Computational methods for large-scale matrix equations: recent advances

V. Simoncini

Dipartimento di Matematica, Università di Bologna (Italy)

valeria.simoncini@unibo.it

Some matrix equations

- Sylvester matrix equation

$$A\mathbf{X} + \mathbf{X}B + D = 0$$

Eigenvalue pbs and tracking, Control, MOR, Assignment pbs, Riccati eqn

Some matrix equations

- Sylvester matrix equation

$$A\mathbf{X} + \mathbf{X}B + D = 0$$

Eigenvalue pbs and tracking, Control, MOR, Assignment pbs, Riccati eqn

- Lyapunov matrix equation

$$A\mathbf{X} + \mathbf{X}A^\top + D = 0, \qquad D = D^\top$$

Stability analysis in Control and Dynamical systems, Signal processing, eigenvalue computations

Some matrix equations

- Sylvester matrix equation

$$A\mathbf{X} + \mathbf{X}B + D = 0$$

Eigenvalue pbs and tracking, Control, MOR, Assignment pbs, Riccati eqn

- Lyapunov matrix equation

$$A\mathbf{X} + \mathbf{X}A^\top + D = 0, \qquad D = D^\top$$

Stability analysis in Control and Dynamical systems, Signal processing, eigenvalue computations

- Algebraic Riccati equation

$$A\mathbf{X} + \mathbf{X}A^\top - \mathbf{X}BB^\top\mathbf{X} + D = 0, \qquad D = D^\top$$

Lancaster-Rodman '95, Konstantinov-Gu-Mehrmann-Petkov, '02, Bini-Iannazzo-Meini '12

Some matrix equations

- Sylvester matrix equation

$$A\mathbf{X} + \mathbf{X}B + D = 0$$

Eigenvalue pbs and tracking, Control, MOR, Assignment pbs, Riccati eqn

- Lyapunov matrix equation

$$A\mathbf{X} + \mathbf{X}A^\top + D = 0, \qquad D = D^\top$$

Stability analysis in Control and Dynamical systems, Signal processing, eigenvalue computations

- Algebraic Riccati equation

$$A\mathbf{X} + \mathbf{X}A^\top - \mathbf{X}BB^\top\mathbf{X} + D = 0, \qquad D = D^\top$$

Lancaster-Rodman '95, Konstantinov-Gu-Mehrmann-Petkov, '02, Bini-Iannazzo-Meini '12

Focus: All or some of the matrices are large (and possibly sparse)

The Lyapunov operator

$$\mathcal{L} : X \mapsto AX + XA^\top \qquad \text{or} \qquad \ell : x \mapsto (I \otimes A + A \otimes I)x$$

- In linear matrix equations: Computational aspects
- A mathematical tool

Solving the Lyapunov equation. The problem

Approximate X in:

$$AX + XA^\top + BB^\top = 0$$

$A \in \mathbb{R}^{n\times n}$ neg.real $\qquad B \in \mathbb{R}^{n\times p}, \qquad 1 \le p \ll n$

Solving the Lyapunov equation. The problem

Approximate X in:

$$AX + XA^\top + BB^\top = 0$$

$A \in \mathbb{R}^{n\times n}$ neg.real $\qquad B \in \mathbb{R}^{n\times p}, \qquad 1 \le p \ll n$

Time-invariant linear system:

$$\mathbf{x}'(t) = A\mathbf{x}(t) + B\mathbf{u}(t), \qquad \mathbf{x}(0) = x_0$$

Closed form solution:

$$X = \frac{1}{2\pi}\int_{-\infty}^{\infty} (\imath\omega I + A)^{-1} BB^\top (\imath\omega I + A)^{-H} \mathrm{d}\omega$$

$\Rightarrow \qquad X$ symmetric semidef.

see, e.g., Antoulas '05, Benner '06

Linear systems vs linear matrix equations

Large linear systems:

$$Ax = b, \quad A \in \mathbb{R}^{n\times n}$$

- Krylov subspace methods (CG, MINRES, GMRES, BiCGSTAB, etc.)
- Preconditioners: find P such that

$$AP^{-1}\widetilde{x} = b \qquad x = P^{-1}\widetilde{x}$$

is easier and fast to solve

Linear systems vs linear matrix equations

Large linear systems:

$$Ax = b, \quad A \in \mathbb{R}^{n\times n}$$

- Krylov subspace methods (CG, MINRES, GMRES, BiCGSTAB, etc.)
- Preconditioners: find P such that

$$AP^{-1}\widetilde{x} = b \qquad x = P^{-1}\widetilde{x}$$

  is easier and fast to solve

Large linear matrix equations:

$$AX + XA^\top + BB^\top = 0$$

- No preconditioning - to preserve symmetry
- X is a large, dense matrix $\Rightarrow$ low rank approximation

$$X \approx \widetilde{X} = ZZ^\top, \quad Z \text{ tall}$$

Linear systems vs linear matrix equations

Large linear systems:

$$Ax = b, \quad A \in \mathbb{R}^{n\times n}$$

- Krylov subspace methods (CG, MINRES, GMRES, BiCGSTAB, etc.)
- Preconditioners: find P such that

$$AP^{-1}\widetilde{x} = b \qquad x = P^{-1}\widetilde{x}$$

is easier and fast to solve

Large linear matrix equations:

$$AX + XA^\top + BB^\top = 0$$

Kronecker formulation:

$$(A \otimes I + I \otimes A)x = b \qquad x = \mathrm{vec}(X)$$

Projection-type methods

Given an approximation space $\mathcal{K}$,

$$X \approx X_m \qquad \mathrm{col}(X_m) \in \mathcal{K}$$

Galerkin condition: $\quad R := AX_m + X_m A^\top + BB^\top \quad \perp \quad \mathcal{K}$

$$V_m^\top R V_m = 0 \qquad\qquad \mathcal{K} = \mathrm{Range}(V_m)$$

Projection-type methods

Given an approximation space $\mathcal{K}$,

$$X \approx X_m \qquad \mathrm{col}(X_m) \in \mathcal{K}$$

Galerkin condition: $\quad R := AX_m + X_m A^\top + BB^\top \quad \perp \quad \mathcal{K}$

$$V_m^\top R V_m = 0 \qquad\qquad \mathcal{K} = \mathrm{Range}(V_m)$$

Assume $V_m^\top V_m = I_m$ and let $X_m := V_m Y_m V_m^\top$.

Projected Lyapunov equation:

$$V_m^\top (AV_m Y_m V_m^\top + V_m Y_m V_m^\top A^\top + BB^\top) V_m \quad = \quad 0$$

Projection-type methods

Given an approximation space $\mathcal{K}$,

$$X \approx X_m \qquad \mathrm{col}(X_m) \in \mathcal{K}$$

Galerkin condition: $\quad R := AX_m + X_m A^\top + BB^\top \quad \perp \quad \mathcal{K}$

$$V_m^\top R V_m = 0 \qquad\qquad \mathcal{K} = \mathrm{Range}(V_m)$$

Assume $V_m^\top V_m = I_m$ and let $X_m := V_m Y_m V_m^\top$.

Projected Lyapunov equation:

$$\begin{aligned} V_m^\top (A V_m Y_m V_m^\top + V_m Y_m V_m^\top A^\top + BB^\top) V_m &= 0 \\ (V_m^\top A V_m) Y_m + Y_m (V_m^\top A^\top V_m) + V_m^\top BB^\top V_m &= 0 \end{aligned}$$

Early contributions: Saad '90, Jaimoukha & Kasenally '94, for $\mathcal{K} = \mathcal{K}_m(A, B) = \mathrm{Range}([B, AB, \ldots, A^{m-1}B])$

More recent options as approximation space

Enrich space to decrease space dimension

- Extended Krylov subspace

$$\mathcal{K} = \mathcal{K}_m(A, B) + \mathcal{K}_m(A^{-1}, A^{-1}B),$$

that is, $\mathcal{K} = \mathrm{Range}([B, A^{-1}B, AB, A^{-2}B, A^2, A^{-3}B, \ldots,])$

(Druskin & Knizhnerman '98, Simoncini '07)

More recent options as approximation space

Enrich space to decrease space dimension

- Extended Krylov subspace

$$\mathcal{K} = \mathcal{K}_m(A, B) + \mathcal{K}_m(A^{-1}, A^{-1}B),$$

that is, $\mathcal{K} = \mathrm{Range}([B, A^{-1}B, AB, A^{-2}B, A^2, A^{-3}B, \ldots,])$

(Druskin & Knizhnerman '98, Simoncini '07)

- Rational Krylov subspace

$$\mathcal{K} = \mathrm{Range}([B, (A - s_1 I)^{-1}B, \ldots, (A - s_m I)^{-1}B])$$

usually, $\{s_1, \ldots, s_m\} \subset \mathbb{C}^+$ chosen a-priori

More recent options as approximation space

Enrich space to decrease space dimension

- Extended Krylov subspace

$$\mathcal{K} = \mathcal{K}_m(A, B) + \mathcal{K}_m(A^{-1}, A^{-1}B),$$

that is, $\mathcal{K} = \mathrm{Range}([B, A^{-1}B, AB, A^{-2}B, A^2, A^{-3}B, \ldots,])$

(Druskin & Knizhnerman '98, Simoncini '07)

- Rational Krylov subspace

$$\mathcal{K} = \mathrm{Range}([B, (A - s_1 I)^{-1}B, \ldots, (A - s_m I)^{-1}B])$$

usually, $\{s_1, \ldots, s_m\} \subset \mathbb{C}^+$ chosen a-priori

In both cases, for Range$(V_m) = \mathcal{K}$, projected Lyapunov equation:

$$(V_m^\top A V_m)Y_m + Y_m(V_m^\top A^\top V_m) + V_m^\top BB^\top V_m = 0$$

$X_m = V_m Y_m V_m^\top$

Rational Krylov Subspaces. A long tradition...

In general,

$$K_m(A, B, \mathbf{s}) = \mathrm{Range}([(A - s_1 I)^{-1} B, (A - s_2 I)^{-1} B, \ldots, (A - s_m I)^{-1} B])$$

- Eigenvalue problems (Ruhe, 1984)
- Model Order Reduction (transfer function evaluation)
- In Alternating Direction Implicit iteration (ADI) for linear matrix equations

The Lyapunov operator

$$\mathcal{L} : X \mapsto AX + XA^\top \qquad \text{or} \qquad \ell : x \mapsto (I \otimes A + A \otimes I)x$$

- In linear matrix equations: Computational aspects
- A mathematical tool

The application. I

Adaptive Legendre-Galerkin discretizations for PDEs:

H_0^1 Tensorized Babuska-Shen basis in $\Omega = (0,1) \times (0,1)$:

$$\eta_{\mathbf{k}}(x_1, x_2) = \eta_{k_1}(x_1)\eta_{k_2}(x_2), \quad k_1, k_2 \geq 2, \quad \mathbf{k} = (k_1, k_2)$$

$\{\eta_{k_i}\}$: k_i-order Legendre polyn (1D BS basis)

Stiffness matrix:

$$(\eta_{\mathbf{k}}, \eta_{\mathbf{m}})_{H_0^1(\Omega)} = (\eta_{k_1}, \eta_{m_1})_{H_0^1(I)}(\eta_{k_2}, \eta_{m_2})_{L^2(I)} + (\eta_{k_1}, \eta_{m_1})_{L^2(I)}(\eta_{k_2}, \eta_{m_2})_{H_0^1(I)}$$

Kronecker structure: $S_\eta^p = M_p \otimes I_p + I_p \otimes M_p$ (max p polyn degree)

The application. I

Adaptive Legendre-Galerkin discretizations for PDEs:

H_0^1 Tensorized Babuska-Shen basis in $\Omega = (0,1) \times (0,1)$:

$$\eta_{\mathbf{k}}(x_1, x_2) = \eta_{k_1}(x_1)\eta_{k_2}(x_2), \quad k_1, k_2 \geq 2, \quad \mathbf{k} = (k_1, k_2)$$

$\{\eta_{k_i}\}$: k_i-order Legendre polyn (1D BS basis)

Stiffness matrix:

$$(\eta_{\mathbf{k}}, \eta_{\mathbf{m}})_{H_0^1(\Omega)} = (\eta_{k_1}, \eta_{m_1})_{H_0^1(I)}(\eta_{k_2}, \eta_{m_2})_{L^2(I)} + (\eta_{k_1}, \eta_{m_1})_{L^2(I)}(\eta_{k_2}, \eta_{m_2})_{H_0^1(I)}$$

Kronecker structure: $S_\eta^p = M_p \otimes I_p + I_p \otimes M_p$ (max p polyn degree)

Note: If higher order polynomial used, then S_η^p simply expands (augmented M_p)

The application. II

Adaptive Legendre-Galerkin discretizations for PDEs:

- Inner product:

$$v = \sum \hat{v}_{\mathbf{k}} \eta_{\mathbf{k}}, \qquad \|v\|^2_{H^1_0} = \hat{v}^T S_\eta \hat{v}$$

The application. II

Adaptive Legendre-Galerkin discretizations for PDEs:

- Inner product:

$$v = \sum \hat{v}_{\mathbf{k}} \eta_{\mathbf{k}}, \qquad \|v\|^2_{H^1_0} = \hat{v}^T S_\eta \hat{v}$$

- (Full!) Orthonormalization: $\{\Phi_{\mathbf{k}}\}$ orth basis,

$$v = \sum \tilde{v}_{\mathbf{k}} \Phi_{\mathbf{k}}, \qquad \|v\|^2_{H^1_0} = \tilde{v}^T G^T S_\eta (G\tilde{v}) = \tilde{v}^T \tilde{v}$$

with $G = L^{-1}$ where $S_\eta = LL^T$

The application. II

Adaptive Legendre-Galerkin discretizations for PDEs:

- Inner product:

$$v = \sum \hat{v}_{\mathbf{k}} \eta_{\mathbf{k}}, \qquad \|v\|^2_{H_0^1} = \hat{v}^T S_\eta \hat{v}$$

- (Full!) Orthonormalization: $\{\Phi_{\mathbf{k}}\}$ orth basis,

$$v = \sum \tilde{v}_{\mathbf{k}} \Phi_{\mathbf{k}}, \qquad \|v\|^2_{H_0^1} = \tilde{v}^T G^T S_\eta (G\tilde{v}) = \tilde{v}^T \tilde{v}$$

  with $G = L^{-1}$ where $S_\eta = LL^T$

- (Cheap!) Quasi-orthonormalization: $\{\Psi_{\mathbf{k}}\}$ quasi-orth basis,

$$v = \sum \check{v}_{\mathbf{k}} \Psi_{\mathbf{k}}, \qquad \|v\|^2_{H_0^1} = \tilde{v}^T \check{G}^T S_\eta (\check{G}\tilde{v}) \approx \tilde{v}^T D \tilde{v}$$

  $\check{G}$ very sparse version of G, D diagonal

The application. II

Adaptive Legendre-Galerkin discretizations for PDEs:

- Inner product:

$$v = \sum \hat{v}_{\mathbf{k}} \eta_{\mathbf{k}}, \qquad \|v\|^2_{H^1_0} = \hat{v}^T S_\eta \hat{v}$$

- (Full!) Orthonormalization: $\{\Phi_{\mathbf{k}}\}$ orth basis,

$$v = \sum \tilde{v}_{\mathbf{k}} \Phi_{\mathbf{k}}, \qquad \|v\|^2_{H^1_0} = \tilde{v}^T G^T S_\eta (G\tilde{v}) = \tilde{v}^T \tilde{v}$$

  with $G = L^{-1}$ where $S_\eta = LL^T$

- (Cheap!) Quasi-orthonormalization: $\{\Psi_{\mathbf{k}}\}$ quasi-orth basis,

$$v = \sum \check{v}_{\mathbf{k}} \Psi_{\mathbf{k}}, \qquad \|v\|^2_{H^1_0} = \tilde{v}^T \check{G}^T S_\eta (\check{G}\tilde{v}) \approx \tilde{v}^T D \tilde{v}$$

  $\check{G}$ very sparse version of G, D diagonal

Q: Does such a $\check{G}$ exist? ...Analyze sparsity of S_η^{-1}

The stiffness matrix

$$S := M \otimes I_n + I_n \otimes M,$$

with M symmetric and positive definite, banded with bandwidth b

- Finite differences: M is second order operator in one space dimension ($b = 1$)

  $\Rightarrow$ for instance, S: 2D Laplace operator in $[a, b]^2$

- Legendre Spectral methods: M spd, nonconstant ($b = 1$)

- ...

More generally,

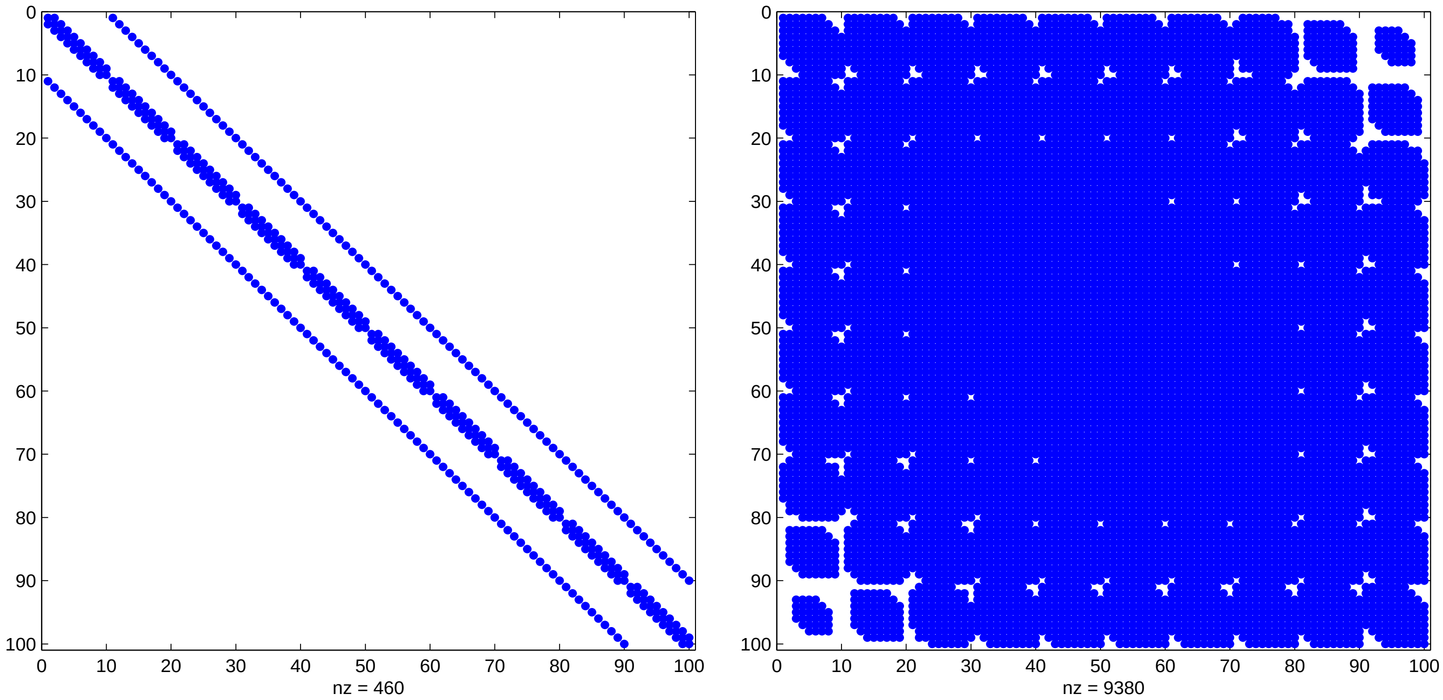
$$S_g := M_1 \otimes I_n + I_n \otimes M_2,$$

with $M_1 \neq M_2$, banded, with not necessarily the same dimensions

The inverse of the 2D Laplace matrix on the unit square

$$S := M \otimes I_n + I_n \otimes M, \qquad M = \mathrm{tridiag}(-1, 2, -1)$$

Sparsity pattern:

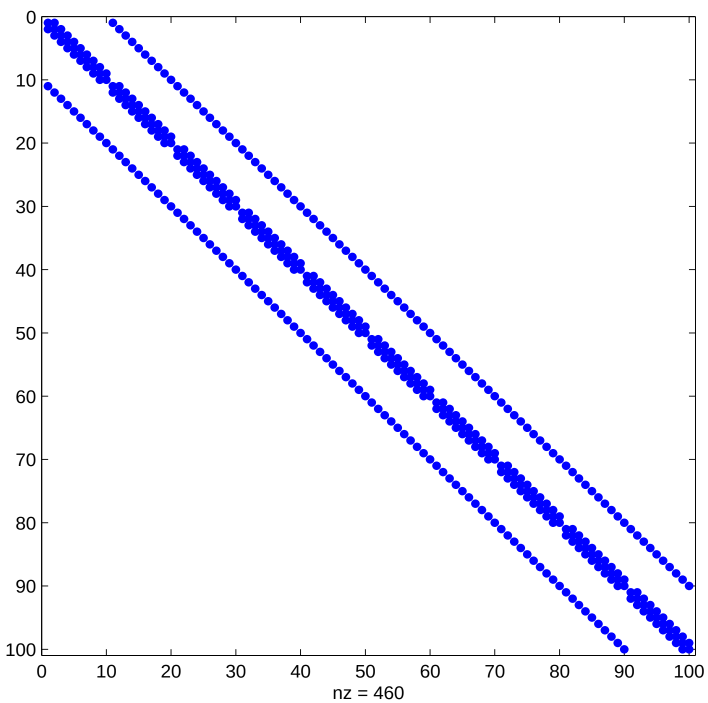


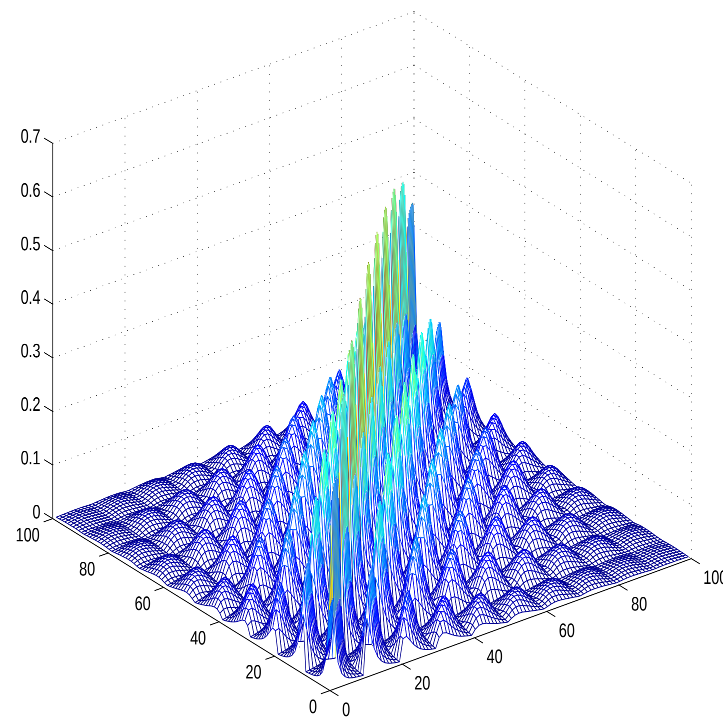
Matrix S $\qquad$ S^{-1}

The inverse of the 2D Laplace matrix on the unit square

$$S := M \otimes I_n + I_n \otimes M, \qquad M = \mathrm{tridiag}(-1, 2, -1)$$

Sparsity pattern:





S $\qquad\qquad$ $|(S^{-1})_{ij}|$

The exponential decay of the entries of S^{-1}

The classical bound (Demko, Moss & Smith):

If S spd is banded with bandwidth b, then

$$|(S^{-1})_{ij}| \leq \gamma q^{\frac{|i-j|}{b}}$$

where

κ: condition number of S

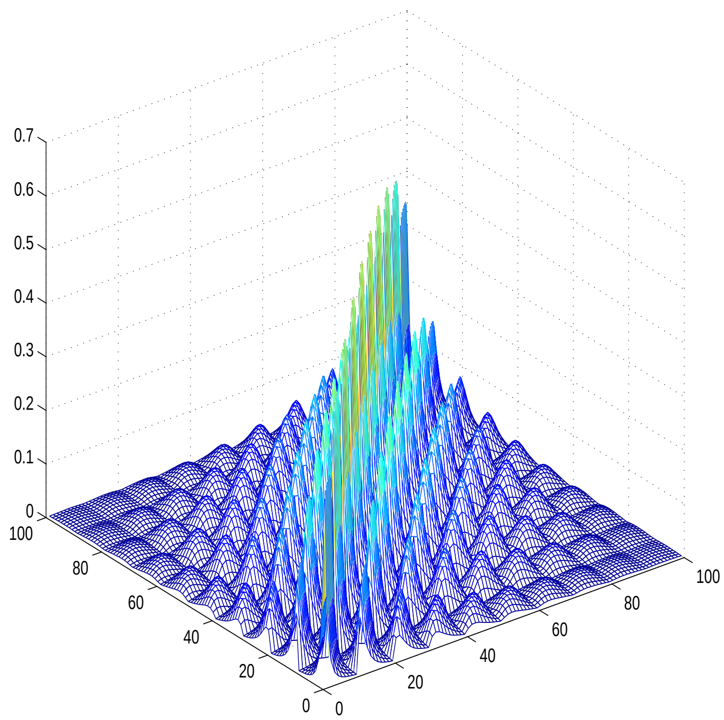
$q := \dfrac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} < 1$

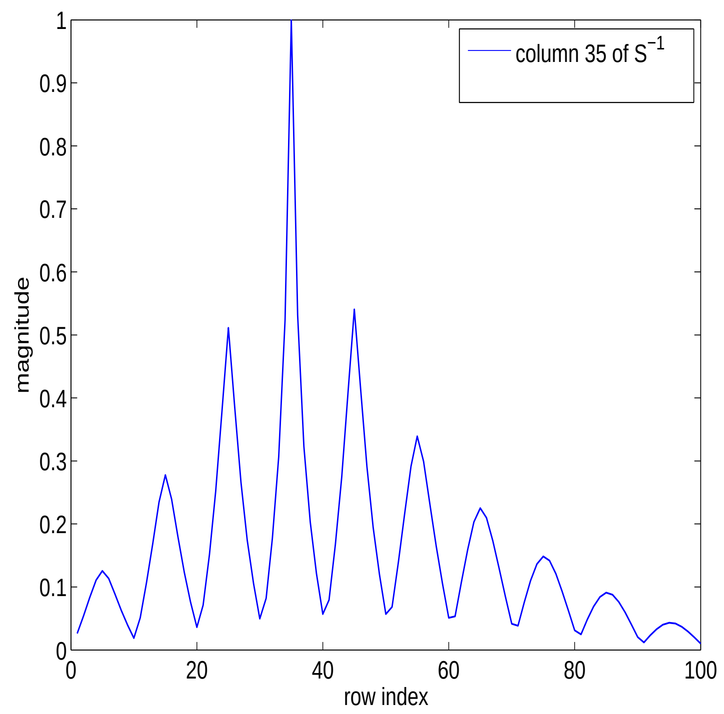
$\gamma := \max\{\lambda_{\min}(S)^{-1}, \hat{\gamma}\}$, and $\hat{\gamma} = \dfrac{(1+\sqrt{\kappa})^2}{2\lambda_{\max}(S)}$

($\lambda_{\min}(\cdot), \lambda_{\max}(\cdot)$ smallest and largest eigenvalues of the given symmetric matrix)

Many contributions: Bebendorf, Hackbusch, Benzi, Boito, Razouk, Golub, Tuma, Concus, Meurant, Mastronardi, Ng, Tyrtyshnikov, Nabben, ...

The true decay





... a very peculiar pattern $\Rightarrow$ much higher sparsity

Where do the repeated peaks come from?

For $S = M \otimes I_n + I_n \otimes M \in \mathbb{R}^{n^2 \times n^2}$:

$$x_t := (S^{-1})_{:,t} = S^{-1} e_t \qquad \Leftrightarrow \qquad \text{Solve}: \; S x_t = e_t$$

Where do the repeated peaks come from?

For $S = M \otimes I_n + I_n \otimes M \in \mathbb{R}^{n^2 \times n^2}$:

$$x_t := (S^{-1})_{:,t} = S^{-1} e_t \qquad \Leftrightarrow \qquad \text{Solve}: \; S x_t = e_t$$

Let

$X_t \in \mathbb{R}^{n \times n}$ be such that $x_t = \mathrm{vec}(X_t)$

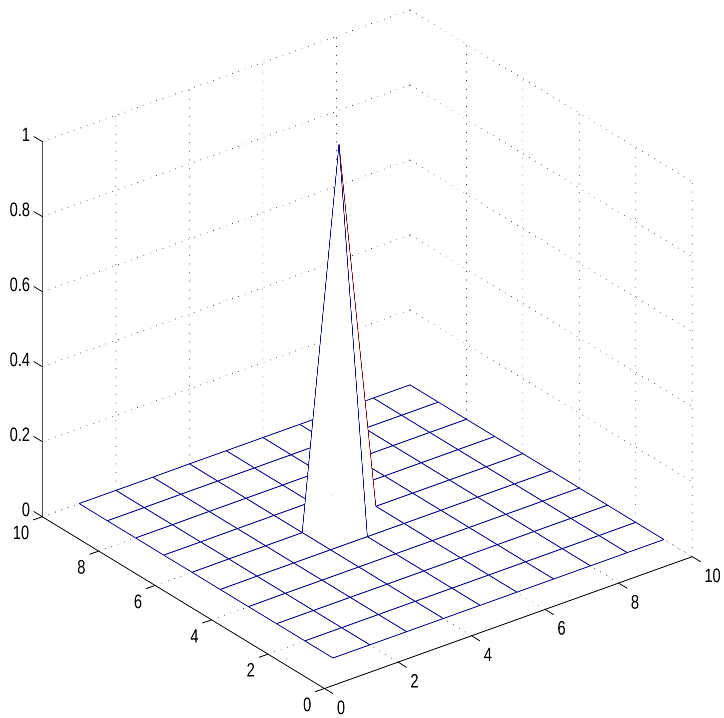
$E_t \in \mathbb{R}^{n \times n}$ be such that $e_t = \mathrm{vec}(E_t)$

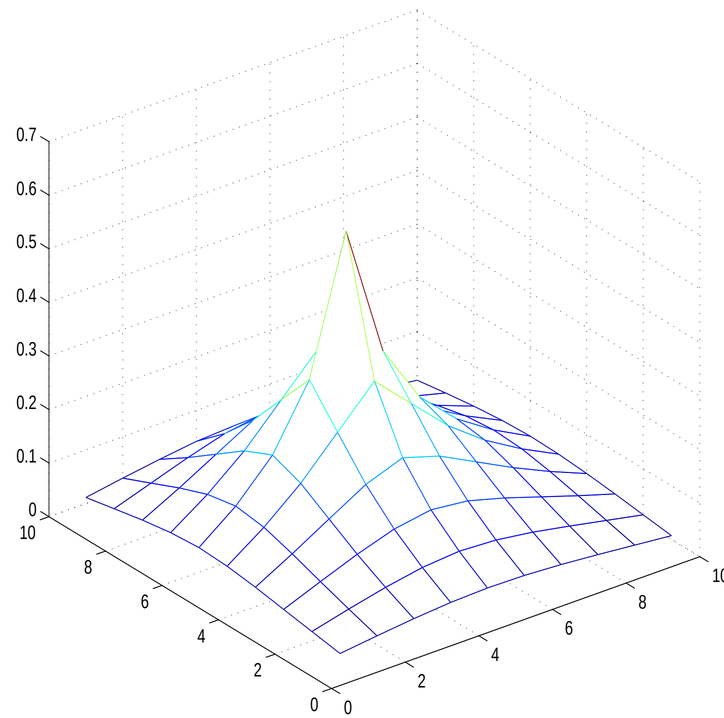
Then

$$S x_t = e_t \qquad \Leftrightarrow \qquad M X_t + X_t M = E_t$$

For S the 2D Laplace operator, $t = 1, \ldots, n^2$

$$t = 35, \qquad Sx_t = e_t \qquad \Leftrightarrow \qquad MX_t + X_tM = E_t$$

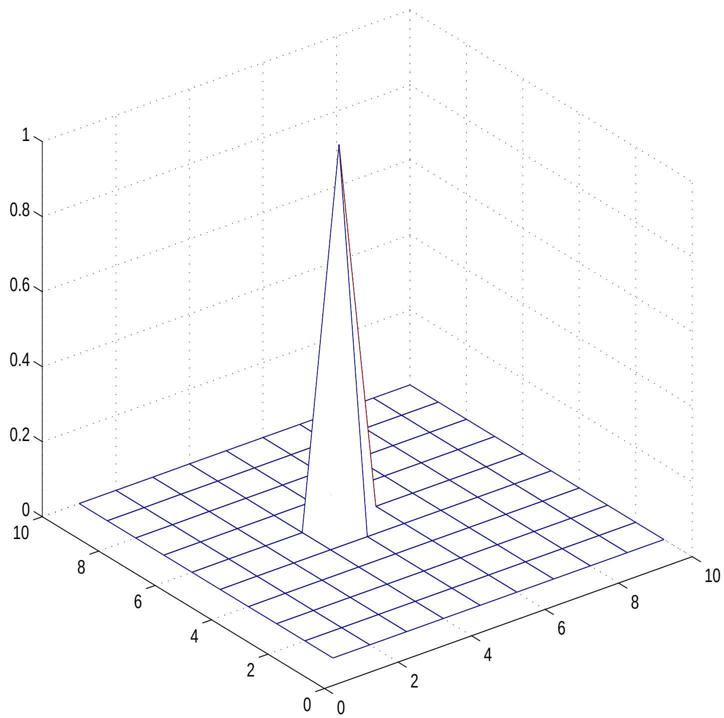


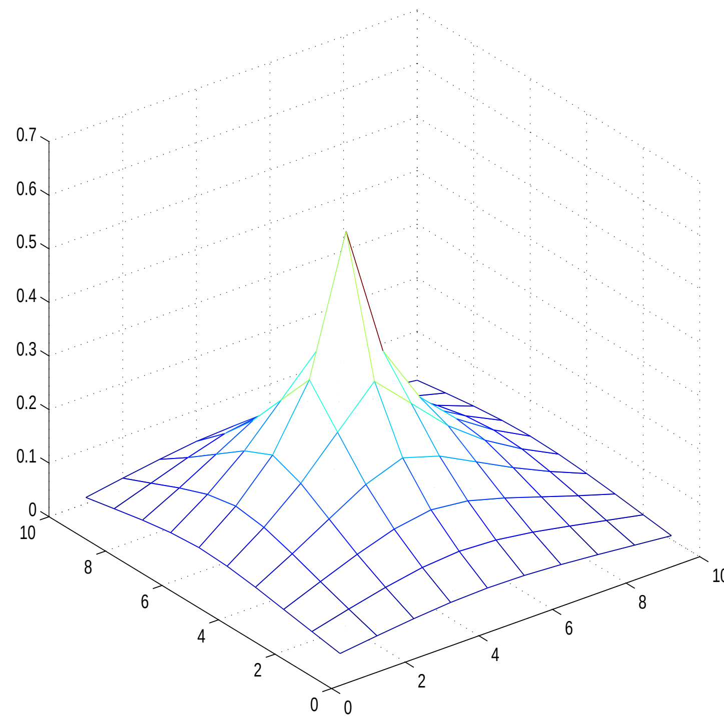


matrix E_t and matrix X_t

For S the 2D Laplace operator, $t = 1, \ldots, n^2$

$$t = 35, \qquad Sx_t = e_t \qquad \Leftrightarrow \qquad MX_t + X_t M = E_t$$

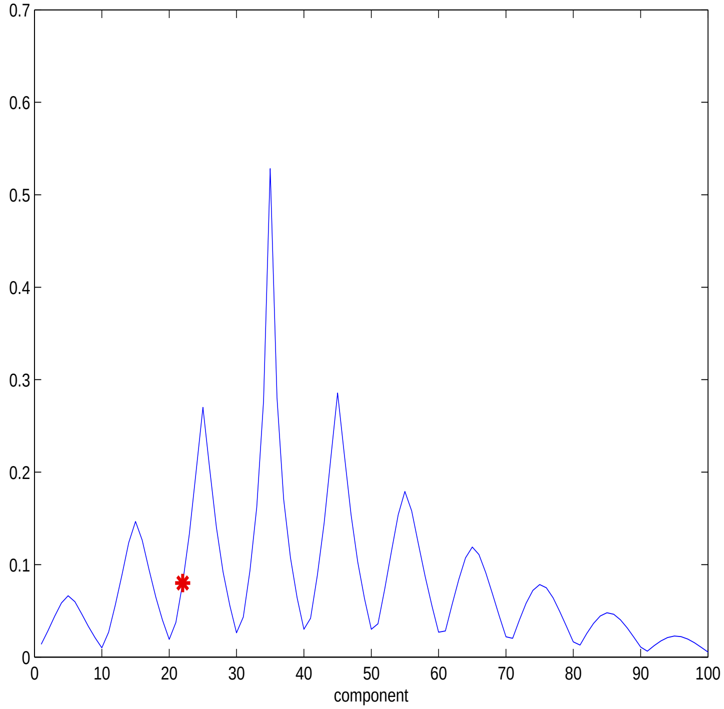


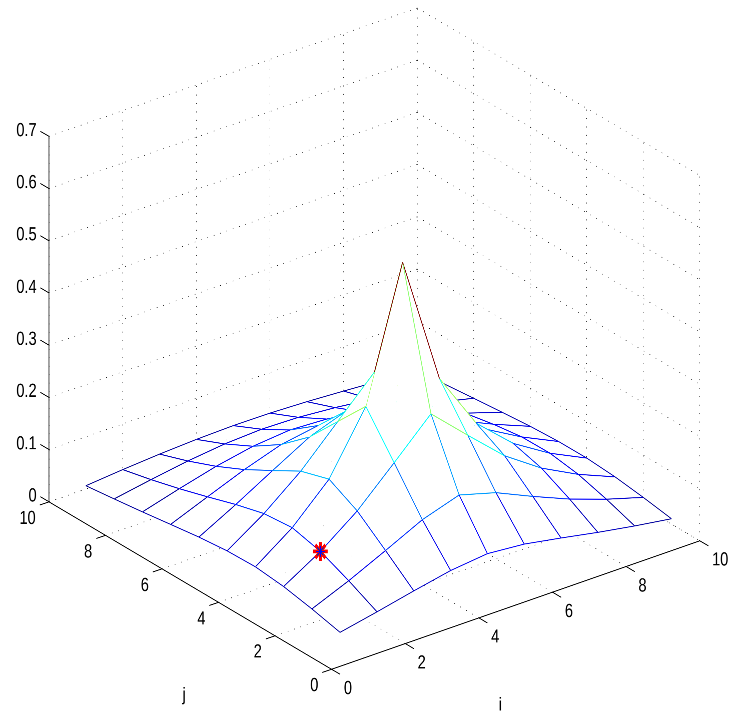


matrix E_t and matrix X_t

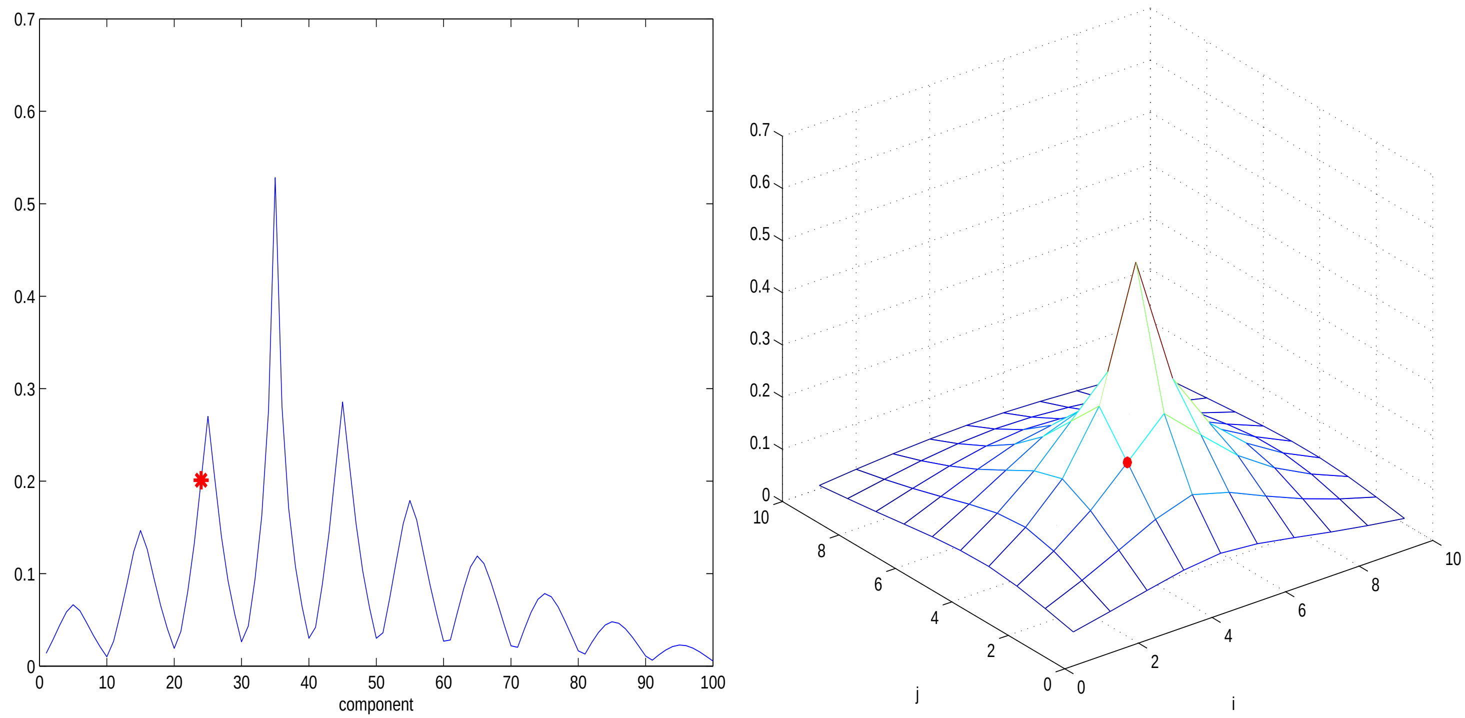
E_t has only one nonzero element

Lexicographic order: $(E_t)_{ij}$, $j = \lfloor (t-1)/n \rfloor + 1$, $i = tn\lfloor (t-1)/n \rfloor$

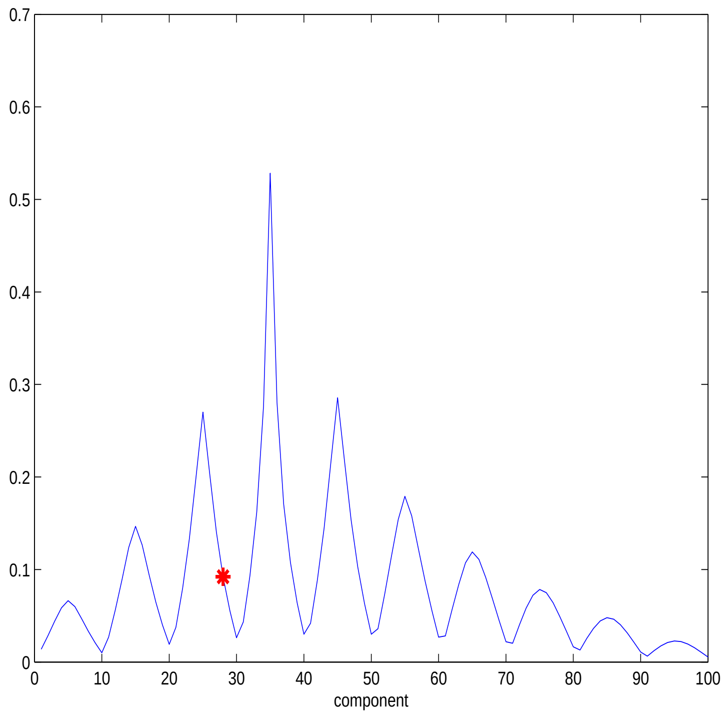


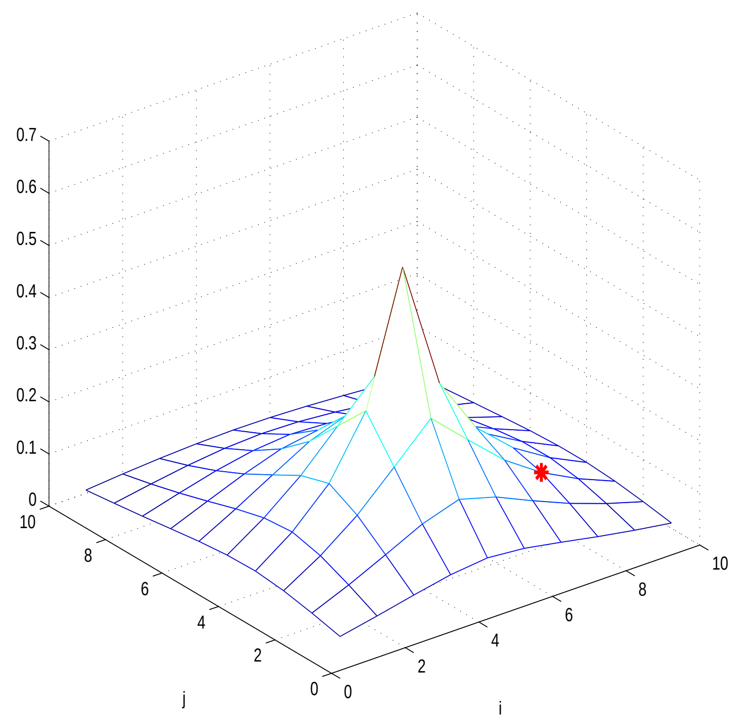


Left: Row of S^{-1} Right: same row on the grid

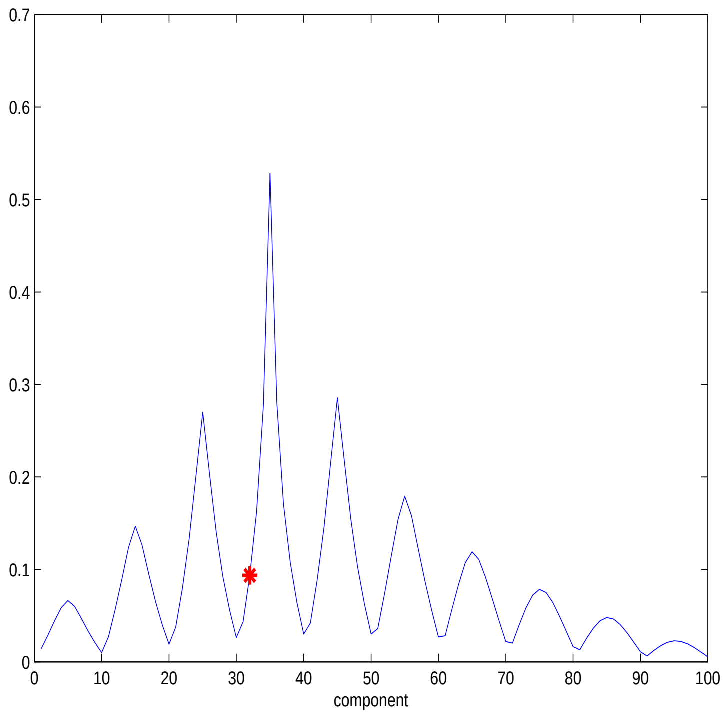


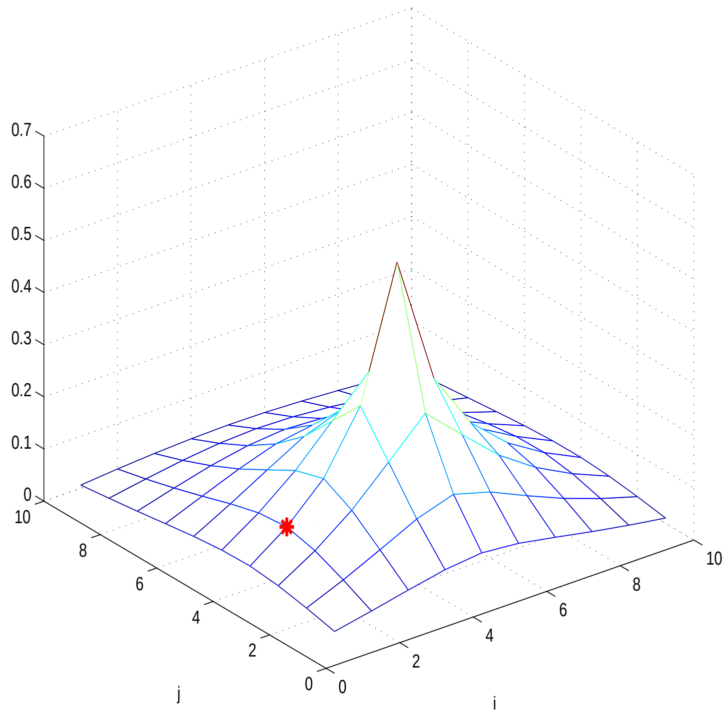
Left: Row of S^{-1} Right: same row on the grid



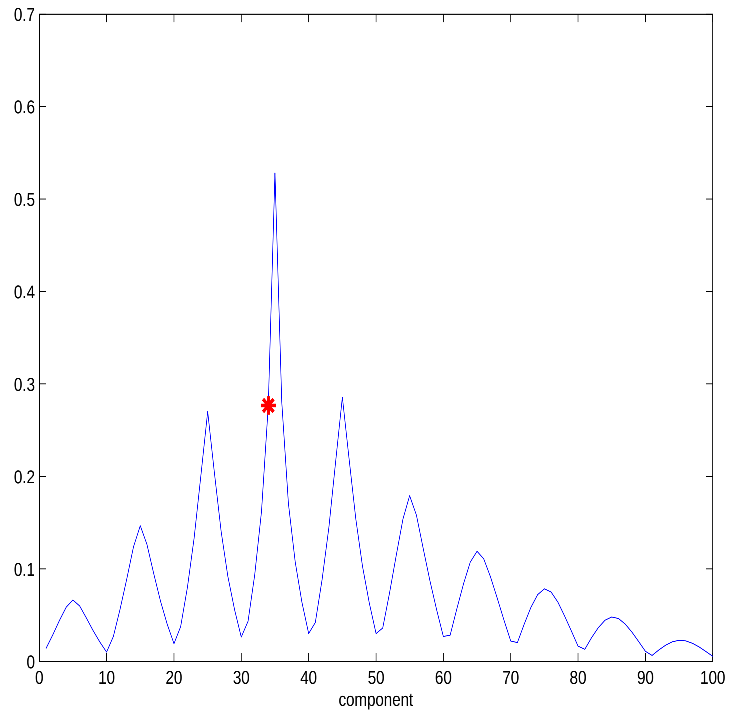


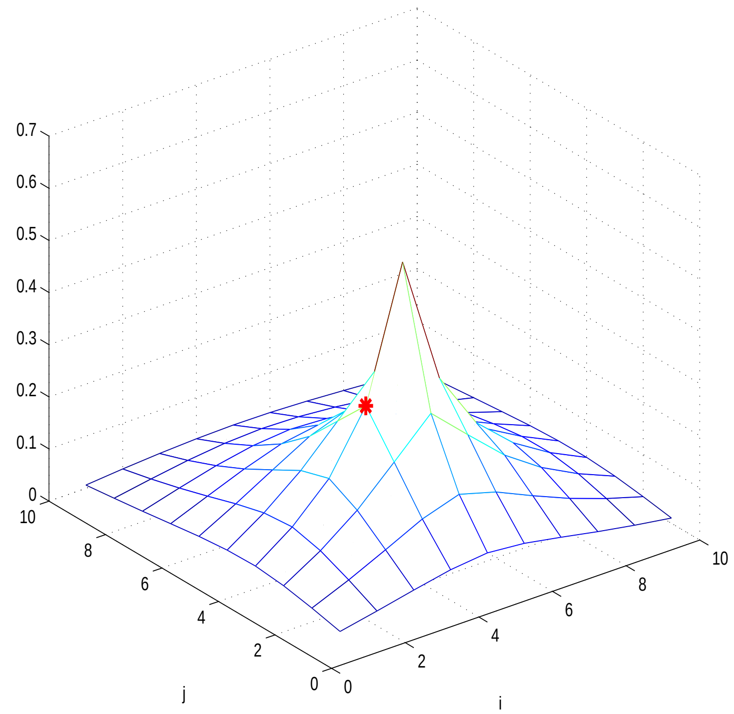
Left: Row of S^{-1} Right: same row on the grid



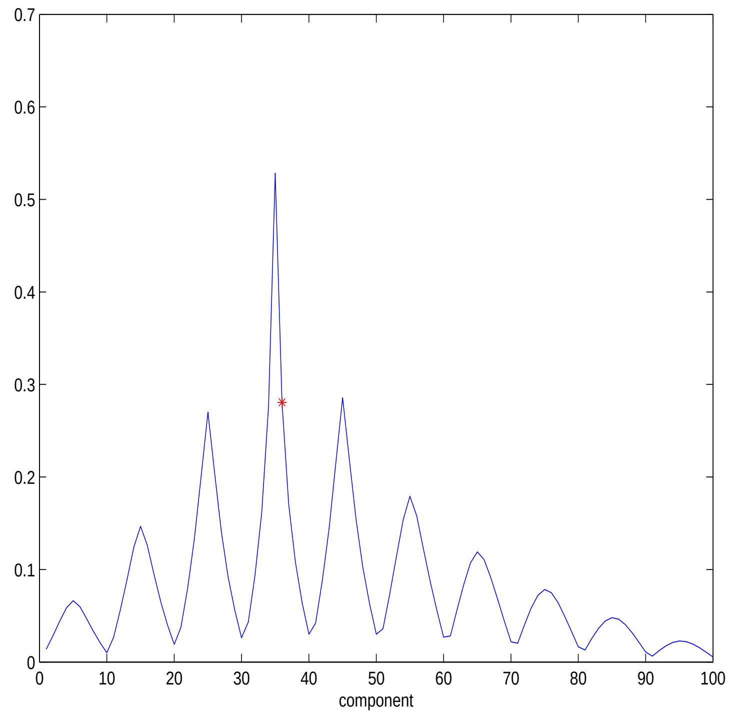


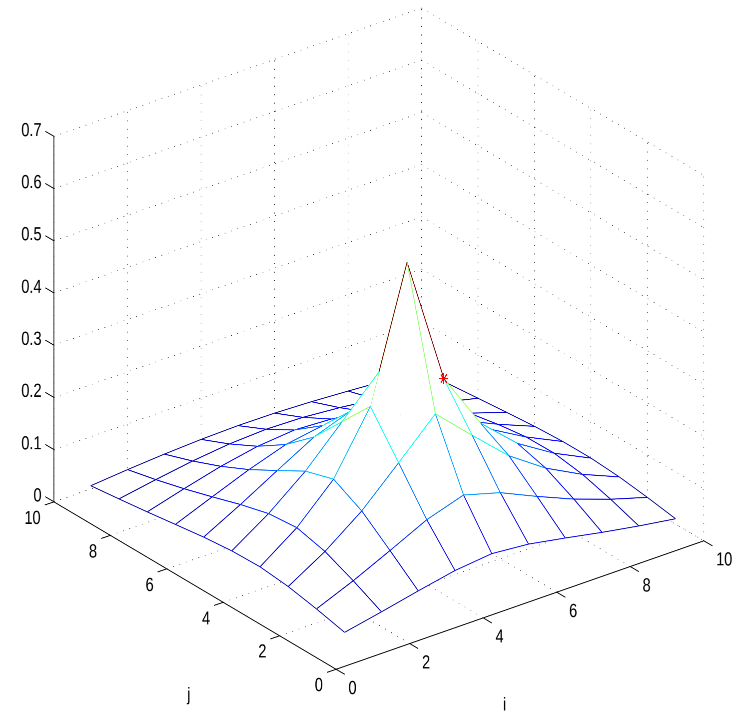
Left: Row of S^{-1} Right: same row on the grid





Left: Row of S^{-1} Right: same row on the grid





Left: Row of S^{-1} Right: same row on the grid

Resolving the entry indexing using $MX_t + X_tM = E_t$

$$(S^{-1})_{k,t} = (S^{-1})_{\ell+n(m-1),t} = e_\ell^\top X_t e_m, \quad \ell, m \in \{1, \ldots, n\}$$

$\Rightarrow$ All the elements of the t-th column, $(S^{-1})_{:,t}$, are obtained by varying $m, \ell \in \{1, \ldots, n\}$

Resolving the entry indexing using $MX_t + X_tM = E_t$

$$(S^{-1})_{k,t} = (S^{-1})_{\ell+n(m-1),t} = e_\ell^\top X_t e_m, \quad \ell, m \in \{1, \ldots, n\}$$

$\Rightarrow$ All the elements of the t-th column, $(S^{-1})_{:,t}$, are obtained by varying $m, \ell \in \{1, \ldots, n\}$

From the Lyapunov equation theory,

$$X_t = \frac{1}{2\pi} \int_{-\infty}^{\infty} (\imath\omega I + M)^{-1} E_t (\imath\omega I + M)^{-*} \mathrm{d}\omega$$

with $E_t = e_i e_j^\top$, $j = \lfloor (t-1)/n \rfloor + 1$, $i = t - n\lfloor (t-1)/n \rfloor$

Therefore,

$$e_\ell^\top \mathcal{X}_t e_m = \frac{1}{2\pi} \int_{-\infty}^{\infty} e_\ell^\top (\imath\omega I + M)^{-1} e_i e_j^\top (\imath\omega I + M)^{-*} e_m \mathrm{d}\omega$$

Qualitative bounds

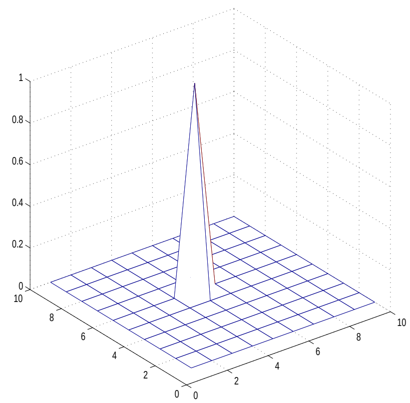
Let $\kappa = \lambda_{\max}/\lambda_{\min} = \mathrm{cond}(M)$

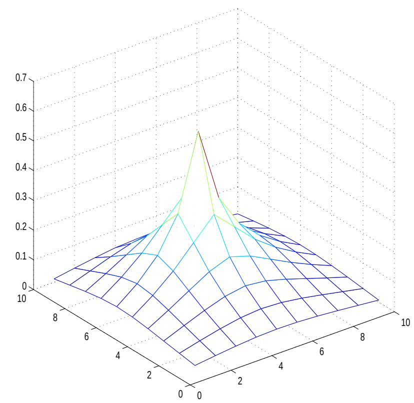
i)Assume $\ell, i, m, j \; : \; \ell \neq i, \; m \neq j$. $\mathfrak{n}_2 := |\ell - i| + |m - j| - 2 > 0$

$$|(S^{-1})_{k,t}| \leq \frac{\sqrt{\kappa^2+1}}{2\lambda_{\min}} \frac{1}{\sqrt{\mathfrak{n}_2}}.$$

ii)Assume $\ell, i, m, j \; : \; \ell = i$ or $m = j$. $\mathfrak{n}_1 := |\ell - i| + |m - j| - 1 > 0$

$$|(S^{-1})_{k,t}| \leq \frac{\kappa\sqrt{\kappa^2+1}}{2} \frac{1}{\sqrt{\mathfrak{n}_1}}.$$

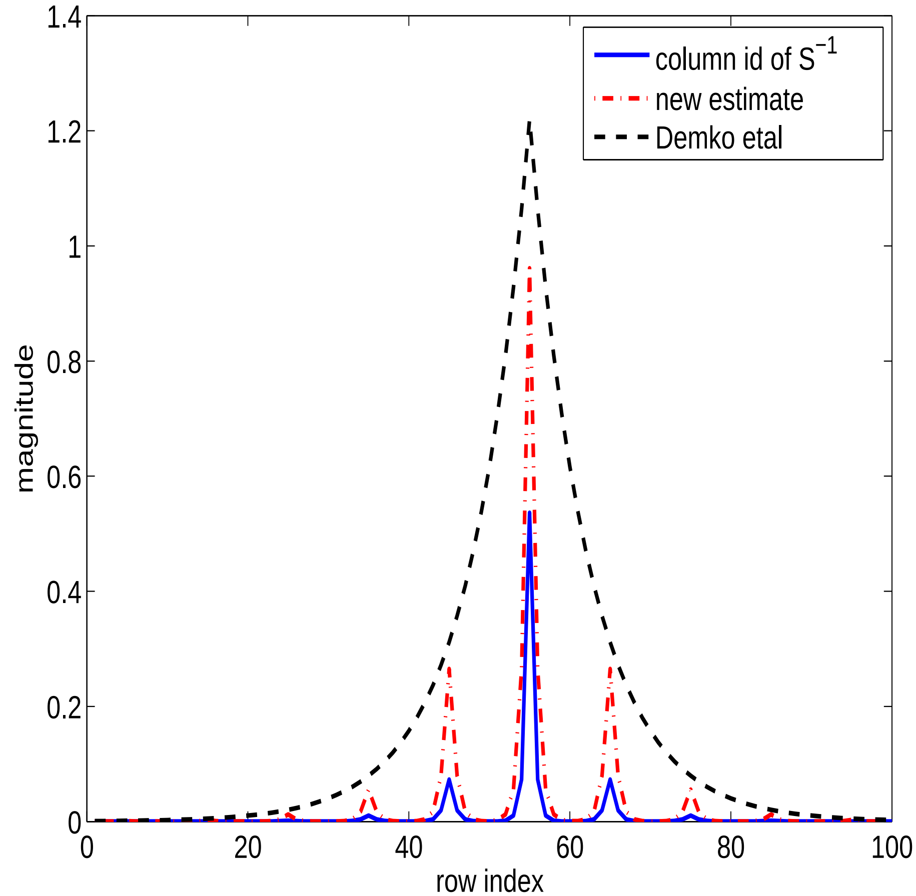




(i, j) and (ℓ, m)

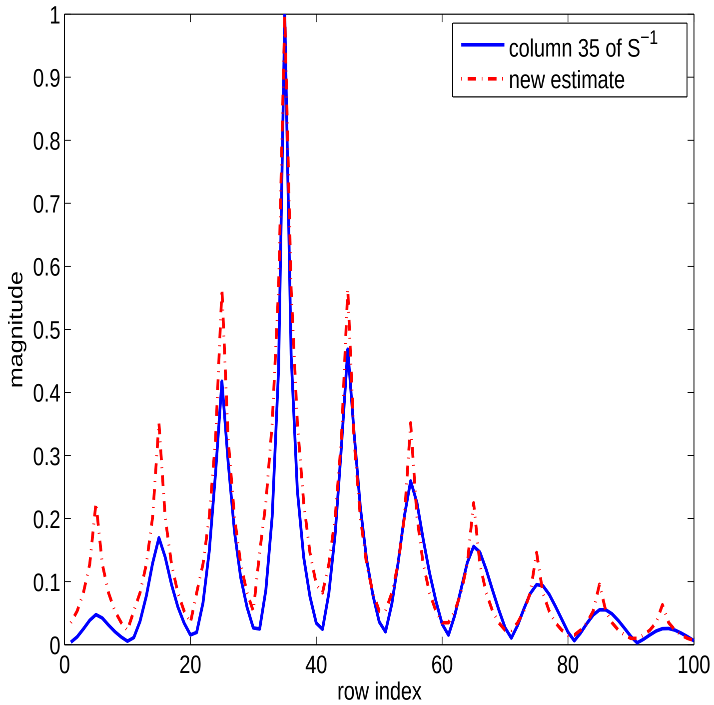
Examples. Symmetric positive definite matrix

$$M = \mathrm{tridiag}(-0.5, \underline{2}, -0.5) \in \mathbb{R}^{10\times 10}$$



Examples. Legendre stiffness matrix (scaled to have peak equal to 1)

$$M = \mathrm{tridiag}(\delta_k, \underline{\gamma_k}, \delta_k)$$



$$\gamma_k = \frac{2}{(4k-3)(4k+1)}$$

$$k = 1, \ldots, n, \quad \text{and}$$

$$\delta_k = \frac{-1}{(4k+1)\sqrt{(4k-1)(4k+3)}}$$

$$k = 1, \ldots, n-1$$

Connections to point-wise estimates for discrete Laplacian

For the discrete Green function G_h on the discrete d-dimensional grid R_h, there exist constants h_0 and C such that for $h \leq h_0$, $x, y \in R_h$,

$$G_h(x,y) \leq \begin{cases} C \log \frac{C}{|x-y|+h} & \text{if } d = 2 \\ \frac{C}{(|x-y|+h)^{d-2}} & \text{if } d \geq 3 \end{cases}$$

(Bramble & Thomee, '69)

Our estimate: entries depend on inverse square root of the distance!

Conclusions and further work ahead

- The Lyapunov operator has a very rich structure
- Appropriate computational devices
- Powerful mathematical tool
- ...this structure is recurrent in many application problems...

REFERENCES:

C. Canuto, V. Simoncini and M. Verani, LAA, v.452, 2014.

C. Canuto, V. Simoncini and M. Verani, J.Sc.Comp, 2014 (online first).

V. Simoncini, tr 2013 (survey on linear matrix equations, submitted)