

Matrix Equations and and Bivariate Function Approximation

D. Kressner Joint work with L. Grasedyck, C. Tobler, N. Truhar.

ETH Zurich, Seminar for Applied Mathematics

Manchester, 17.06.2009



Sylvester matrix equation linear matrix equation

$$AX + XB = D \quad (\text{SYLV})$$

with real $n \times n$ coefficient matrices A, B, D .

Applications:

- controllability/observability Gramians of linear systems
($B = A^T$, $-D$ symmetric pos semidef)
- balanced truncation model reduction
- Newton methods for solving algebraic Riccati equation
- stability analysis
- 2D PDE with separable coefficients
- ...

Outline

Connection to bivariate functions

- $X \Leftrightarrow 1/(\alpha + \beta)$
- Approximation to $X \Leftrightarrow$ Approximation to $1/(\alpha + \beta)$

Applications

- Singular value decay of X
- Convergence of Krylov subspace methods
- Convergence of extended Krylov subspace methods

Extension to higher dimensions

Connection to bivariate functions

The diagonalizable case

Assume A, B diagonalizable: $P^{-1}AP = \Lambda_A$, $Q^{-1}BQ = \Lambda_B$ with

$$\Lambda_A = \text{diag}(\alpha_1, \dots, \alpha_m), \quad \Lambda_B = \text{diag}(\beta_1, \dots, \beta_n).$$

Then

$$AX + XA^T = D \quad \rightsquigarrow \quad \Lambda_A \tilde{X} + \tilde{X} \Lambda_B = P^{-1}DQ =: \tilde{D}.$$

with explicit solution

$$\tilde{x}_{ij} = \frac{\tilde{d}_{ij}}{\alpha_i + \beta_j}$$

The diagonalizable case

Assume A, B diagonalizable: $P^{-1}AP = \Lambda_A$, $Q^{-1}BQ = \Lambda_B$ with

$$\Lambda_A = \text{diag}(\alpha_1, \dots, \alpha_m), \quad \Lambda_B = \text{diag}(\beta_1, \dots, \beta_n).$$

Then

$$AX + XA^T = D \quad \rightsquigarrow \quad \Lambda_A \tilde{X} + \tilde{X} \Lambda_B = P^{-1}DQ =: \tilde{D}.$$

with explicit solution

$$\tilde{x}_{ij} = \frac{\tilde{d}_{ij}}{\alpha_i + \beta_j}$$

$$\tilde{X} = \tilde{D} \circ \left[\frac{1}{\alpha_i + \beta_j} \right]$$

\tilde{X} is the Hadamard product between \tilde{D} and the bivariate function $1/(\alpha + \beta)$ evaluated at $\Lambda(A) \times \Lambda(B)$.

Separable approximations to bivariate functions

Bivariate function of the form $f(\alpha)g(\beta)$ is called **separable**.

$1/(\alpha + \beta)$ is not separable but...

Separable approximations to bivariate functions

Bivariate function of the form $f(\alpha)g(\beta)$ is called **separable**.

$1/(\alpha + \beta)$ is not separable but...

... might be well approximated by short sums of separable functions:

$$\frac{1}{\alpha + \beta} \approx s_r(y, z) := \sum_{k=1}^r f_k(\alpha)g_k(\beta).$$

Separable approximations to bivariate functions

Bivariate function of the form $f(\alpha)g(\beta)$ is called **separable**.

$1/(\alpha + \beta)$ is not separable but...

... might be well approximated by short sums of separable functions:

$$\frac{1}{\alpha + \beta} \approx s_r(y, z) := \sum_{k=1}^r f_k(\alpha)g_k(\beta).$$

Set $X_r = \sum_{k=1}^r f_k(A)Dg_k(B)$, then

$$\tilde{X}_r = \tilde{D} \circ [s_r(\alpha_i, \beta_j)]$$

Hadamard product between \tilde{D} and $s_r(\alpha, \beta)$ evaluated at $\Lambda(A) \times \Lambda(B)$.

Approximation to $X \Leftrightarrow$ Approximation to $1/(\alpha + \beta)$

$$AX + XB = D$$

Let $\Lambda(A) \subseteq \Omega_A$, $\Lambda(B) \subseteq \Omega_B$ and define

$$X_r = \sum_{k=1}^r f_k(A)Dg_k(B), \quad s_r(\alpha, \beta) = \sum_{k=1}^r f_k(\alpha)g_k(\beta),$$

with functions $f_k : \Omega_A \rightarrow \mathbb{C}$, $g_k : \Omega_B \rightarrow \mathbb{C}$.

$$\|X - X_r\|_F \leq \kappa(P)\kappa(Q) \left\| \frac{1}{\alpha + \beta} - s_r(\alpha, \beta) \right\|_{\infty, \Omega_A \times \Omega_B} \|D\|_F.$$

Idea of proof

If $A = \Lambda_A, B = \Lambda_B$ are diagonal \rightsquigarrow solution of $\Lambda_A X + X \Lambda_B = D$:

$$X = \begin{bmatrix} \frac{d_{11}}{\alpha_1 + \beta_1} & \frac{d_{12}}{\alpha_1 + \beta_2} & \cdots \\ \frac{d_{21}}{\alpha_2 + \beta_1} & \frac{d_{22}}{\alpha_2 + \beta_2} & \cdots \\ \vdots & \vdots & \end{bmatrix}$$

$$X - X_r = \begin{bmatrix} \frac{1}{\alpha_1 + \beta_1} - \sum f_k(\alpha_1)g_k(\beta_1) & \frac{1}{\alpha_1 + \beta_2} - \sum f_k(\alpha_1)g_k(\beta_2) & \cdots \\ \frac{1}{\alpha_2 + \beta_1} - \sum f_k(\alpha_2)g_k(\beta_1) & \frac{1}{\alpha_2 + \beta_2} - \sum f_k(\alpha_2)g_k(\beta_2) & \cdots \\ \vdots & \vdots & \end{bmatrix} \circ D$$

The non-diagonalizable case

Consider contours $\Gamma_A, \Gamma_B \subset \mathbb{C}$ enclosing $\Lambda(A), \Lambda(B)$ with $\Gamma_A \cap -\Gamma_B = \emptyset$. Then

$$X = \frac{1}{4\pi^2} \oint_{\Gamma_A} \oint_{\Gamma_B} \frac{1}{\alpha + \beta} (\alpha I - A)^{-1} D(\beta I - B)^{-1} d\alpha d\beta$$

solves $AX + XB = D$.

The non-diagonalizable case

Consider contours $\Gamma_A, \Gamma_B \subset \mathbb{C}$ enclosing $\Lambda(A), \Lambda(B)$ with $\Gamma_A \cap -\Gamma_B = \emptyset$. Then

$$X = \frac{1}{4\pi^2} \oint_{\Gamma_A} \oint_{\Gamma_B} \frac{1}{\alpha + \beta} (\alpha I - A)^{-1} D(\beta I - B)^{-1} d\alpha d\beta$$

solves $AX + XB = D$.

Similarly, for

$$X_r = \sum_{k=1}^r f_k(A) D g_k(A^T), \quad s_r(\alpha, \beta) = \sum_{k=1}^r f_k(\alpha) g_k(\beta),$$

we have

$$X_r = \frac{1}{4\pi^2} \oint_{\Gamma_A} \oint_{\Gamma_B} s_r(\alpha, \beta) (\alpha I - A)^{-1} D (\beta I - B)^{-1} d\alpha d\beta.$$

The non-diagonalizable case

$$\|X - X_r\|_F \leq \kappa(\Omega_A)\kappa(\Omega_B) \left\| \frac{1}{\alpha + \beta} - s_r(\alpha, \beta) \right\|_{\infty, \Omega_A \times \Omega_B} \|D\|_F.$$

“Constants”:

$$\kappa(\Omega_A) = \frac{1}{2\pi} \oint_{\Gamma_A} \|(\alpha I - A)^{-1}\|_2 d\alpha,$$

$$\kappa(\Omega_B) = \frac{1}{2\pi} \oint_{\Gamma_B} \|(\beta I - B)^{-1}\|_2 d\beta.$$

See discussion in [Beattie, Embree, Sorensen, SIAM Review'05] on $\kappa(\Omega_A)$ and $\kappa(\Omega_B)$.

General notion of bivariate matrix functions?

Consider two matrices A, B and contours $\Gamma_A, \Gamma_B \subset \mathbb{C}$ enclosing $\Lambda(A), \Lambda(B)$.

For *any* function $f(\alpha, \beta)$ analytic on $\Gamma_A \times \Gamma_B$, one *could* define

$$f(A, B)[D] = \frac{1}{4\pi^2} \oint_{\Gamma_A} \oint_{\Gamma_B} f(\alpha, \beta) (\alpha I - A)^{-1} D (\beta I - B)^{-1} d\alpha d\beta,$$

as the evaluation of f at (A, B) applied to C .

Such an abstract framework might be of use in other applications (Stein equations, Fréchet derivative of matrix functions).

Applications

Singular value decay

Singular value decay

Approximation results

$$\|X - X_r\|_F \lesssim \left\| \frac{1}{\alpha + \beta} - s_r(\alpha, \beta) \right\|_{\infty, \Omega_A \times \Omega_B} \|D\|_F.$$

If D has rank one $\rightsquigarrow X_r = \sum f_k(A)Dg_k(B)$ has rank r .

$$\sigma_{r+1}(X) \lesssim \left\| \frac{1}{\alpha + \beta} - s_r(\alpha, \beta) \right\|_{\infty, \Omega_A \times \Omega_B} \|D\|_2$$

No restriction on the choice of functions f_k, g_k .

Example: The symmetric case

Assume A, B symmetric with eigenvalues contained in interval $[\lambda_{\min}, \lambda_{\max}]$ with $\lambda_{\min} \leq \lambda_{\max} < 0$.

[Braess'86], [Braess/Hackbusch'05]: Let

$$s_r(\alpha, \beta) = \sum_{k=1}^r \omega_k \exp(\gamma_k \alpha) \exp(\gamma_k \beta),$$

$\exists \omega_k < 0, \gamma_k > 0$ s.t.

$$\|1/(\alpha + \beta) - s_r(\alpha, \beta)\|_{[\lambda_{\min}, \lambda_{\max}]^2} \leq \frac{8}{|\lambda_{\max}|} \exp \left[-\frac{r\pi^2}{\log(8\frac{\lambda_{\min}}{\lambda_{\max}})} \right].$$

Example: The symmetric case

If D has rank one:

$$\sigma_{r+1}(X) \leq 8 \|A^{-1}\|_2 \exp\left[-\frac{r\pi^2}{\log(8\kappa(A))}\right] \|D\|_2.$$

- This matches an existing bound in [Sabino'06] for Lyapunov equations.
- Results for complex $\Lambda(A), \Lambda(B)$ can be obtained by sinc interpolation [Grasedyck/Hackbusch/Khoromskij'02, Grasedyck/K.'09].
- Isolated eigenvalues close to zero can be matched by exact polynomial interpolation.

Lyapunov equation: Singular value decay

Other existing decay bounds:

- [Penzl'00]:

$$\frac{\sigma_{r+1}(X)}{\|X\|_2} \lesssim \left(\prod_{j=0}^{r-1} \frac{\kappa^{(2j+1)/(2k)} - 1}{\kappa^{(2j+1)/(2k)} + 1} \right)^2 = \exp \left(-\mathcal{O} \left(\frac{r}{\log \kappa(A)} \right) \right).$$

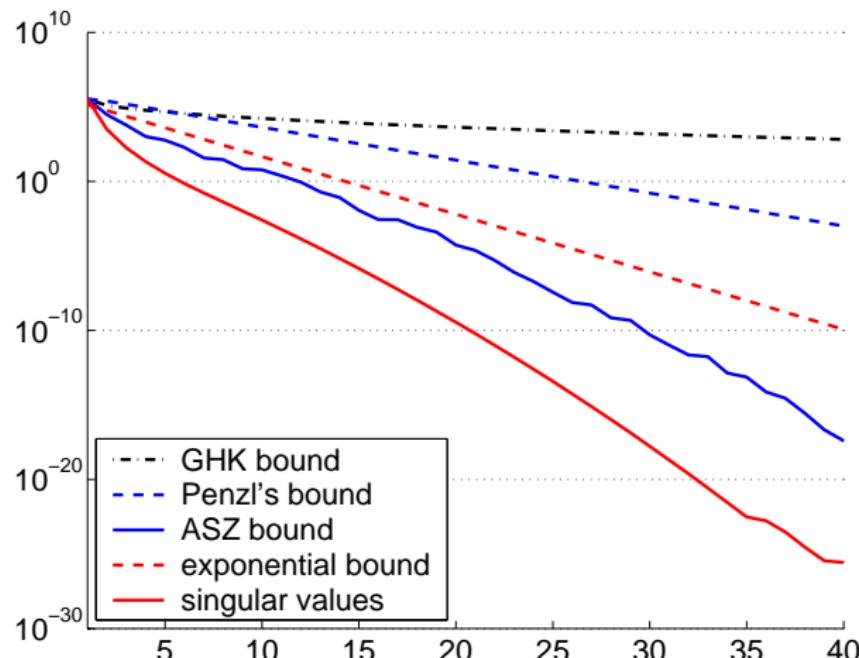
- [Antoulas/Sorensen/Zhou'02]:

$$\sigma_r(X) \lesssim \prod_{j=1}^{r-1} \left| \frac{\lambda_r - \lambda_j}{\overline{\lambda_r} + \lambda_j} \right|^2.$$

- [Grasedyck/Hackbusch/Khoromskij'02]:

$$\frac{\sigma_{r+1}(X)}{\|X\|_2} \lesssim \exp(-\sqrt{r}).$$

Example: $A \in \mathbb{R}^{200 \times 200}$ second-difference operator and $D = bb^T$ with $b = [1, \dots, 1]^T$.



Applications

Subspace projection methods

Basic idea of subspace projection methods

Approximate solution to $AX + XA^T = -bb^T$ using a subspace $\mathcal{U} \subset \mathbb{R}^n$ that contains

- Krylov subspace for A :

$$\mathcal{K}_r(A, b) = \text{span}\{b, Ab, A^2b, \dots, A^{r-1}b\}$$

Basic idea of subspace projection methods

Approximate solution to $AX + XA^T = -bb^T$ using a subspace $\mathcal{U} \subset \mathbb{R}^n$ that contains

- Krylov subspace for A :

$$\mathcal{K}_r(A, b) = \text{span}\{b, Ab, A^2b, \dots, A^{r-1}b\}$$

- Krylov subspace for A^{-1} :

$$\mathcal{K}_r(A^{-1}, b) = \text{span}\{b, A^{-1}b, \dots, A^{-r+1}b\}$$

Basic idea of subspace projection methods

Approximate solution to $AX + XA^T = -bb^T$ using a subspace $\mathcal{U} \subset \mathbb{R}^n$ that contains

- Krylov subspace for A :

$$\mathcal{K}_r(A, b) = \text{span}\{b, Ab, A^2b, \dots, A^{r-1}b\}$$

- Krylov subspace for A^{-1} :

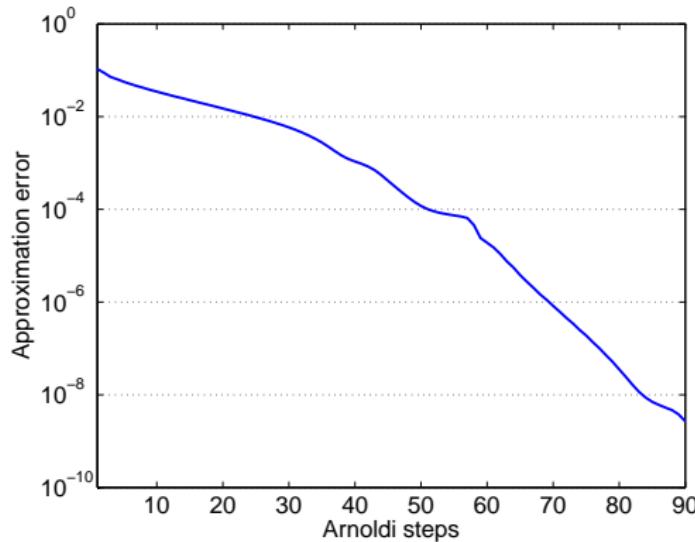
$$\mathcal{K}_r(A^{-1}, b) = \text{span}\{b, A^{-1}b, \dots, A^{-r+1}b\}$$

- Krylov subspaces for $(A - \sigma I)^{-1}$:

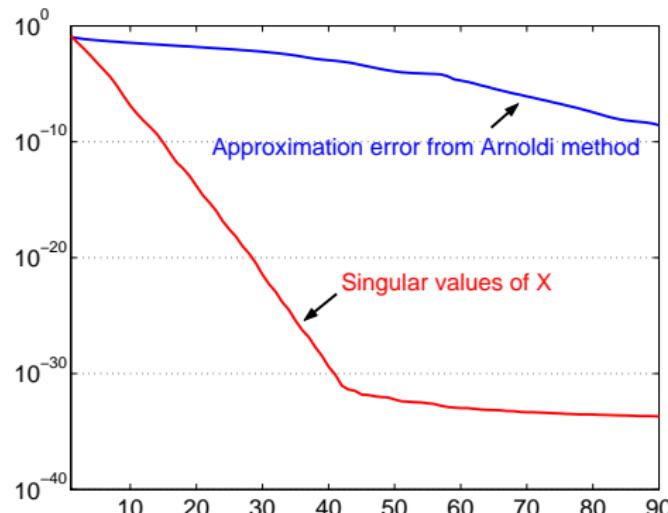
$$\mathcal{K}_r((A - \sigma I)^{-1}, b) = \text{span}\{b, (A - \sigma I)^{-1}b, \dots, (A - \sigma I)^{-r+1}b\}$$

Convergence of Arnoldi for Lyapunov

2D instationary heat equation on the unit square with homogeneous kind boundary conditions; 30×30 finite difference discretization.



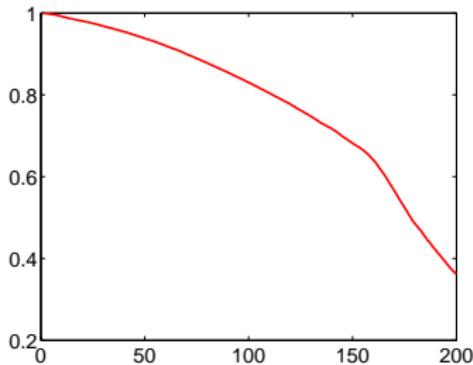
Convergence does not capture singular value decay



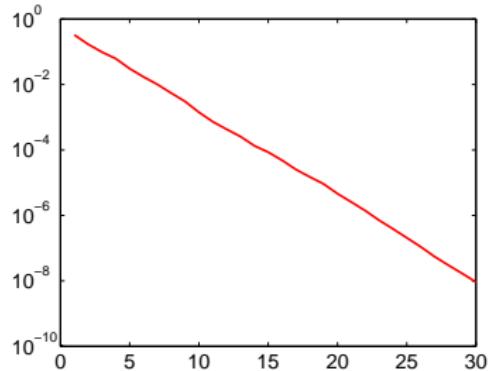
Example:

- $\text{svd}(X)$ tells there is rank 12 approximation with error 10^{-8} ;
- Arnoldi yields only rank 84 approximation with error 10^{-8} .

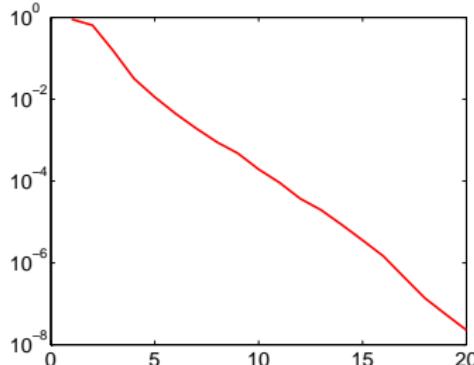
Arnoldi



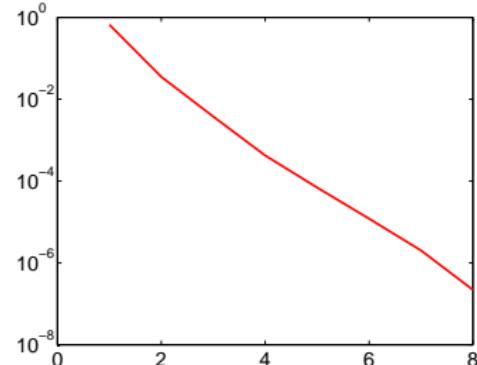
Extended Arnoldi



Rational Arnoldi (6 shifts)



Rational Arnoldi (13 shifts)



Convergence for Steel example (Oberwolfach collection).

Galerkin condition

From now on: A is symmetric and $\Lambda(A) \subset [\lambda_{\min}, \lambda_{\max}]$ with $\lambda_{\min} \leq \lambda_{\max} < 0$.

Approximation X_r satisfies **Galerkin condition**

$$AX_r + X_r A + bb^T \perp \mathcal{U} \otimes \mathcal{U}.$$

Equivalent to

$$\|X - X_r\|_{\mathcal{A}} = \min_{Z \in \mathcal{U} \otimes \mathcal{U}} \|X - Z\|_{\mathcal{A}},$$

in the “weighted Frobenius norm”

$$\|C\|_{\mathcal{A}} := \|\text{vec}(C)\|_{-(I \otimes A + A \otimes I)}.$$

Convergence of Arnoldi method

If $\mathcal{U} = \mathcal{K}_r(A, b)$, every $Z \in \mathcal{U} \times \mathcal{U}$ takes the form

$$Z = \sum_{i,j=0}^{r-1} \gamma_{ij} A^i b b^T A^j, \quad \gamma_{ij} \in \mathbb{R}.$$

Hence,

$$\|X - X_r\|_A \leq \sqrt{\|A\|_2} \min_{\gamma_{ij} \in \mathbb{R}} \left\| \frac{1}{\alpha + \beta} - \sum_{i,j=1}^{r-1} \gamma_{ij} \alpha^i \beta^j \right\|_{[\lambda_{\min}, \lambda_{\max}]^2} \|b\|_2^2.$$

Convergence of Arnoldi method

Error for separable polynomial approximation of $1/(\alpha + \beta)$:

$$\frac{1}{|\lambda_{\max}|} \cdot \frac{\sqrt{\kappa} + 1}{\sqrt{\kappa}} \cdot \frac{(\sqrt{\kappa} - 1)^r}{(\sqrt{\kappa} + 1)^r}$$

with $\kappa = (|\lambda_{\min}| + |\lambda_{\max}|)/(2|\lambda_{\min}|)$. Proof based on careful analysis of Chebyshev approximation to exponentials.

Convergence of Arnoldi method

Error for separable polynomial approximation of $1/(\alpha + \beta)$:

$$\frac{1}{|\lambda_{\max}|} \cdot \frac{\sqrt{\kappa} + 1}{\sqrt{\kappa}} \cdot \frac{(\sqrt{\kappa} - 1)^r}{(\sqrt{\kappa} + 1)^r}$$

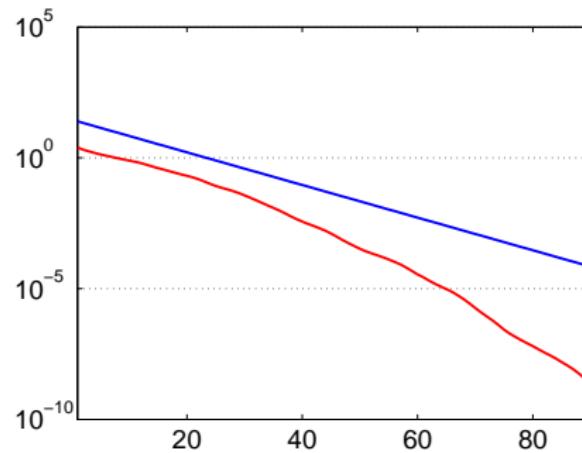
with $\kappa = (|\lambda_{\min}| + |\lambda_{\max}|)/(2|\lambda_{\min}|)$. Proof based on careful analysis of Chebyshev approximation to exponentials.

↔ Convergence bound for Arnoldi method:

$$\|X - X_r\|_{\mathcal{A}} \leq \sqrt{\|A\|_2} \|A^{-1}\|_2 \|b\|_2^2 \cdot \frac{\sqrt{\kappa} + 1}{\sqrt{\kappa}} \cdot \frac{(\sqrt{\kappa} - 1)^r}{(\sqrt{\kappa} + 1)^r},$$

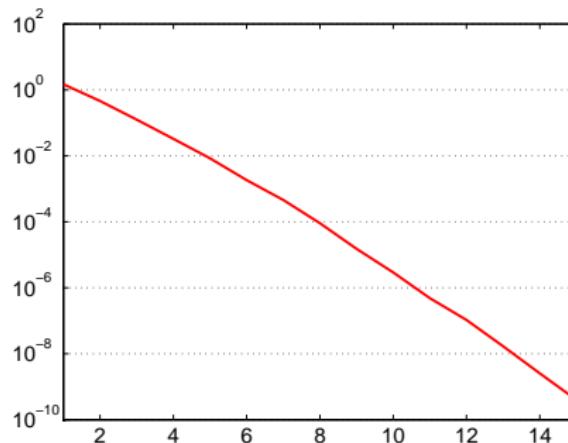
with $\kappa \approx \kappa(A)/2$. Matches a bound in [Simoncini/Druskin'09].

Convergence of Arnoldi method



Observed convergence and convergence bound.

Convergence of Extended Arnoldi method



Observed convergence for $\mathcal{U} = \text{span}(\mathcal{K}_r(A, b) \cup \mathcal{K}_r(A^{-1}, b))$.

Convergence of Extended Arnoldi method

$$\mathcal{U} = \text{span}(\mathcal{K}_r(A, b) \cup \mathcal{K}_r(A^{-1}, b)) = \{p(A)b : p \in \mathbb{L}_r\}$$

where \mathbb{L}_r denotes the set of Laurent polynomials

$$\gamma_{-r}\alpha^{-r} + \cdots + \gamma_{-1}\alpha^{-1} + \gamma_0 + \gamma_1\alpha + \cdots + \gamma_{r-1}\alpha^{r-1}.$$

$$\|X - X_r\|_A \leq \sqrt{\|A\|_2} \min_{\gamma_{ij} \in \mathbb{R}} \left\| \frac{1}{\alpha + \beta} - \sum_{i,j=-r}^{r-1} \gamma_{ij} \alpha^i \beta^j \right\| \|b\|_2^2.$$

Convergence of extended Arnoldi method

Error for separable Laurent polynomial approximation of $1/(\alpha + \beta)$:

$$\frac{1}{|\lambda_{\max}|} \cdot \frac{\kappa^{1/4} + 1}{\kappa^{1/4}} \cdot \frac{(\kappa^{1/4} - 1)^r}{(\kappa^{1/4} + 1)^r}$$

with $\kappa = (|\lambda_{\min}| + |\lambda_{\max}|)/(2|\lambda_{\max}|)$. Proof based on polynomial approximation: If p is polynomial, $p(\alpha + \gamma/\alpha)$ is Laurent polynomial.

Convergence of extended Arnoldi method

Error for separable Laurent polynomial approximation of $1/(\alpha + \beta)$:

$$\frac{1}{|\lambda_{\max}|} \cdot \frac{\kappa^{1/4} + 1}{\kappa^{1/4}} \cdot \frac{(\kappa^{1/4} - 1)^r}{(\kappa^{1/4} + 1)^r}$$

with $\kappa = (|\lambda_{\min}| + |\lambda_{\max}|)/(2|\lambda_{\max}|)$. Proof based on polynomial approximation: If p is polynomial, $p(\alpha + \gamma/\alpha)$ is Laurent polynomial.

~~ Convergence bound for extended Arnoldi method:

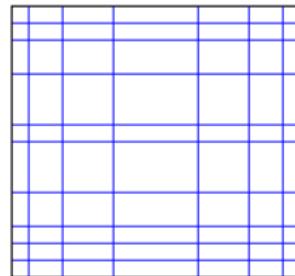
$$\|X - X_r\|_A \leq \sqrt{\|A\|_2} \|A^{-1}\|_2 \|b\|_2^2 \cdot \frac{\kappa^{1/4} + 1}{\kappa^{1/4}} \cdot \frac{(\kappa^{1/4} - 1)^r}{(\kappa^{1/4} + 1)^r},$$

with $\kappa \approx \kappa(A)/2$ [K./Tobler'09].

High-dimensional Extensions

A toy PDE

$$-\Delta \mathbf{u} = f \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0,$$



Tensorize 1D FE bases

$$\begin{aligned}\mathbb{V} &= \{v_1(x_1), \dots, v_m(x_1)\} \\ \mathbb{W} &= \{w_1(x_2), \dots, w_n(x_2)\}\end{aligned}\quad \mathbb{V} \otimes \mathbb{W} = \left\{ \sum \alpha_{ij} v_i(x_1) w_j(x_2) \right\}$$

~≈ Galerkin discretization

$$(K_V \otimes M_W + M_V \otimes K_W) \mathbf{u} = \mathbf{f},$$

with 1D mass/stiffness matrices M_V, M_W, K_V, K_W ;
⊗ denotes Kronecker product.

reshape \mathbf{u}

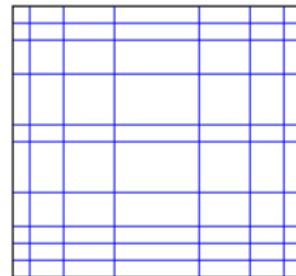
$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \end{bmatrix}, \mathbf{u}_j \in \mathbb{R}^m \quad \Leftrightarrow \quad U = [\mathbf{u}_1, \mathbf{u}_2, \dots]$$

MATLAB: `U = reshape(u,m,n);`

$$(K_V \otimes M_W)\mathbf{u} \quad \Leftrightarrow \quad K_V U M_W.$$

A toy PDE

$$-\Delta u = f \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0,$$



Tensorize 1D FE bases

$$\begin{aligned}\mathbb{V} &= \{v_1(x_1), \dots, v_m(x_1)\} \\ \mathbb{W} &= \{w_1(x_2), \dots, w_n(x_2)\}\end{aligned}\quad \mathbb{V} \otimes \mathbb{W} = \left\{ \sum \alpha_{ij} v_i(x_1) w_j(x_2) \right\}$$

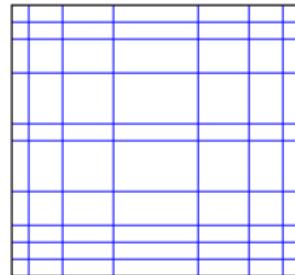
\rightsquigarrow Galerkin discretization

$$K_V \textcolor{red}{U} M_W + M_V \textcolor{red}{U} K_W = F,$$

with 1D mass/stiffness matrices M_V, M_W, K_V, K_W ;
 $\mathbf{u} = \text{vec}(U)$, $\mathbf{f} = \text{vec}(F)$.

A toy PDE

$$-\Delta u = f \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0,$$



Tensorize 1D FE bases

$$\begin{aligned}\mathbb{V} &= \{v_1(x_1), \dots, v_m(x_1)\} \\ \mathbb{W} &= \{w_1(x_2), \dots, w_n(x_2)\}\end{aligned}\quad \mathbb{V} \otimes \mathbb{W} = \left\{ \sum \alpha_{ij} v_i(x_1) w_j(x_2) \right\}$$

↔ Galerkin discretization

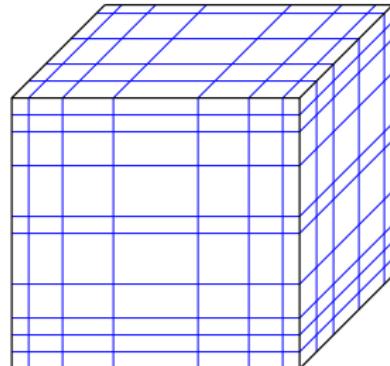
$$(M_V^{-1/2} K_V M_V^{-1/2}) X + X (M_W^{-1/2} K_W M_W^{-1/2}) = M_V^{-1/2} F M_W^{-1/2},$$

with 1D mass/stiffness matrices M_V, M_W, K_V, K_W ;

$$X = M_V^{1/2} U M_W^{1/2}, \mathbf{f} = \text{vec}(F).$$

The toy PDE in 3D

$$-\Delta u = f \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0,$$



Tensorize 1D FE bases

$$\mathbb{V} = \{v_1(x_1), \dots, v_m(x_1)\}$$

$$\mathbb{W} = \{w_1(x_2), \dots, w_n(x_2)\}$$

$$\mathbb{Z} = \{z_1(x_3), \dots, z_n(x_3)\}$$

$$\mathbb{V} \otimes \mathbb{W} \otimes \mathbb{Z} = \left\{ \sum \alpha_{ijk} v_i(x_1) w_j(x_2) z_k(x_3) \right\}$$

\leadsto Galerkin discretization

$$(K_V \otimes M_W \otimes M_Z + M_V \otimes K_W \otimes M_Z + M_V \otimes M_W \otimes K_Z) \mathbf{u} = \mathbf{f},$$

with 1D mass/stiffness matrices $M_V, M_W, M_Z, K_V, K_W, K_Z$.

Higher-dimensional matrix equations

“Kroneckerized” Sylvester equation

$$(A_1 \otimes I + I \otimes A_2)x = b.$$

Generalization to dimension $d = 3$:

$$(A_1 \otimes I \otimes I + I \otimes A_2 \otimes I + I \otimes I \otimes A_3)x = b.$$

Generalization to arbitrary dimension:

$$\mathcal{A}x = b, \quad \mathcal{A} = \sum_{j=1}^d I \otimes \cdots \otimes I \otimes A_j \otimes I \otimes \cdots \otimes I.$$

Properties

$$\mathcal{A}x = b, \quad \mathcal{A} = \sum_{j=1}^d I \otimes \cdots \otimes I \otimes A_j \otimes I \otimes \cdots \otimes I.$$

- If $A_j \in \mathbb{R}^{n \times n}$ then $\mathcal{A} \in \mathbb{R}^{n^d \times n^d}$. **Curse of dimension!**
- $\Lambda(\mathcal{A}) = \{\lambda^{(1)} + \cdots + \lambda^{(d)} : \lambda^{(j)} \in \Lambda(A_j)\}$.
- If $\kappa(A_j) \equiv \kappa$ then $\kappa(\mathcal{A}) = \kappa$.
- If $b = b_1 \otimes b_2 \otimes \cdots \otimes b_d$ (rank-1 tensor) and all A_j spd
 \rightsquigarrow solution x can be well approximated by short sum of rank-1 tensors [Grasedyck'04, K./Tobler'09].

Tensor Krylov subspace method

Arnoldi method for Lyapunov equations extends to $d > 2$ for $b = b_1 \otimes \cdots \otimes b_d$.

- Works with $\mathcal{K}_r(A_j, b_j)$ (Krylov subspace in each individual direction).
- Approximates x by x_r satisfying Galerkin condition

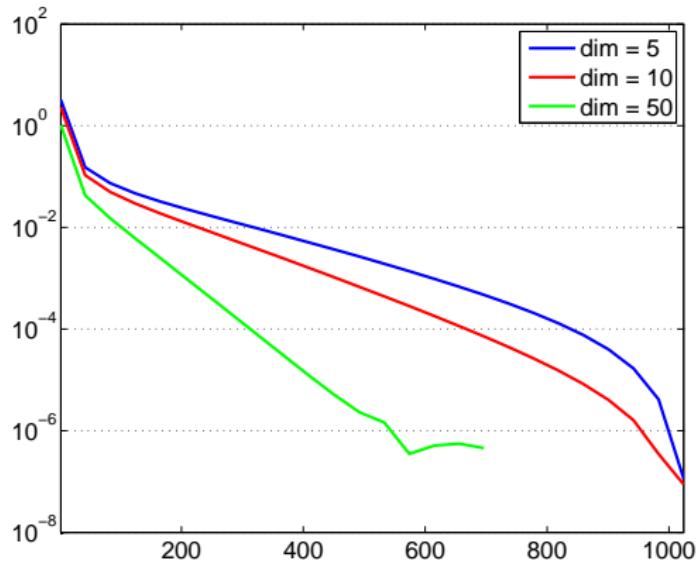
$$\mathcal{A}x_r - b \perp \text{span}(\mathcal{K}_r(A_1, b_1) \otimes \cdots \otimes \mathcal{K}_r(A_d, b_d)).$$

- If all A_j spd:

$$\|x - x_r\|_2 \lesssim \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^r, \quad \kappa \approx \frac{\kappa(\mathcal{A})}{d}.$$

CG vs. Tensor Krylov

	Comp. Complexity	Storage requ.
CG	$O(r^2 n^d)$	$O(rn^d)$
Tensor Krylov	$O(r^2 dn)$	$O(rdn)$



Example: Laplace on $[0, 1]^d$, $n = 1024$, $n^{50} \approx 3 \cdot 10^{150}$.

Tensor Krylov requires 300 iterations and $\approx 1\text{h}$ to attain residual 10^{-6} for $d = 50$.

Conclusions

- Singular value decay/projection methods for Sylvester equations \Leftrightarrow bivariate function approximation.
- Low tensor rank approximation/projection methods for linear systems with tensor product structure \Leftrightarrow multivariate function approximation.

See <http://www.math.ethz.ch/~kressner/> for further details.

Workshop on Matrix Equations

Fr, 9th October 2009
TU Braunschweig, Germany

Organizers:

- Peter Benner, TU Chemnitz
- Heike Faßbender, TU Braunschweig
- Lars Grasedyck, MPI Leipzig
- Daniel Kressner, ETH Zurich

Contact: h.fassbender@tu-bs.de