Optimal design, orthogonal polynomials and random matrices

Holger Dette [5]

joint work with F. Bretz [1], S. Delvaux[2], L. Imhof[3], W.J. Studden[4]

[1]Novartis, Basel
[2]Katholieke Universiteit Leuven
[3]University of Bonn
[4]Purdue University
[5]Ruhr-University Bochum

September 2011, Köln

SFB | TR12

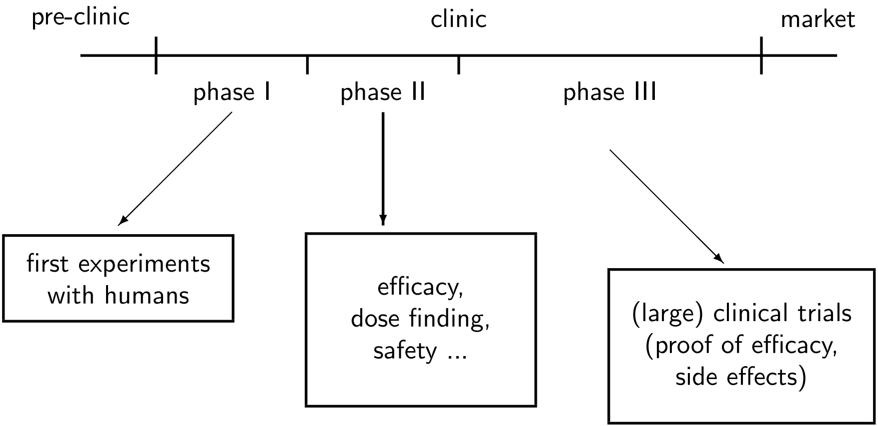
Contents

- Motivating example: dose finding experiment
- Some optimal design theory
- Optimal design for weighted polynomial regression
- Weak asymptotics of optimal designs

Contents

- Motivating example: dose finding experiment
- Some optimal design theory
- Optimal design for weighted polynomial regression
- Weak asymptotics of optimal designs
- Random matrices - the Gaussian ensemble
- Random band matrices
- Matrix orthogonal polynomials
- The limiting spectrum of random band matrices

Motivating example: drug development (clinical phase)

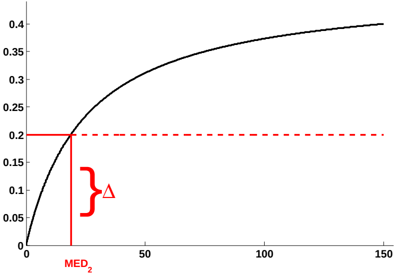


- Phase I: 20 – 40 patients
- Phase II: 100 – 300 patients
- Phase III: 1000 – 10000 patients

What dose level should be used in the the phase III trial?

Motivating Example: drug development

- Confirmatory trial (phase II) to determine the appropriate target dose
- Main goal: estimation of the minimum effective dose level (target dose), which produces at least the clinically relevant effect
- Mathematical (extremely simplified) description of the dose response relationship **(Michaelis Menten model)**



(Nonlinear) regression model

$$Y \;=\; \eta(x,\theta) \;+\; \sigma(x,\theta)\varepsilon, \quad x \in \mathcal{X}$$

- $\mathcal{X}$ denotes the **design space**
- ε random error, $E[\varepsilon] = 0$, $E[\varepsilon^2] = 1$
- m independent observations $Y_1, \ldots, Y_m$ at experimental conditions $x_1, \ldots, x_m$ to estimate the vector of **parameters** θ
- Expectation of Y (at experimental condition x) is given by $\eta(x,\theta)$
- Variance of Y (at experimental condition x) is given by $\sigma^2(x,\theta)$
- **Example:** Michaelis Menten model

$$\eta(x,\theta) = \frac{\theta_1 x}{x+\theta_2} \;, \quad \sigma(x,\theta) = \frac{\theta_1 x}{x+\theta_2} \;, \quad x \in \mathcal{X} = (0,\infty)$$

Problem: At which points x_i should we take observations ?

Definition: An **approximate design** ξ is a probability measure on the design space $\mathcal{X}$.

Example:

$$\xi = \begin{pmatrix} 25 & 80 & 150 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$\Rightarrow$ 1/3 of the total observations at each point 25, 80 and 150

- $m = 30 \to 10, 10, 10$
- $m = 40 \to 13, 14, 13$

Measuring the quality of designs

- **Weighted** least squares estimator: $\hat{\theta}$

$$\Rightarrow \text{Cov}(\hat{\theta}) \sim \frac{1}{m} M^{-1}(\xi)$$

where

$$\begin{aligned} M(\xi) &= \int_{\mathcal{X}} \frac{1}{\sigma^2(x,\theta)} \left(\frac{\partial \eta(x,\theta)}{\partial \theta}\right)^T \frac{\partial \eta(x,\theta)}{\partial \theta} \\ &\quad + \frac{1}{2\sigma^4(x,\theta)} \left(\frac{\partial \sigma^2(x,\theta)}{\partial \theta}\right)^T \frac{\partial \sigma^2(x,\theta)}{\partial \theta} d\xi(x) \end{aligned}$$

denotes the **information matrix** of the design ξ (this measure refers to the normality assumption).

Measuring the quality of designs

- **Weighted** least squares estimator: $\hat{\theta}$

$$\Rightarrow \text{Cov}(\hat{\theta}) \sim \frac{1}{m} M^{-1}(\xi)$$

where

$$M(\xi) = \int_{\mathcal{X}} \frac{1}{\sigma^2(x,\theta)} \left(\frac{\partial \eta(x,\theta)}{\partial \theta} \right)^T \frac{\partial \eta(x,\theta)}{\partial \theta} + \frac{1}{2\sigma^4(x,\theta)} \left(\frac{\partial \sigma^2(x,\theta)}{\partial \theta} \right)^T \frac{\partial \sigma^2(x,\theta)}{\partial \theta} d\xi(x)$$

denotes the **information matrix** of the design ξ (this measure refers to the normality assumption).

Goal:

- Maximize $M(\xi)$ w.r.t. the choice of the design ξ (impossible!!)

Optimality criteria

- Only a partial ordering in the space of nonnegative definite matrices
- Maximize real valued (statistical meaningful) functions of $M(\xi) \rightarrow$ **optimality criteria**

Optimality criteria

- Only a partial ordering in the space of nonnegative definite matrices
- Maximize real valued (statistical meaningful) functions of $M(\xi) \rightarrow$ **optimality criteria**
- The application determines the criterion
  - *c*-**optimality** (MED-estimation)

    $$\xi^* = \arg\max_{\xi} (c^T M^{-1}(\xi) c)^{-1}$$

    where c is a vector determined by the regression model.
  - *D*-**optimality** (precise estimation of all parameters)

    $$\xi^* = \arg\max_{\xi} |M(\xi)|$$

- In this talk we will only consider *D*-**optimal designs** and **polynomial models**!

Classical (weigthed) polynomial regression model

- Polynomial regression model $[\theta = (\theta_0, \ldots, \theta_{n-1})^T,\ x \in (-\infty, \infty)]$

$$\eta(x,\theta) = \sum_{j=0}^{n-1} \theta_j x^j$$

$$\sigma^2(x,\theta) = e^{x^2}$$

- Example: $n = 2$, linear regression model (with heteroscedastic error)

- $$\frac{\partial}{\partial\theta}\eta(x,\theta) = (1, x, \ldots, x^{n-1})^T\ ,\quad \frac{\partial}{\partial\theta}\sigma^2(x,\theta) = 0$$

D-optimal design problem (weighted polynomial regression)

A D-optimal design maximizes the determinant

$$
\begin{aligned}
|M(\xi)| &= \left| \left(\int_{\mathbb{R}} x^{i+j} e^{-x^2} d\xi(x) \right)_{i,j=0,\dots,n-1} \right| \\
&= \begin{vmatrix}
\int_{\mathbb{R}} e^{-x^2} d\xi(x) & \int_{\mathbb{R}} x e^{-x^2} d\xi(x) & \dots & \int_{\mathbb{R}} x^{n-1} e^{-x^2} d\xi(x) \\
\int_{\mathbb{R}} x e^{-x^2} d\xi(x) & \int_{\mathbb{R}} x^2 e^{-x^2} d\xi(x) & \dots & \int_{\mathbb{R}} x^{n} e^{-x^2} d\xi(x) \\
\vdots & \ddots & \ddots & \vdots \\
\int_{\mathbb{R}} x^{n-1} e^{-x^2} d\xi(x) & \int_{\mathbb{R}} x^2 e^{-x^2} d\xi(x) & \dots & \int_{\mathbb{R}} x^{2n-2} e^{-x^2} d\xi(x)
\end{vmatrix}
\end{aligned}
$$

in the class of all probability measures of $\mathbb{R}$.

D-optimal design problem

Theorem 1: The D-optimal design ξ^* is a uniform distribution on the set

$$\Big\{z \mid H_n(z) \;=\; 0\Big\}$$

where H_n denotes the n-th Hermite polynomial, orthogonal with respect to the measure

$$e^{-x^2}dx$$

Two Proofs:

- Equivalence theorems (from design theory) and second order differential equations (Stieltjes)
- Moment theory

Proof; Step 1 (idea): identification of the weights

- **Equivalence theorem for D-optimality (Kiefer and Wolfowitz, 1960):** ξ^* is D-optimal if and only if

$$\forall x \in \mathbb{R} \quad e^{-x^2}(1, x, \ldots, x^{n-1})M^{-1}(\xi^*)(1, x, \ldots, x^{n-1})^T \leq n$$

Moreover, there is equality for all support points of the D-optimal design.

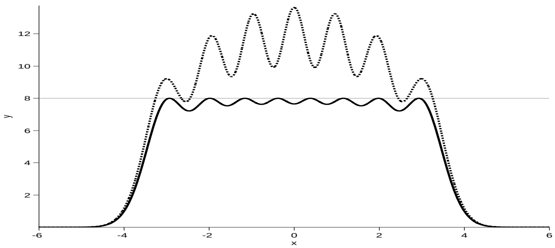
Proof; Step 1 (idea): identification of the weights

- **Equivalence theorem for D-optimality (Kiefer and Wolfowitz, 1960):** ξ^* is D-optimal if and only if

$$\forall x \in \mathbb{R} \quad e^{-x^2}(1, x, \ldots, x^{n-1}) M^{-1}(\xi^*)(1, x, \ldots, x^{n-1})^T \ \leq \ n$$

  Moreover, there is equality for all support points of the D-optimal design.

- **Example:** weighted polynomial regression of degree 7 ($n = 8$)
  - D-optimal design (solid curve)
  - Equidistant design on 10 points in the interval $[-4, 4]$
  - **Note:** D-optimal design has 8 support points (saturated)



Proof; Step 1 (idea): identification of the weights

- **Equivalence theorem for D-optimality:** ξ^* is D-optimal if and only if

$$\forall x \in \mathbb{R} \quad e^{-x^2}(1, x, \ldots, x^{n-1})M^{-1}(\xi^*)(1, x, \ldots, x^{n-1})^T \;\leq\; n$$

  Moreover, there is equality for all support points of the D-optimal design.

- The optimal design has n support points

$$\Rightarrow \quad \xi^* = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ w_1 & w_2 & \ldots & w_n \end{pmatrix}$$

Proof; Step 1 (idea): identification of the weights

- **Equivalence theorem for D-optimality:** ξ^* is D-optimal if and only if

$$\forall x \in \mathbb{R} \quad e^{-x^2}(1, x, \ldots, x^{n-1})M^{-1}(\xi^*)(1, x, \ldots, x^{n-1})^T \ \leq \ n$$

  Moreover, there is equality for all support points of the D-optimal design.

- The optimal design has n support points

$$\Rightarrow \quad \xi^* = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ w_1 & w_2 & \ldots & w_n \end{pmatrix}$$

-

$$|\mathbf{M}(\xi^*)| \quad = \quad \prod_{1 \leq i < j \leq n} (\mathbf{x_i} - \mathbf{x_j})^2 \prod_{i=1}^{n} \mathbf{e}^{-\mathbf{x_i^2}} \prod_{i=1}^{n} \mathbf{w_i}$$

$$\longrightarrow \quad \max_{x_i, w_i}$$

$$\longrightarrow \quad w_i \ = \ \frac{1}{n} \ , \quad i = 1, \ldots, n$$

Proof; Step 2 (idea): identification of the support

- Let
$$f(x) = (x - x_1)\dots(x - x_n)$$
denote the **supporting polynomial**.
- The necessary condition for an extremum yields a system of n non-linear equations
$$f''(x_j) - 2x_j f'(x_j) = 0 \quad j = 1, \dots n$$

Proof; Step 2 (idea): identification of the support

- Let

$$f(x) = (x - x_1)\dots(x - x_n)$$

  denote the **supporting polynomial**.

- The necessary condition for an extremum yields a system of n non-linear equations

$$f''(x_j) - 2x_j f'(x_j) = 0 \quad j = 1, \dots n$$

- Derive a differential equation for the supporting polynomial

$$f''(x) - 2xf'(x) = -2nf(x)$$

- This differential equation has exactly **one** polynomial solution

$$f(x) = cH_n(x)$$

Weak asymptotics of roots of Hermite polynomials:

- **Theorem 2:**

$$\xi_n^*((0,t]) \;=\; \frac{1}{n}\#\Big\{ z \le t \mid H_n\big(\sqrt{n}z\big) \;=\; 0\Big\}$$

If $n \to \infty$, then : ξ_n^* converges weakly to an absolute continuous measure μ^* with density

$$\frac{d\mu^*}{dx} \;=\; \frac{1}{\pi}\sqrt{2-x^2}I_{[-\sqrt{2},\sqrt{2}]}(x)$$

- μ^* is called the **Wigner semi-circle law**

Proof (idea):

- Use the differential equation for Hermite polynomials to derive a recurrence relation for the moments of the uniform distribution ξ_n^* on the set

$$\left\{ z \leq t \mid H_n\left(\sqrt{n}z\right) = 0\right\}$$

that is

$$\mu_{\mathbf{2r,n}} = \frac{\mathbf{1}}{\mathbf{2}}\left\{\sum_{\nu=\mathbf{0}}^{\mathbf{r-1}} \mu_{\mathbf{2r-2}\nu\mathbf{-2,n}}\mu_{\mathbf{2}\nu\mathbf{,n}} - \frac{\mathbf{2r-1}}{\mathbf{n}}\mu_{\mathbf{2r-2,n}}\right\}$$

- Recurrence relation in the limit $(n \to \infty)$

$$\mu^*_{\mathbf{2r}} = \frac{\mathbf{1}}{\mathbf{2}}\sum_{\nu=\mathbf{0}}^{\mathbf{r-1}} \mu^*_{\mathbf{2r-2}\nu\mathbf{-2}}\mu^*_{\mathbf{2}\nu}$$

- Identify the moments and the limit distribution

$$\mu^*_{2r} = \frac{1}{r+1}\left(\frac{1}{2}\right)^r\binom{2r}{r} = \frac{1}{\pi}\int_{-\sqrt{2}}^{\sqrt{2}} x^{2r}\sqrt{2-x^2}dx$$

Elementary random matrix theory

- $M_n \in \mathbb{R}^{n\times n}$ symmetric matrix with i.i.d. entries $M_n(i,j) \sim \mathcal{N}(0, \frac{1}{2})$
- **Problem:** location of the eigenvalues of the random matrix M_n ?

Elementary random matrix theory

- $M_n \in \mathbb{R}^{n\times n}$ symmetric matrix with i.i.d. entries $M_n(i,j) \sim \mathcal{N}(0, \frac{1}{2})$
- **Problem:** location of the eigenvalues of the random matrix M_n ?
- The joint density of the (random) eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ of the matrix M_n is given by

$$\mathbf{h}(\lambda) = \mathbf{c} \prod_{1\leq i<j\leq n} |\lambda_i - \lambda_j| \prod_{i=1}^{n} \mathbf{e}^{-\frac{\lambda_i^2}{2}} ,$$

- (Maximum likelihood) Typical locations are the points where the density is maximal!

Elementary random matrix theory

- $M_n \in \mathbb{R}^{n\times n}$ symmetric matrix with i.i.d. entries $M_n(i,j) \sim \mathcal{N}(0, \frac{1}{2})$
- **Problem:** location of the eigenvalues of the random matrix M_n ?
- The joint density of the (random) eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ of the matrix M_n is given by

$$\mathbf{h}(\lambda) = \mathbf{c} \prod_{1\leq i<j\leq n} |\lambda_i - \lambda_j| \prod_{i=1}^{n} \mathbf{e}^{-\frac{\lambda_i^2}{2}} ,$$

- (Maximum likelihood) Typical locations are the points where the density is maximal!
- *D*-**optimal design theory tells us:** look at roots of the Hermite polynomial $H_n(z)$
- **Note:** If $n \to \infty$ the roots of $H_n(\sqrt{n}z)$ become dense in $[-\sqrt{2}, \sqrt{2}]$.

Semi-circle law for the Gaussian ensemble

Theorem 3 Let $\lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \ldots \leq \lambda_n^{(n)}$ denote the eigenvalues of the random matrix

$$\frac{1}{\sqrt{n}} M_n$$

and by

$$\mu_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j^{(n)}}$$

the empirical eigenvalue distribution (δ_x is the Dirac measure), then for any $t \in [-\sqrt{2}, \sqrt{2}]$

$$\lim_{n\to\infty} \mu_n((-\sqrt{2}, t]) = \frac{1}{\pi} \int_{-\sqrt{2}}^{t} \sqrt{2 - x^2} dx \quad a.s.$$

Eigenvalues of a 5000×5000 matrix

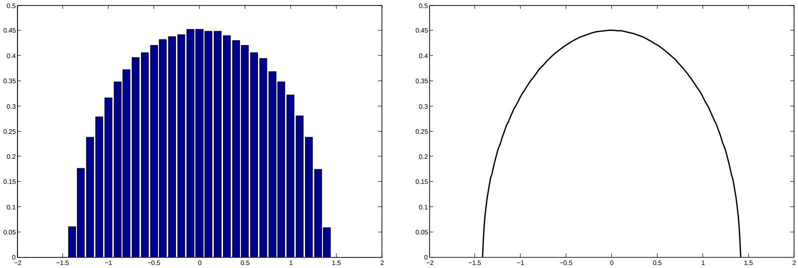


Figure: *Left panel: histogram of the simulated eigenvalues. Right panel: asymptotic distribution*

Eigenvalues of a 5000×5000 matrix

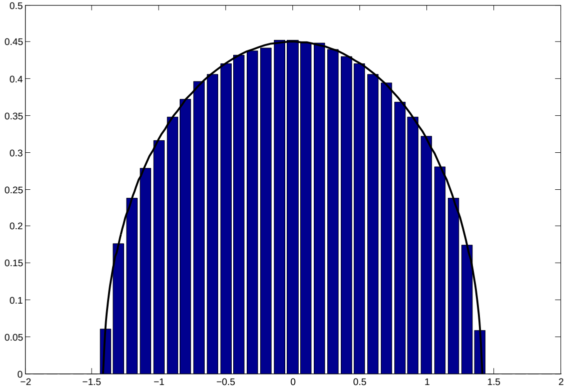


Figure: *Histogram of the simulated eigenvalues and the asymptotic distribution*

β-ensembles

- The β-ensemble ($\beta > 0$) is defined by the density

$$\mathbf{h}(\lambda) = \mathbf{c} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^{n} e^{-\frac{\lambda_i^2}{2}} , \tag{1}$$

Density of the eigenvalues of a $n \times n$ matrix with normally distributed random variables [Dyson (1962)], where

$\beta = 1$: real entries
$\beta = 2$: complex entries
$\beta = 4$: quaternion entries

β-ensembles

- The β-ensemble ($\beta > 0$) is defined by the density

$$\mathbf{h}(\lambda) = \mathbf{c} \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^{\beta} \prod_{i=1}^{n} e^{-\frac{\lambda_i^2}{2}} , \tag{1}$$

  Density of the eigenvalues of a $n \times n$ matrix with normally distributed random variables [Dyson (1962)], where

$\beta = 1$:	real entries
$\beta = 2$:	complex entries
$\beta = 4$:	quaternion entries

- **Is there any random matrix whose eigenvalue distribution is given by (1) for any $\beta > 0$?**

β-ensembles

- The β-ensemble ($\beta > 0$) is defined by the density

$$\mathbf{h}(\lambda) = \mathbf{c} \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^{\beta} \prod_{i=1}^{n} e^{-\frac{\lambda_i^2}{2}}, \tag{1}$$

  Density of the eigenvalues of a $n \times n$ matrix with normally distributed random variables [Dyson (1962)], where

  $\beta = 1$: real entries
  $\beta = 2$: complex entries
  $\beta = 4$: quaternion entries

- **Is there any random matrix whose eigenvalue distribution is given by** (1) **for any** $\beta > 0$**?**
- The answer is positive [Dumitriu and Edelman, 2004]
- The matrix can be chosen in a tridiagonal form (Householder transformations)!

Tridiagonal matrix representation for the β-ensemble

$$G_n^{(1)} = \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{2}N_1 & \mathcal{X}_{(n-1)\beta} & & & \\ \mathcal{X}_{(n-1)\beta} & \sqrt{2}N_2 & \mathcal{X}_{(n-2)\beta} & & \\ & \ddots & \ddots & \ddots & \\ & & \mathcal{X}_{2\beta} & \sqrt{2}N_{n-1} & \mathcal{X}_{\beta} \\ & & & \mathcal{X}_{\beta} & \sqrt{2}N_n \end{bmatrix}$$

Note:

- $N_1, N_2, \ldots, N_n$ are standard normal distributed ($N_j \sim \mathcal{N}(0,1)$)
- For $j = 1, \ldots, n-1$ the random variable $\mathcal{X}_{j\beta}^2$ is chi-square distributed with "$j\beta$ degrees of freedom" ($\mathcal{X}_{j\beta}^2 \sim \chi^2(j\beta)$)
- **All** random variables are independent

Eigenvalues are "close" to roots of orthogonal polynomials

Theorem 4: [D., Imhof, 2007] If

$$\lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \ldots \leq \lambda_n^{(n)}$$

denote the eigenvalues of the matrix $\frac{1}{\sqrt{n}} G_n^{(1)}$ and

$$\xi_1^{(n)} < \xi_2^{(n)} < \cdots < \xi_n^{(n)}$$

denote the zeros of the polynomial $H_n(\sqrt{n\beta} z)$, then $(n \to \infty)$

$$\max_{1 \leq j \leq n} |\lambda_j^{(n)} - \xi_j^{(n)}| = O\Big(\Big(\frac{\log n}{n}\Big)^{1/2}\Big) \quad a.s.$$

Idea of a proof of Theorem 4

- Expectation of chi-square distribution $\mathbf{E}[\mathcal{X}^{\mathbf{2}}_{\mathbf{j}\beta}] = \mathbf{j}\beta$. Approximate

$$\mathbf{E}[\mathcal{X}_{\mathbf{j}\beta}] \approx \sqrt{\mathbf{j}\beta}$$

- Consider the (non-random) matrix

$$E[G_n^{(1)}] \approx F_n = \sqrt{\frac{\beta}{2}} \begin{bmatrix} 0 & \sqrt{n-1} & & & \\ \sqrt{n-1} & 0 & \sqrt{n-2} & & \\ & \ddots & \ddots & \ddots & \\ & & \sqrt{2} & 0 & 1 \\ & & & 1 & 0 \end{bmatrix} \quad (2)$$

- Note: by the three term recurrence relation for Hermite polynomials we have:

$$\det(\mathbf{xI_n} - \mathbf{F_n}) = \left(\frac{\sqrt{\beta}}{\mathbf{2}}\right)^{\mathbf{n}} \mathbf{H_n}\left(\frac{\mathbf{x}}{\sqrt{\beta}}\right)$$

Idea of a proof of Theorem 4

- $\lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \ldots \leq \lambda_n^{(n)}$: eigenvalues of the matrix $\frac{1}{\sqrt{n}} G_n^{(1)}$
- $\xi_1^{(n)} < \cdots < \xi_n^{(n)}$: roots of the polynomial $H_n(\sqrt{n\beta} z)$
- **Weyl's inequality**

$$\sqrt{n} \max_{1 \leq j \leq n} |\lambda_j^{(n)} - \xi_j^{(n)}| \leq \| G_n^{(1)} - F_n \|_\infty \ = \ \max_{1 \leq i \leq n} \sum_{j=1}^{n} |(G_n^{(1)} - F_n)_{ij}|$$

Idea of a proof of Theorem 4

- $\lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \ldots \leq \lambda_n^{(n)}$: eigenvalues of the matrix $\frac{1}{\sqrt{n}} G_n^{(1)}$
- $\xi_1^{(n)} < \cdots < \xi_n^{(n)}$: roots of the polynomial $H_n(\sqrt{n\beta} z)$
- **Weyl's inequality**

$$\sqrt{n} \max_{1 \leq j \leq n} |\lambda_j^{(n)} - \xi_j^{(n)}| \leq \| G_n^{(1)} - F_n \|_\infty \;=\; \max_{1 \leq i \leq n} \sum_{j=1}^{n} |(G_n^{(1)} - F_n)_{ij}|$$

- **Large deviations:**

$$P\left\{ \frac{|\mathcal{X}_{j\beta} - \sqrt{j\beta}|}{\sqrt{2}} \geq \frac{\epsilon}{3} \right\} \leq 2e^{-\epsilon^2/9}, \quad P\left\{ |N_j| \geq \frac{\epsilon}{3} \right\} \leq 2e^{-\epsilon^2/18}$$

$$\Longrightarrow \quad P\left\{ \max_{1 \leq j \leq n} \left| \lambda_j^{(n)} - \xi_j^{(n)} \right| \geq \epsilon \right\} \leq 4ne^{-n\beta\epsilon^2/9}$$

Idea of a proof of Theorem 4

- $\lambda_1^{(n)} \le \lambda_2^{(n)} \le \ldots \le \lambda_n^{(n)}$: eigenvalues of the matrix $\frac{1}{\sqrt{n}} G_n^{(1)}$
- $\xi_1^{(n)} < \cdots < \xi_n^{(n)}$: roots of the polynomial $H_n(\sqrt{n\beta} z)$
- **Weyl's inequality**

$$\sqrt{n} \max_{1\le j\le n} |\lambda_j^{(n)} - \xi_j^{(n)}| \le \|G_n^{(1)} - F_n\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |(G_n^{(1)} - F_n)_{ij}|$$

- **Large deviations:**

$$P\left\{ \frac{|\mathcal{X}_{j\beta} - \sqrt{j\beta}|}{\sqrt{2}} \ge \frac{\epsilon}{3} \right\} \le 2e^{-\epsilon^2/9}, \quad P\left\{ |N_j| \ge \frac{\epsilon}{3} \right\} \le 2e^{-\epsilon^2/18}$$

$$\Longrightarrow \quad P\left\{ \max_{1\le j\le n} \left|\lambda_j^{(n)} - \xi_j^{(n)}\right| \ge \epsilon \right\} \le 4ne^{-n\beta\epsilon^2/9}$$

- **Borel Cantelli**

$$\max_{1\le j\le n} |\lambda_j^{(n)} - \xi_j^{(n)}| = O\left(\left(\frac{\log n}{n} \right)^{1/2} \right) \quad a.s.$$

Idea of a proof of Theorem 3 ($\beta = 1$)

- The random eigenvalues of the matrix $\frac{1}{\sqrt{n}} G_n^{(1)}$ can be (uniformly, almost surely) approximated by roots of the Hermite polynomial $H_n(\sqrt{n\beta} z)$.
- The uniform distribution on the roots Hermite polynomial $H_n(\sqrt{n\beta} z)$ converges weakly to Wigner's semi-circle law.
- The empirical eigenvalue distribution of the random matrix $\frac{1}{\sqrt{n}} G_n^{(1)}$ converges weakly to Wigner's semi-circle law (almost surely).

Random band matrices - tridiagonal ($r = 1$, $\beta_1 > 0$)

$$G_n^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\, N_1 & \mathcal{X}_{(n-1)\beta_1} & & & & & \\ \mathcal{X}_{(n-1)\beta_1} & \sqrt{2}\, N_2 & \mathcal{X}_{(n-2)\beta_1} & & & & \\ & \mathcal{X}_{(n-2)\beta_1} & \sqrt{2} N_3 & \mathcal{X}_{(n-3)\beta_1} & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

All random variables are independent!

Random 5-band matrices ($r = 2$, $\beta_1, \beta_2 > 0$)

$$G_n^{(2)} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\,N_1 & \mathcal{X}_{(n-1)\beta_1} & \mathcal{X}_{(n-2)\beta_2} & & & & \\ \mathcal{X}_{(n-1)\beta_1} & \sqrt{2}\,N_2 & \mathcal{X}_{(n-2)\beta_1} & \mathcal{X}_{(n-3)\beta_2} & & & \\ \mathcal{X}_{(n-2)\beta_2} & \mathcal{X}_{(n-2)\beta_1} & \sqrt{2}N_3 & \mathcal{X}_{(n-3)\beta_1} & \mathcal{X}_{(n-4)\beta_2} & \ddots & \ddots \\ & \mathcal{X}_{(n-3)\beta_2} & \mathcal{X}_{(n-3)\beta_1} & \sqrt{2}N_4 & \mathcal{X}_{(n-4)\beta_1} & \ddots & \ddots \\ & & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

All random variables are independent!

Random 7-band matrices ($r = 3,\ \beta_1, \beta_2, \beta_3 > 0$)

$$G_n^{(3)} =$$

$$\frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\ N_1 & \mathcal{X}_{(n-1)\beta_1} & \mathcal{X}_{(n-2)\beta_2} & \mathcal{X}_{(n-3)\beta_3} & & & \\ \mathcal{X}_{(n-1)\beta_1} & \sqrt{2}\ N_2 & \mathcal{X}_{(n-2)\beta_1} & \mathcal{X}_{(n-3)\beta_2} & \mathcal{X}_{(n-4)\beta_3} & & \\ \mathcal{X}_{(n-2)\beta_2} & \mathcal{X}_{(n-2)\beta_1} & \sqrt{2}N_3 & \mathcal{X}_{(n-3)\beta_1} & \mathcal{X}_{(n-4)\beta_2} & \mathcal{X}_{(n-5)\beta_3} & \ddots \\ \mathcal{X}_{(n-3)\beta_3} & \mathcal{X}_{(n-3)\beta_2} & \mathcal{X}_{(n-3)\beta_1} & \sqrt{2}N_4 & \mathcal{X}_{(n-4)\beta_1} & \mathcal{X}_{(n-5)\beta_2} & \ddots \\ & \mathcal{X}_{(n-4)\beta_3} & \mathcal{X}_{(n-4)\beta_2} & \mathcal{X}_{(n-4)\beta_1} & \sqrt{2}N_4 & \mathcal{X}_{(n-5)\beta_1} & \ddots \\ & & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

All random variables are independent!

Random $2r+1$ band matrices $(\beta_1, \ldots, \beta_r > 0)$

$$\sqrt{2}G_n^{(r)} =$$

$$\left|\begin{array}{cccccccc}
\sqrt{2}\ N_1 & \mathcal{X}_{(n-1)\beta_1} & \cdots & \mathcal{X}_{(n-r)\beta_r} & & & & \\
\mathcal{X}_{(n-1)\beta_1} & \sqrt{2}\ N_2 & \cdots & \mathcal{X}_{(n-r)\beta_{r-1}} & \mathcal{X}_{(n-r-1)\beta_r} & & & \\
\mathcal{X}_{(n-2)\beta_2} & \mathcal{X}_{(n-2)\beta_1} & \ddots & \ddots & \ddots & \ddots & & \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \\
\mathcal{X}_{(n-r)\beta_r} & \mathcal{X}_{(n-r)\beta_{r-1}} & \ddots & \ddots & \ddots & \ddots & \ddots & \\
 & \mathcal{X}_{(n-r-1)\beta_r} & & \ddots & \ddots & \ddots & \ddots & \ddots \\
 & & & \ddots & \ddots & \ddots & \ddots & \ddots \\
 & & & \ddots & \ddots & \ddots & \ddots & \\
 & & & & & & \mathcal{X}_{2\beta_1} & \mathcal{X}_{\beta_2} \\
 & & & & & & \sqrt{2}\ N_{n-1} & \mathcal{X}_{\beta_1} \\
 & & & & & & \mathcal{X}_{\beta_1} & \sqrt{2}\ N_n
\end{array}\right|$$

5-band and tridiagonal block matrices (2×2 blocks)

$$G_n^{(2)} =$$

$$\frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\ N_1 & \mathcal{X}_{(n-1)\beta_1} & \mathcal{X}_{(n-2)\beta_2} & 0 & & & \\ \mathcal{X}_{(n-1)\beta_1} & \sqrt{2}\ N_2 & \mathcal{X}_{(n-2)\beta_1} & \mathcal{X}_{(n-3)\beta_2} & & & \\ \mathcal{X}_{(n-2)\beta_2} & \mathcal{X}_{(n-2)\beta_1} & \sqrt{2}N_3 & \mathcal{X}_{(n-3)\beta_1} & \mathcal{X}_{(n-4)\beta_2} & 0 & \\ 0 & \mathcal{X}_{(n-3)\beta_2} & \mathcal{X}_{(n-3)\beta_1} & \sqrt{2}N_4 & \mathcal{X}_{(n-4)\beta_1} & \mathcal{X}_{(n-5)\beta_2} & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

7-band and tridiagonal block matrices (3×3 blocks)

$$G_n^{(3)} =$$

$$\frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\,N_1 & \mathcal{X}_{(n-1)\beta_1} & \mathcal{X}_{(n-2)\beta_2} & \mathcal{X}_{(n-3)\beta_3} & 0 & 0 & \ddots \\ \mathcal{X}_{(n-1)\beta_1} & \sqrt{2}\,N_2 & \mathcal{X}_{(n-2)\beta_1} & \mathcal{X}_{(n-3)\beta_2} & \mathcal{X}_{(n-4)\beta_3} & 0 & \ddots \\ \mathcal{X}_{(n-2)\beta_2} & \mathcal{X}_{(n-2)\beta_1} & \sqrt{2}N_3 & \mathcal{X}_{(n-3)\beta_1} & \mathcal{X}_{(n-4)\beta_2} & \mathcal{X}_{(n-5)\beta_3} & \ddots \\ \mathcal{X}_{(n-3)\beta_3} & \mathcal{X}_{(n-3)\beta_2} & \mathcal{X}_{(n-3)\beta_1} & \sqrt{2}N_4 & \mathcal{X}_{(n-4)\beta_1} & \mathcal{X}_{(n-5)\beta_2} & \ddots \\ 0 & \mathcal{X}_{(n-4)\beta_3} & \mathcal{X}_{(n-4)\beta_2} & \mathcal{X}_{(n-4)\beta_1} & \sqrt{2}N_5 & \mathcal{X}_{(n-5)\beta_1} & \ddots \\ 0 & 0 & \mathcal{X}_{(n-5)\beta_3} & \mathcal{X}_{(n-5)\beta_2} & \mathcal{X}_{(n-5)\beta_1} & \sqrt{2}N_6 & \ddots \\ & & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

$2r+1$-band and tridiagonal block matrices ($r \times r$ blocks)

$$G_n^{(r)} = \begin{pmatrix} B_0 & A_1 & & & \\ A_1^T & B_1 & A_2 & & \\ & \ddots & \ddots & \ddots & \\ & & A_{m-2}^T & B_{m-2} & A_{m-1} \\ & & & A_{m-1}^T & B_{m-1} \end{pmatrix} \in \mathbb{R}^{n\times n}$$

where

- $n = mr$
- B_i are symmetric random matrices
- A_i are lower random triangular matrices

Problem: location of the eigenvalues?

Excursion: matrix orthogonal polynomials

- Matrix polynomials [Krein (1969), Damanik, Killip, Pushnitski, Simon (2008,2010)]

$$P_n(x) = D_n x^n + D_{n-1} x^{n-1} + \ldots + D_1 x + D_0$$

  where $D_0, \ldots, D_n$ are $r \times r$ matrices with real entries

- **Example:**

$$P_3(x) = \begin{pmatrix} x^3 + x - 1 & 2x + 1 \\ x - 1 & 3x^2 \end{pmatrix}$$

Excursion: matrix orthogonal polynomials

- Matrix polynomials [Krein (1969), Damanik, Killip, Pushnitski, Simon (2008,2010)]

$$P_n(x) = D_n x^n + D_{n-1}x^{n-1} + \ldots + D_1 x + D_0$$

  where $D_0, \ldots, D_n$ are $r \times r$ matrices with real entries
- **Example:**

$$P_3(x) = \begin{pmatrix} x^3 + x - 1 & 2x + 1 \\ x - 1 & 3x^2 \end{pmatrix}$$

- Roots of a matrix polynomial are defined by $\det P_n(x) = 0$
- Matrix measure ψ is a matrix of signed Borel measures on the real line such for any Borel set A the matrix $\psi(A)$ is nonnegative definite (spectral measure of multivariate stationary processes)
- "Inner product" with respect to the matrix measure ψ

$$\langle P_n, P_m \rangle \ := \ \int_{\mathbb{R}} P_n(x) d\psi(x) P_m^T(x) \ \in \ \mathbb{R}^{r \times r}$$

Excursion: matrix orthogonal polynomials

- Matrix polynomials are called orthonormal if and only if

$$\langle P_n, P_m \rangle \;=\; \delta_{n,m} I_r \in \mathbb{R}^{r\times r}$$

- Some properties of the scalar case are still valid
  - All roots of orthogonal matrix polynomials are real
  - **Favard's Theorem:** $\{P_n\}_{n\in\mathbb{N}}$ defines a sequence of matrix orthonormal polynomials if and only

$$\mathbf{xP_n(x) = A_{n+1}P_{n+1}(x) + B_nP_n(x) + A_n^TP_{n-1}(x),} \qquad n \geq 0,$$

for symmetric matrices B_n and arbitrary non singular matrices A_n [D. and Studden (2002)]

Excursion: matrix orthogonal polynomials

- Matrix multiplication is not commutative
- Orthonormal matrix polynomials are **not** uniquely determined
- The roots of matrix orthogonal polynomials are **not** interlacing
- Characterization of the boundary of the moment space corresponding to matrix measures?
- There exists no example of matrix orthogonal polynomials, which has been completely understood

Excursion: matrix orthogonal polynomials

- Matrix multiplication is not commutative
- Orthonormal matrix polynomials are **not** uniquely determined
- The roots of matrix orthogonal polynomials are **not** interlacing
- Characterization of the boundary of the moment space corresponding to matrix measures?
- There exists no example of matrix orthogonal polynomials, which has been completely understood
  - **Example:** Scalar Chebyshev polynomials (first kind)

$$T_{-1}(x) = 0 \ , \ \ T_0(x) = 1 \ , \ \ T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

  - Trigonometric representation: $T_n(x) = \cos(n \arccos x)$
  - Measure of orthogonality: arcsine distribution with density

$$\frac{1}{\pi} \frac{1}{\sqrt{1-x^2}} I_{[-1,1]}(x)$$

Excursion: matrix Chebyshev polynomials

- $A \in \mathbb{R}^{r \times r}$ non singular; $B \in \mathbb{R}^{r \times r}$ symmetric
- Recurrence relation $T_0^{A,B}(x) = I_p$,

$$\begin{aligned} T_1^{A,B}(x) &= (\sqrt{2}A)^{-1}(xI_p - B) \\ xT_1^{A,B}(x) &= AT_2^{A,B}(x) + BT_1^{A,B}(x) + \sqrt{2}A^T T_0^{A,B}(x) \\ xT_n^{A,B}(x) &= AT_{n+1}^{A,B}(x) + BT_n^{A,B}(x) + A^T T_{n-1}^{A,B}(x), \; n \geq 2, \end{aligned}$$

- If $r = 1$ the measure of orthogonality is given by a linear transformation of the arcsine distribution with density

$$\frac{1}{\pi}\frac{1}{\sqrt{1 - x^2}} I_{[-1,1]}(x)$$

- **Open problem:** The matrix measure $X_{A,B}$ of orthogonality in the case $r > 1$?

Return to random of block matrices

We are interested in the eigenvalues of the matrix

$$G_n^{(r)} = \begin{pmatrix} B_0 & A_1 & & & & \\ A_1^T & B_1 & A_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & A_{m-2}^T & B_{m-2} & A_{m-1} & \\ & & & A_{m-1}^T & B_{m-1} & \end{pmatrix} \in \mathbb{R}^{n\times n}$$

where

- $n = mr$
- B_i are symmetric random matrices
- A_i are lower random triangular matrices

Problem: location of the eigenvalues?

The structure of the blocks $(r \times r)$

$$B_i = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}N_{ir+1} & \mathcal{X}_{(n-ir-1)\beta_1} & \cdots & \mathcal{X}_{(n-(i+1)r+1)\beta_{r-1}} \\ \mathcal{X}_{(n-ir-1)\beta_1} & \sqrt{2}N_{ir+2} & \cdots & \mathcal{X}_{(n-(i+1)r+1)\beta_{r-2}} \\ \vdots & \ddots & \ddots & \vdots \\ \mathcal{X}_{(n-(i+1)r+1)\beta_{r-1}} & \cdots & \mathcal{X}_{(n-(i+1)r+1)\beta_1} & \sqrt{2}N_{(i+1)r} \end{pmatrix}$$

$$A_i = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathcal{X}_{(n-ir)\beta_r} & 0 & 0 & \cdots & 0 \\ \mathcal{X}_{(n-ir)\beta_{r-1}} & \mathcal{X}_{(n-ir-1)\beta_r} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathcal{X}_{(n-ir)\beta_1} & \mathcal{X}_{(n-ir-1)\beta_{r-1}} & \cdots & \cdots & \mathcal{X}_{(n-(i+1)r+1)\beta_r} \end{pmatrix},$$

The structure of the blocks in the case $r = 3$:

$$B_i = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\, N_{3i+1} & \mathcal{X}_{(n-3i-1)\beta_1} & \mathcal{X}_{(n-3i-2)\beta_2} \\ \mathcal{X}_{(n-3i-1)\beta_1} & \sqrt{2}\, N_{3i+2} & \mathcal{X}_{(n-3i-2)\beta_1} \\ \mathcal{X}_{(n-3i-2)\beta_2} & \mathcal{X}_{(n-3i-2)\beta_1} & \sqrt{2} N_{3i+3} \end{pmatrix}$$

$$A_i = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathcal{X}_{(n-3i)\beta_3} & 0 & 0 \\ \mathcal{X}_{(n-3i)\beta_2} & \mathcal{X}_{(n-3i-1)\beta_3} & 0 \\ \mathcal{X}_{(n-3i)\beta_1} & \mathcal{X}_{(n-3i-1)\beta_2} & \mathcal{X}_{(n-3i-2)\beta_3} \end{pmatrix}$$

The structure of the blocks in the case $r = 3$:

$$B_i = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\, N_{3i+1} & \mathcal{X}_{(n-3i-1)\beta_1} & \mathcal{X}_{(n-3i-2)\beta_2} \\ \mathcal{X}_{(n-3i-1)\beta_1} & \sqrt{2}\, N_{3i+2} & \mathcal{X}_{(n-3i-2)\beta_1} \\ \mathcal{X}_{(n-3i-2)\beta_2} & \mathcal{X}_{(n-3i-2)\beta_1} & \sqrt{2} N_{3i+3} \end{pmatrix}$$

$$A_i = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathcal{X}_{(n-3i)\beta_3} & 0 & 0 \\ \mathcal{X}_{(n-3i)\beta_2} & \mathcal{X}_{(n-3i-1)\beta_3} & 0 \\ \mathcal{X}_{(n-3i)\beta_1} & \mathcal{X}_{(n-3i-1)\beta_2} & \mathcal{X}_{(n-3i-2)\beta_3} \end{pmatrix}$$

Note: in the following discussion we will explain the structure always in the case $r = 3$!

Eigenvalues of block matrices and roots of polynomials

Theorem 5: [D., Reuther, 2010] Let

$$\lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \ldots \leq \lambda_n^{(n)}$$

denote the eigenvalues of the random block matrix

$$\frac{1}{\sqrt{n}} G_n^{(r)},$$

then as $n \to \infty$:

$$\max_{1 \leq j \leq n} |\lambda_j^{(n)} - \xi_j^{(n)}| = O\left(\left(\frac{\log n}{n}\right)^{1/2}\right) \quad a.s.$$

where

$$\xi_1^{(n)} \leq \xi_2^{(n)} \leq \ldots \leq \xi_n^{(n)}$$

are the roots of the $m = (n/r)$th matrix orthonormal polynomial $\mathbf{R_{m,n}(x)}$ defined by $\mathbf{R_{-1,n}(x) = 0}$, $\mathbf{R_{0,n}(x) = I_r}$

$$\mathbf{xR_{k,n}(x) = A_{k+1,n}R_{k+1,n}(x) + B_{k,n}R_{k,n}(x) + A^T_{k,n}R_{k-1,n}(x)}; \quad \mathbf{k \geq 0},$$

Coefficients in the recurrence relation (here for $r = 3$):

$$\mathbf{A_{k,n}} = \frac{1}{\sqrt{2n}} \begin{pmatrix} \sqrt{(3k-2)\beta_3} & 0 & 0 \\ \sqrt{(3k-1)\beta_2} & \sqrt{(3k-1)\beta_3} & 0 \\ \sqrt{3k\beta_1} & \sqrt{3k\beta_2} & \sqrt{3k\beta_3} \end{pmatrix}$$

$$\mathbf{B_{k,n}} = \frac{1}{\sqrt{2n}} \begin{pmatrix} 0 & \sqrt{(3k+1)\beta_1} & \sqrt{(3k+1)\beta_2} \\ \sqrt{(3k+1)\beta_1} & 0 & \sqrt{(3k+2)\beta_1} \\ \sqrt{(3k+1)\beta_2} & \sqrt{3k+2\beta_1} & 0 \end{pmatrix}$$

Coefficients in the recurrence relation (here for $r = 3$):

Note: If $n \to \infty$ and $\frac{k}{n} \to u \in (0,1)$, then

$$\mathbf{A_{k,n}} = \frac{1}{\sqrt{2n}} \begin{pmatrix} \sqrt{(3k-2)\beta_3} & 0 & 0 \\ \sqrt{(3k-1)\beta_2} & \sqrt{(3k-1)\beta_3} & 0 \\ \sqrt{3k\beta_1} & \sqrt{3k\beta_2} & \sqrt{3k\beta_3} \end{pmatrix}$$

$$\longrightarrow A(u) := \sqrt{\frac{3u}{2}} \begin{pmatrix} \sqrt{\beta_3} & 0 & 0 \\ \sqrt{\beta_2} & \sqrt{\beta_3} & 0 \\ \sqrt{\beta_1} & \sqrt{\beta_2} & \sqrt{\beta_3} \end{pmatrix}$$

$$\mathbf{B_{k,n}} = \frac{1}{\sqrt{2n}} \begin{pmatrix} 0 & \sqrt{(3k+1)\beta_1} & \sqrt{(3k+1)\beta_2} \\ \sqrt{(3k+1)\beta_1} & 0 & \sqrt{(3k+2)\beta_1} \\ \sqrt{(3k+1)\beta_2} & \sqrt{3k+2\beta_1} & 0 \end{pmatrix}$$

$$\longrightarrow B(u) := \sqrt{\frac{3u}{2}} \begin{pmatrix} 0 & \sqrt{\beta_1} & \sqrt{\beta_2} \\ \sqrt{\beta_1} & 0 & \sqrt{\beta_1} \\ \sqrt{\beta_2} & \sqrt{\beta_1} & 0 \end{pmatrix}$$

Matrix orthogonal polynomials with **varying coefficients**

- **Problem:** For $n \in \mathbb{N}$ let $\{R_{k,n}(x)\}_{k\in\mathbb{N}_0}$ denote a sequence of matrix orthonormal polynomials defined by $\mathbf{R_{-1,n}(x) = 0_r}$, $\mathbf{R_{0,n}(x) = I_r}$

$$\mathbf{xR_{k,n}(x) = A_{k+1,n}R_{k+1,n}(x) + B_{k,n}R_{k,n}(x) + A_{k,n}^{T}R_{k-1,n}(x)}; \quad \mathbf{k \geq 0},$$

where

$$\lim_{\frac{k}{n}\to u} \mathbf{B_{k,n} = B(u)}, \quad \lim_{\frac{k}{n}\to u} \mathbf{A_{k,n} = A(u)}$$

whenever $u \in (0,1)$. What is the behavior of the roots of the polynomials

$$\mathbf{Q_{k,n}(x)}$$

if $n \to \infty$?

Matrix orthogonal polynomials with **varying coefficients**

- **Problem:** For $n \in \mathbb{N}$ let $\{R_{k,n}(x)\}_{k\in\mathbb{N}_0}$ denote a sequence of matrix orthonormal polynomials defined by $\mathbf{R_{-1,n}(x) = 0_r}$, $\mathbf{R_{0,n}(x) = I_r}$

$$\mathbf{xR_{k,n}(x) = A_{k+1,n}R_{k+1,n}(x) + B_{k,n}R_{k,n}(x) + A^T_{k,n}R_{k-1,n}(x); \quad k \geq 0},$$

where

$$\lim_{\frac{k}{n}\to u} \mathbf{B_{k,n} = B(u)}, \quad \lim_{\frac{k}{n}\to u} \mathbf{A_{k,n} = A(u)}$$

whenever $u \in (0,1)$. What is the behavior of the roots of the polynomials

$$\mathbf{Q_{k,n}(x)}$$

if $n \to \infty$?

- **Note:** By Theorem 5 we expect that the eigenvalues of the random band matrix have similar properties ($k = m; n = mr \to u = 1/r$)!

An algebraic equation (Widom, 1974)

Define the equation ($x, z \in \mathbb{C}$)

$$\mathbf{0} = \mathbf{f_u}(\mathbf{z}, \mathbf{x}) := \det(\mathbf{A}(\mathbf{u})^{\mathbf{T}}\mathbf{z} + \mathbf{B}(\mathbf{u}) + \mathbf{A}(\mathbf{u})\mathbf{z}^{-\mathbf{1}} - \mathbf{x}\mathbf{I_r}) \qquad (3)$$

Note:

- For fixed $x \in \mathbb{C}$ there exist $2r$ roots $z_1(x, u), \ldots z_{2r}(x, u)$ of equation (3), which can ordered according to

$$|z_1(x, u)| \leq |z_2(x, u)| \ldots \leq |z_{2r}(x, u)|$$

- For any $u \in (0, 1)$

$$\Gamma_0(u) = \{x \in \mathbb{C} \mid |z_r(x, u)| = |z_{r+1}(x, u)|\} \subset \mathbb{R}$$

  is a union of at most r disjoint intervals.

Weak asymptotics for matrix orthonormal polynomials

Theorem 6 [Delvaux, D., 2011] Let

$$\nu_n = \frac{1}{n}\sum_{j=1}^{n} \delta_{\xi_j^{(n)}}$$

denote empirical distribution function of the roots of the polynomial $\mathbf{R_{k,n}(x)}$ defined by

$$\mathbf{xR_{k,n}(x) = A_{k+1,n}R_{k+1,n}(x) + B_{k,n}R_{k,n}(x) + A^T_{k,n}R_{k-1,n}(x)}; \quad \mathbf{k \geq 0},$$

where

$$\lim_{\frac{k}{n}\to u} \mathbf{B_{k,n} = B(u)}, \quad \lim_{\frac{k}{n}\to u} \mathbf{A_{k,n} = A(u)}.$$

Then ν_n converges weakly to a measure $\mu_{0,u}$, with logarithmic potential

$$\frac{1}{ru}\int_0^u \log|z_1(x,t)\dots z_r(x,t)|\,dt + C_u, \qquad x \in \mathbb{C} \setminus \bigcup_{0\leq t\leq u} \Gamma_0(t),$$

(here C_u is some constant).

Identification of the limit distribution

Theorem 7 [Delvaux, D. 2011] The measure $\mu_{0,u}$ with logarithmic potential

$$\frac{1}{ru}\int_0^u \log|z_1(x,t)\dots z_r(x,t)|\,dt + C_u, \qquad x\in\mathbb{C}\setminus\bigcup_{0\le t\le u}\Gamma_0(t),$$

is absolute continuous with density given by

$$\frac{\mathbf{d}\mu_{\mathbf{0},\mathbf{u}}(\mathbf{x})}{\mathbf{dx}} = \frac{\mathbf{1}}{\mathbf{2}\pi\mathbf{ur}}\int_{\mathbf{0}}^{\mathbf{u}}\sum_{\mathbf{k}:|\mathbf{z_k}(\mathbf{x},\mathbf{s})|=\mathbf{1}}\left|\frac{\frac{\partial}{\partial\mathbf{x}}\mathbf{z_k}(\mathbf{x},\mathbf{s})}{\mathbf{z_k}(\mathbf{x},\mathbf{s})}\right|\mathbf{ds}$$

Application to random block matrices

- By Theorem 5 it can be shown that the eigenvalue distribution has the same asymptotic properties as the distribution of the roots of matrix orthogonal polynomials $Q_{m,n}(x)$, where $m = n/r$
- This means
$$\lim_{n\to\infty} \frac{m}{n} = \frac{1}{r}$$
- Theorem 7 yields for the limiting distribution
$$\frac{\mathbf{d}\mu_{\mathbf{0},\mathbf{1}/\mathbf{r}}(\mathbf{x})}{\mathbf{dx}} = \frac{\mathbf{1}}{\mathbf{2}\pi}\int_{\mathbf{0}}^{\mathbf{1}/\mathbf{r}} \frac{\mathbf{1}}{\sqrt{\mathbf{s}}} \sum_{\mathbf{k}:|\mathbf{z_k}(\mathbf{x}/\sqrt{\mathbf{s}})|=\mathbf{1}} \left|\frac{\mathbf{z}'_{\mathbf{k}}(\mathbf{x}/\sqrt{\mathbf{s}})}{\mathbf{z_k}(\mathbf{x}/\sqrt{\mathbf{s}})}\right| \mathbf{ds}$$

where $z_1(x), z_2(x), \ldots, z_{2r}(x)$ are the (ordered) roots of the equation
$$\mathbf{0} = \mathbf{f}(\mathbf{z},\mathbf{x}) := \det(\mathbf{A}^{\mathbf{T}}(\mathbf{1})\mathbf{z} + \mathbf{B}(\mathbf{1}) + \mathbf{A}(\mathbf{1})\mathbf{z}^{-\mathbf{1}} - \mathbf{x}\mathbf{I}_{\mathbf{r}})$$

Application to random block matrices

$$\mathbf{A(1)} \;:=\; \sqrt{\frac{r}{2}}\begin{pmatrix} \sqrt{\beta_r} & 0 & 0 & \cdots & 0 \\ \sqrt{\beta_{r-1}} & \sqrt{\beta_r} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \sqrt{\beta_2} & \cdots & \sqrt{\beta_{r-1}} & \sqrt{\beta_r} & 0 \\ \sqrt{\beta_1} & \cdots & \sqrt{\beta_{r-2}} & \sqrt{\beta_{r-1}} & \sqrt{\beta_r} \end{pmatrix} \in \mathbb{R}^{r\times r},$$

$$\mathbf{B(1)} \;:=\; \sqrt{\frac{r}{2}}\begin{pmatrix} 0 & \sqrt{\beta_1} & \sqrt{\beta_2} & \cdots & \sqrt{\beta_{r-1}} \\ \sqrt{\beta_1} & 0 & \sqrt{\beta_1} & \cdots & \sqrt{\beta_{r-2}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \sqrt{\beta_{r-2}} & \cdots & \sqrt{\beta_1} & 0 & \sqrt{\beta_1} \\ \sqrt{\beta_{r-1}} & \cdots & \sqrt{\beta_2} & \sqrt{\beta_1} & 0 \end{pmatrix} \in \mathbb{R}^{r\times r},$$

Eigenvalues of a 5000×5000 matrix ($\beta_1 = \beta_2 = 1$)

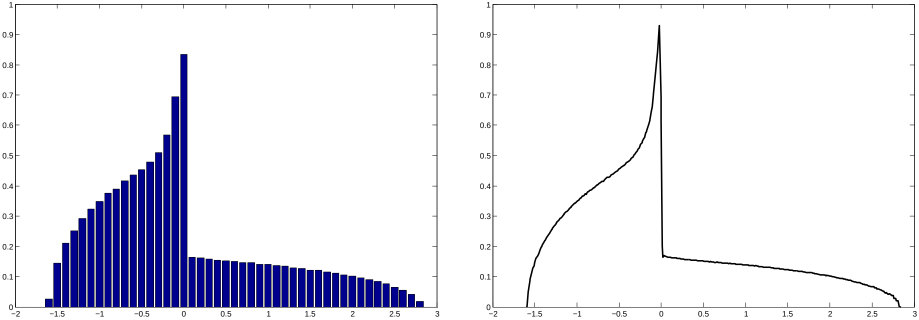


Figure: *Left panel: histogram of the simulated eigenvalues*
Right panel: asymptotic distribution

Eigenvalues of a 5000 × 5000 matrix ($\beta_1 = 5$; $\beta_2 = 1$)

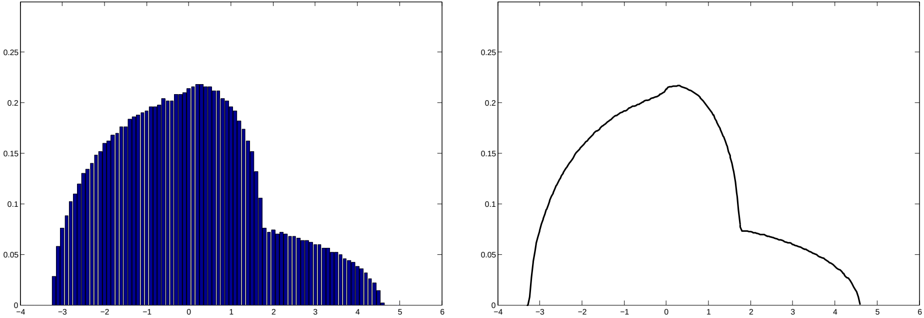


Figure: *Left panel: histogram of the simulated eigenvalues*
Right panel: asymptotic distribution

Eigenvalues of a 5000×5000 matrix

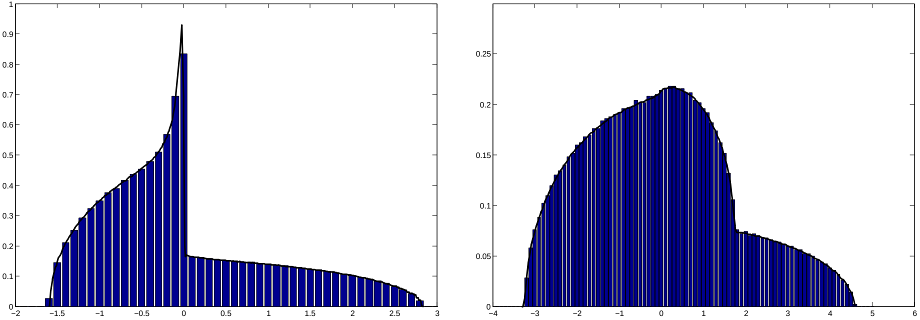


Figure: *Left panel: histogram and density* $(\beta_1 = 1;\ \beta_2 = 1)$
Right panel: histogram and density $(\beta_1 = 5;\ \beta_2 = 1)$

Conclusions and further research

- Optimal designs - (matrix) orthogonal polynomials - random matrices
- I did **not** present a solution of the design problem for the dose finding trial (it is too complicated)!

Conclusions and further research

- Optimal designs - (matrix) orthogonal polynomials - random matrices
- I did **not** present a solution of the design problem for the dose finding trial (it is too complicated)!
- Possible future research:
  - Measure of orthogonality for matrix Chebyshev polynomials?
  - Wigner block matrices (there seem to exist relations to free probability)?
  - Distribution of the eigenvalues of the band matrices considered here?
  - Use matrix orthogonal polynomials for solving optimal design problems?
  - Matrix measures and stationary processes?

Some selected references

- **Optimal design**

  H. Dette, F. Bretz, A. Pepelyshev and J. Pinheiro. *Optimal designs for dose finding studies.* J. Amer. Statist. Assoc. 103, (2008), 1225-1237.

- **β-ensembles**

  I. Dumitriu and A. Edelmann. *Matrix models for beta ensembles.* J. Math. Phys. 43 (2002), 5830-5847.

  H. Dette and L. Imhof. *Uniform approximation of eigenvalues in Laguerre and Hermite-ensembles by roots of orthogonal polynomials.* Trans. Amer. Math. Soc. 359, (2007), 4999-5018.

  F.J. Dyson. *The threefold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics.* J. Math. Phys. 3, (1962), 1199-1215.

- **Matrix polynomials**

  D. Damanik, A. Pushnitski and B. Simon. *The analytic theory of matrix orthogonal polynomials.* Surv. Approx. Theory 4, (2008), 1-85.

  D. Damanik, R. Killip and B. Simon. *Perturbations of orthogonal polynomials with periodic recursion coefficients.* Ann. of Math. 171, (2010), 1931-2010.

  H. Dette and W.J. Studden. *Matrix measures, moment spaces and Favard's theorem for the interval* $[0, 1]$ *and* $[0, \infty]$. Linear Algebra Appl. 345 (2002), 169-193.

  M.G. Krein. *Infinite J-matrices and a matrix moment problem.* Dokl. Akad. Nauk SSSR 69, (1949), 125-128 (in Russian).

- **(Random) block matrices**

  A. Böttcher and B. Silbermann (1998) *Introduction to Large Truncated Toeplitz Matrices.* Universitext, Springer-Verlag, New York, 1998.

  H. Dette and B. Reuther. *Random block matrices and matrix orthogonal polynomials.* J. Theor. Probab. 23, (2010), 378-400.

  S. Delvaux, and H. Dette. *Zeros and ratio asymptotics for matrix orthogonal polynomials* Submitted for publication, (2011). arXiv:1108.5155v2

  H. Widom. *Asymptotic behavior of block Toeplitz matrices and determinants.* Advances in Math. 13, (1974), 284-322.