On low-rank methods for large-scale matrix equations and application to PDEs

V. Simoncini

Dipartimento di Matematica

Alma Mater Studiorum - Università di Bologna

valeria.simoncini@unibo.it

Some matrix equations - large scale

- Sylvester matrix equation $\quad A\mathbf{X} + \mathbf{X}B + D = 0$

  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, ...

Survey article: V.Simoncini, SIAM Review 2016.

More matrix equations - large scale

- Systems of linear matrix equations:

$$\begin{aligned} A_2\mathbf{X} + \mathbf{X}A_1 + B^\top \mathbf{P} &= F_1 \\ A_1\mathbf{Y} + \mathbf{Y}A_2 + \mathbf{P}B &= F_2 \\ B\mathbf{X} + \mathbf{Y}B^\top &= F_3 \end{aligned}$$

(Simoncini, 2019 to appear in IMA Num.Anal.)

More matrix equations - large scale

- Systems of linear matrix equations:

$$\begin{aligned} A_2\mathbf{X} + \mathbf{X}A_1 + B^\top\mathbf{P} &= F_1 \\ A_1\mathbf{Y} + \mathbf{Y}A_2 + \mathbf{P}B &= F_2 \\ B\mathbf{X} + \mathbf{Y}B^\top &= F_3 \end{aligned}$$

(Simoncini, 2019 to appear in IMA Num.Anal.)

- Riccati equation: Find $\mathbf{X} \in \mathbb{R}^{n\times n}$ such that

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

workhorse in Control Theory

More matrix equations - large scale

- Systems of linear matrix equations:

$$\begin{aligned} A_2\mathbf{X} + \mathbf{X}A_1 + B^\top \mathbf{P} &= F_1 \\ A_1\mathbf{Y} + \mathbf{Y}A_2 + \mathbf{P}B &= F_2 \\ B\mathbf{X} + \mathbf{Y}B^\top &= F_3 \end{aligned}$$

(Simoncini, 2019 to appear in IMA Num.Anal.)

- Riccati equation: Find $\mathbf{X} \in \mathbb{R}^{n\times n}$ such that

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

workhorse in Control Theory

- Tensor equation: Find $\mathbf{X} \in \mathbb{R}^{n\times n\times n}$ such that

$$(H \otimes M \otimes A + M \otimes A \otimes H + A \otimes H \otimes M)\mathbf{x} + c = 0 \quad \mathbf{x} = \mathrm{vec}(\mathbf{X})$$

Discretization of parameter-dependent PDEs

Projection-type methods

Approximate $\mathbf{X}$ in:

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

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

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

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

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

Projection-type methods

Approximate $\mathbf{X}$ in:

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

Given an low dimensional 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:

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

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

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

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

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

  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

$$\mathcal{K} = \mathbb{K} := \mathrm{Range}([B, (A - s_2 I)^{-1}B, \ldots, (A - s_m I)^{-1}B])$$

  usually, $\{s_2, \ldots, s_m\} \subset \mathbb{C}^+$ chosen either a-priori or dynamically

  (Used in different contexts, since Ruhe '84)

More recent options as approximation space

Enrich space to decrease space dimension

- Extended Krylov subspace

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

  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

$$\mathcal{K} = \mathbb{K} := \mathrm{Range}([B, (A - s_2 I)^{-1}B, \ldots, (A - s_m I)^{-1}B])$$

  usually, $\{s_2, \ldots, s_m\} \subset \mathbb{C}^+$ chosen either a-priori or dynamically
  (Used in different contexts, since Ruhe '84)

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

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

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:

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

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, Szyld, 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:

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

Alternative approaches:

- Projection onto rich approximation space
- Compression to two-term matrix equation
- Splitting strategy towards two-term matrix equation
- ...

PDEs on uniform grids and separable coeffs

$$-\varepsilon\Delta u+\phi_1(x)\psi_1(y)u_x+\phi_2(x)\psi_2(y)u_y+\gamma_1(x)\gamma_2(y)u=f \quad (x,y)\in\Omega$$

ϕ_i, ψ_i, γ_i, $i=1,2$ sufficiently regular functions + b.c.

Problem discretization by means of a tensor basis:

Finite differences, isogeometric analysis, spectral methods, etc.

PDEs on uniform grids and separable coeffs

$$-\varepsilon\Delta u+\phi_1(x)\psi_1(y)u_x+\phi_2(x)\psi_2(y)u_y+\gamma_1(x)\gamma_2(y)u=f \quad (x,y)\in\Omega$$

ϕ_i, ψ_i, γ_i, $i=1,2$ sufficiently regular functions + b.c.

Problem discretization by means of a tensor basis:

Finite differences, isogeometric analysis, spectral methods, etc.

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

(see, e.g., Palitta & Simoncini, '16)

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 function)

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

$$\Downarrow$$

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

(Possibly extending results of Grasedyck, 2004)

Matrix Galerkin approximation of the deterministic part. 1

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, SISC 2017)

More on Kronecker connection for low-rank Galerkin approximation

$$A_1XB_1 + A_2XB_2 + \ldots + A_\ell XB_\ell = F$$

- Operators: $\mathcal{S} : \mathbb{R}^{n\times p} \to \mathbb{R}^{n\times p}$,

$$\mathcal{S} : \quad X \mapsto \sum_{j=1}^{\ell} A_j X B_j,$$

and $\mathcal{S}_\ell := \sum_{j=1}^{\ell} B_j^\top \otimes A_j$. So that

$$\mathcal{S}(\mathbf{X}) = F \quad \Leftrightarrow \quad \mathcal{S}_\ell \mathrm{vec}(\mathbf{X}) = \mathrm{vec}(F)$$

More on Kronecker connection for low-rank Galerkin approximation

$$A_1 X B_1 + A_2 X B_2 + \ldots + A_\ell X B_\ell = F$$

- Operators: $\mathcal{S} : \mathbb{R}^{n\times p} \to \mathbb{R}^{n\times p}$,

$$\mathcal{S} : \quad X \mapsto \sum_{j=1}^{\ell} A_j X B_j,$$

and $\mathcal{S}_\ell := \sum_{j=1}^{\ell} B_j^\top \otimes A_j$. So that

$$\mathcal{S}(\mathbf{X}) = F \quad \Leftrightarrow \quad \mathcal{S}_\ell \mathrm{vec}(\mathbf{X}) = \mathrm{vec}(F)$$

- Galerkin condition: $V_k^\top R_k W_k = 0 \quad \Leftrightarrow \quad (W_k \otimes V_k)^\top r_k = 0$

where $r_k = \mathrm{vec}(R_k)$ and $\mathcal{V}_m = \mathrm{range}(W_k \otimes V_k)$

Optimality properties of low-rank Galerkin approximation

For $\mathcal{S}$: $\quad X \mapsto \sum_{j=1}^{\ell} A_j X B_j$ and $\mathcal{S}_\ell := \sum_{j=1}^{\ell} B_j^\top \otimes A_j$:

DEF: $\mathcal{S}$ is symmetric and positive definite if for any $0 \neq x \in \mathbb{R}^{np}$, $x = \mathrm{vec}(X)$, with $X \in \mathbb{R}^{n \times p}$, it holds

- $\mathcal{S}_\ell = \mathcal{S}_\ell^\top$
- $x^\top \mathcal{S}_\ell x > 0$, where $x^\top \mathcal{S}_\ell x = \mathrm{trace}\left(\sum_{j=1}^{\ell} X^\top A_j X B_j\right)$

$\Rightarrow$ We use $\|X\|_{\mathcal{S}}^2 := x^\top \mathcal{S}_\ell x$. (see, e.g., Vandereycken & Vandewalle, '10)

Optimality properties of low-rank Galerkin approximation

For $\mathcal{S}$: $\quad X \mapsto \sum_{j=1}^{\ell} A_j X B_j$ and $\mathcal{S}_\ell := \sum_{j=1}^{\ell} B_j^\top \otimes A_j$:

DEF: $\mathcal{S}$ is symmetric and positive definite if for any $0 \neq x \in \mathbb{R}^{np}$, $x = \mathrm{vec}(X)$, with $X \in \mathbb{R}^{n\times p}$, it holds

- $\mathcal{S}_\ell = \mathcal{S}_\ell^\top$
- $x^\top \mathcal{S}_\ell x > 0$, where $x^\top \mathcal{S}_\ell x = \mathrm{trace}\left(\sum_{j=1}^{\ell} X^\top A_j X B_j\right)$

$\Rightarrow$ We use $\|X\|_{\mathcal{S}}^2 := x^\top \mathcal{S}_\ell x$. (see, e.g., Vandereycken & Vandewalle, '10)

PROP: Let $\mathcal{S}(X) = F$ with $\mathcal{S} : X \mapsto \sum_j A_j X B_j$ sym.pos.def., and let $X_k = V_k Y_k W_k^\top$ be the Galerkin approximate solution. Then

$$\|X - X_k\|_{\mathcal{S}} = \min_{\substack{Z = V_k Y W_k^\top \\ Y \in \mathbb{R}^{k\times k}}} \|X - Z\|_{\mathcal{S}}$$

(Palitta & Simoncini, tr2019; see also Kressner & Tobler, '10 for related results)

Optimality properties of Galerkin approximation. Lyapunov equation.

$$A^\top X + XA + F = 0$$

For A sym.pos.def.,

$$\|X\|_{\mathcal{S}}^2 = 2\,\mathrm{trace}(X^\top AX).$$

Let $E_k = X - X_k$. Then

$$\|E_k\|_{\mathcal{S}}^2 = \min_{\substack{Z = V_k Y W_k^\top \\ Y \in \mathbb{R}^{k\times k}}} \|X - Z\|_{\mathcal{S}}^2 = 2\,\mathrm{trace}(E_k A E_k).$$

Optimality properties of Galerkin approximation. Lyapunov equation.

$$A^\top X + XA + F = 0$$

For A sym.pos.def.,

$$\|X\|_{\mathcal{S}}^2 = 2\,\mathrm{trace}(X^\top AX).$$

Let $E_k = X - X_k$. Then

$$\|E_k\|_{\mathcal{S}}^2 = \min_{\substack{Z=V_kYW_k^\top \\ Y\in\mathbb{R}^{k\times k}}} \|X - Z\|_{\mathcal{S}}^2 = 2\,\mathrm{trace}(E_kAE_k).$$

♣: If F is sym, we can choose $W_k = V_k$

Optimality properties of low-rank Petrov-Galerkin approximation

$\mathcal{K}_k = \mathrm{range}(W_k \otimes V_k)$ approximation space, $\mathcal{L}_k$ test space.

In vector form: $\mathcal{L}_k^\top r_m = 0$

Optimality properties of low-rank Petrov-Galerkin approximation

$\mathcal{K}_k = \mathrm{range}(W_k \otimes V_k)$ approximation space, $\mathcal{L}_k$ test space.

In vector form: $\mathcal{L}_k^\top r_m = 0$

For $\mathcal{L}_k = \mathcal{S}_\ell \mathcal{K}_k$, minimization of residual norm:

$$\min_{x \in \mathsf{range}(W_k \otimes V_k)} \|\mathsf{vec}(F) - \mathcal{S}_\ell x\|_2 = \min_{y \in \mathbb{R}^{k^2}} \|\mathsf{vec}(F) - \mathcal{S}_\ell (W_k \otimes V_k) y\|_2$$

Optimality properties of low-rank Petrov-Galerkin approximation

$\mathcal{K}_k = \mathrm{range}(W_k \otimes V_k)$ approximation space, $\mathcal{L}_k$ test space.

In vector form: $\mathcal{L}_k^\top r_m = 0$

For $\mathcal{L}_k = \mathcal{S}_\ell \mathcal{K}_k$, minimization of residual norm:

$$\min_{x \in \mathsf{range}(W_k \otimes V_k)} \|\mathsf{vec}(F) - \mathcal{S}_\ell x\|_2 = \min_{y \in \mathbb{R}^{k^2}} \|\mathsf{vec}(F) - \mathcal{S}_\ell (W_k \otimes V_k) y\|_2$$

In matrix form:

$$\min_{X = V_k Y W_k^\top} \|F - \mathcal{S}(X)\|_F = \min_{Y \in \mathbb{R}^{k \times k}} \|F - \mathcal{S}(V_k Y W_k^\top)\|_F$$

For the Lyapunov eqn, Lin & Simoncini, 2013

Optimality properties of low-rank Petrov-Galerkin approximation

$\mathcal{K}_k = \mathrm{range}(W_k \otimes V_k)$ approximation space, $\mathcal{L}_k$ test space.

In vector form: $\mathcal{L}_k^\top r_m = 0$

For $\mathcal{L}_k = \mathcal{S}_\ell \mathcal{K}_k$, minimization of residual norm:

$$\min_{x \in \mathrm{range}(W_k \otimes V_k)} \|\mathrm{vec}(F) - \mathcal{S}_\ell x\|_2 = \min_{y \in \mathbb{R}^{k^2}} \|\mathrm{vec}(F) - \mathcal{S}_\ell (W_k \otimes V_k) y\|_2$$

In matrix form:

$$\min_{X = V_k Y W_k^\top} \|F - \mathcal{S}(X)\|_F = \min_{Y \in \mathbb{R}^{k \times k}} \|F - \mathcal{S}(V_k Y W_k^\top)\|_F$$

For the Lyapunov eqn, Lin & Simoncini, 2013

A problem: Even if exact solution is definite, Y is not necessarily definite!

A constrained optimality of low-rank Petrov-Galerkin approximation

Impose definiteness as constraint:

$$\min_{\substack{Y \in \mathbb{R}^{k \times k} \\ Y \leq 0}} \|F - \mathcal{S}(V_k Y W_k^\top)\|_F$$

A constrained optimality of low-rank Petrov-Galerkin approximation

Impose definiteness as constraint:

$$\min_{\substack{Y \in \mathbb{R}^{k \times k} \\ Y \leq 0}} \|F - \mathcal{S}(V_k Y W_k^\top)\|_F$$

Optimization context: linear matrix inequalities

$$Y \leq 0, \qquad \begin{bmatrix} I & \mathrm{vec}(F - \mathcal{S}(V_k Y W_k^\top)) \\ \mathrm{vec}(F - \mathcal{S}(V_k Y W_k^\top))^\top & \gamma \end{bmatrix} \geq 0$$

for the unknown matrix Y and scalar $\gamma > 0$

A constrained optimality of low-rank Petrov-Galerkin approximation

Impose definiteness as constraint:

$$\min_{\substack{Y \in \mathbb{R}^{k \times k} \\ Y \le 0}} \|F - \mathcal{S}(V_k Y W_k^\top)\|_F$$

Optimization context: linear matrix inequalities

$$Y \le 0, \qquad \begin{bmatrix} I & \mathrm{vec}(F - \mathcal{S}(V_k Y W_k^\top)) \\ \mathrm{vec}(F - \mathcal{S}(V_k Y W_k^\top))^\top & \gamma \end{bmatrix} \ge 0$$

for the unknown matrix Y and scalar $\gamma > 0$

♣: In practice, optimization problem solved in the reduced space!

Palitta & Simoncini, tr2019

Low-rank Tensor equation

Find the unique $\mathbf{X} \in \mathbb{R}^{n \times n \times n}$ such that

$$(H \otimes M \otimes A + M \otimes A \otimes H + A \otimes H \otimes M)\text{vec}(\mathbf{X}) + f_3 \otimes f_2 \otimes f_1 = 0$$

Low-rank Tensor equation

Find the unique $\mathbf{X} \in \mathbb{R}^{n\times n\times n}$ such that

$$(H \otimes M \otimes A + M \otimes A \otimes H + A \otimes H \otimes M)\mathrm{vec}(\mathbf{X}) + f_3 \otimes f_2 \otimes f_1 = 0$$

PROP. Let $A^\top H^{-\top} = Q\Lambda Q^{-1}$ be the eigendecomposition of $A^\top H^{-\top}$. Then for each $k = 1, \ldots, n$, the solution $\mathbf{X}$ is obtained as $\mathbf{X}(:, :, k) = \mathbf{Z}_k Q^{-\top}$ where $\mathbf{Z}_k$ solves

$$(A + \lambda_k H)\mathbf{Z}M^\top + M\mathbf{Z}A^\top + f_2 g_k f_3^\top = 0$$

with $g^\top = f_1^\top H^{-\top} Q$.

Low-rank Tensor equation

Find the unique $\mathbf{X} \in \mathbb{R}^{n \times n \times n}$ such that

$$(H \otimes M \otimes A + M \otimes A \otimes H + A \otimes H \otimes M)\mathrm{vec}(\mathbf{X}) + f_3 \otimes f_2 \otimes f_1 = 0$$

PROP. Let $A^\top H^{-\top} = Q \Lambda Q^{-1}$ be the eigendecomposition of $A^\top H^{-\top}$. Then for each $k = 1, \ldots, n$, the solution $\mathbf{X}$ is obtained as $\mathbf{X}(:,:,k) = \mathbf{Z}_k Q^{-\top}$ where $\mathbf{Z}_k$ solves

$$(A + \lambda_k H)\mathbf{Z} M^\top + M \mathbf{Z} A^\top + f_2 g_k f_3^\top = 0$$

with $g^\top = f_1^\top H^{-\top} Q$.

An example: (Random data)

n	9	25	49	81	121
CPU Time (secs)	0.0054	0.0442	0.3069	1.679	6.7465

(Work in progress)

♣ For $n = 121 \quad \Rightarrow \quad n^3 = 1,771,561$

Conclusions and Outlook

Large-scale (Multiterm) linear equations are a new computational tool

- Reduced Order methods are a key ingredient and may show optimality properties
- Matrix equation challenges rely on strength and maturity of linear system solvers
- Low-rank tensor equations require new thinking

Webpage: `www.dm.unibo.it/~simoncin`

Reference for linear matrix equations:

⋆ V. Simoncini,

Computational methods for linear matrix equations,

SIAM Review, Sept. 2016.