Matrix equation techniques for a class of PDE problems with data uncertainty

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria.simoncini@unibo.it

The matrix equation problem

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

$A_i \in \mathbb{R}^{n \times n}$, $B_i \in \mathbb{R}^{m \times m}$, $\mathbf{X}$ unknown matrix

Large dimensions, sparse coefficient matrices

The matrix equation problem

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

$A_i \in \mathbb{R}^{n\times n}$, $B_i \in \mathbb{R}^{m\times m}$, $\mathbf{X}$ unknown matrix

Large dimensions, sparse coefficient matrices

The problem in its full generality is far from tractable, although the transformation to a matrix-vector equation [...] allows us to use the considerable arsenal of numerical weapons currently avaiable for the solution of such problems.

Peter Lancaster, SIAM Rev. 1970

Multiterm linear matrix equation. Classical device

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

Kronecker formulation

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

Iterative methods: matrix-matrix multiplications and rank truncation

(Benner, Breiten, Bouhamidi, Chehab, Damm, Grasedyck, Jbilou, Kressner, Matthies, Nagy, Onwunta, Raydan, Stoll, Tobler, Wedderburn, Zander, ...)

Multiterm linear matrix equation. Classical device

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

Kronecker formulation

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

Iterative methods: matrix-matrix multiplications and rank truncation

(Benner, Breiten, Bouhamidi, Chehab, Damm, Grasedyck, Jbilou, Kressner, Matthies, Nagy, Onwunta, Raydan, Stoll, Tobler, Wedderburn, Zander, ...)

Applications:

- Stochastic dynamical systems
- (Stochastic) PDEs
- Inverse problems and optimization
- ...

Multiterm linear matrix equation

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

Alternative approaches to the Kronecker form:

- Fixed point iterations (an "evergreen"...)
- Projection-type methods $\Rightarrow$ low rank approximation
- Ad-hoc problem-dependent procedures
- etc.

Multiterm linear matrix equation

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

Alternative approaches to the Kronecker form:

- Fixed point iterations (an “evergreen”...)
- Projection-type methods $\Rightarrow$ low rank approximation
- Ad-hoc problem-dependent procedures
- etc.

A sample of these methodologies on different problems:

- Control problem with stochastic parameters
- PDEs on uniform discretizations
- Stochastic PDE

A class of generalized Lyapunov equations

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

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

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

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

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

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

Stationary iterative methods by splitting

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

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

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

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

Assuming that (A, B) is controllable and $\mathbf{X}$ sym positive semi-def then

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

Stationary iteration:

$$\mathcal{M}(\mathbf{X}_k) = \mathcal{N}(\mathbf{X}_{k-1}) - BB^T, \quad k = 1, 2, \ldots.$$

(Shank & Simoncini & Szyld, 2016)

Stationary iterative methods by splitting. Cont'd

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

Stationary iteration:

$$\mathcal{M}(\mathbf{X}_k) = \mathcal{N}(\mathbf{X}_{k-1}) - BB^T, \qquad k = 1, 2, \ldots.$$

> **Approximately Solve** $A\mathbf{X} + \mathbf{X}A^T + BB^T = 0$ for $\mathbf{X}_1 = Z_1 Z_1^T$
>
> for $k = 2, 3, \ldots$
>
>   **Set** $B_k = [N_1 Z_{k-1}, \ \cdots \ , N_m Z_{k-1}, \ B]$
>
>   **Approximately Solve** $A\mathbf{X} + \mathbf{X}A^T + B_k B_k^T = 0$ for $\mathbf{X}_k = Z_k Z_k^T$
>
>   **If** sufficiently accurate **then stop**

Stationary iterative methods by splitting. Cont'd

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

Stationary iteration:

$$\mathcal{M}(\mathbf{X}_k) = \mathcal{N}(\mathbf{X}_{k-1}) - BB^T, \qquad k = 1, 2, \ldots.$$

Approximately Solve $A\mathbf{X} + \mathbf{X}A^T + BB^T = 0$ for $\mathbf{X}_1 = Z_1 Z_1^T$
for $k = 2, 3, \ldots$
Set $B_k = [N_1 Z_{k-1}, \ \cdots \ , N_m Z_{k-1}, \ B]$
Approximately Solve $A\mathbf{X} + \mathbf{X}A^T + B_k B_k^T = 0$ for $\mathbf{X}_k = Z_k Z_k^T$
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

Multi-term linear matrix equations in PDEs

– uniform grids and separable coeffs –

$$-\varepsilon \boldsymbol{\Delta} \mathbf{u} + \phi_{\mathbf{1}}(\mathbf{x})\psi_{\mathbf{1}}(\mathbf{y})\mathbf{u}_{\mathbf{x}} + \phi_{\mathbf{2}}(\mathbf{x})\psi_{\mathbf{2}}(\mathbf{y})\mathbf{u}_{\mathbf{y}} + \gamma_{\mathbf{1}}(\mathbf{x})\gamma_{\mathbf{2}}(\mathbf{y})\mathbf{u} = \mathbf{f}$$

$(x, y) \in \Omega \subset \mathbb{R}^2$, ϕ_i, ψ_i, γ_i, $i = 1, 2$ sufficiently regular func's + b.c.

Problem discretization by means of a tensor basis

Multi-term linear matrix equations in PDEs

– uniform grids and separable coeffs –

$$-\varepsilon\boldsymbol{\Delta}\mathbf{u} + \phi_{\mathbf{1}}(\mathbf{x})\psi_{\mathbf{1}}(\mathbf{y})\mathbf{u}_{\mathbf{x}} + \phi_{\mathbf{2}}(\mathbf{x})\psi_{\mathbf{2}}(\mathbf{y})\mathbf{u}_{\mathbf{y}} + \gamma_{\mathbf{1}}(\mathbf{x})\gamma_{\mathbf{2}}(\mathbf{y})\mathbf{u} = \mathbf{f}$$

$(x, y) \in \Omega \subset \mathbb{R}^2$, ϕ_i, ψ_i, γ_i, $i = 1, 2$ sufficiently regular func's + b.c.

Problem discretization by means of a tensor basis

Multiterm linear equation:

$$-\varepsilon T_1\mathbf{U} - \varepsilon\mathbf{U}T_2 + \Phi_1 B_1\mathbf{U}\Psi_1 + \Phi_2\mathbf{U}B_2^\top\Psi_2 + \Gamma_1\mathbf{U}\Gamma_2 = F$$

Finite Diff.: $\mathbf{U}_{i,j} = \mathbf{U}(x_i, y_j)$ approximate solution at the nodes

Multi-term linear matrix equations in PDEs

– uniform grids and separable coeffs –

$$-\varepsilon\boldsymbol{\Delta}\mathbf{u} + \phi_{\mathbf{1}}(\mathbf{x})\psi_{\mathbf{1}}(\mathbf{y})\mathbf{u_x} + \phi_{\mathbf{2}}(\mathbf{x})\psi_{\mathbf{2}}(\mathbf{y})\mathbf{u_y} + \gamma_{\mathbf{1}}(\mathbf{x})\gamma_{\mathbf{2}}(\mathbf{y})\mathbf{u} = \mathbf{f}$$

$(x, y) \in \Omega \subset \mathbb{R}^2$, ϕ_i, ψ_i, γ_i, $i = 1, 2$ sufficiently regular func's + b.c.

Problem discretization by means of a tensor basis

Multiterm linear equation:

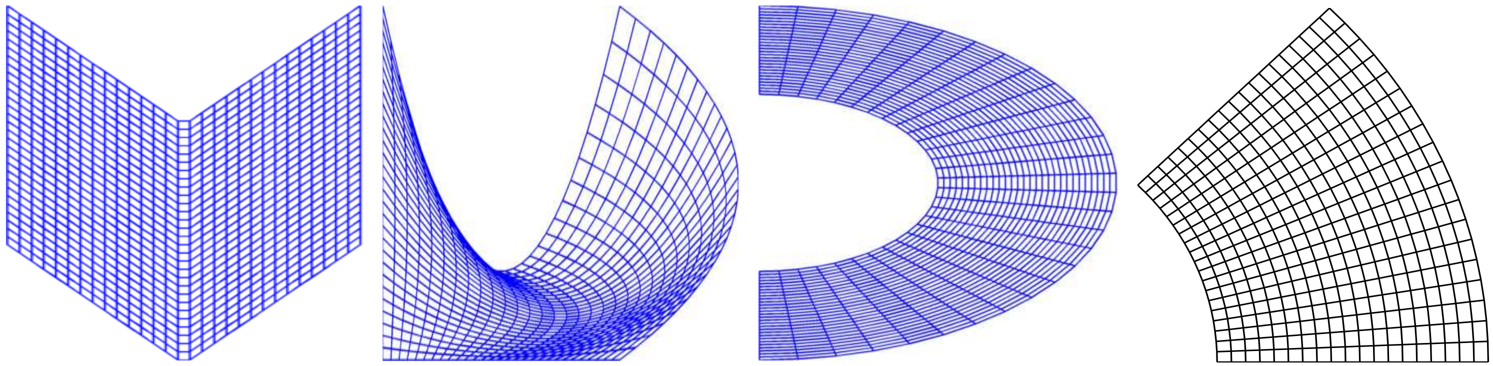
$$-\varepsilon T_1\mathbf{U} - \varepsilon\mathbf{U}T_2 + \Phi_1 B_1\mathbf{U}\Psi_1 + \Phi_2\mathbf{U}B_2^\top\Psi_2 + \Gamma_1\mathbf{U}\Gamma_2 = F$$

Finite Diff.: $\mathbf{U}_{i,j} = \mathbf{U}(x_i, y_j)$ approximate solution at the nodes

$\boxed{\text{More than finite differences and rectangular domains!}}$

Multi-term linear matrix equations in PDEs

Some very classical domains



$\Rightarrow$ FD w/transfinite interpolation, Isogeometric Analysis, etc.

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:

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

$$\Leftrightarrow \quad \mathcal{A}\mathbf{U}+\mathbf{U}\mathcal{B}=F$$

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)

- 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: $Find\ u : D \times \Omega \to \mathbb{R}\ s.t.\ \mathbb{P}\text{-}a.s.,$

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

f: deterministic;

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

PDEs with random inputs

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

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

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.)

Stochastic Galerkin discretization. The SPDE-practitioner approach.

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

$$\Downarrow$$

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

Matrix Galerkin approximation of the deterministic part.

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

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

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 !

(Powell & Silvester & Simoncini 2017)

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)
	2	45	17	9.8	128	45	32.1	13.4 (8)
8	3	165	21	12.2	160	129	41.4	56.6 (10)
87%	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)
	2	91	15	9.9	165	89	47.8	30.0 (8)
12	3	455	18	12.2	201	196	61.6	175.0 (10)
89%	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)
	2	231	16	9.4	281	206	111.0	94.7 (8)
20	3	1,771	23	12.3	399	399	197.0	845.0 (10)
93%	4	10,626	26	15.4	454	454	556.0	Out of Mem

% of variance integral of a

Reduction of the stochastic space. 1

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

$\{K_i\}$ $\Rightarrow$ Range(V_k) reduced spatial space

$\{G_i\}$ $\Rightarrow$ Range(W_k) reduced stocastic space ?

Goal: $\mathbf{X} \;\approx\; \widetilde{\mathbf{X}} = V_k Y_k W_k^T$

Reduction of the stochastic space. 1

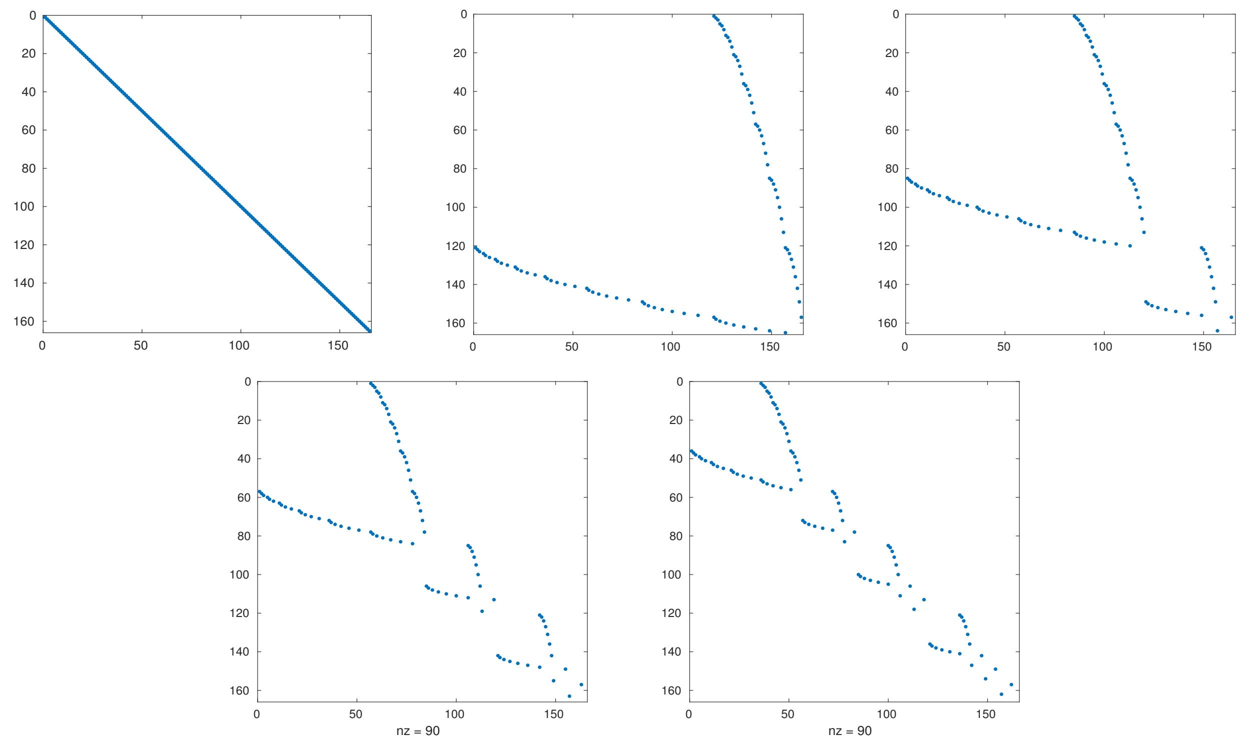
$$K_0\mathbf{X}G_0 + K_1\mathbf{X}G_1 + \ldots + K_m\mathbf{X}G_m = F, \qquad F = \mathbf{f}_0\mathbf{g}_0^\top$$

$\{K_i\}$ $\Rightarrow$ Range(V_k) reduced spatial space

$\{G_i\}$ $\Rightarrow$ Range(W_k) reduced stocastic space ?

Goal: $\mathbf{X} \;\approx\; \widetilde{\mathbf{X}} = V_k Y_k W_k^T$

Some structural properties of the $\{G_i\}$:



Reduction of the stochastic space. 1

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

$\{K_i\}$ $\Rightarrow$ Range(V_k) reduced spatial space

$\{G_i\}$ $\Rightarrow$ Range(W_k) reduced stocastic space ?

Goal: $\mathbf{X} \approx \widetilde{\mathbf{X}} = V_k Y_k W_k^T$

Some algebraic properties of the $\{G_i\}$:

- G_i sym of size $n_\xi = \frac{(m+p)!}{m!p!}$ with only two nonzeros per row
- Eigenvalues are known and bounded (symmetric to the origin, many zeros)
- $G_0 = I$, and G_i, $i = 1, \ldots, m$ are permutation-similar to the same block-diagonal matrix $G^\star$ with $\binom{M-1+p}{p}$ diagonal blocks.
- Each of these blocks is a leading principal submatrix of a certain known $(p+1) \times (p+1)$ matrix

(Powell & Elman 2009, Ernst & Ullmann 2010)

Reduction of the stochastic space. 2

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

$$\mathbf{X} \;\approx\; \widetilde{\mathbf{X}} = V_k Y_k W_k^T$$

- Choice of reduced basis for nested spatial space:

$$V_{k+1} = \left[V_k, [(K_1 + s_1K_0)^{-1}v_k, \ldots, (K_1 + s_mK_0)^{-1}v_k]\right]$$

(Powell & Silvester & Simoncini 2017)

Reduction of the stochastic space. 2

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

$$\mathbf{X} \approx \widetilde{\mathbf{X}} = V_k Y_k W_k^T$$

- Choice of reduced basis for nested spatial space:

$$V_{k+1} = \left[V_k, [(K_1 + s_1 K_0)^{-1} v_k, \ldots, (K_1 + s_m K_0)^{-1} v_k] \right]$$

  (Powell & Silvester & Simoncini 2017)

- Choice of reduced (sparse) basis for nested stochastic space:

$$W_k = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k]$$

  where

  $\mathbf{w}_1$ eigvecs of largest block (leading eigenvalues)

  $\mathbf{w}_2$ eigvecs of next largest blocks

  etc.

  (Locatelli & Simoncini, work in progress)

Reduction of the stochastic space. 3

n_x: spatial dimension n_ξ: stochastic dimension

slow decay in KL expansion:

n_x	(m,p)	n_ξ	dim V	rank($\widetilde{X}$)	Unreduced W_k	reduced W_k
16129	(5,3)	56	69	43	1.80	1.88
	(7,5)	792	78	63	3.65	2.40
	(9,7)	11440	95	77	494.63	29.11
65025	(5,3)	56	69	43	8.74	8.95
	(7,5)	792	78	63	12.44	11.16
	(9,7)	11440	95	77	548.60	40.02

Reduction of the stochastic space. 3

n_x: spatial dimension $\qquad n_\xi$: stochastic dimension

slow decay in KL expansion:

n_x	(m,p)	n_ξ	dim V	rank($\widetilde{X}$)	Unreduced W_k	reduced W_k
16129	(5,3)	56	69	43	1.80	1.88
	(7,5)	792	78	63	3.65	2.40
	(9,7)	11440	95	77	494.63	29.11
65025	(5,3)	56	69	43	8.74	8.95
	(7,5)	792	78	63	12.44	11.16
	(9,7)	11440	95	77	548.60	40.02

fast decay in KL expansion:

n_x	(m,p)	n_ξ	dim V	rank($\widetilde{X}$)	Unreduced W_k	reduced W_k
16129	(5,3)	56	107	30	3.07	3.17
	(7,5)	792	133	45	8.88	5.90
	(9,7)	11440	166	57	1432.3	92.33
65025	(5,3)	56	107	30	14.61	14.37
	(7,5)	792	142	46	27.77	24.17
	(9,7)	11440	162	58	1543.90	112.77

Not discussed but in this category

- Bilinear systems of matrix equations. E.g.,

$$\begin{aligned} A_1X + XA_2 + YB &= C_1 \\ B^TX &= C_2 \end{aligned}$$

...very few numerical procedures available

Not discussed but in this category

- Bilinear systems of matrix equations. E.g.,

$$\begin{aligned} A_1X + XA_2 + YB &= C_1 \\ B^TX &= C_2 \end{aligned}$$

  ...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 equations. E.g.,

$$\begin{aligned} A_1X + XA_2 + YB &= C_1 \\ B^TX &= C_2 \end{aligned}$$

  ...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 level $matrix$ challenges rely on strength and maturity of linear system solvers
- Low-rank tensor formats is the new generation of approximations

Conclusions

Multiterm (Kron) linear equations is the new challenge

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

Some references cited in the text:

⋆ D. Palitta and V. Simoncini, *Matrix-equation-based strategies for convection-diffusion equations*, BIT Numer. Math. 2016

⋆ C. E. Powell, D. Silvester and V. Simoncini, *An efficient reduced basis solver for stochastic Galerkin matrix equations*, SISC, 2017

⋆ S. D. Shank, V. Simoncini and D. B. Szyld, *Efficient low-rank solutions of Generalized Lyapunov equations*, Numerische Mathematik, 2016

⋆ V. Simoncini, *Computational methods for linear matrix equations*, SIAM Rev. 2016