sigma}_{12} \hat{\Sigma}_{22}^{-1} \hat{\Sigma}_{21})$$

What is this $\textstyle{_1F_1}$? The confluent hypergeometric function which appears at the beginning of Johnstone’s Ann. Statist. paper on principal components for high-dimensional random matrices.

Resistance is futile: You will be assimilated into the area of world of matrix-argument hypergeometric functions.
Wan-Ying and still can establish optimality properties of classical statistical hypothesis tests

Testing $H_0 : \mu = \mu_0$ against $H_a : \mu \neq \mu_0$: We have modified likelihood ratio test statistics which are unbiased

Testing $H_0 : (\mu, \Sigma) = (\mu_0, \Sigma_0)$ against $H_a : (\mu, \Sigma) \neq (\mu_0, \Sigma_0)$

Ditto

The sphericity test: $H_0 : \Sigma \propto I$

We have the distribution of the LRT statistic, but unbiasedness continues to evade us

Testing $\Sigma_{12} = 0$: Eaton and Kariya (1983)