Computational methods for large-scale linear matrix equations: recent advances

V. Simoncini

Dipartimento di Matematica, Università di Bologna

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

- Multiterm matrix equation

$$A_1\mathbf{X}B_1 + A_2\mathbf{X}B_2 + \ldots + A_\ell\mathbf{X}B_\ell = C$$

Control, (Stochastic) PDEs, ...

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

- Multiterm matrix equation

$$A_1\mathbf{X}B_1 + A_2\mathbf{X}B_2 + \ldots + A_\ell\mathbf{X}B_\ell = C$$

Control, (Stochastic) PDEs, ...

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

Solving the Lyapunov equation. The problem

Approximate $\mathbf{X}$ in:

$$A\mathbf{X} + \mathbf{X}A^\top + BB^\top = 0$$

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

Solving the Lyapunov equation. The problem

Approximate $\mathbf{X}$ in:

$$A\mathbf{X} + \mathbf{X}A^\top + BB^\top = 0$$

$A \in \mathbb{R}^{n\times n}$ 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:

$$\mathbf{X} = \int_0^\infty e^{-tA} BB^\top e^{-tA^\top} dt$$

$\Rightarrow \qquad \mathbf{X}$ 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:

$$A\mathbf{X} + \mathbf{X}A^\top + BB^\top = 0$$

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

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

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:

$$A\mathbf{X} + \mathbf{X}A^\top + BB^\top = 0$$

Kronecker formulation:

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

Projection-type methods

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

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

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 $\mathcal{K}$,

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

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

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

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

Projected Lyapunov equation:

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

Projection-type methods

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

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

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

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

(Druskin & Knizhnerman '98, Simoncini '07)

- Rational Krylov subspace

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

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

(Druskin & Knizhnerman '98, Simoncini '07)

- Rational Krylov subspace

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

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

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

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

$X_m = V_m Y_m V_m^\top$

Multiterm linear matrix equation

$$A_1 \mathbf{X} B_1 + A_2 \mathbf{X} B_2 + \ldots + A_\ell \mathbf{X} B_\ell = C$$

Applications:

- Matrix least squares
- Control
- (Stochastic) PDEs
- ...

Multiterm linear matrix equation

$$A_1\mathbf{X}B_1 + A_2\mathbf{X}B_2 + \ldots + A_\ell\mathbf{X}B_\ell = C$$

Applications:

- Matrix least squares
- Control
- (Stochastic) PDEs
- ...

Main device: Kronecker formulation

$$\left(B_1^\top \otimes A_1 + \ldots + B_\ell^\top \otimes A_\ell\right) x = c$$

Iterative methods: matrix-matrix multiplications and rank truncation

(Benner, Breiten, Bouhamidi, Chehab, Damm, Grasedyck, Jbilou, Kressner, Matthies, Onwunta, Raydan, Stoll, Tobler, Zander, and **many** others...)

Multiterm linear matrix equation

$$A_1 \mathbf{X} B_1 + A_2 \mathbf{X} B_2 + \ldots + A_\ell \mathbf{X} B_\ell = C$$

Applications:

- Matrix least squares
- Control
- (Stochastic) PDEs
- ...

Alternative approaches:

low-rank approx in the problem space. Some examples:

- Control problem

- PDEs on uniform discretizations

- Stochastic PDE

A class of generalized Lyapunov equations

$$AX + XA^T + \sum_{j=1}^{m} N_j X N_j^T + BB^T = 0$$

* $A \in \mathbb{R}^{n\times n}$ nonsing
* $N_j \in \mathbb{R}^{n\times n}$ low rank
* $B \in \mathbb{R}^{n\times \ell}$, $\ell \ll n$

Typical applications:

- Model order reduction of bilinear control systems
- Linear parameter-varying systems
- Stability analysis of linear stochastic differential equations

Stationary iterative methods by splitting

$$AX + XA^T + \sum_{j=1}^{m} N_j X N_j^T + BB^T = 0$$

$$\mathcal{M}(X) - \mathcal{N}(X) + BB^T = 0,$$

where $\mathcal{M}(X) = AX + XA^T$ (Lyapunov operator)

$$-\mathcal{N}(X) = \sum_{i=1}^{m} N_j X N_j^T$$

Stationary iterative methods by splitting

$$AX + XA^T + \sum_{j=1}^{m} N_j X N_j^T + BB^T = 0$$

$$\mathcal{M}(X) - \mathcal{N}(X) + BB^T = 0,$$

where $\mathcal{M}(X) = AX + XA^T$ (Lyapunov operator)

$$-\mathcal{N}(X) = \sum_{i=1}^{m} N_j X N_j^T$$

Assuming that (A, B) is controllable and X sym positive semi-def then

$$\operatorname{spec}(A) \subset \mathbb{C}^-, \qquad \rho(\mathcal{M}^{-1}\mathcal{N}) < 1$$

Stationary iteration:

$$\mathcal{M}(X_k) = \mathcal{N}(X_{k-1}) - BB^T, \quad k = 1, 2, \ldots.$$

(Shank & Simoncini & Szyld, 2016)

Stationary iterative methods by splitting. Cont'd

$$AX + XA^T + \sum_{j=1}^{m} N_j X N_j^T + BB^T = 0$$

Stationary iteration:

$$\mathcal{M}(X_k) = \mathcal{N}(X_{k-1}) - BB^T, \qquad k = 1, 2, \ldots.$$

In practice:

Approximately Solve $AX + XA^T + BB^T = 0$ for $X_1 = Z_1 Z_1^T$

for $k = 2, 3, \ldots$

Set $B_k = [N_1 Z_{k-1}, \ \cdots \ , N_m Z_{k-1}, \ B]$

Approximately Solve $AX + XA^T + B_k B_k^T = 0$ for $X_k = Z_k Z_k^T$

If sufficiently accurate **then stop**

Stationary iterative methods by splitting. Cont'd

Approximately Solve $AX + XA^T + BB^T = 0$ for $X_1 = Z_1 Z_1^T$

for $k = 2, 3, \ldots$

  Set $B_k = [N_1 Z_{k-1}, \; \cdots \; , N_m Z_{k-1}, \; B]$

  Approximately Solve $AX + XA^T + B_k B_k^T = 0$ for $X_k = Z_k Z_k^T$

  If sufficiently accurate **then stop**

Challenges:

- Inexact solves of Lyapunov equation at each step k
- Increase of B_k's rank
- Computational cost of Lyapunov solves
- Memory effective stopping criterion

Matrix equations in PDEs

The Poisson equation - revisited

$$-u_{xx} - u_{yy} = f, \quad \text{in} \quad \Omega = (0,1)^2$$

$+$ Dirichlet b.c. (zero b.c. for simplicity)

Usual discretization $\Rightarrow$ $\quad Au = b \quad$ (with $A = T \otimes I + I \otimes T$)

Matrix equations in PDEs

The Poisson equation - revisited

$$-u_{xx} - u_{yy} = f, \quad \text{in} \quad \Omega = (0,1)^2$$

$+$ Dirichlet b.c. (zero b.c. for simplicity)

Usual discretization $\Rightarrow$ $\quad Au = b \quad$ (with $A = T \otimes I + I \otimes T$)

Discretization: $U_{i,j} \approx u_{x_i,y_j}$, with (x_i, y_j) interior nodes, so that $\qquad h$: meshsize

$$u_{xx}(x_i, y_j) \approx \frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{h^2} = \frac{1}{h^2}[1, -2, 1] \begin{bmatrix} U_{i-1,j} \\ U_{i,j} \\ U_{i+1,j} \end{bmatrix}$$

$$u_{yy}(x_i, y_j) \approx \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{h^2} = \frac{1}{h^2}[U_{i,j-1}, \, U_{i,j}, \, U_{i,j+1}] \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

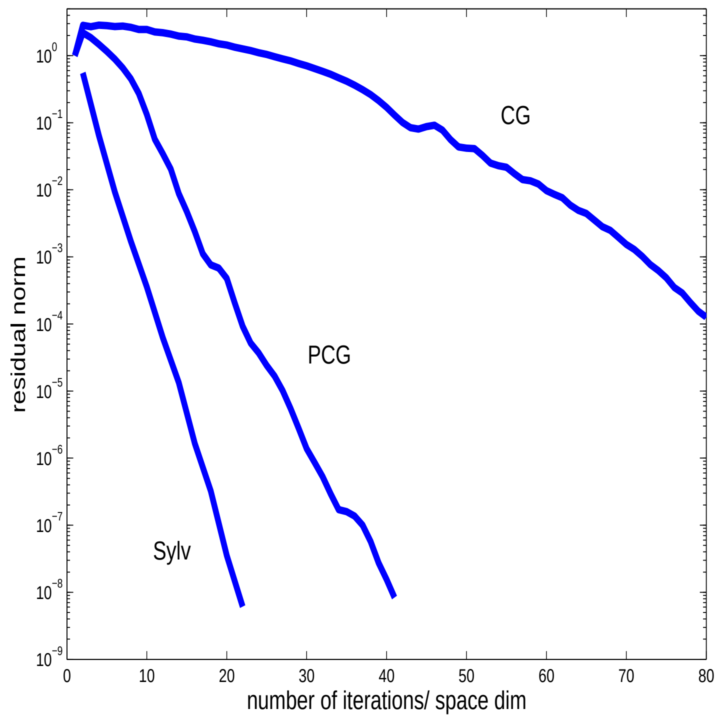
$$T\mathbf{U} + \mathbf{U}T = F, \qquad b = \text{vec}(F)$$

$$-\Delta u = 1, \quad \Omega = (0,1)^3 \quad \Rightarrow \quad A = (T \otimes I \otimes I + I \otimes T \otimes I + I \otimes I \otimes T)$$

$$-\Delta u = 1, \quad \Omega = (0,1)^3 \quad \Rightarrow \quad A = (T\otimes I\otimes I + I\otimes T\otimes I + I\otimes I\otimes T)$$

CG for $Ax = b$ vs Iterative solver for $(I \otimes T + T \otimes I)\mathbf{U} + \mathbf{U}T = F$

$T \in \mathbb{R}^{n\times n}$, $A \in \mathbb{R}^{n^3\times n^3}$, $\qquad n = 50$



	CG	PCG	Matrix Eqn solver
Elapsed Time	2.91	0.56	0.08

A 3D convection-diffusion equation

$-\epsilon\Delta u + \mathbf{w}\cdot\nabla u = 1$, in $\Omega = (0,1)^3$, with convection term

$$\mathbf{w} = (x\sin x, y\cos y, e^{z^2-1})$$

Sylvester equation:

$$\left[I \otimes (T_1 + \Phi_1 B_1) + (T_2 + \Psi_2 B_2)^\top \otimes I\right] \mathbf{U} + \mathbf{U}\ (T_3 + B_3 \Upsilon_3) = \mathbf{1}\mathbf{1}^\top$$

ϵ	n_x	FGMRES+AGMG CPU time (# its)	GMRES+MI20 CPU time (# its)	Sylv Solver CPU time (# its)
0.0050	100	8.0207 (15)	9.7207 (7)	0.5677 (22)
0.0010	100	7.6815 (14)	9.4935 (7)	0.5446 (22)
0.0005	100	7.3914 (14)	9.6274 (7)	0.5927 (24)

- Also for more general, separable coeff., operators on uniform grids
- If not separable coeff., use as preconditioner

(Palitta & Simoncini 2016)

... A classical approach

Matrix formulation is not new...

- Bickley & McNamee, 1960: Early literature on difference equations
- Wachspress, 1963: Model problem for ADI algorithm
- Ellner & Wachspress (1980's): interplay between the matrix and vector formulations (via preconditioning)

Novel solvers for matrix equations allow faster convergence

PDEs with random inputs

Stochastic steady-state diffusion eqn: *Find* $u : D \times \Omega \to \mathbb{R}$ *s.t.* $\mathbb{P}$-*a.s.*,

$$\begin{cases} -\nabla \cdot (a(\mathbf{x},\omega)\nabla u(\mathbf{x},\omega)) = & f(\mathbf{x}) \quad in \quad D \\ u(\mathbf{x},\omega) = & 0 \qquad on \quad \partial D \end{cases}$$

f: deterministic;

a: random field, linear function of finite no. of real-valued random variables $\xi_r : \Omega \to \Gamma_r \subset \mathbb{R}$

Common choice: truncated Karhunen–Loève (KL) expansion,

$$a(\mathbf{x},\omega) = \mu(\mathbf{x}) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r}\phi_r(\mathbf{x})\xi_r(\omega),$$

$\mu(\mathbf{x})$: expected value of diffusion coef. σ: std dev.

$(\lambda_r, \phi_r(\mathbf{x}))$ eigs of the integral operator $\mathcal{V}$ wrto $V(\mathbf{x},\mathbf{x}') = \frac{1}{\sigma^2}C(\mathbf{x},\mathbf{x}')$

($\lambda_r \searrow$ $C : D \times D \to \mathbb{R}$ covariance fun.)

Discretization by stochastic Galerkin

Approx with space in tensor product form[a] $\mathcal{X}_h \times S_p$

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \qquad \mathcal{A} = G_0 \otimes K_0 + \sum_{r=1}^{m} G_r \otimes K_r, \quad \mathbf{b} = \mathbf{g}_0 \otimes \mathbf{f}_0,$$

$\mathbf{x}$: expansion coef. of approx to u in the tensor product basis $\{\varphi_i \psi_k\}$

$K_r \in \mathbb{R}^{n_x \times n_x}$, FE matrices (sym)

$G_r \in \mathbb{R}^{n_\xi \times n_\xi}$, $r = 0, 1, \ldots, m$ Galerkin matrices associated w/ S_p (sym.)

$\mathbf{g}_0$: first column of G_0

$\mathbf{f}_0$: FE rhs of deterministic PDE

$$n_\xi = \dim(S_p) = \frac{(m+p)!}{m!p!} \qquad \Rightarrow \boxed{n_x \cdot n_\xi} \text{ huge}$$

---

[a] S_p set of multivariate polyn of total degree $\leq p$

The matrix equation formulation

$$(G_0 \otimes K_0 + G_1 \otimes K_1 + \ldots + G_m \otimes K_m)\,\mathbf{x} = \mathbf{g}_0 \otimes \mathbf{f}_0$$

transforms into

$$K_0 \mathbf{X} G_0 + K_1 \mathbf{X} G_1 + \ldots + K_m \mathbf{X} G_m = F, \qquad F = \mathbf{f}_0 \mathbf{g}_0^\top$$

$(G_0 = I)$

Solution strategy. Conjecture:

- $\{K_r\}$ from trunc'd Karhunen–Loève (KL) expansion

$$\Downarrow$$

$$\mathbf{X} \approx \widetilde{X} \text{ low rank, } \widetilde{X} = X_1 X_2^T$$

(Possibly extending results of Gradesyk, 2004)

Matrix Galerkin approximation of the deterministic part. 1

Approximation space $\mathcal{K}_k$ and basis matrix V_k: $\qquad \mathbf{X} \approx X_k = V_k Y$

$$V_k^\top R_k = 0, \qquad R_k := K_0 X_k + K_1 X_k G_1 + \ldots + K_m X_k G_m - \mathbf{f}_0 \mathbf{g}_0^\top$$

Computational challenges:

- Generation of $\mathcal{K}_k$ involved $m+1$ different matrices $\{K_r\}$!
- Matrices K_r have different spectral properties
- n_x, n_ξ so large that X_k, R_k should not be formed !

Joint project with Catherine Powell, David Silvester, Univ. Manchester

Example 2. $-\nabla \cdot (a \nabla u) = 1, \quad D = (-1,1)^2$. KL expansion.

$\mu = 1$, $\xi_r \sim U(-\sqrt{3}, \sqrt{3})$ and $C(\vec{x}_1, \vec{x}_2) = \sigma^2 \exp\left(-\frac{\|\vec{x}_1 - \vec{x}_2\|_1}{2}\right)$, $n_x = 65,025$, $\sigma = 0.3$

m	p	n_ξ	k	inner its	n_k $\mathcal{K}_k$	rank $\widetilde{\mathbf{X}}$	time secs	CG time (its)
8 87%	2	45	17	9.8	128	45	32.1	13.4 (8)
	3	165	21	12.2	160	129	41.4	56.6 (10)
	4	495	24	14.5	183	178	51.1	197.0 (12)
	5	1,287	27	16.9	207	207	64.0	553.0 (13)
12 89%	2	91	15	9.9	165	89	47.8	30.0 (8)
	3	455	18	12.2	201	196	61.6	175.0 (10)
	4	1,820	21	15.0	236	236	86.4	821.0 (12)
	5	6,188	25	18.6	281	281	188.0	3070.0 (13)
20 93%	2	231	16	9.4	281	206	111.0	94.7 (8)
	3	1,771	23	12.3	399	399	197.0	845.0 (10)
	4	10,626	26	15.4	454	454	556.0	Out of Mem

% of variance integral of a

Not discussed but in this category

- Bilinear systems of matrix equation

$$A_1 X + Y B_1 = C_1$$

$$A_2 X + Y B_2 = C_2$$

...very few numerical procedures available

Not discussed but in this category

- Bilinear systems of matrix equation

$$A_1X + YB_1 = C_1$$
$$A_2X + YB_2 = C_2$$

...very few numerical procedures available

- Sylvester-like linear matrix equations

$$AX + f(X)B = C$$

typically (but not only!): $f(X) = \bar{X}$, $f(X) = X^\top$, or $f(X) = X^*$
(Bevis, Braden, Byers, Chiang, De Terán, Dopico, Duan, Feng, Gonzalez, Guillery, Hall, Hartwig, Ikramov, Kressner, Montealegre, Reyes, Schröder, Vorntsov, Watkins, Wu, ...)

Not discussed but in this category

- Bilinear systems of matrix equation

$$A_1X + YB_1 = C_1$$
$$A_2X + YB_2 = C_2$$

...very few numerical procedures available

- Sylvester-like linear matrix equations

$$AX + f(X)B = C$$

typically (but not only!): $f(X) = \bar{X},\ f(X) = X^\top$, or $f(X) = X^*$
(Bevis, Braden, Byers, Chiang, De Terán, Dopico, Duan, Feng, Gonzalez, Guillery, Hall, Hartwig, Ikramov, Kressner, Montealegre, Reyes, Schröder, Vorntsov, Watkins, Wu, ...)

- Linear systems with complex tensor structure

$$\mathcal{A}\mathbf{x} = b \quad \text{with} \quad \mathcal{A} = \sum_{j=1}^{k} I_{n_1} \otimes \cdots \otimes I_{n_{j-1}} \otimes A_j \otimes I_{n_{j+1}} \cdots \otimes I_{n_k}.$$

Dolgov, Grasedyck, Khoromskij, Kressner, Oseledets, Tobler, Tyrtyshnikov, and many more...

Conclusions

Multiterm (Kron) linear equations is the new challenge

- Great advances in solving really large linear matrix equations
- Second order (matrix) challenges rely on strength and maturity of linear system solvers
- Low-rank tensor formats is the new generation of approximations

Reference for linear matrix equations:

⋆ V. Simoncini,

Computational methods for linear matrix equations,

SIAM Review, Sept. 2016.