Optimal design, matrix polynomials and random matrices

Holger Dette [5]

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

[1]Katholieke Universiteit Leuven
[2]University of Bonn
[3]Technical University of Munich
[4]Purdue University
[5]Ruhr-University Bochum

August 2011, Cambridge

SFB | TR12

Contents

Classical (weigthed) polynomial regression model

$$Y = \sum_{j=0}^{n-1} \theta_j x^j + \frac{\varepsilon}{\sqrt{\lambda(x)}}$$

$$\lambda(x) = e^{-x^2} \qquad x \in (-\infty, \infty)$$

- ε random error, $E[\varepsilon] = 0$, $E[\varepsilon^2] = 1$
- m independent observations $Y_1, \ldots, Y_m$ at experimental conditions $x_1, \ldots, x_m$
- Variance is proportional to e^{x^2}
- Example: $n = 2$, linear regression model (with heteroscedastic error)

Optimal designs

- Consider approximate designs on $(-\infty, \infty)$
- Weighted least squares estimator: $\hat{\theta} = (\hat{\theta}_0, \ldots, \hat{\theta}_{n-1})$

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

where

$$M(\xi) \;=\; \left(\int_0^\infty x^{i+j} e^{-x^2} d\xi(x) \right)_{i,j=0}^{n-1}$$

denotes the information matrix of the design ξ.

Goal: ***D*-optimal designs**

- Maximize the determinant $|M(\xi)|$ with respect to the choice of the design ξ

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

D-optimal design

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

$$\left\{ z \mid H_n(z) = 0 \right\}$$

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

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

Two Proofs:

- Equivalence theorems (Fedorov, 1972, Stieltjes)
- Moment theory

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

- The optimal design has n support points [equivalence theorem for D-optimality]

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

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

- The optimal design has n support points [equivalence theorem for D-optimality]

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

-

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

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

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

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

- The optimal design has n support points [equivalence theorem for D-optimality]

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

-

$$\begin{aligned} |\mathbf{M}(\xi^*)| \quad &= \quad \prod_{1 \le i < j \le n} (\mathbf{x_i} - \mathbf{x_j})^2 \prod_{i=1}^{n} \mathbf{e}^{-\mathbf{x_i^2}} \prod_{i=1}^{n} \mathbf{w_i} \\ &\longrightarrow \quad \max_{x_i, w_i} \\ &\longrightarrow \quad w_i \;=\; \frac{1}{n} \;\;, \quad i = 1, \dots, n \end{aligned}$$

- $x_1, \dots, x_n$ are the roots of the polynomial $H_n(z)$ [Stieltjes $\rightarrow$ second order differential equation]

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

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

- In other words for any $a, b \in [-\sqrt{2}, \sqrt{2}]$

$$\lim_{n\to\infty} \xi_n^*((a,b]) \;=\; \frac{1}{\pi}\int_a^b \sqrt{2-x^2}dx$$

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

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

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

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 $H_n(z)$

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 $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ of the matrix M_n is given by

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

- (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

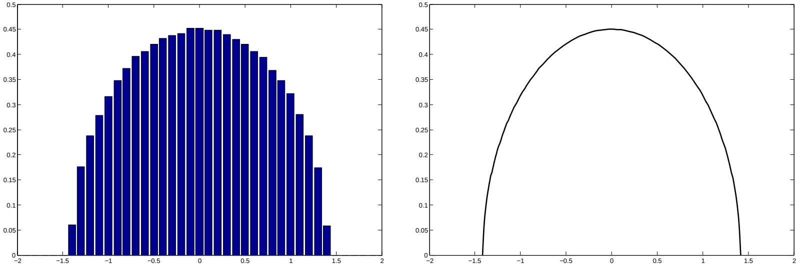

Figure: *Left panel: histogram of the simulated eigenvalues.*
Right panel: asymptotic distribution

Eigenvalues of a 5000×5000 matrix

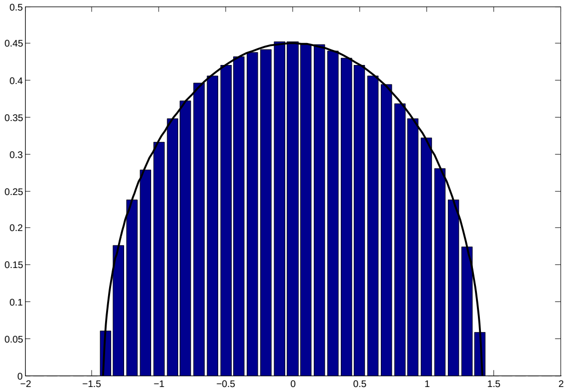

Figure: *Histogram of the simulated eigenvalues and the asymptotic distribution*

β-ensembles

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

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

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

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

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

  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:

- $N_1, N_2, \dots, N_n$ are standard normal distributed ($N_j \sim \mathcal{N}(0,1)$)
- For $j = 1, \dots, n-1$ 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: [Dette, Imhof, Trans. AMS, 2007] If

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

denote the eigenvalues of 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 $(n \to \infty)$

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

Idea of a proof of Theorem 4

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

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

- Consider the (non-random) matrix

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

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

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

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

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

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

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

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

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

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

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

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

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

5-diagonal random 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-diagonal random block matrices (3×3 blocks)

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

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

$2r + 1$-diagonal random 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 matrices
- A_i are lower tridiagonal matrices

Problem: location of the eigenvalues?

Excursion: matrix orthogonal polynomials

- Matrix polynomials

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

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

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

  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 (statistical applications in multivariate time series)
- "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

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

- Orthonormal matrix polynomials are **not** uniquely determined

Excursion: matrix orthogonal polynomials

- Matrix polynomials are called orthonormal if and only if

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

- Orthonormal matrix polynomials are **not** uniquely determined
- 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 [Dette and Studden (2002)]
- The structure of matrix orthogonal polynomials is complicated, because matrix multiplication is not comutative

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 matrices
- A_i are lower tridiagonal matrices

The structure of the blocks

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

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

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

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

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

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

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

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

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: 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 $n \to \infty$:

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

where

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

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

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

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

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

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

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

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

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

Matrix orthogonal polynomials with varying coefficients

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

$$\mathbf{x}\mathbf{R}_{\mathbf{k},\mathbf{n}}(\mathbf{x}) = \mathbf{A}_{\mathbf{k}+\mathbf{1},\mathbf{n}}\mathbf{R}_{\mathbf{k}+\mathbf{1},\mathbf{n}}(\mathbf{x}) + \mathbf{B}_{\mathbf{k},\mathbf{n}}\mathbf{R}_{\mathbf{k},\mathbf{n}}(\mathbf{x}) + \mathbf{A}^{\mathbf{T}}_{\mathbf{k},\mathbf{n}}\mathbf{R}_{\mathbf{k}-\mathbf{1},\mathbf{n}}(\mathbf{x}); \quad \mathbf{k} \geq \mathbf{0},$$

  where

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

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

$$\mathbf{Q}_{\mathbf{k},\mathbf{n}}(\mathbf{x})$$

  if $n \to \infty$?

- **Note:** By Theorem 5 we expect that the eigenvalues of the random band matrix have similar properties!

An algebraic equation

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

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

and 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

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

Weak asymptotics for matrix orthonormal polynomials

Theorem 6 Let

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

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

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

where

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

Then ν_n converges weakly to an absolute continuous measure $\mu_{0,u}$, with

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

Application to random block matrices

- By Theorem 5 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 6 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

$$\mathbf{0} = \mathbf{f}(\mathbf{z}, \mathbf{x}) := \det(\mathbf{A}^\mathbf{T}\mathbf{z} + \mathbf{B} + \mathbf{A}\mathbf{z}^{-\mathbf{1}} - \mathbf{x}\mathbf{I}_\mathbf{r})$$

Application to random block matrices

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

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

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

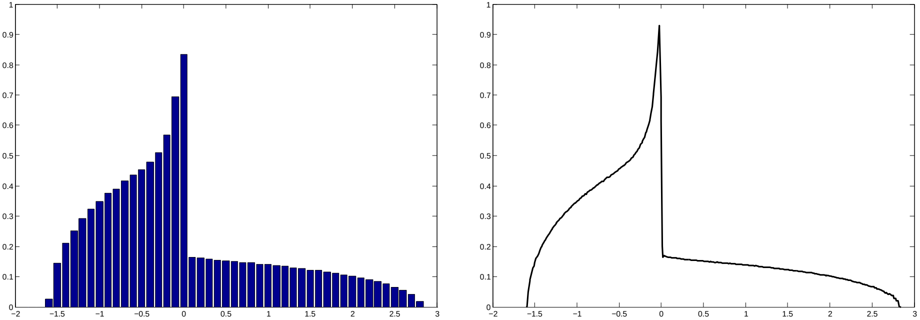

Figure: *Left panel: histogram of the simulated eigenvalues*
Right panel: asymptotic distribution

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

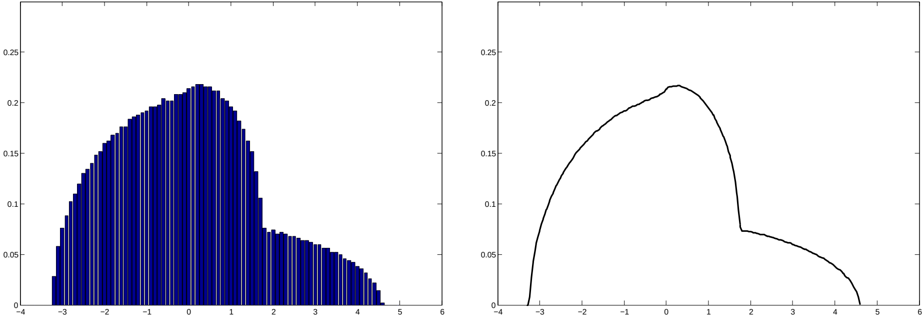

Figure: *Left panel: histogram of the simulated eigenvalues*
Right panel: asymptotic distribution

Eigenvalues of a 5000×5000 matrix

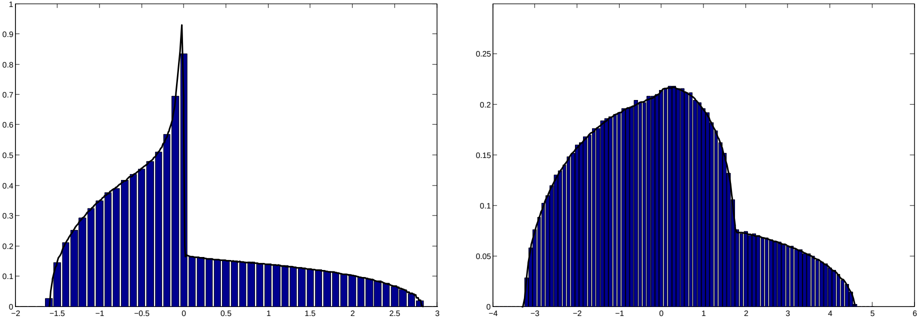

Figure: *Left panel: histogram and density* $(\beta_1 = 1;\ \beta_2 = 1)$
Right panel: histogram and density $(\beta_1 = 5;\ \beta_2 = 1)$