Random Vandermonde Matrices and Covariance Estimates

Gabriel H. Tucci

Bell Labs

Erwin Schrödinger International Institute

April 2011

Random Vandermonde Matrices

Let V_n be the $n \times m$ random matrix of the form

$$V_n = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & \cdots & 1 \\ e^{2\pi i \theta_1} & \cdots & e^{2\pi i \theta_m} \\ \vdots & \ddots & \vdots \\ e^{2\pi i (n-1)\theta_1} & \cdots & e^{2\pi i (n-1)\theta_m} \end{bmatrix}$$

where the $\{\theta_1, \ldots, \theta_m\}$ are random variables in $[0, 1]$.

Random Vandermonde Matrices

Let V_n be the $n \times m$ random matrix of the form

$$V_n = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & \dots & 1 \\ e^{2\pi i \theta_1} & \dots & e^{2\pi i \theta_m} \\ \vdots & \ddots & \vdots \\ e^{2\pi i (n-1)\theta_1} & \dots & e^{2\pi i (n-1)\theta_m} \end{bmatrix}$$

where the $\{\theta_1, \dots, \theta_m\}$ are random variables in $[0, 1]$.

We are interested in the limit eigenvalue distribution of the random matrix $V_n^* V_n$ under the conditions that:

- The phases are i.i.d. in the interval $[0, 1]$ with probability distribution ν
- $\lim_{n\to\infty} \frac{n}{m} = c \in (0, \infty)$

d-fold Random Vandermonde Matrices

More generally, we will be studying the random matrices $V_n^{(d)}$ for $d \geq 1$

Let $\ell = (\ell_1, \ldots, \ell_d) \in \{0, 1, \ldots, n-1\}^d$ and consider the function

$$f(\ell) = \sum_{j=1}^{d} n^{j-1} \ell_j \quad \text{a bijection to } \{0, 1, \ldots, n^d - 1\}$$

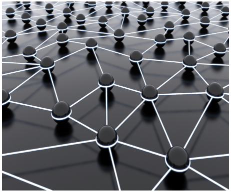
Consider $x_1, \ldots, x_m$ random vectors in $[0,1]^d$ independent and identically distributed.

We know define the matrix

$$V^{(d)}_{f(\ell),q} = n^{-\frac{d}{2}} \exp(2\pi i \langle \ell, x_q \rangle)$$

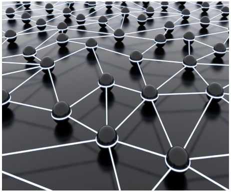
We are interested in the limit eigenvalue distribution of the random matrix $(V_n^{(d)})^*(V_n^{(d)})$ under the condition $\lim_{n\to\infty} \frac{n^d}{m} = c \in (0, \infty)$

Why do we care?



- Let m wireless sensors measure the value of a spatially finite physical field (air temperature, pressure, etc) defined over a d dimensional compact space
    - One can think of sensor nodes randomly deployed over the geographical region
    - We want to reconstruct the field from a collection of samples that are noisy
    - Measure reconstruction accuracy by the MSE

Why do we care?



- Let m wireless sensors measure the value of a spatially finite physical field (air temperature, pressure, etc) defined over a d dimensional compact space
  - One can think of sensor nodes randomly deployed over the geographical region
  - We want to reconstruct the field from a collection of samples that are noisy
  - Measure reconstruction accuracy by the MSE

The sensors $X = \{x_1, x_2, \dots, x_m\} \in [0,1]^d$. By truncating the field's Fourier series expansion:

$$s(x) \approx n^{-\frac{d}{2}} \sum_{\ell} a_\ell \; e^{2\pi j \langle x, \ell \rangle}$$

where n is the approximate one-sided bandwidth (per dimension) of the field $\ell = (\ell_1, \dots, \ell_d)$.

Multidimensional Signal Processing and Sensor Networks cont'd

We can write the vector s as a function of the field spectrum:

$$s = \beta_{n,m}^{-\frac{1}{2}} V^* a$$

where V is the $n^d \times m$ d-fold Vandermonde matrix with sensors iid and $\beta_{n,m} = \frac{n^d}{m}$

Multidimensional Signal Processing and Sensor Networks cont'd

We can write the vector s as a function of the field spectrum:

$$s = \beta_{n,m}^{-\frac{1}{2}} V^* a$$

where V is the $n^d \times m$ d-fold Vandermonde matrix with sensors iid and $\beta_{n,m} = \frac{n^d}{m}$

Let

$$p = (s(x_1) + n_1, s(x_2) + n_2, \dots, s(x_m) + n_m)$$

be the sensor measurements.

Using the Sherman-Morrison-Woodbury identity, we can obtain the MSE as

$$MSE^{(m)} = \mathrm{snr}^{-1} \beta_{n,m} \cdot \mathbb{E}\Big(\mathrm{tr}\Big\{(V_n^{(d)})(V_n^{(d)})^* + \mathrm{snr}^{-1}\beta_{n,m}\Big\}\Big)$$

The bottom line is that we need to understand the spectrum of $(V_n^{(d)})^*(V_n^{(d)})$ when n and m are big

Some natural polytopes

Let $k \geq 2$ and consider a partition $\rho \in \mathcal{P}(k)$

$$\rho = \{B_1, \dots, B_{|\rho|}\}$$

with $|\rho|$ blocks.

Some natural polytopes

Let $k \geq 2$ and consider a partition $\rho \in \mathcal{P}(k)$

$$\rho = \{B_1, \ldots, B_{|\rho|}\}$$

with $|\rho|$ blocks.

For each $i \in \{1, 2, \ldots, k\}$ consider a variable $0 \leq x_i \leq 1$ and for each block j consider the equation

$$E_j \ : \ \sum_{i \in B_j} x_{i-1} = \sum_{i \in B_j} x_i$$

Some natural polytopes

Let $k \geq 2$ and consider a partition $\rho \in \mathcal{P}(k)$

$$\rho = \{B_1, \ldots, B_{|\rho|}\}$$

with $|\rho|$ blocks.

For each $i \in \{1, 2, \ldots, k\}$ consider a variable $0 \leq x_i \leq 1$ and for each block j consider the equation

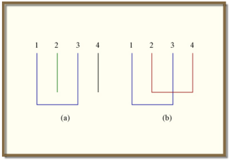
$$E_j \;:\; \sum_{i \in B_j} x_{i-1} = \sum_{i \in B_j} x_i$$

The solution set has $k + 1 - |\rho|$ free variables. We define

$$K_\rho = \text{volume of the solution set in } [0,1]^{k+1-|\rho|}$$

Example

Consider $k = 4$ and the following two partitions :



- Number of free variables is 2

$$E_1 \ : x_1 + x_3 = x_2 + x_4 \quad \& \quad E_2 \ : x_2 = x_1 \quad \& \quad E_3 \ : x_4 = x_3$$

  then $x_1 = x_2$ and $x_3 = x_4$ therefore

$$K_\rho = 1$$

- The second partition is $\rho = \{\{1, 3\}, \{2, 4\}\}$, the number of free variables is 3 and

$$E_1 = E_2 \ : \ x_1 + x_3 = x_2 + x_4$$

  then

$$K_\rho = \mathrm{vol}\Big(\{(x_1, x_2, x_3) \in [0, 1]^3 \ : \ 0 \leq x_1 + x_3 - x_2 \leq 1\}\Big) = \frac{2}{3}$$

Moments' formulas

Theorem (Ryan and Debbah)

Assume that ν has a continuous density p. Then the asymptotic p-th moment of $V_n^* V_n$

$$m_k = \lim_{m \to \infty} \mathbb{E}\left[\mathrm{tr}_m (V_n^* V_n)^k)\right] = \sum_{\rho \in \mathcal{P}(k)} K_{\rho,\nu} c^{|\rho|-1}$$

exists when $\frac{n}{m} \to c \in (0, \infty)$ where

$$K_{\rho,\nu} = K_\rho \cdot \int_0^1 p(x)^{|\rho|} \, dx$$

and $|\rho|$ is the number of blocks of ρ.

Moments' formulas

Theorem (Ryan and Debbah)

Assume that ν *has a continuous density* p*. Then the asymptotic* p*-th moment of* $V_n^* V_n$

$$m_k = \lim_{m \to \infty} \mathbb{E}\Big[\mathrm{tr}_m\big((V_n^* V_n)^k\big)\Big] = \sum_{\rho \in \mathcal{P}(k)} K_{\rho,\nu} c^{|\rho|-1}$$

exists when $\frac{n}{m} \to c \in (0, \infty)$ *where*

$$K_{\rho,\nu} = K_\rho \cdot \int_0^1 p(x)^{|\rho|} \, dx$$

and $|\rho|$ *is the number of blocks of* ρ*.*

- For every partition ρ we have $0 < K_\rho \leq 1$ and corresponds to the uniform distribution

Moments' formulas

Theorem (Ryan and Debbah)

Assume that ν has a continuous density p. Then the asymptotic p-th moment of $V_n^* V_n$

$$m_k = \lim_{m \to \infty} \mathbb{E}\left[\mathrm{tr}_m (V_n^* V_n)^k)\right] = \sum_{\rho \in \mathcal{P}(k)} K_{\rho,\nu} c^{|\rho|-1}$$

exists when $\frac{n}{m} \to c \in (0, \infty)$ where

$$K_{\rho,\nu} = K_\rho \cdot \int_0^1 p(x)^{|\rho|}\, dx$$

and $|\rho|$ is the number of blocks of ρ.

- For every partition ρ we have $0 < K_\rho \leq 1$ and corresponds to the uniform distribution
- K_ρ is a rational number for every $\rho \in \mathcal{P}(k)$

Moments' formulas

Theorem (Ryan and Debbah)

Assume that ν has a continuous density p. Then the asymptotic p-th moment of $V_n^* V_n$

$$m_k = \lim_{m\to\infty} \mathbb{E}\left[\mathrm{tr}_m(V_n^* V_n)^k)\right] = \sum_{\rho\in\mathcal{P}(k)} K_{\rho,\nu} c^{|\rho|-1}$$

exists when $\frac{n}{m} \to c \in (0,\infty)$ where

$$K_{\rho,\nu} = K_\rho \cdot \int_0^1 p(x)^{|\rho|}\, dx$$

and $|\rho|$ is the number of blocks of ρ.

- For every partition ρ we have $0 < K_\rho \leq 1$ and corresponds to the uniform distribution
- K_ρ is a rational number for every $\rho \in \mathcal{P}(k)$
- $K_\rho = 1$ if and only if ρ is non–crossing

Formula to compute K_ρ

Proposition (T., Whiting)

For each partition $\rho \in \mathcal{P}(k)$ with

$$\rho = \{B_1, \dots, B_{|\rho|}\}$$

let $\rho : \{1, 2, \dots, k\} \to \{1, 2, \dots, |\rho|\}$ given by $\rho(i) = j$ if i is in block j. Then

$$K_\rho = \int_{\mathbb{R}^{(|\rho|-1)}} G_\rho(t_1, \dots, t_{|\rho|})\, dt_1 \dots dt_{|\rho|-1}$$

where

$$G_\rho(t_1, \dots, t_{|\rho|}) := \prod_{j=1}^{k} G(t_{\rho(j)} - t_{\rho(j+1)})$$

and

$$G(t) = \frac{\sin(\pi t)}{\pi t}$$

d-fold case

Consider a random Vandermonde matrix defined as before with phase distribution ν such that:

d-fold case

Consider a random Vandermonde matrix defined as before with phase distribution ν such that:

- ν is abs. cont. wrt Lebesgue measure with density p

d-fold case

Consider a random Vandermonde matrix defined as before with phase distribution ν such that:

- ν is abs. cont. wrt Lebesgue measure with density p
- $p \in \cap_{r \geq 1} L^r([0,1]^d)$

d-fold case

Consider a random Vandermonde matrix defined as before with phase distribution ν such that:

- ν is abs. cont. wrt Lebesgue measure with density p
- $p \in \cap_{r \geq 1} L^r([0,1]^d)$
- other extra condition

d-fold case

Consider a random Vandermonde matrix defined as before with phase distribution ν such that:

- ν is abs. cont. wrt Lebesgue measure with density p
- $p \in \cap_{r\geq 1} L^r([0,1]^d)$
- other extra condition

Theorem (T., Whiting)

For every $d \geq 1$ we have that

$$m_{\nu,k}^{(d)} = \lim_{n\to\infty} \mathbb{E}\Big[\mathrm{tr}_m\Big\{\Big((V_\nu^{(d)})^*(V_\nu^{(d)})\Big)^k\Big\}\Big] = \sum_{\rho\in\mathcal{P}(k)} K_{\rho,\nu}^d c^{|\rho|-1}$$

Moreover, there exists a unique limit measure $\mu_{\nu,c}^{(d)}$ with these moments and $\mu_{\nu,c}^{(d)}$ has unbounded support.

Note that for the uniform distribution on $[0,1]^d$

$$m_k^{(d)} = \sum_{\rho\in\mathcal{P}(k)} K_\rho^d c^{|\rho|-1} \xrightarrow[d\to\infty]{} \sum_{\rho\in\mathcal{P}(k)} c^{|\rho|-1}$$

which is the k-th moment of the Marchenko-Pastur distribution!

Maximum eigenvalue

Let V be a square $n \times n$ random Vandermonde matrix with phase distribution ν with density p

$$\{\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n\}$$

the e-values of V^*V

Maximum eigenvalue

Let V be a square $n \times n$ random Vandermonde matrix with phase distribution ν with density p

$$\{\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n\}$$

the e-values of V^*V

- $\lambda_n \in [0, n]$

Maximum eigenvalue

Let V be a square $n \times n$ random Vandermonde matrix with phase distribution ν with density p

$$\{\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n\}$$

the e-values of V^*V

- $\lambda_n \in [0, n]$
- $\lambda_n = n$ iff $\theta_1 = \theta_2 = \ldots = \theta_n$

Maximum eigenvalue

Let V be a square $n \times n$ random Vandermonde matrix with phase distribution ν with density p

$$\{\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n\}$$

the e-values of V^*V

- $\lambda_n \in [0, n]$
- $\lambda_n = n$ iff $\theta_1 = \theta_2 = \ldots = \theta_n$

What is the behaviour of $\mathbb{E}[\lambda_n]$ as a function of n?

Maximum eigenvalue cont'd

Theorem (T., Whiting)

Assume $p \in L^\infty([0,1])$, *then for every* $\epsilon > 0$ *and every* $u \geq 0$ *we have that*

$$\mathbb{P}\Big(\lambda_n \geq (C+\epsilon)\log n + u\Big) \leq K\frac{e^{-u}}{n^\epsilon}$$

where $K > 0$ *is a constant independent on* ϵ, u *and* n *and* $C = \big(4\pi\|p\|_\infty(e-1)+1\big)$.

Maximum eigenvalue cont'd

Theorem (T., Whiting)

Assume $p \in L^\infty([0,1])$, then for every $\epsilon > 0$ and every $u \geq 0$ we have that

$$\mathbb{P}\Big(\lambda_n \geq (C+\epsilon)\log n + u\Big) \leq K\frac{e^{-u}}{n^\epsilon}$$

where $K > 0$ is a constant independent on ϵ, u and n and $C = \big(4\pi\|p\|_\infty(e-1)+1\big)$.

Corollary

$$\mathbb{E}(\lambda_n) \leq \Big(4\pi\|p\|_\infty(e-1)+1\Big)\log n + o(1)$$

Maximum eigenvalue cont'd

Theorem (T., Whiting)

Let ν be a phase distribution which is abs. cont. wrt Lebesgue measure. For any $0 < \alpha < 1$

$$\mathbb{P}\left(\lambda_n \geq \frac{\alpha \cdot \log n}{\log\log n}\right) = 1 - o(1)$$

Corollary

$$\mathbb{E}(\lambda_n) \geq \frac{\alpha \cdot \log n}{\log\log n}(1 - o(1))$$

We have similar results for the d-fold case.

Questions:

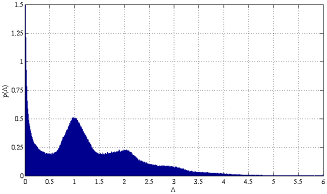


Figure: Spectral distribution for $d = 1$ and $n = m = 1000$

For simplicity consider the square case with uniform distribution.

Questions:

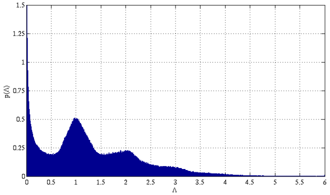


Figure: Spectral distribution for $d = 1$ and $n = m = 1000$

For simplicity consider the square case with uniform distribution.

- Does the limit distribution have an atom at zero? We can prove that $(V_n^* V_n)^{-1}$ has no finite moments.

Questions:

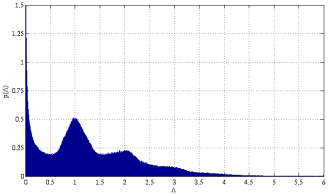


Figure: Spectral distribution for $d = 1$ and $n = m = 1000$

For simplicity consider the square case with uniform distribution.

- Does the limit distribution have an atom at zero? We can prove that $(V_n^* V_n)^{-1}$ has no finite moments.
- Can we find a lower bound for K_ρ ?

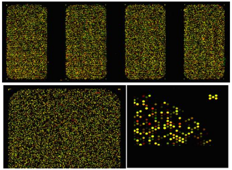
Conjecture: For every $\rho \in \mathcal{P}(k)$ with $|\rho|$ blocks

$$K_\rho \geq \left(\frac{6(|\rho| - 1)}{\pi k} \right)^{\frac{|\rho|-1}{2}}$$

Multivariate Statistics: Covariance Estimates

We have m random variables

Correlation between random variables? from n observations



- Weather Forecast: Sensor Network where each sensor is measuring temperature, pressure or other field property
- Military Applications (adaptative sensor array)
- Gene Expression Arrays
- High Dimensional Problems with $n \leq m$

We can't perform as many observations as the number of variables!

Mathematical Formulation

Let Σ be the true $m \times m$ covariance matrix

$\{x_1, x_2, \ldots, x_n\} \sim CN(0, \Sigma)$ observations or measurements

Sample covariance Matrix: $K_x = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^* \longrightarrow \Sigma$

Mathematical Formulation

Let Σ be the true $m \times m$ covariance matrix

$\{x_1, x_2, \ldots, x_n\} \sim CN(0, \Sigma)$ observations or measurements

Sample covariance Matrix: $K_x = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^* \longrightarrow \Sigma$

We want to estimate Σ or Σ^{-1} from the n samples ($n \leq m$)

- K_x has at least $m - n$ zero eigenvalues

Mathematical Formulation

Let Σ be the true $m \times m$ covariance matrix

$\{x_1, x_2, \ldots, x_n\} \sim CN(0, \Sigma)$ observations or measurements

Sample covariance Matrix: $K_x = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^* \longrightarrow \Sigma$

We want to estimate Σ or Σ^{-1} from the n samples ($n \leq m$)

- K_x has at least $m - n$ zero eigenvalues
- In many applications we need Σ^{-1} (e.g. linear estimation)

Mathematical Formulation

Let Σ be the true $m \times m$ covariance matrix

$\{x_1, x_2, \ldots, x_n\} \sim CN(0, \Sigma)$ observations or measurements

Sample covariance Matrix: $K_x = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^* \longrightarrow \Sigma$

We want to estimate Σ or Σ^{-1} from the n samples ($n \leq m$)

- K_x has at least $m - n$ zero eigenvalues
- In many applications we need Σ^{-1} (e.g. linear estimation)
- regularization method to find an invertible estimate

Classical solution : ridge regression method or diagonal loading

$$\alpha K_x + \beta I_m$$

where $\alpha, \beta > 0$

New Idea: Dimensional Reduction (Marzetta, Simon, T.)

Estimate Σ or Σ^{-1} from $\{x_1, x_2, \ldots, x_n\}$

New Idea: Dimensional Reduction (Marzetta, Simon, T.)

Estimate Σ or Σ^{-1} from $\{x_1, x_2, \ldots, x_n\}$

$$K_x = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^*$$

New Idea: Dimensional Reduction (Marzetta, Simon, T.)

Estimate Σ or Σ^{-1} from $\{x_1, x_2, \ldots, x_n\}$

$$K_x = \frac{1}{n}\sum_{i=1}^n x_i x_i^*$$

Fix a parameter $p \leq n$ (to be tuned later) and consider the Stiefel manifold

$$\Omega_{p,m} := \left\{ \Phi \in M_{p,m}(\mathbb{C}) \, : \, \Phi\Phi^* = I_p \right\}$$

with isotropically random prob. measure

New Idea: Dimensional Reduction (Marzetta, Simon, T.)

Estimate Σ or Σ^{-1} from $\{x_1, x_2, \dots, x_n\}$

$$K_x = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^*$$

Fix a parameter $p \leq n$ (to be tuned later) and consider the Stiefel manifold

$$\Omega_{p,m} := \left\{ \Phi \in M_{p,m}(\mathbb{C}) \ : \ \Phi\Phi^* = I_p \right\}$$

with isotropically random prob. measure

Haar Integration Method

$$\mathrm{cov}_p(K_x) := \mathbb{E}\Big(\Phi^* (\Phi K_x \Phi^*) \Phi \Big)$$

$$\mathrm{invcov}_p(K_x) := \mathbb{E}\Big(\Phi^* (\Phi K_x \Phi^*)^{-1} \Phi \Big)$$

Mathematical properties of cov_p

K_x is $m \times m$, rank $n \leq m$ (sample covariance matrix)

$$\mathrm{cov}_p(K_x) := \mathbb{E}\Big(\Phi^*(\Phi K_x \Phi^*)\Phi\Big)$$

Mathematical properties of cov_p

K_x is $m \times m$, rank $n \leq m$ (sample covariance matrix)

$$\mathrm{cov}_p(K_x) := \mathbb{E}\Big(\Phi^*(\Phi K_x \Phi^*)\Phi\Big)$$

- Closed–form expression for the expectation

Mathematical properties of cov_p

K_x is $m \times m$, rank $n \leq m$ (sample covariance matrix)

$$\mathrm{cov}_p(K_x) := \mathbb{E}\Big(\Phi^*(\Phi K_x \Phi^*)\Phi\Big)$$

- Closed–form expression for the expectation
- $\mathrm{cov}_p(K_x) = \frac{p}{m(m^2-1)}\Big((mp-1)\cdot K_x + (m-p)\cdot \mathrm{Tr}(K_x)\cdot I_m\Big)$

Mathematical properties of cov_p

K_x is $m \times m$, rank $n \leq m$ (sample covariance matrix)

$$\mathrm{cov}_p(K_x) := \mathbb{E}\Big(\Phi^*(\Phi K_x \Phi^*)\Phi\Big)$$

- Closed–form expression for the expectation
- $\mathrm{cov}_p(K_x) = \frac{p}{m(m^2-1)}\Big((mp-1) \cdot K_x + (m-p) \cdot \mathrm{Tr}(K_x) \cdot I_m\Big)$
- Equivalent to diagonal loading!

Basic properties of invcov_p

Decompose K_x as

$$K_x = U\,D\,U^* \quad \text{where} \quad D = \mathrm{diag}(d_1, \ldots, d_n, 0_{m-n}) \quad \text{and} \quad U \text{ unitary}$$

Basic properties of invcov_p

Decompose K_x as

$$K_x = U\,D\,U^* \quad \text{where} \quad D = \operatorname{diag}(d_1, \ldots, d_n, 0_{m-n}) \quad \text{and} \quad U \text{ unitary}$$

Then

$$\operatorname{invcov}_p(K_x) = U\,\mathbb{E}\Big(\Phi^*(\Phi D \Phi^*)^{-1}\Phi\Big)\,U^*$$

Basic properties of invcov_p

Decompose K_x as

$$K_x = U\,D\,U^* \quad \text{where} \quad D = \mathrm{diag}(d_1, \dots, d_n, 0_{m-n}) \quad \text{and} \quad U \text{ unitary}$$

Then

$$\mathrm{invcov}_p(K_x) = U\,\mathbb{E}\Big(\Phi^*(\Phi D \Phi^*)^{-1}\Phi\Big)\,U^*$$

- This method preserves the eigenvectors
- $\mathrm{invcov}_p = \mathbb{E}\Big(Z^*(ZK_xZ^*)^{-1}Z : Z \;\; p \times m \text{ stand. Gaussian}\Big)$
- Tom and Steve have formulas for the entries of invcov_p but are too complicated!
- $\mathbb{E}\Big(\Phi^*(\Phi D \Phi^*)^{-1}\Phi\Big)$ is diagonal and moreover

$$\mathbb{E}\Big(\Phi^*(\Phi D \Phi^*)^{-1}\Phi\Big) = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n, \mu I_{m-n})$$

- We obtained asymptotic formulas for the entries λ_k's and μ (using Free Probability)
- The e-values $\lambda_k = \lambda_k(d_1, \dots, d_n)$ are a function of the non–zero entries. We also have a functional equation.

Basic properties of invcov_p

Decomposing $Z = [X, Y]$ where X is $p \times n$ and Y is $p \times (m - n)$ independent standard Gaussian then

$$\mathbb{E}\Big[Z^*\Big(Z\begin{bmatrix} D_n & 0 \\ 0 & 0_{m-n} \end{bmatrix}Z^*\Big)^{-1}Z\Big] = \begin{bmatrix} \mathbb{E}(X^*(XD_nX^*)^{-1}X) & 0 \\ 0 & \mathbb{E}(Y^*(XD_nX^*)^{-1}Y) \end{bmatrix}$$

Let $\{d_1, d_2, \ldots, d_n, 0, 0, \ldots, 0\}$ be the e-values of K_x

Basic properties of invcov_p

Decomposing $Z = [X, Y]$ where X is $p \times n$ and Y is $p \times (m-n)$ independent standard Gaussian then

$$\mathbb{E}\Big[Z^*\Big(Z \left[\begin{array}{cc} D_n & 0 \\ 0 & 0_{m-n} \end{array}\right] Z^*\Big)^{-1} Z\Big] = \left[\begin{array}{cc} \mathbb{E}(X^*(XD_nX^*)^{-1}X) & 0 \\ 0 & \mathbb{E}(Y^*(XD_nX^*)^{-1}Y) \end{array}\right]$$

Let $\{d_1, d_2, \ldots, d_n, 0, 0, \ldots, 0\}$ be the e-values of K_x

The matrix invcov_p has eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_n, \mu_n, \mu_n, \ldots, \mu_n\}$ where

$$\mathbb{E}(Y^*(XD_nX^*)^{-1}Y) = \mathrm{Tr}\; \mathbb{E}\left[(XD_nX^*)^{-1}\right] \cdot I_{m-n} = \mu_n I_{m-n}$$

Basic properties of invcov_p

Decomposing $Z = [X, Y]$ where X is $p \times n$ and Y is $p \times (m-n)$ independent standard Gaussian then

$$\mathbb{E}\Big[Z^*\Big(Z\left[\begin{array}{cc} D_n & 0 \\ 0 & 0_{m-n}\end{array}\right]Z^*\Big)^{-1}Z\Big] = \left[\begin{array}{cc} \mathbb{E}(X^*(XD_nX^*)^{-1}X) & 0 \\ 0 & \mathbb{E}(Y^*(XD_nX^*)^{-1}Y)\end{array}\right]$$

Let $\{d_1, d_2, \ldots, d_n, 0, 0, \ldots, 0\}$ be the e-values of K_x

The matrix invcov_p has eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_n, \mu_n, \mu_n, \ldots, \mu_n\}$ where

$$\mathbb{E}(Y^*(XD_nX^*)^{-1}Y) = \mathrm{Tr}\ \mathbb{E}\left[(XD_nX^*)^{-1}\right] \cdot I_{m-n} = \mu_n I_{m-n}$$

Assume $D_n \longrightarrow \nu$ in distribution and $\lim \frac{n}{p} = \beta$ then $\frac{1}{p}XD_nX^*$ converges in distribution also to a measure γ

$$\frac{1}{p}XD_nX^* \longrightarrow \gamma.$$

Properties of invcov_p

Using classical relations between the Cauchy transform of ν and γ it is not difficult to prove

$$\lim_{n\to\infty} \mu_n = \mu := \int_0^\infty t^{-1}\, d\gamma(t)$$

and

$$\lim_{n\to\infty} \frac{n-p}{n} = \frac{\beta-1}{\beta} = \int_0^\infty \frac{1}{1+\mu t}\, d\nu(t).$$

Asymptotics : μ is uniquely determined by the relation

$$\frac{n-p}{n} \approx \frac{1}{n}\sum_{k=1}^{n} \frac{1}{1+\mu d_k} \qquad : \qquad \mu_n \approx \mu \quad \text{for } n \text{ large !}$$

Basic properties of invcov_p cont'd

It is not difficult to prove that $\lambda_k = \lambda_k(d_1, \ldots, d_n)$ is a function of the non–zero entries. We have asymptotic formulas also!

Define $\Lambda_p(D_n) := \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ then

$$\mathbb{E}\Big(\Phi^*(\Phi D \Phi^*)^{-1}\Phi\Big) = \mathrm{diag}(\Lambda_p(D_n), \mu I_{m-n})$$

Functional Equation

$$\Lambda_p(D_n)D_n + \Lambda_{n-p}(D_n^{-1})D_n^{-1} = I_n$$

Basic properties of invcov_p cont'd

It is not difficult to prove that $\lambda_k = \lambda_k(d_1, \ldots, d_n)$ is a function of the non–zero entries. We have asymptotic formulas also!

Define $\Lambda_p(D_n) := \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ then

$$\mathbb{E}\Big(\Phi^*(\Phi D \Phi^*)^{-1}\Phi\Big) = \mathrm{diag}(\Lambda_p(D_n), \mu I_{m-n})$$

Functional Equation

$$\Lambda_p(D_n)D_n + \Lambda_{n-p}(D_n^{-1})D_n^{-1} = I_n$$

Basic properties of invcov_p cont'd

It is not difficult to prove that $\lambda_k = \lambda_k(d_1, \ldots, d_n)$ is a function of the non–zero entries. We have asymptotic formulas also!

Define $\Lambda_p(D_n) := \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ then

$$\mathbb{E}\Big(\Phi^*(\Phi D \Phi^*)^{-1}\Phi\Big) = \mathrm{diag}(\Lambda_p(D_n), \mu I_{m-n})$$

Functional Equation

$$\Lambda_p(D_n)D_n + \Lambda_{n-p}(D_n^{-1})D_n^{-1} = I_n$$

What if we are more ambitious?

Basic properties of invcov_p cont'd

It is not difficult to prove that $\lambda_k = \lambda_k(d_1, \ldots, d_n)$ is a function of the non–zero entries. We have asymptotic formulas also!

Define $\Lambda_p(D_n) := \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ then

$$\mathbb{E}\Big(\Phi^*(\Phi D \Phi^*)^{-1}\Phi\Big) = \mathrm{diag}(\Lambda_p(D_n), \mu I_{m-n})$$

Functional Equation

$$\Lambda_p(D_n)D_n + \Lambda_{n-p}(D_n^{-1})D_n^{-1} = I_n$$

What if we are more ambitious?

Q: Given the matrix K can we compute exactly

$$\mathbb{E}\Big(\Phi^* g(\Phi K \Phi^*)\Phi\Big) \quad \text{for } g \text{ sufficiently nice ?}$$

This will be an estimate for $g(\Sigma)$.

Basic properties of invcov_p cont'd

It is not difficult to prove that $\lambda_k = \lambda_k(d_1, \ldots, d_n)$ is a function of the non–zero entries. We have asymptotic formulas also!

Define $\Lambda_p(D_n) := \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ then

$$\mathbb{E}\Big(\Phi^*(\Phi D \Phi^*)^{-1}\Phi\Big) = \mathrm{diag}(\Lambda_p(D_n), \mu I_{m-n})$$

Functional Equation

$$\Lambda_p(D_n)D_n + \Lambda_{n-p}(D_n^{-1})D_n^{-1} = I_n$$

What if we are more ambitious?

Q: Given the matrix K can we compute exactly

$$\mathbb{E}\Big(\Phi^* g(\Phi K \Phi^*)\Phi\Big) \quad \text{for } g \text{ sufficiently nice ?}$$

This will be an estimate for $g(\Sigma)$.

A: Yes! Using results from representation theory we have exact closed–formed expressions for every g continuous.

Exact Solution

Let

$$\Omega_{p,n} = \{\Phi \in \mathbb{C}^{n\times p} \;:\; \Phi^*\Phi = I_p\}$$

be the Stiefel manifold with the isotropic measure $d\phi$.

Then for any to Hermitian matrices A and B we have that:

$$\int_{\Omega_{p,n}} s_\lambda(\Phi^* B \Phi A) d\phi = \frac{s_\lambda(B) s_\lambda(A)}{s_\lambda(I_n)}$$

where s_λ is the Schur polynomial associated with the partition λ.

$$s_\lambda(X) = \frac{\det(x_i^{n+\lambda_j - j})_{i,j=1}^n}{\det(x_i^{n-j})_{i,j=1}^n}, \qquad \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$$

Exact Solution cont'd

Theorem

Let D be a diagonal matrix of rank n. For any continuous function $f \in C[d_{min}, d_{max}]$

$$\int_{\Omega_{p,n}} \mathrm{Tr}(f(\Phi^* D\Phi))\, d\phi = \sum_{k=0}^{p-1} \frac{(n-(k+1))!}{(p-(k+1))!} \cdot \frac{\det(G_k)}{\det(\Delta(D))} \tag{2.1}$$

where $\Delta(D)$ is the Vandermonde matrix associated to D and G_k is the matrix defined by replacing the $(k+1)^{th}$–row of the Vandermonde matrix by the row

$$\left(I^{(n-p)}(x^{(p-(k+1))} f(x))|_{x=d_1}, \ldots, I^{(n-p)}(x^{(p-(k+1))} f(x))|_{x=d_n} \right).$$

Exact Solution cont'd

Theorem

Let D be a diagonal matrix of rank n. For any continuous function $f \in C[d_{min}, d_{max}]$

$$\int_{\Omega_{p,n}} \mathrm{Tr}(f(\Phi^* D \Phi))\, d\phi = \sum_{k=0}^{p-1} \frac{(n-(k+1))!}{(p-(k+1))!} \cdot \frac{\det(G_k)}{\det(\Delta(D))} \tag{2.1}$$

where $\Delta(D)$ is the Vandermonde matrix associated to D and G_k is the matrix defined by replacing the $(k+1)^{th}$–row of the Vandermonde matrix by the row

$$\left(I^{(n-p)}(x^{(p-(k+1))} f(x))|_{x=d_1}, \ldots, I^{(n-p)}(x^{(p-(k+1))} f(x))|_{x=d_n} \right).$$

$$G_k = \begin{pmatrix} d_1^{n-1} & d_2^{n-1} & \ldots & d_n^{n-1} \\ d_1^{n-2} & d_2^{n-2} & \ldots & d_n^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{(k)} & a_2^{(k)} & \ldots & a_n^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ d_1 & d_2 & \ldots & d_n \\ 1 & 1 & \ldots & 1 \end{pmatrix}$$

Exact Solution cont'd

Theorem

Let D be a diagonal matrix of rank n. For any continuous function $f \in C[d_{min}, d_{max}]$

$$\int_{\Omega_{p,n}} \mathrm{Tr}(f(\Phi^* D\Phi))\, d\phi = \sum_{k=0}^{p-1} \frac{(n-(k+1))!}{(p-(k+1))!} \cdot \frac{\det(G_k)}{\det(\Delta(D))} \tag{2.1}$$

where $\Delta(D)$ is the Vandermonde matrix associated to D and G_k is the matrix defined by replacing the $(k+1)^{th}$–row of the Vandermonde matrix by the row

$$\left(I^{(n-p)}(x^{(p-(k+1))} f(x))|_{x=d_1}, \ldots, I^{(n-p)}(x^{(p-(k+1))} f(x))|_{x=d_n} \right).$$

$$G_k = \begin{pmatrix} d_1^{n-1} & d_2^{n-1} & \ldots & d_n^{n-1} \\ d_1^{n-2} & d_2^{n-2} & \ldots & d_n^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{(k)} & a_2^{(k)} & \ldots & a_n^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ d_1 & d_2 & \ldots & d_n \\ 1 & 1 & \ldots & 1 \end{pmatrix}$$

$$a_i^{(k)} = I^{(n-p)}(x^{(p-k)} f(x))|_{x=d_i}$$

Exact Solution cont'd

Lemma

Let f be a differentiable function in the interval $[d_{min}, d_{max}]$. Then

$$\frac{\partial}{\partial d_k} \int_{\Omega_{p,n}} \mathrm{Tr}\Big(f(\Phi^* D \Phi) \Big) \, d\phi = \left[\int_{\Omega_{p,n}} \Phi f'(\Phi^* D \Phi) \Phi^* \, d\phi \right]_{kk}$$

As an application let us compute explicitly the elements λ_k and μ as a function of the d_i's:

$$\Lambda_p(D) = \int_{\Omega_{p,n}} \Phi(\Phi^* D \Phi)^{-1} \Phi^* \, d\phi = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

with

$$\lambda_k = \frac{\partial}{\partial d_k} \int_{\Omega_{p,n}} \mathrm{Tr} \log(\Phi^* D \Phi) \, d\phi$$

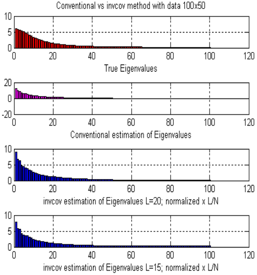
and

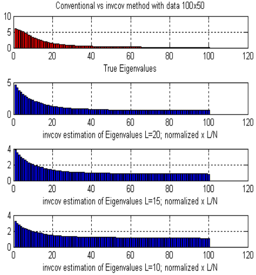
$$\mu = \int_{\Omega_{p,n}} \mathrm{Tr}(\Phi^* D \Phi)^{-1} \, d\phi$$

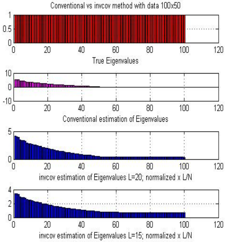
For both these terms the main Theorem gives us a closed form expression.

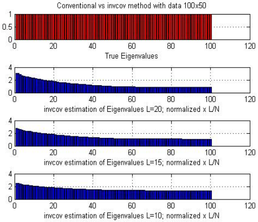
Simulations Results: Toeplitz and White Noise

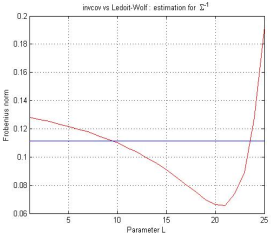
Simulations Results: Toeplitz and White Noise











Thanks!