A Simple Formula for the Moments of Unitarily Invariant Matrix Distributions

Stephen Howard[1] **and Ali Pezeshki**[2]

[1]Defence Science & Technology Group, Australia
[2]Colorado State University, USA

CoE Meeting
June 2022

Complex Unitarily Invariant Distributions

- Complex multivariate Gaussian: Modeling baseband signal measurements.

- Complex Wishart and complex inverse Wishart: Sample covariances and inverse sample covariances

$$R = XX^H \text{ and } R^{-1} = (XX^H)^{-1}$$

- Complex beta: Likelihood ratios for detection

$$y = \mu Hx + Jz + n \sim \begin{cases} \mathcal{CN}[Jz, \sigma^2 I], & \mathcal{H}_0 \ (\mu = 0) \\ \mathcal{CN}[\mu Hx + Jz, \sigma^2 I], & \mathcal{H}_1 \ (\mu > 0) \end{cases}$$

$$\ell = \frac{y^H P_J^\perp E_{JH} P_J^\perp y}{y^H P_J^\perp (I - E_{JH}) P_J^\perp y}$$

Complex Unitarily Invariant Distributions

- Complex multivariate Gaussian: Modeling baseband signal measurements.
- Complex Wishart and complex inverse Wishart: Sample covariances and inverse sample covariances

$$R = XX^H \text{ and } R^{-1} = (XX^H)^{-1}$$

- Complex beta: Likelihood ratios for detection

$$y = \mu Hx + Jz + n \sim \begin{cases} \mathcal{CN}[Jz, \sigma^2 I], & \mathcal{H}_0 \ (\mu = 0) \\ \mathcal{CN}[\mu Hx + Jz, \sigma^2 I], & \mathcal{H}_1 \ (\mu > 0) \end{cases}$$

$$\ell = \frac{y^H P_J^{\perp} E_{JH} P_J^{\perp} y}{y^H P_J^{\perp} (I - E_{JH}) P_J^{\perp} y}$$

Complex Unitarily Invariant Distributions

- Complex multivariate Gaussian: Modeling baseband signal measurements.
- Complex Wishart and complex inverse Wishart: Sample covariances and inverse sample covariances

$$R = XX^H \text{ and } R^{-1} = (XX^H)^{-1}$$

- Complex beta: Likelihood ratios for detection

$$y = \mu Hx + Jz + n \sim \begin{cases} \mathcal{CN}[Jz, \sigma^2 I], & \mathcal{H}_0 \ (\mu = 0) \\ \mathcal{CN}[\mu Hx + Jz, \sigma^2 I], & \mathcal{H}_1 \ (\mu > 0) \end{cases}$$

$$\ell = \frac{y^H P_J^\perp E_{JH} P_J^\perp y}{y^H P_J^\perp (I - E_{JH}) P_J^\perp y}$$

Complex Unitarily Invariant Distributions

- Complex matrix beta:
  - Fisher Information in compressed sensing (Pakrooh et al. 2015)

$$y = x(\theta) + n \quad \text{vs} \quad z = \Phi y$$

$$G = \left[\frac{\partial}{\partial \theta_1} x(\theta) \mid \cdots \mid \frac{\partial}{\partial \theta_p} x(\theta) \right]$$

$$F = \frac{1}{\sigma^2} G^H P_{\Phi^H} G \quad \text{and} \quad \text{CRLB} = F^{-1}.$$

  - Detection statistics in passive radar (Howard et al. 2015)

$$\gamma = \frac{|G|}{\prod_{i=1}^{M} |G_{ii}|}; \quad G = X^H X$$

Complex Unitarily Invariant Distributions

- Complex matrix beta:
  - Fisher Information in compressed sensing (Pakrooh et al. 2015)

$$y = x(\theta) + n \quad \text{vs} \quad z = \Phi y$$

$$G = \left[\frac{\partial}{\partial \theta_1} x(\theta) \mid \cdots \mid \frac{\partial}{\partial \theta_p} x(\theta) \right]$$

$$F = \frac{1}{\sigma^2} G^H P_{\Phi^H} G \quad \text{and} \quad \text{CRLB} = F^{-1}.$$

  - Detection statistics in passive radar (Howard et al. 2015)

$$\gamma = \frac{|G|}{\prod_{i=1}^{M} |G_{ii}|}; \quad G = X^H X$$

Calculating Moments

- Select prior work:
  - Multivariate normal: Isserlis (1918)
  - Complex multivariate normal: James (1953); Muirhead (book)
  - Real Wishart and matrix beta: Muirhead (book)
  - Complex Wishart: Maiwald & Kraus (2000); Graczyk *et al.* (2003); Nagar & Gupta (2011).
  - Complex matrix beta: Gupta *et al.* (2009), Pakrooh *et al.* (2015)
- Goal: Derive a formula for moments of *all* complex matrix distributions that can be transformed to a unitarily invariant distribution via conjugation by a fixed matrix.

Overview

- Methodology: Exploiting the relationship between the moments of unitarily invariant distributions and the joint action of the unitary and symmetric groups on $(\mathbb{C}^N)^{\otimes d}$.
- Key: *Schur-Weyl duality*, which relates the irreducible representations of unitary and symmetric groups.
- Advantage: A moment formula for all complex multivariate unitarily invariant distribution that separates the combinatorial aspect of moment computation from the calculation of a small number of specific distribution dependent moments.
- Closest prior work: Graczyk *et al.* (2003); Exploits the relationship between moments of the Wishart distribution and the symmetric group.

The Moment Problem

- Q is a complex $N \times N$ matrix-valued random variable with probability distribution $F(Q)$.
- The distribution is unitarily invariant, if for all $V \in \mathcal{U}(N)$, the group of $N \times N$ unitary matrices, $VQV^\dagger$ has the same distribution as Q.
- The notation $Q^{\otimes d}$ denotes the tensor power $Q \otimes Q \otimes \cdots \otimes Q$, with d factors.
- View Q as a random linear operator on $\mathbb{C}^N$ and $Q^{\otimes d}$ as a random linear operator on $(\mathbb{C}^N)^{\otimes d}$.
- $\mathsf{E}\left(Q^{\otimes d}\right)$ is a fixed linear operator on $(\mathbb{C}^N)^{\otimes d}$. Its matrix elements include the moments corresponding to all products of matrix elements of Q with d factors.
- Problem: Compute $\mathsf{E}\left(Q^{\otimes d}\right)$.

Group Actions

- Two important group actions can be defined on $(\mathbb{C}^N)^{\otimes d}$.
  - The first is the action of the symmetric group $\mathcal{S}_d$ (group of permutations on d letters) on $(\mathbb{C}^N)^{\otimes d}$ defined by

$$L_\sigma(\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_d) = \boldsymbol{v}_{\sigma(1)} \otimes \cdots \otimes \boldsymbol{v}_{\sigma(d)}.$$

  - The other is the action of the unitary group $\mathcal{U}(N)$ on $(\mathbb{C}^N)^{\otimes d}$ defined by

$$U^{\otimes d}(\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_d) = U\boldsymbol{v}_1 \otimes \cdots \otimes U\boldsymbol{v}_d.$$

- Under these two group actions $(\mathbb{C}^N)^{\otimes d}$ becomes a unitary representation of the product group $\mathcal{U}(N) \times \mathcal{S}_d$.

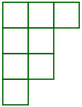
Decomposition into Irreducible Representations

- This representation decomposes into irreducible representations $\{W_{\boldsymbol{\lambda}}\}$:

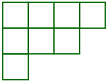
$$(\mathbb{C}^N)^{\otimes d} = \bigoplus_{\boldsymbol{\lambda}} W_{\boldsymbol{\lambda}}$$

where the sum is over all partitions $\boldsymbol{\lambda}$ of d with no more than N parts. That is, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_\ell)$, $\ell \leq N$, with the λ_i integers satisfying $\lambda_i \geq \lambda_{i+1}$ and $\lambda_1 + \cdots + \lambda_\ell = d$.

- Each $\boldsymbol{\lambda}$ can be associated with a Young diagram:



$\boldsymbol{\lambda} = (3, 2^2, 1)$



$\boldsymbol{\lambda}' = (4, 3, 1)$

Decomposition into Irreducible Representations

- The decomposition $(\mathbb{C}^N)^{\otimes d} = \bigoplus_{\boldsymbol{\lambda}} W_{\boldsymbol{\lambda}}$ means that linear operators in $\mathcal{U}(N) \times \mathcal{S}_d$ can be simultaneously block diagonalized with the blocks labeled by $\boldsymbol{\lambda}$.

- Schur-Weyl duality states that

$$W_{\boldsymbol{\lambda}} = V_{\boldsymbol{\lambda}} \otimes S_{\boldsymbol{\lambda}}$$

where $V_{\boldsymbol{\lambda}}$ are non-equivalent unitary irreducible representations of $\mathcal{U}(N)$ and $S_{\boldsymbol{\lambda}}$ are non-equivalent unitary irreducible representations of $\mathcal{S}_d$.

Decomposition into Irreducible Representations

- Associated with each $W_{\boldsymbol{\lambda}}$ there is a Young orthogonal projection operator $Y_{\boldsymbol{\lambda}}$ onto that subspace. These
  - satisfy

$$\sum_{\boldsymbol{\lambda}} Y_{\boldsymbol{\lambda}} = I, \quad Y_{\boldsymbol{\lambda}_1} Y_{\boldsymbol{\lambda}_2} = \delta_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2} Y_{\boldsymbol{\lambda}_1}$$

  - and take the form

$$Y_{\boldsymbol{\lambda}} = \frac{\chi_{\boldsymbol{\lambda}}(e)}{d!} \sum_{\sigma \in \mathcal{S}_d} \chi_{\boldsymbol{\lambda}}(\sigma) L_{\sigma}.$$

  The $\chi_{\boldsymbol{\lambda}}(\sigma)$ are real integer coefficients which comprise the irreducible *characters* of $\mathcal{S}_d$. These can be constructed by standard methods (Hamermesh (1962)).

A Key Observation

- The linear operator $\mathsf{E}(Q^{\otimes d})$ commutes with the action of $\mathcal{U}(N) \times \mathcal{S}_d$ on $(\mathbb{C}^N)^{\otimes d}$.
- So if $\mathsf{E}(Q^{\otimes d})$ is restricted to one of the irreducible representations $W_{\boldsymbol{\lambda}}$, then Schur's Lemma implies that $\mathsf{E}(Q^{\otimes d})$ acts as a multiple, $m_{\boldsymbol{\lambda}}(Q)$, of the identity.
- Then we simply need to calculate this $m_{\boldsymbol{\lambda}}(Q)$, which is a coefficient moment of Q.
- Choose a convenient unit vector $\boldsymbol{w}_{\boldsymbol{\lambda}} \in W_{\lambda}$ in each irreducible representation and compute

$$m_{\boldsymbol{\lambda}} = \langle \boldsymbol{w}_{\boldsymbol{\lambda}}, \mathsf{E}\left(Q^{\otimes d}\right) \boldsymbol{w}_{\boldsymbol{\lambda}} \rangle.$$

The Main Result

- The d^{th} order moments of Q are given by

$$\mathsf{E}\left(Q^{\otimes d}\right) = \sum_{\boldsymbol{\lambda}} m_{\boldsymbol{\lambda}} Y_{\boldsymbol{\lambda}}$$

where the coefficient moments $m_{\boldsymbol{\lambda}}(Q)$ are given by

$$m_{\boldsymbol{\lambda}} = \mathsf{E}\left(M_{\lambda'_1}^{\alpha'_1} M_{\lambda'_2}^{\alpha'_2} \ldots M_{\lambda'_k}^{\alpha'_k}\right)$$

where M_ℓ denotes the leading principal minor of order ℓ for Q, i.e., the determinant of its top left $\ell \times \ell$ block:

$$\det[\langle \boldsymbol{e}_i, Q\boldsymbol{e}_j \rangle]_{i,j=1,\ldots,\ell},$$

and $\boldsymbol{\lambda}' = ({\lambda'}_1^{\alpha'_1}, {\lambda'}_2^{\alpha'_2}, \ldots, {\lambda'}_k^{\alpha'_k})$ is the transpose of $\boldsymbol{\lambda}$.

The Main Result (Continued)

- If X is a complex matrix-valued random variable with the property that $Q = G^{-1}X(G^{-1})^{\dagger}$ has a unitarily invariant probability distribution for some fixed non-singular matrix G, then

$$\mathsf{E}(X^{\otimes d}) = \sum_{\boldsymbol{\lambda}} m_{\boldsymbol{\lambda}}(Q)\, R_{\boldsymbol{\lambda}}$$

where $R_{\boldsymbol{\lambda}} = Y_{\boldsymbol{\lambda}} R^{\otimes d} Y_{\boldsymbol{\lambda}}$, with $R = GG^{\dagger}$.

Computing Moments

- Calculate a d^{th} order moment of elements of $X = GQG^{\dagger}$.
- For $R = GG^{\dagger}$, matrix elements of $R_{\boldsymbol{\lambda}}$ are

$$(R_{\boldsymbol{\lambda}})_{IJ} = \frac{\chi_{\boldsymbol{\lambda}}(e)}{d!} \sum_{\sigma \in \mathcal{S}_d} \chi_{\boldsymbol{\lambda}}(\sigma) (R^{\otimes d})_{I\sigma(J)}$$

where $I = (i_1, \dots, i_d)$ and $J = (j_1, \dots, j_d)$.

- We have

$$\mathsf{E}(x_{i_1 j_1} \cdots x_{i_d j_d}) = \mathsf{E}(X^{\otimes d})_{IJ} = \sum_{\boldsymbol{\lambda}} m_{\boldsymbol{\lambda}}(Q)\ (R_{\boldsymbol{\lambda}})_{IJ}$$

- For $R = I$,

$$\mathsf{E}(q_{i_1 j_1} \cdots q_{i_d j_d}) = \mathsf{E}(Q^{\otimes d})_{IJ} = \sum_{\boldsymbol{\lambda}} m_{\boldsymbol{\lambda}}(Q)\ \frac{\chi_{\boldsymbol{\lambda}}(e)}{d!} \sum_{\sigma : I = \sigma(J)} \chi_{\boldsymbol{\lambda}}(\sigma)$$

Minor Moments for Some Distributions

- For some complex matrix distributions, the coefficient moments $m_{\boldsymbol{\lambda}}(Q)$ can be computed from just a knowledge of the normalization factors of the distributions.
- Two important examples:

Distribution	$m_{\boldsymbol{\lambda}}$
Beta	$\prod_{\ell=1}^{k} \frac{B_{\lambda'_\ell}(a+\sum_{j=1}^{\ell} \alpha'_j, b)}{B_{\lambda'_\ell}(a+\sum_{j=1}^{\ell-1} \alpha'_j, b)}$
Wishart/Gamma	$\prod_{\ell=1}^{k} \frac{\Gamma_{\lambda'_\ell}(a+\sum_{j=1}^{\ell} \alpha'_j)}{\Gamma_{\lambda'_\ell}(a+\sum_{j=1}^{\ell-1} \alpha'_j)}$

- Also applies to the type II complex beta distribution, the inverse complex matrix beta distribution, and the inverse complex gamma/Wishart distributions.

Generalization of Isserlis' Theorem

- Suppose $\boldsymbol{z}$ is a unitarily invariant vector-valued random variable.
- If $\boldsymbol{y} = R^{1/2}\boldsymbol{z}$, with R a positive-definite hermitian matrix, and $Q = \boldsymbol{z}\boldsymbol{z}^\dagger$, then there is only a single term $\boldsymbol{\lambda} = (d)$ in the expansion with $m_{(d)}(Q) = \mathsf{E}(|z_1|^{2d})$:

$$\begin{aligned} \mathsf{E}(y_{i_1}\overline{y}_{j_1}\cdots y_{i_d}\overline{y}_{j_d}) &= R^{\otimes d}\mathsf{E}(Q^{\otimes d}) \\ &= \frac{1}{d!}\mathsf{E}(|z_1|^{2d})\sum_{\sigma\in\mathcal{S}_d}\left(R^{\otimes d}\right)_{I\sigma(J)}. \end{aligned}$$

- This is a generalization of Isserlis' Theorem (1918) to unitarily invariant vector distributions.

Concluding Remarks

- The formula for the moments of complex multivariate unitarily invariant distributions derived here provides a powerful tool for
  - dealing with the combinatorial aspect of moment computation
  - and separating them from calculations of a small number of specific distribution dependent moments.
- Our approach could be generalized to compute moments of other classes of distributions.