James' Five Fold Way and spiked models in multivariate analysis

Iain Johnstone

Statistics, Stanford

Joint with Alexei Onatski, Prathapa Dharmawansa,

Hong Kong, January 8, 2015

Basic equation of classical multivariate statistics

$$\det(H - xE) = 0$$

with $p \times p$ matrices

$$n_1 H = \sum_{\nu=1}^{n_1} \mathbf{x}_\nu \mathbf{x}_\nu' \qquad \text{`hypothesis' SS}$$

$$n_2 E = \sum_{\nu=1}^{n_2} \mathbf{z}_\nu \mathbf{z}_\nu' \qquad \text{`error' SS}$$

(Invariant) methods use (generalized) eigenvalues $\{x_i\}_{i=1}^{p}$

$\Leftrightarrow$ eigenvalues of "F-ratio" $E^{-1}H$.

Textbook topics using $E^{-1}H$

- Canonical correlation analysis
- Discriminant analysis
- Factor analysis*
- Multidimensional scaling*
- Multivariate Analysis of Variance – MANOVA
- Multivarate regression analysis
- Principal Component analysis*
- Signal detection (equality of covariance matrices)

* use limiting form $\det(H - xI) = 0$ with $E = I_p$, $(n_2 \to \infty)$

100–75–50–25 years ago

1915 Fisher publishes distribution of correlation coefficient

1939 Fisher, Girshick, Hsu, Roy, Mood *independently* obtain null distribution of roots of $E^{-1}H$

1964 James classifies non-null distributions of $E^{-1}H$

1987 Bootstrap era: Efron publishes BCa paper

100–75–50–25 years ago

1915 Fisher publishes distribution of correlation coefficient

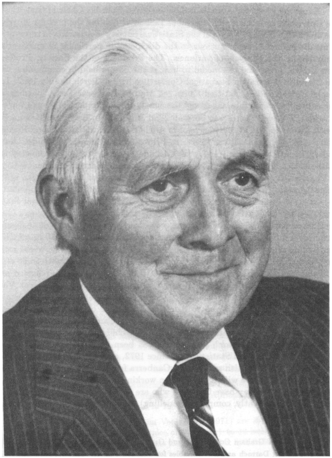
1939 Fisher, Girshick, Hsu, Roy, Mood *independently* obtain null distribution of roots of $E^{-1}H$

1964 James classifies non-null distributions of $E^{-1}H$

1987 Bootstrap era: Efron publishes BCa paper

Hamlet: "The time is out of joint ..."

A.T. James (1924 - 2013)



DISTRIBUTIONS OF MATRIX VARIATES AND LATENT ROOTS DERIVED FROM NORMAL SAMPLES[1]

By Alan T. James

Yale University

1. Summary. The paper is largely expository, but some new results are included to round out the paper and bring it up to date.

The following distributions are quoted in Section 7.

1. Type ${}_0F_0$, exponential: (i) χ^2, (ii) Wishart, (iii) latent roots of the covariance matrix.
2. Type ${}_1F_0$, binomial series: (i) variance ratio, F, (ii) latent roots with unequal population covariance matrices.
3. Type ${}_0F_1$, Bessel: (i) noncentral χ^2, (ii) noncentral Wishart, (iii) noncentral means with known covariance.
4. Type ${}_1F_1$, confluent hypergeometric: (i) noncentral F, (ii) noncentral multivariate F, (iii) noncentral latent roots.
5. Type ${}_2F_1$, Gaussian hypergeometric: (i) multiple correlation coefficient, (ii) canonical correlation coefficients.

Some questions suggested by James (1964)

- remarkable systematization of distributions
- focus on zonal polynomials – hard to compute
- what about approximations, guided by James' framework?
- current interest: p large
  - consider approximations with $p \propto n$
  - connections to random matrix theory
  - focus on low rank alternatives, a.k.a. "spiked models"

Our talks:

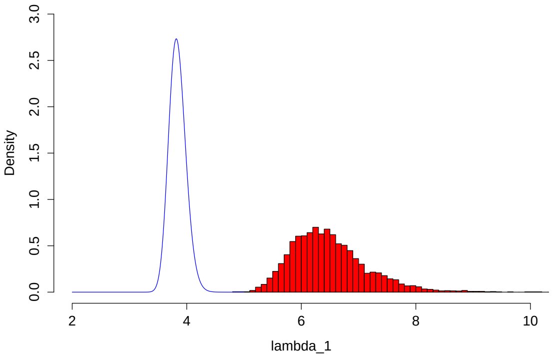
- explore likelihood ratios, phase transitions in James' setting
- behavior above transition:
  - Gaussian limits,
  - Limiting experiments for likelihood ratios,
  - Confidence intervals
- behavior below transition:
  - Limiting experiments for likelihood ratios,

Aside: why not just resample??

$X, \ n \times p$ with independent rows $N_p(0, I)$. $\quad H = n^{-1}X'X$.

Largest eigenvalue $\lambda_1(H) \approx$ (scaled) Tracy-Widom.

Resample rows of $X \to$ histogram for λ_1^* $\quad [n = p = 100]$



Outline

James' Five Fold Way: review classical examples,
(plus a limiting case), focus on

- eigenvalue equations
- low rank alternatives

Common structure in joint density of eigenvalues

- null case: random matrix connection
- alternatives: matrix hypergeometric functions
- integral representations for low rank cases

Principal Components Analysis (PCA) $[{}_0F_0]$

Data

$$X = [\mathbf{x}_1 \cdots \mathbf{x}_n] \qquad p \times n$$

Covariance structure:

$$\Sigma = \mathrm{Cov}(\mathbf{x}_\nu) = \Sigma_0 + \Phi$$

Low rank:

$$\Phi = \sum_{k=1}^{r} \theta_k \gamma_k \gamma_k'$$

Sample covariance matrix:

$$S = n^{-1} H, \qquad H = XX'$$

Eigenvalues:

$$\det(n^{-1} H - \lambda_i I) = 0$$

Regression - Known Covariance (REG_0) $[{}_0F_1]$

p-variate response:

$$\mathbf{y}_\nu = B'\mathbf{x}_\nu + \mathbf{z}_\nu, \qquad \Sigma_0 = \mathrm{Cov}(\mathbf{z}_\nu)$$

$$H_0 : CB = 0 \qquad C = \text{contrast matrix}$$

Σ_0 known:

$$H = YP_H Y' \qquad n \text{ hypothesis d.f.}$$

Eigenvalues:

$$\det(n^{-1}H - \lambda_i I) = 0.$$

Low rank: noncentrality (e.g. some MANOVA)

$$\Phi = \Sigma_0^{-1} MM'/n_1 = \Sigma_0^{-1} \sum_{k=1}^{r} \theta_k \gamma_k \gamma_k'$$

Signal Detection (SigDet) $[{}_1F_0]$

Data: $\mathbf{x}_\nu = \sum_1^r \sqrt{\theta}_k u_{\nu,k} \gamma_k + \mathbf{z}_\nu$

$$u_{\nu k} \overset{ind}{\sim} (0,1), \quad \mathrm{Cov}(\mathbf{z}_\nu) = \Sigma$$

Low rank structure: test $H_0 : \theta = 0$

$$\mathrm{Cov}(\mathbf{x}_\nu) = \Phi + \Sigma \qquad \Phi = \sum_1^r \theta_k \gamma_k \gamma_k'$$

Eigenvalues: $H = \sum_1^{n_1} \mathbf{x}_\nu \mathbf{x}_\nu' \qquad E = \sum_{n_1}^{n_1+n_2} \mathbf{z}_\nu \mathbf{z}_\nu'$

$$\det(n_1^{-1} H - \tilde{\lambda}_i n_2^{-1} E) = 0$$
$$\Leftrightarrow \det(H - \lambda_i (E + H)) = 0$$

Regression - Multiple Response (REG) $[{}_1F_1]$

$$\mathbf{y}_\nu = B'\mathbf{x}_\nu + \mathbf{z}_\nu, \qquad \Sigma = \mathrm{Cov}(\mathbf{z}_\nu)$$

Sums of squares matrices:

$$H = YP_H Y' \qquad n_1 \text{ hypothesis d.f.}$$
$$E = YP_E Y' \qquad n_2 \text{ error d.f.}$$

Eigenvalues:

$$\det(n_1^{-1}H - \tilde{\lambda}_i n_2^{-1}E) = 0 \qquad \text{multivariate } F$$
$$\Leftrightarrow \det(H - \lambda_i(E + H)) = 0 \qquad \text{multivariate Beta}$$

Low rank: noncentrality

$$\Phi = \Sigma^{-1}MM'/n_1 = \Sigma^{-1}\sum_{k=1}^{r}\theta_k\gamma_k\gamma_k'$$

Canonical Correlation Analysis (CCA) $\quad [{}_2F_1]$

$$\mathbf{x}_\nu \in \mathbb{R}^p \qquad \mathbf{y}_\nu \in \mathbb{R}^{n_1} \qquad \nu = 1, \ldots, n_1 + n_2 + 1$$

Look for maximally correlated $a'\mathbf{x}_\nu,\ b'\mathbf{y}_\nu$

$$\mathrm{Cov}\begin{pmatrix}\mathbf{x}_\nu \\ \mathbf{y}_\nu\end{pmatrix} = \begin{pmatrix}\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{pmatrix} \qquad \text{sample} \begin{pmatrix}S_{11} & S_{12} \\ S_{21} & S_{22}\end{pmatrix}$$

Eigenvalues:

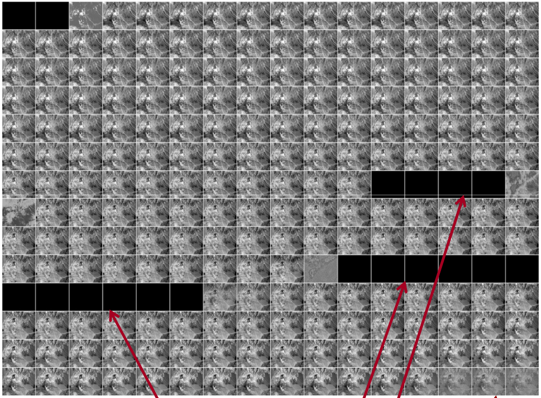
$$\det(S_{11}^{-1} S_{12} S_{22}^{-1} S_{21} - \lambda_i I_p) = 0$$

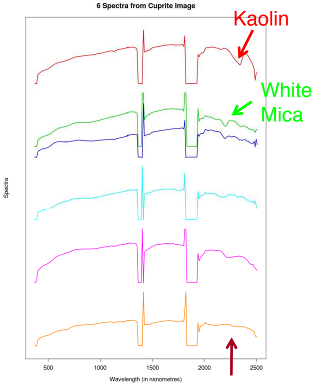
Low rank:

$$\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2} = \Phi^{1/2} = \sum_1^r \sqrt{\theta}_k \gamma_k \eta_k' \overset{e.g.}{=} \begin{bmatrix}\mathrm{diag}(\sqrt{\theta}_1, \ldots, \sqrt{\theta}_r) & 0 \\ 0 & 0\end{bmatrix}$$

Hyperspectral image example: Cuprite, Nevada

(224 AVIRIS images, [370, 2507] nm, ~ 9.6 nm apart, 614 x 512 (= 314,368) pixels, atmospherically corrected)





Water absorption bands

Noisy bands

Most mineral diagnostic information for minerals in [2000, 2500]

Gaussian assumptions

Assume $\mathbf{x}_\nu$, resp $\mathbf{z}_\nu$ are (zero-mean) Gaussian (⇒ formulas)

Why eigenvalues?

Group structure $\implies$ (λ_i) are maximal invariants.

O.K. for low rank alternatives if subspaces are unknown.

[**Wishart definition:** If $\underset{n \times p}{Z} \sim N(M, I \otimes \Sigma)$ is a normal data matrix, then

$$A = Z'Z = \sum_1^n \mathbf{z}_\nu \mathbf{z}_\nu' \sim W_p(n, \Sigma, \Omega),$$

with degrees of freedom n , and non-centrality $\Omega = \Sigma^{-1/2} M' M \Sigma^{-1/2}$]

Five Fold Way - Gaussian assumptions

PCA $[{}_0F_0]$

$$H \sim W_p(n, \Sigma_0 + \Phi)$$

REG$_0$ $[{}_0F_1]$

$$H \sim W_p(n, \Sigma_0, n\,\Phi)$$

SigDet $[{}_1F_0]$

$$H \sim W_p(n_1, \Sigma + \Phi)$$
$$E \sim W_p(n_2, \Sigma)$$

REG $[{}_1F_1]$

$$H \sim W_p(n_1, \Sigma, n_1\Phi)$$
$$E \sim W_p(n_2, \Sigma)$$

CCA $[{}_2F_1]$

$$H \sim W_p(n_1, I - \Phi, \Omega(\Phi))$$
$$E \sim W_p(n_2, I - \Phi)$$

$\Omega(\Phi)$ random

Symmetric Matrix Denoising (SMD)

$$G = \Phi + Z \qquad\qquad Z \text{ symmetric } p \times p$$

$$Z_{ij} \overset{ind}{\sim} N(0, 1 + \delta_{ij}) \qquad\qquad \mathrm{GOE}_p$$

Low rank: $\Phi = \sum_1^r \theta_k \gamma_k \gamma_k'$

Eigenvalues: $\det(G - \lambda_i I_p) = 0$

Limiting Case: $H_n \sim W_p(n, \Sigma_n) \qquad \Sigma_n = I_p + \Phi/\sqrt{n}.$

For p fixed, as $n \to \infty$ PCA $\to$ SMD:

$$\sqrt{n}(n^{-1} H_n - I_p) \overset{\mathcal{D}}{\Rightarrow} \Phi + Z, \qquad Z \sim \mathrm{GOE}_p$$

SMD as the limiting "simple" case

$$
\begin{tikzcd}
& G = \Phi + Z & \\
W_p(n, \Sigma_0 + n^{-1/2}\Phi) \arrow[ur] & & W_p(n, \Sigma_0, n^{1/2}\Phi) \arrow[ul] \\
\begin{array}{c} W_p(n_1, \Sigma + n_1^{-1/2}\Phi) \\ W_p(n_2, \Sigma) \end{array} \arrow[u] & & \begin{array}{c} W_p(n_1, \Sigma, n_1^{1/2}\Phi) \\ W_p(n_2, \Sigma) \end{array} \arrow[u] \\
& \begin{array}{c} W_p(n_1, I - \Phi_n, \Omega(\Phi_n)) \\ W_p(n_2, I - \Phi_n) \\ \Phi_n = n_2^{-1/2}\Phi \end{array} \arrow[uuu] &
\end{tikzcd}
$$

SMD as the limiting "simple" case

$$
\begin{tikzcd}
 & G = \Phi + Z & \\
W_p(n, \Sigma_0 + n^{-1/2}\Phi) \arrow[ur, "n \to \infty"] & & W_p(n, \Sigma_0, n^{1/2}\Phi) \arrow[ul, "n \to \infty"'] \\
\begin{array}{c} W_p(n_1, \Sigma + n_1^{-1/2}\Phi) \\ W_p(n_2, \Sigma) \end{array} \arrow[u, "n_2 \to \infty"] & & \begin{array}{c} W_p(n_1, \Sigma, n_1^{1/2}\Phi) \\ W_p(n_2, \Sigma) \end{array} \arrow[u, "n_2 \to \infty"'] \\
 & \begin{array}{c} W_p(n_1, I - \Phi_n, \Omega(\Phi_n)) \\ W_p(n_2, I - \Phi_n) \\ \Phi_n = n_2^{-1/2}\Phi \end{array} \arrow[uuu, "n_2 \to \infty"] &
\end{tikzcd}
$$

Outline

James' Five Fold Way: review classical examples, focus on

- eigenvalue equations
- low rank alternatives

Common structure in joint density of eigenvalues

- null case: random matrix connection
- alternatives: matrix hypergeometric functions
- integral representations for low rank cases

Links to RMT: Null Hypothesis

SMD			Gaussian	$\lambda(G/\sqrt{p})$
PCA	REG_0	$\xrightarrow{\Phi=0}$	Laguerre	$\lambda(H/n)$
SigDet	REG		Jacobi	$\lambda((E+H)^{-1}H)$
CCA				

$H_0 : \theta = 0 \rightarrow$ classical matrix ensembles. Joint eigenvalue density

$$p_0(\lambda) = \pi(\lambda)\Delta(\lambda)$$

$$\pi(\lambda) = \prod_{i=1}^{p} \pi(\lambda_i), \qquad \Delta(\lambda) = \prod_{i>j}(\lambda_i - \lambda_j)$$

Empirical spectral distributions converge:

$$F_n(\lambda) = p^{-1}\#\{i : \lambda_i \leq \lambda\} \quad \stackrel{a.s.}{\rightarrow} \quad F(\lambda)$$

Links to RMT - Three Fold Way

weight	$\pi(\lambda)$	Spectral Law	$F(\lambda)$
Gaussian	$e^{-\lambda^2/2}$	Semi-circle	$\propto \sqrt{4-\lambda^2}$
Laguerre	$\lambda^\alpha e^{-\lambda/2}$	Marcenko-Pastur	$\propto \frac{\sqrt{(b_+-\lambda)(\lambda-b_-)}}{\lambda}$
	$\alpha = \frac{1}{2}(n-p-1)$		$b_\pm = (1 \pm \sqrt{c})^2$
Jacobi	$\lambda^a(1-\lambda)^b$	Wachter	$\propto \frac{\sqrt{(b_+-\lambda)(\lambda-b_-)}}{\lambda(1-\lambda)}$
	$a = (n_1-p-1)/2$ $b = (n_2-p-1)/2$		$b_\pm = c_2(1 \pm r)^2/(r \pm c_2)^2$ $r = \sqrt{c_1+c_2-c_1c_2}$

Warmup: Bulk Distribution (Wachter)

Spectral density of limit $F(d\lambda) = \lim p^{-1} \sum_i \delta_{(n_2/n_1)\lambda_i(E^{-1}H)}$:

$$f(\lambda) = \frac{1-c_2}{2\pi} \frac{\sqrt{(b_+ - \lambda)(\lambda - b_-)}}{\lambda(c_1 + c_2\lambda)}$$

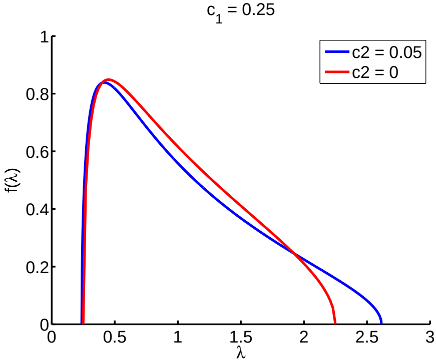
Let $r = \sqrt{c_1 + c_2 - c_1 c_2}$.

Support limits:

$$b_\pm = \left(\frac{1 \pm r}{1 - c_2} \right)^2$$

$$\to (1 \pm \sqrt{c_1})^2$$

as $c_2 \to 0$. [Marčenko-Pastur]



Outline

James' Five Fold Way: review classical examples, focus on

- eigenvalue equations
- low rank alternatives

Common structure in joint density of eigenvalues

- null case: random matrix connection
- alternatives: matrix hypergeometric functions
- integral representations for low rank cases

Hypergeometric functions

Scalar:
$$ {}_{\mathrm{p}}F_{\mathrm{q}}(a, b; x) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_{\mathrm{p}})_k}{(b_1)_k \cdots (a_{\mathrm{q}})_k} \frac{x^k}{k!} $$

Single matrix argument: S symmetric, usually diagonal

$$ {}_{\mathrm{p}}F_{\mathrm{q}}(a, b; S) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \cdots (a_{\mathrm{p}})_\kappa}{(b_1)_\kappa \cdots (a_{\mathrm{q}})_\kappa} \frac{C_\kappa(S)}{k!} $$

$$ {}_0F_0(S) = e^{\mathrm{tr} S}, \qquad {}_1F_0(a, S) = |I - S|^{-a} $$

Two matrix arguments: S, T symmetric

$$ {}_{\mathrm{p}}F_{\mathrm{q}}(a, b; S, T) = \int_{O(p)} {}_{\mathrm{p}}F_{\mathrm{q}}(a, b; SUTU')(dU) $$

Joint density: rewriting James

$\Lambda = \mathrm{diag}(\lambda_i)$ eigenvalues of $p^{-1/2}G, n^{-1}H$ or $(E+H)^{-1}H$.

$\Phi = \Gamma\Theta\Gamma'$ low rank alternative, $\Theta = \mathrm{diag}(\theta_1, \ldots, \theta_r)$.

General structure for joint density of eigenvalues:

$$p(\lambda; \Theta) = \rho(\alpha; \Psi) \cdot {}_{\mathrm{p}}F_{\mathrm{q}}(a, b; c\Psi, \Lambda)\, \pi(\lambda)\Delta(\lambda)$$

$$\Psi(\Theta) = \begin{cases} \Theta(I+\Theta)^{-1} & q = 0 \\ \Theta & q = 1, \mathrm{SMD} \end{cases}$$

Under $H_0 : \Theta = 0$, have $\Psi = 0$ and $\rho \cdot {}_{\mathrm{p}}F_{\mathrm{q}} = 1$.

$\rho(\alpha, \Psi) = \exp\{-\alpha \sum_1^r \beta_{\mathrm{p,q}}(\theta_i)\}$ has product structure.

Joint Density - parameter table

$$p(\lambda;\Theta) = \rho(\alpha;\Psi)\,{}_{\mathrm{p}}F_{\mathrm{q}}(a,b;c\Psi,\Lambda)\,\pi(\lambda)\Delta(\lambda)$$

		$2a$	$2b$	$2c$	2α
SMD	${}_0F_0$	.	.	p	$p/4$
PCA	${}_0F_0$	.	.	n	n
REG_0	${}_0F_1$	.	n	$n^2/2$	n
SigDet	${}_1F_0$	n_1+n_2	.	2	n_1
REG	${}_1F_1$	n_1+n_2	$p \vee n_1$	n_1	n_1
CCA	${}_2F_1$	n_1+n_2, n_1+n_2	n_1	2	n_1+n_2

p

$$\rho(\alpha,\Psi) = \begin{cases} |I-\Psi|^{\alpha} & \mathrm{p}>\mathrm{q} \text{ or } \mathrm{p}=\mathrm{q}=0 \\ e^{-\alpha \mathrm{tr}\Psi^2} & \text{SMD} \\ e^{-\alpha \mathrm{tr}\Psi} & \text{o/w} \end{cases}$$

Contour Integral Representation - Rank 1

Suppose $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_r)$, and let $m = \frac{r}{2} - 1$.

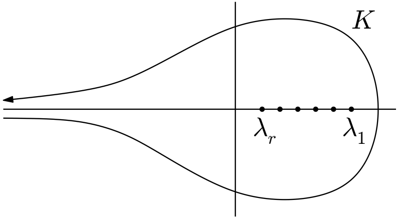
Assume $\Psi = \mathrm{diag}(\psi, 0, \ldots, 0)$ has rank one. Then DJ (14)

$$ {}_{\mathrm{p}}F_{\mathrm{q}}(a, b; \Psi, \Lambda) = \frac{c_m}{\psi^m} \frac{1}{2\pi i} \int_K {}_{\mathrm{p}}F_{\mathrm{q}}(a-m, b-m; \psi s) \prod_{i=1}^{r} (s-\lambda_i)^{-1/2} ds $$

Univariate integral and ${}_{\mathrm{p}}F_{\mathrm{q}}$!

$$ c_m = \frac{\Gamma(m+1)}{\rho_m(a-m, b-m)} $$

$$ \rho_k(a, b) = \frac{(a_1)_k \cdots (a_{\mathrm{p}})_k}{(b_1)_k \cdots (a_{\mathrm{q}})_k} $$



rank$(\Psi) = r$: r-fold contour integral Passemier-McKay-Chen (14).

Double Scaling

$$p/n_1 \to c_1 > 0, \qquad\qquad p/n_2 \to c_2 \in [0, 1)$$

- today: $c_1, c_2 < 1$, [but $c_1 \geq 1$ is relevant]
- $c_2 = 0 \quad \leftrightarrow \quad$ single Wishart H
- $c_2 \geq 1$ is a singular case – E not invertible

Remarks for statisticians:

- high dimensional statistics ...
- new perspectives on 'small p'
- role of c_2: even small $c_2 > 0$ can have quite large effect:
  - $c_2 = .03 \quad \Rightarrow \quad \kappa(E) = \lambda_1(E)/\lambda_p(E) \doteq 2.0$
  - $c_2 = .10 \quad \Rightarrow \quad \kappa(E) \doteq 3.7$

Double Scaling and small p

$$\log\det H \approx \begin{cases} N(\,0, 2p/n\,) & p \text{ fixed} \\ N(\,pd_c, \ell_c\,) & p/n \to c \quad \text{(Bai-Silverstein)} \end{cases}$$

$$\left[\qquad \ell_c = 2\log(1-c)^{-1} \qquad\qquad d_c = \frac{1-c}{2c}\ell_c - 1 \qquad \right]$$

Double scaling (pBaiS) gives better approximation for $p = 2$:

n	p	qtile	pBaiS	pFix
100	2	0.90	0.899	0.923
100	2	0.95	0.951	0.965

[10000 reps, 2SE $\leq$ 0.006]

Outline

James' Five Fold Way: review classical examples, focus on

- eigenvalue equations
- low rank alternatives

Common structure in joint density of eigenvalues

- null case: random matrix connection
- alternatives: matrix hypergeometric functions
- integral representations for low rank cases

Some consequences: [Alexei]

- phase transition for largest eigenvalue
- likelihood ratios below & above the transition

Conclusion

James' (1964) representations $\Rightarrow$

$$p(\lambda; \Theta) = \rho(\alpha; \Psi)\, {}_{\mathrm{p}}F_{\mathrm{q}}(a, b; c\Psi, \Lambda)\, \pi(\lambda)\Delta(\lambda)$$

- powerful systematization for multivariate distributions
- leads to simple approximations in low rank cases *via* double scaling limit

THANK YOU!

Same year as James(1964)

PHYSICAL REVIEW VOLUME 134, NUMBER 5B 8 JUNE 1964

Higher Order Spacing Distributions for a Class of Unitary Ensembles*

DAVID FOX AND PETER B. KAHN
State University of New York at Stony Brook, Stony Brook, New York
(Received 22 January 1964)

We consider the nth-order spacing distribution, $P^n(s)$, in the statistical theory of energy levels of complex systems. Each P^n is written as a sum of multiple integrals over correlation functions. This procedure is used to establish the identity of the spacing distributions for all members of a class of Hamiltonian unitary ensembles. A power-series expansion of $P^n(s)$, valid for all n, is developed.

I. INTRODUCTION

A STATISTICAL theory has been developed[1–6] which has been applied to the problem of level spacing in heavy nuclei in a region of the excitation spectrum where the level density is approximately constant over, say, a hundred levels. A suitably chosen ensemble of N-dimensional Hamiltonian matrices is introduced, and one studies the distribution of the eigenvalues of ensemble members.

We are interested in developing approximation procedures for the calculation of energy level spacing distributions for a class of Hamiltonian matrix ensembles. To date, nearest-neighbor spacing distributions, $P^0(s)$ determine the distribution of elements in the Hamiltonian matrix ensembles.

Members of the class of Hamiltonian ensembles in which $f(x_1,\cdots,x_n)$ is a product, $\prod_i [g(x_i)]^2$, have been extensively studied.[1–3,5,6,8] For example, the choices

$$\begin{aligned}
[g(x)]^2 &= \exp(-x^2) && -\infty < x < \infty\,,\\
&= (1-x)^\mu(1+x)^\nu && \mu,\,\nu > -1;\quad |x| \leq 1\,,\\
&= x^\alpha e^{-x} && \alpha > -1;\quad 0 \leq x < \infty\,,\\
&= 1 && x = e^{i\theta},\quad 0 \leq \theta \leq 2\pi\,,
\end{aligned}$$

lead to the so-called Gaussian, Jacobi, Laguerre, and circular ensembles, respectively.[9] The circular[5] and Gaussian[4–6] ensembles have been shown to have identical nearest-neighbor spacing distributions for $\beta = 1, 2, 4$.