Maximal Invariants Over Symmetric Cones

Applications of Jordan Algebra

Ahmad S. Yasamin

Indiana University

August 31, 2008

Outline

- Consider a statistical model consists of a sample space $\mathcal{X}$ and unknown probability measures, parameterized by a parameter set Θ as $\theta \mapsto P_\theta$.

- Consider a statistical model consists of a sample space $\mathcal{X}$ and unknown probability measures, parameterized by a parameter set Θ as $\theta \mapsto P_\theta$.
- In this setting, the statistical inference about the model is often focused on parameter estimation, and hypothesis testing:

$$H_0 : \theta \in \Theta_0 \subset \Theta \text{ vs. } H : \theta \in \Theta.$$

- Consider a statistical model consists of a sample space $\mathcal{X}$ and unknown probability measures, parameterized by a parameter set Θ as $\theta \mapsto P_\theta$.
- In this setting, the statistical inference about the model is often focused on parameter estimation, and hypothesis testing:

$$H_0 : \theta \in \Theta_0 \subset \Theta \text{ vs. } H : \theta \in \Theta.$$

- Systematically, this means to find a suitable test statistic t from $\mathcal{X}$ to a measurable space Y, and the distribution of $t(P_\theta)$, the transformed measure under t.

Examples

For the *Gaussian model*

$$(\mathcal{N}_n(0,\Sigma)\in\mathcal{P}(\mathbb{R}^n)|\Sigma\in \mathrm{P}(n,\mathbb{R}))$$

some classical examples of such hypotheses are:

Spherical Test

$$H_0 : \Sigma = \begin{pmatrix} \sigma^2 & & \\ & \ddots & \\ & & \sigma^2 \end{pmatrix} \quad \text{vs.} \quad H : \Sigma \neq \begin{pmatrix} \sigma^2 & & \\ & \ddots & \\ & & \sigma^2 \end{pmatrix} \quad \sigma^2 \in \mathbb{R}_+. \tag{1}$$

Testing For Complex Structure

$$H_0 : \Sigma = \Sigma_0 \triangleq \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \quad \text{vs.} \quad \Sigma \neq \Sigma_0, \tag{2}$$

for some real square matrices A, B.

Testing For Quaternion Structure

$$H_0 : \Sigma = \Sigma_1 \triangleq \begin{pmatrix} A & -B & -C & -D \\ B & A & -D & C \\ C & D & A & -B \\ D & -C & B & A \end{pmatrix} \quad \text{vs.} \quad H : \Sigma \neq \Sigma_1, \tag{3}$$

An Observation

In each of these hypotheses, both model (over $\mathrm{P}(n,\mathbb{R})$) and submodel (over a subcone of $\mathrm{P}(n,\mathbb{R})$) are invariant under an action of a subgroup G of $\mathrm{O}(n,\mathbb{R})$, the group of orthogonal matrices over $\mathbb{R}$:

$$\mathbf{x} \sim \mathcal{N}_n(0,\Sigma) \Longrightarrow A\mathbf{x} \sim \mathcal{N}_n(0, A\Sigma A^T) \quad \forall A \in G.$$

Classical Approach

- The maximal invariants that arise in testing problems like those stated in (1), (2) and (3) are functions of the eigenvalues of the sample covariance matrix.

Classical Approach

- The maximal invariants that arise in testing problems like those stated in (1), (2) and (3) are functions of the eigenvalues of the sample covariance matrix.
- Therefore, the problem is reduced to finding the distribution of the ordered eigenvalues of a (Wishart, often) distribution over a classical symmetric cone (the cone of positive definite matrices over real, complex or quaternion).

Setup

- The model
  $$\mathcal{M} : (P_\sigma \in \mathcal{P}(\mathcal{X}) | \sigma \in \Omega)$$
  is a statistical model parameterized by cone Ω.

Setup

- The model
$$\mathcal{M} : (P_\sigma \in \mathcal{P}(\mathcal{X}) | \sigma \in \Omega)$$
is a statistical model parameterized by cone Ω.
- The submodel
$$\mathcal{M}_0 : (P_\sigma \in \mathcal{P}(\mathcal{X}) | \sigma \in \Omega_0)$$
is parameterized over Ω_0, a symmetric subcone of Ω.

General Testing Problem

$$H_0 : \sigma \in \Omega_0 \quad \text{vs.} \quad H : \sigma \in \Omega \backslash \Omega_0 .$$

Reduction of The Problem

Similar to classical cases, the problem is reduced to finding the distribution of the ordered eigenvalues.

The Distribution of The Ordered Eigenvalues

Lemma

Suppose $\mathbf{x} : \mathcal{X} \to \Omega$ *is a random vector, where* Ω *is an irreducible symmetric cone in a Euclidean Vector space* $(V, \langle\cdot|\cdot\rangle)$. *Let define*

$$\begin{aligned} \mathrm{Eign} : \Omega \to \mathbb{R}^r_+ &\triangleq \{(\xi_1, \ldots, \xi_r) : 0 < \xi_1 \leq \cdots \leq \xi_r\} \\ x \mapsto (\xi_1(x), \ldots, \xi_r(x)), \end{aligned}$$

where $0 < \xi(x) \leq \cdots \xi_r(x)$ *are the eigenvalues of* $x \in \Omega$. *Then the density of* $\mathrm{Eign}(\mathbf{x})$ *is given by*

$$\left(\frac{(2\pi)^{n-r}\Gamma_\Omega(\frac{d}{2})^r}{\Gamma_\Omega(\frac{rd}{2})} \right) \prod_{j>i} (\xi_j - \xi_i)^d \int_K p_{\mathbf{x}}(k \sum_{j=1}^{r} \xi_j c_j) d\mu_K(k).$$

$\mathbf{x} \sim \mathcal{W}_\Omega(\eta, \sigma)$

If $\mathbf{x} \sim \mathcal{W}_\Omega(\eta, \sigma)$, that is $\mathbf{x}$ is a random vector in Ω having the Wishart distribution with parameters $\eta > d(r-1)$ and $\sigma \in \Omega$, then the density of the maximal invariant $\mathrm{Eign}(\mathbf{x})$, thus the joint density of the distribution of the ordered eigenvalues of $\mathbf{x}$, is given by

$$c_0 \frac{\prod_{j>i}(\xi_j - \xi_i)^d \prod_{j=1}^{r} \xi_j^{\frac{1}{2}\eta - \frac{n}{r}}}{2^{\frac{1}{2}\eta r}\Gamma_\Omega(\frac{\eta}{2}) \det(\sigma)^{\frac{1}{2}\eta}} {}_1F_1^d(x, \frac{1}{2}\sigma^{-1}).$$

$\mathbf{x} \sim \mathcal{W}_\Omega(\eta, \zeta e)$

In particular, if $\sigma = \zeta e$ for some $\zeta > 0$, then this density is

$$\left(\frac{(2\pi)^{n-r}\Gamma(\frac{d}{2})^r}{(2\zeta)^{\frac{1}{2}\eta r}\Gamma_\Omega(\frac{1}{2}\eta)\Gamma_\Omega(\frac{rd}{2})} \right) \exp\left\{ -\frac{1}{2\zeta}\sum_{j=1}^{r}\xi_j \right\} \prod_{j=1}^{r} \xi_j^{\frac{1}{2}\eta - \frac{n}{r}} \prod_{j>i}(\xi_j - \xi_i)^d .$$

$\mathbf{x} \sim \mathcal{W}_\Omega(\eta, \zeta e, \epsilon)$

If $\sigma = \zeta e$, then the density of the maximal invariant $\mathrm{Eign}(\mathbf{x})$ is given by

$$=c_0 \frac{\prod_{j>i}(\xi_j - \xi_i)^d \prod_{j=1}^r \xi_j^{\frac{1}{2}\eta - \frac{n}{r}}}{(2\zeta)^{\frac{1}{2}\eta r}\Gamma_\Omega(\frac{\eta}{2})} \exp\left\{-\frac{1}{2\zeta}\sum_{j=1}^r \xi_j\right\}$$

$$\cdot \exp\left\{-\frac{1}{2}tr(\epsilon)\right\} {}_0F_1^d(\frac{1}{2}\eta; x; \frac{1}{4\zeta}\epsilon).$$

$\mathbf{u} \sim \mathcal{B}_{\Omega}^{I}(\eta_1, \eta_2, \epsilon)$

For the non-central beta distribution $\mathcal{B}_{\Omega}^{I}(\eta_1, \eta_2, \epsilon)$, the joint density of the distribution of the ordered eigenvalues of $\mathbf{u}$, $\mathrm{Eign}(\mathbf{u})$ is given by

$$c_0 \frac{1}{B_{\Omega}(\eta_1, \eta_2)} \exp\{-\mathrm{tr}(\epsilon)\} \prod_{j>i} (\beta_j - \beta_i)^d \prod_{j=1}^{r} \beta_j^{\eta_1 - \frac{n}{r}} \prod_{j=1}^{r} (1 - \beta_j)^{\eta_2 - \frac{n}{r}}$$

$$\cdot \, {}_1F_1^d \left(\eta_1 + \eta_2; \eta_2; u, \epsilon\right),$$

Algebras

Definition

Let $\mathbb{F}$ be the field of real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$.

- An *algebra* over $\mathbb{F}$ is a vector space A over $\mathbb{F}$ equipped with a bilinear mapping

$$\begin{aligned} A \times A &\rightarrow A \\ (x, y) &\mapsto xy. \end{aligned}$$

Algebras

Definition

Let $\mathbb{F}$ be the field of real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$.

- An *algebra* over $\mathbb{F}$ is a vector space A over $\mathbb{F}$ equipped with a bilinear mapping

$$\begin{array}{ccc} A \times A & \rightarrow & A \\ (x, y) & \mapsto & xy. \end{array}$$

- For any element $x \in A$ the linear representation

$$\begin{array}{rcl} L : A & \rightarrow & \mathrm{L}(A) \\ x & \mapsto & L(x)(y \mapsto xy) \end{array}$$

  is called the (left) *regular representation* of A.

Definition

A *Jordan algebra* J is an algebra satisfying

1. $xy = yx$ for any $x, y \in J$, and

Example

The vector space $\mathbb{R} \times \mathbb{R}^n$ with multiplication

$$(\zeta, x)(\zeta', x') \triangleq (\zeta\zeta' + \langle x|x'\rangle, \zeta x' + \zeta' x).$$

Definition

A *Jordan algebra* J is an algebra satisfying

1. $xy = yx$ for any $x, y \in J$, and
2. $x(x^2y) = x^2(xy)$ for any $x, y \in J$.

Example

The vector space $\mathbb{R} \times \mathbb{R}^n$ with multiplication

$$(\zeta, x)(\zeta', x') \triangleq (\zeta\zeta' + \langle x|x'\rangle, \zeta x' + \zeta' x).$$

Definition

A *Jordan algebra* J is an algebra satisfying

1. $xy = yx$ for any $x, y \in J$, and
2. $x(x^2y) = x^2(xy)$ for any $x, y \in J$.

- A real Jordan algebra J is called *Euclidean* if there is an inner product $\langle \cdot | \cdot \rangle$ on J such that

$$\langle x|yz \rangle = \langle xz|y \rangle.$$

Example

The vector space $\mathbb{R} \times \mathbb{R}^n$ with multiplication

$$(\zeta, x)(\zeta', x') \triangleq (\zeta\zeta' + \langle x|x' \rangle, \zeta x' + \zeta' x).$$

Quadratic Representation Map

- The map

$$\begin{array}{rcl} P : J & \longrightarrow & \mathrm{End}(J) \\ x & \mapsto & 2L(x)^2 - L(x^2) \quad x \in J \end{array}$$

is called the *quadratic representation* of J.

Quadratic Representation Map

- The map

$$\begin{array}{rcl} P : J & \longrightarrow & \mathrm{End}(J) \\ x & \longmapsto & 2L(x)^2 - L(x^2) \quad x \in J \end{array}$$

  is called the *quadratic representation* of J.
- The cone of positive elements of J is defined by

$$J_+ \triangleq \{x^2 : x \in J^\times\}.$$

Quadratic Representation Map

- The map
$$\begin{aligned} P : J &\rightarrow \operatorname{End}(J) \\ x &\mapsto 2L(x)^2 - L(x^2) \quad x \in J \end{aligned}$$
is called the *quadratic representation* of J.
- The cone of positive elements of J is defined by
$$J_+ \triangleq \{x^2 : x \in J^\times\}.$$

Example

For the Jordan Algebra $\mathbb{R} \times \mathbb{R}^n$ we have

$$(\mathbb{R} \times \mathbb{R}^n)_+ = \{(\zeta, x) : \zeta > \|x\|\}.$$

Quadratic Representation Map

- The map

$$\begin{array}{rcl} P : J & \rightarrow & \mathrm{End}(J) \\ x & \mapsto & 2L(x)^2 - L(x^2) \quad x \in J \end{array}$$

  is called the *quadratic representation* of J.
- The cone of positive elements of J is defined by

$$J_+ \triangleq \{x^2 : x \in J^\times\}.$$

Example

For the Jordan Algebra $\mathbb{R} \times \mathbb{R}^n$ we have

$$(\mathbb{R} \times \mathbb{R}^n)_+ = \{(\zeta, x) : \zeta > \|x\|\}.$$

Notation

For $x \in J_+$ and $y \in J$ $P(x^{\frac{1}{2}})(y)$ is denoted by $x \star y$.

Basic properties of J_+

- The group of automorphisms of J_+,

$$G \triangleq \{g \in \mathrm{GL}(J) : gJ_+ = J_+\},$$

  is a Lie group.

Basic properties of J_+

- The group of automorphisms of J_+,

$$G \triangleq \{g \in \mathrm{GL}(J) : gJ_+ = J_+\},$$

  is a Lie group.
- If e is the identity element of J, then

$$K \triangleq G_e = \mathrm{O}(J) \cap G,$$

  and $J_+ \cong G \backslash K$.

Table of Irreducible Symmetric Cones

J	J_+	G	K	$dimJ$	r
$\mathrm{Sym}(n, \mathbb{R})$	$\mathrm{P}(n, \mathbb{R})$	$GL^+(n, \mathbb{R})$	$\mathrm{SO}(n, \mathbb{R})$	$\frac{1}{2}n(n+1)$	
$\mathrm{Herm}(n, \mathbb{C})$	$\mathrm{P}(n, \mathbb{C})$	$\mathrm{GL}(n, \mathbb{C})$	$U(n, \mathbb{C})$	n^2	
$\mathrm{Herm}(n, \mathbb{H})$	$\mathrm{P}(n, \mathbb{H})$	$\mathrm{GL}(n, \mathbb{H})$	$\mathrm{Sp}(2n, \mathbb{C})$	$n(2n-1)$	
$\mathbb{R} \times \mathbb{R}^{n-1}$	L_n	$\mathbb{R} \times \mathrm{SO}^+(1, n-1)$	$\mathrm{SO}(n-1, \mathbb{R})$	n	
$\mathrm{Herm}(3, \mathbb{O})$	$\mathrm{P}(3, \mathbb{O})$	$\mathbb{R} \times E_6$	F_4	27	

Table: Classification of irreducible symmetric cones

(Spectral theorem, second version)

.

- If $x \in J$, then there is a Jordan frame $c_1, \dots, c_r$, depending on x, in J such that

$$x = \xi_1 c_1 + \cdots + \xi_r c_r \tag{4}$$

  where $\xi_1, \dots, \xi_r \in \mathbb{R}$ are (not necessarily distinct) eigenvalues of x.

(Spectral theorem, second version)

.

- If $x \in J$, then there is a Jordan frame $c_1, \ldots, c_r$, depending on x, in J such that
$$x = \xi_1 c_1 + \cdots + \xi_r c_r \tag{4}$$
where $\xi_1, \ldots, \xi_r \in \mathbb{R}$ are (not necessarily distinct) eigenvalues of x.
- Furthermore

$$\begin{aligned} \det(x) &= \xi_1 \xi_2 \cdots \xi_r \\ tr(x) &= \xi_1 + \cdots + \xi_r \end{aligned}$$

Gamma Functions

- For $(s_1, \ldots, s_r) \in \mathbb{C}^r$ the *Gamma function* over J_+ is defined by

$$\Gamma_{J_+}(s) \triangleq (2\pi)^{\frac{n-r}{2}} \prod_{j=1}^{r} \Gamma(s_j - (j-1)\frac{d}{2}).$$

Gamma Functions

- For $(s_1, \dots, s_r) \in \mathbb{C}^r$ the *Gamma function* over J_+ is defined by

$$\Gamma_{J_+}(s) \triangleq (2\pi)^{\frac{n-r}{2}} \prod_{j=1}^{r} \Gamma(s_j - (j-1)\frac{d}{2}).$$

- In particular, identifying $z \in \mathbb{C}$ with $(z, \dots, z) \in \mathbb{C}^r$ we get

$$\Gamma_{J_+}(z) = \int_{J_+} \exp\{-\mathrm{tr}(x)\} \det(x)^{z-\frac{n}{r}} dx.$$

Spherical polynomials

- A finite sequence $\lambda = (\lambda_1, \ldots, \lambda_r)$ of non-increasing positive integers is called a *partition* of length r.

Spherical polynomials

- A finite sequence $\lambda = (\lambda_1, \ldots, \lambda_r)$ of non-increasing positive integers is called a *partition* of length r.
- For the partition λ, the *spherical polynomial* Φ_λ is defined by

$$\Phi_\lambda(x) \triangleq \int_K \Delta_\lambda(kx) d\mu_K(k), \quad x \in J^{\mathbb{C}}$$

Zonal Polynomials

Definition

The *zonal polynomial* Z^d_λ is a homogeneous K-invariant polynomial defined by

$$Z^d_\lambda(x) \triangleq d_\lambda \frac{|\lambda|!}{\left(\frac{n}{r}\right)^d_\lambda} \Phi_\lambda(x), \tag{5}$$

Zonal polynomials are K-invariant polynomials normalized by the property

$$\mathrm{tr}(x)^k = \sum_{|\lambda|=k} Z^d_\lambda(x) \quad x \in J_+. \tag{6}$$

Relationship With Symmetric Functions

Proposition

Let J_+ *be an irreducible symmetric cone with rank* r *and the Peirce constant* d. *Then for each* x *in* J_+ *and any partition* λ *we have*

$$Z_\lambda^d(x) = C_\lambda^{(\frac{2}{d})}(\xi_1, \ldots, \xi_r),$$

where $\xi_1, \ldots, \xi_r$ *are the eigenvalues of* x, *and* $C_\lambda^{(\frac{2}{d})}(\xi_1 \ldots, \xi_r)$ *are the Jack symmetric functions of* C*-type.*

Hypergeometric Functions

Definition

Let $a_1, \dots, a_p$ and $b_1, \dots, b_q$ be real numbers with

$$a_i - \tfrac{d}{2}(i-1) \geq 0 \quad \text{and} \quad b_j - \tfrac{d}{2}(j-1) \geq 0,$$

and x and y be in $J^{\mathbb{C}}$. The hypergeometric function ${}_pF_q^d$ is defined by

$$\begin{gathered} {}_pF_q^d(a_1, \dots, a_p; b_1, \dots, b_q; x, y) \triangleq \\ \sum_{k=1}^{\infty} \sum_{|\lambda|=k} \frac{(a_1)_\lambda^d \cdots (a_p)_\lambda^d}{(b_1)_\lambda^d \cdots (b_q)_\lambda^d} \frac{Z_\lambda^d(x)}{k!} \frac{Z_\lambda^d(y)}{Z_\lambda^d(e)}. \end{gathered}$$

For $y = e$ we set

$${}_pF_q^d(a_1, \dots, a_p; b_1, \dots, b_q, x) \triangleq {}_pF_q^d(a_1, \dots, a_p; b_1, \dots, b_q, x, e).$$

Noncentral Wishart Distribution

A random vector $\mathbf{x}$ with values in an irreducible symmetric cone Ω is said to have the *Wishart distribution* with parameters $\eta > (r-1)d$, $\sigma \in \Omega$ and $\epsilon \in \overline{\Omega}$ if its probability density (with respect to the restriction of the standard Lebesgue measure of J on Ω) is given by

$$\frac{1}{2^{\frac{1}{2}\eta r}\Gamma_\Omega(\frac{\eta}{2})\det(\sigma)^{\frac{1}{2}\eta}}\exp\left\{-\frac{1}{2}\mathrm{tr}(\sigma^{-1}x)\right\}\exp\left\{-\frac{1}{2}\mathrm{tr}(\epsilon)\right\}$$
$$.\det(x)^{\frac{1}{2}\eta-\frac{n}{r}}{}_0F_1^d\left(\frac{1}{2}\eta;\frac{1}{4}(\sigma^{-1}\star\epsilon)\star x\right)\mathbf{1}_\Omega(x).$$

This is denoted by $\mathbf{x} \sim \mathcal{W}_\Omega(\eta,\sigma,\epsilon)$.

Special case $\epsilon = 0$

In particular, if $\epsilon = 0$, the distribution is called the *(central) Wishart distribution* and is denoted by

$$\mathbf{x} \sim \mathcal{W}_{\Omega}(\eta, \sigma),$$

otherwise, is called the *non-central Wishart distribution*.

Bartlett's Test

- Suppose $\eta > (r-1)d$ is a known parameter, and consider the statistical model

$$(\mathcal{W}_\Omega(\eta,\sigma_1) \otimes \mathcal{W}_\Omega(\eta,\sigma_2) \in \mathcal{P}(\Omega \times \Omega) | (\sigma_1,\sigma_2) \in \Omega \times \Omega)$$

and its submodel

$$(\mathcal{W}_\Omega(\eta,\sigma) \otimes \mathcal{W}_\Omega(\eta,\sigma,\epsilon) \in P(\Omega \times \Omega) | \sigma \in \Omega)\,.$$

Bartlett's Test

- Suppose $\eta > (r-1)d$ is a known parameter, and consider the statistical model

  $$(\mathcal{W}_\Omega(\eta,\sigma_1)\otimes\mathcal{W}_\Omega(\eta,\sigma_2)\in\mathcal{P}(\Omega\times\Omega)|(\sigma_1,\sigma_2)\in\Omega\times\Omega)$$

  and its submodel

  $$(\mathcal{W}_\Omega(\eta,\sigma)\otimes\mathcal{W}_\Omega(\eta,\sigma,\epsilon)\in P(\Omega\times\Omega)|\sigma\in\Omega)\,.$$

- We wish to test the null hypothesis

  $$H_0:\sigma_1=\sigma_2=\sigma \quad \text{vs.} \quad H:\sigma_1\neq\sigma_2.$$

Theorem

For the observation $(x,y) \in \Omega \times \Omega$ the maximum likelihood estimator of σ under H_0 is $s(x,y) \triangleq \frac{x+y}{2}$, and the likelihood ratio statistic for testing H_0 vs. H is

$$\prod_{j=1}^{r} (1 - {l_j}^2)^{\frac{1}{2}\eta},$$

where $l_1 \leq \cdots \leq l_r$ are the eigenvalues of $\mathbf{r} \triangleq \frac{\mathbf{x}-\mathbf{y}}{2}$ with respect to $\mathbf{s}$.

Theorem

Furthermore, under the null hypothesis H_0 the statistics $s(\mathbf{x},\mathbf{y})$ and $\pi(\mathbf{x},\mathbf{y}) \triangleq (l_1, \cdots, l_r)$ are independently distributed, $s(\mathbf{x},\mathbf{y}) \sim \mathcal{W}_\Omega(\eta,\sigma)$ and the density of $\pi(x,y)$ is given by

$$c_0 \frac{2^{n-\eta}}{B_\Omega(\frac{1}{2}\eta; \frac{1}{2}\eta)} \prod_{j>i}(l_j - l_i)^d \prod_{j=1}^{r}(1 - l_j^2)^{\frac{1}{2}\eta - \frac{n}{r}}.$$