On the versatility of Krylov subspaces in modern NLA

Valeria Simoncini

Dipartimento di Matematica
Alma Mater Studiorum - Università di Bologna
valeria.simoncini@unibo.it

The framework

It is given an operator $v \to \mathcal{A}_\epsilon(v)$.

Efficiently solve the given problem in the approximation space

$$\mathcal{K}_m = \mathrm{span}\{v, \mathcal{A}_{\epsilon_1}(v), \mathcal{A}_{\epsilon_2}(\mathcal{A}_{\epsilon_1}(v)), \ldots\}, \quad v \in \mathbb{C}^n$$

with $\dim(\mathcal{K}_m) = m$, where $\mathcal{A}_\epsilon \to \mathcal{A}$ for $\epsilon \to 0$ (ϵ may be tuned)

$\star$ for $\mathcal{A} = A$, $\epsilon = 0 \Rightarrow \quad \mathcal{K}_m = \mathrm{span}\{v, Av, A^2 v, \ldots, A^{m-1} v\}$

Examples of $\mathcal{A}$:

- Solution of (preconditioned) large linear systems,

$$Ax = b \qquad n \times n \qquad \mathcal{A} = A$$

- Shift-and invert eigensolvers

$$Ax = \lambda Mx, \qquad \|x\| = 1, \qquad \mathcal{A} = (\sigma M - A)^{-1}$$

- Preconditioned exponential approximation

$$x = \exp(A)v, \qquad \mathcal{A} = (\gamma I - A)^{-1}$$

- ...

Goal: Achieve approximation x_m to x within a fixed tolerance, by using $\mathcal{A}_\epsilon$ (and *not* $\mathcal{A}$), with variable ϵ

Many applications in Scientific Computing

$\mathcal{A}(v)$ function (linear in v):

- Structured problems (e.g., Schur complement)
- Krylov-based approximations
  1. Matrix functions evaluations
  2. Matrix equations
- Preconditioned system: $AP^{-1}x = b$, where $P^{-1}v_i \approx P_i^{-1}v_i$
- etc.

Other inexact computations for which the same setting holds

- Round-off error analysis
- Mixed-precision computations (e.g., Gratton, Simon, Titley-Peloquin, Toint)
- Truncated Matrix/Tensor computations

Many applications in Scientific Computing

$\mathcal{A}(v)$ function (linear in v):

- Structured problems (e.g., Schur complement)
- Krylov-based approximations
  1. Matrix functions evaluations
  2. Matrix equations
- Preconditioned system: $AP^{-1}x = b$, where $P^{-1}v_i \approx P_i^{-1}v_i$
- etc.

Other inexact computations for which the same setting holds

- Round-off error analysis
- Mixed-precision computations (e.g., Gratton, Simon, Titley-Peloquin, Toint)
- Truncated Matrix/Tensor computations

The exact approach

To focus our attention: $\mathcal{A} = A$.

$\mathcal{K}_m$ Krylov subspace $\qquad V_m$ orthogonal basis

Key relation in Krylov subspace methods:

$$AV_m = V_{m+1}\underline{H}_m \qquad v = V_{m+1}e_1\beta \qquad \underline{H}_m = \begin{bmatrix} H_m \\ h_{m+1,m}e_m^T \end{bmatrix}$$

System: $x_m \in \mathcal{K}_m \quad \Rightarrow \quad x_m = V_m y_m \qquad (x_0 = 0)$

Eigenpb: (θ, y) eigenpair of $H_m \quad \Rightarrow \quad (\theta, V_m y)$ Ritz pair for (λ, x)

The inexact key relation

$$\mathcal{A} = A \quad \rightarrow \quad \mathcal{A}_\epsilon \approx A$$

e.g., $\mathcal{A}_\epsilon v := \mathcal{A}v + w, \qquad \|w\| = \epsilon$

$$AV_m = V_{m+1}\underline{H}_m + \underbrace{F_m}_{[f_1, f_2, \ldots, f_m]} \qquad F_m \text{ error matrix}, \ \|f_j\| = O(\epsilon_j)$$

How large is F_m allowed to be?

system:

$$\begin{aligned} r_m &= b - AV_m y_m = b - V_{m+1}\underline{H}_m y_m - F_m y_m \\ &= \underbrace{V_{m+1}(e_1\beta - \underline{H}_m y_m)}_{\text{computed residual} =: \tilde{r}_m} - F_m y_m \end{aligned}$$

eigenproblem: $(\theta, V_m y)$

$$r_m = \theta V_m y - AV_m y = v_{m+1} h_{m+1,m} e_m^T y - F_m y$$

The inexact key relation

$$\mathcal{A} = A \quad \to \quad \mathcal{A}_\epsilon \approx A$$

e.g., $\mathcal{A}_\epsilon v := \mathcal{A}v + w, \qquad \|w\| = \epsilon$

$$AV_m = V_{m+1}\underline{H}_m + \underbrace{F_m}_{[f_1, f_2, \ldots, f_m]} \qquad F_m \text{ error matrix, } \|f_j\| = O(\epsilon_j)$$

How large is F_m allowed to be?

system:

$$\begin{aligned} r_m &= b - AV_m y_m = b - V_{m+1}\underline{H}_m y_m - F_m y_m \\ &= \underbrace{V_{m+1}(e_1\beta - \underline{H}_m y_m)}_{\text{computed residual} =: \tilde{r}_m} - F_m y_m \end{aligned}$$

eigenproblem: $\quad (\theta, V_m y)$

$$r_m = \theta V_m y - AV_m y = v_{m+1} h_{m+1,m} e_m^T y - F_m y$$

The inexact key relation

$$\mathcal{A} = A \quad \rightarrow \quad \mathcal{A}_\epsilon \approx A$$

e.g., $\mathcal{A}_\epsilon v := \mathcal{A}v + w, \qquad \|w\| = \epsilon$

$$AV_m = V_{m+1}\underline{H}_m + \underbrace{F_m}_{[f_1, f_2, \ldots, f_m]} \qquad F_m \text{ error matrix, } \|f_j\| = O(\epsilon_j)$$

How large is F_m allowed to be?

system:

$$\begin{aligned} r_m &= b - AV_m y_m = b - V_{m+1}\underline{H}_m y_m - F_m y_m \\ &= \underbrace{V_{m+1}(e_1\beta - \underline{H}_m y_m)}_{\text{computed residual} =: \tilde{r}_m} - F_m y_m \end{aligned}$$

eigenproblem: $(\theta, V_m y)$

$$r_m = \theta V_m y - AV_m y = v_{m+1} h_{m+1,m} e_m^T y - F_m y$$

A dynamic setting

$$F_m y = [f_1, f_2, \ldots, f_m] \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} = \sum_{i=1}^{m} f_i \eta_i$$

◇ The terms $f_i \eta_i$ need to be small:

$$\|f_i \eta_i\| < \frac{1}{m}\epsilon \quad \forall i \quad \Rightarrow \quad \|F_m y\| < \epsilon$$

◇ If η_i small $\quad \Rightarrow \quad$ f_i is allowed to be large

Linear systems: The solution pattern

$y_m = [\eta_1; \eta_2; \ldots; \eta_m]$ depends on the chosen method, e.g.

- Petrov-Galerkin (e.g. GMRES): $y_m = \mathrm{argmin}_y \|e_1\beta - \underline{H}_m y\|$,

$$|\eta_i| \le \frac{1}{\sigma_{\min}(\underline{H}_m)} \|\tilde{r}_{i-1}\|$$

$\tilde{r}_{i-1}$: GMRES computed residual at iteration $i-1$.

Simoncini & Szyld, '03 (see also Sleijpen & van den Eshof, '04, Bouras-Frayssé '05)

Analogous result for Galerkin methods (e.g. FOM)

Relaxing the inexactness in A

$A \cdot v_i$ not performed exactly $\quad \Rightarrow \quad (A + E_i) \cdot v_i$

True (unobservable) vs. computed residuals:

$$r_m = b - AV_m y_m = V_{m+1}(e_1\beta - \underline{H}_m y_m) - F_m y_m$$

GMRES: If (Similar result for FOM)

$$\|E_i\| \le \frac{\sigma_{\min}(\underline{H}_m)}{m} \frac{1}{\|\tilde{r}_{i-1}\|}\varepsilon \quad i = 1, \ldots, m$$

then $\|F_m y_m\| \le \varepsilon \quad \Rightarrow \quad \|r_m - V_{m+1}(e_1\beta - \underline{H}_m y_m)\| \le \varepsilon$

$\tilde{r}_{i-1}$: GMRES computed residual at iteration $i - 1$

An example: Schur complement

$$\underbrace{B^T S^{-1} B}_{A} x = b \qquad\qquad y_i \leftarrow B^T S^{-1} B v_i$$

Inexact matrix-vector product:

$$\begin{cases} \text{Solve } S w_i = B v_i \\ \text{Compute } y_i = B^T w_i \end{cases} \quad \overset{\text{Inexact}}{\Rightarrow} \quad \begin{cases} \text{Approx solve } S w_i = B v_i & \Rightarrow \widehat{w}_i \\ \text{Compute } \widehat{y}_i = B^T \widehat{w}_i \end{cases}$$

$w_i = \widehat{w}_i + \epsilon_i$ $\qquad$ ϵ_i error in inner solution $\qquad$ so that

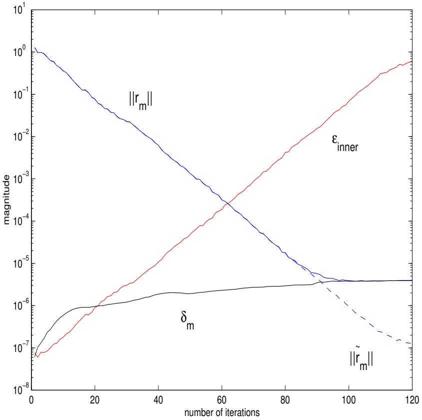
$$A v_i \quad \rightarrow \quad B^T \widehat{w}_i = \underbrace{B^T w_i}_{A v_i} - \underbrace{B^T \epsilon_i}_{-E_i v_i} = (A + E_i) v_i$$

Numerical experiment

$$\underbrace{B^T S^{-1} B}_{A} x = b \qquad \text{at each it. } i \text{ solve} \quad Sw_i = Bv_i$$

Inexact FOM

$\delta_m = \|r_m - (b - V_{m+1}\underline{H}_m y_m)\|$



Back to the inexact key relation

$$AV_m = V_{m+1}\underline{H}_m + \underbrace{F_m}_{[f_1,f_2,\ldots,f_m]} \qquad F_m \text{ error matrix, } \|f_j\| = O(\epsilon_j)$$

$$F_m y = [f_1, f_2, \ldots, f_m]\begin{bmatrix}\eta_1\\ \eta_2\\ \vdots \\ \eta_m\end{bmatrix} = \sum_{i=1}^{m} f_i\eta_i$$

◇ The terms $f_i\eta_i$ need to be small:

$$\|f_i\eta_i\| < \frac{1}{m}\epsilon \quad \forall i \quad \Rightarrow \quad \|F_m y\| < \epsilon$$

◇ If η_i small $\quad\Rightarrow\quad$ f_i is allowed to be large and the "residual" remains unaltered

This applies to any problem/method involving a component-wise decaying y in the residual norm

Back to the inexact key relation

$$AV_m = V_{m+1}\underline{H}_m + \underbrace{F_m}_{[f_1, f_2, \ldots, f_m]} \qquad F_m \text{ error matrix}, \ \|f_j\| = O(\epsilon_j)$$

$$F_m y = [f_1, f_2, \ldots, f_m] \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} = \sum_{i=1}^{m} f_i \eta_i$$

◇ The terms $f_i \eta_i$ need to be small:

$$\|f_i \eta_i\| < \frac{1}{m}\epsilon \quad \forall i \quad \Rightarrow \quad \|F_m y\| < \epsilon$$

◇ If η_i small $\quad \Rightarrow \quad$ f_i is allowed to be large and the "residual" remains unaltered

This applies to any problem/method involving a component-wise decaying y in the residual norm

Approximating the evaluation of a matrix function

Given $V_m \in \mathbb{R}^{n\times m}$ whose columns are an orthogonal basis of some approximation space, $0 \neq t \in \mathbb{R}$,

$$f(tA)v \approx \mathbf{y}_m := V_m f(tH_m)e_1, \qquad \text{with} \quad H_m = V_m^\top A V_m, v = V_m e_1$$

"Residual" evaluation:

$$r_m(t) := |h_{m+1,m}\mathbf{e}_m^T e^{-tH_m}\mathbf{e}_1|, \qquad h_{m+1,m} = v_{m+1}^\top A V_m$$

If $y(t) = f(tA)v$ is the solution to the differential equation $y^{(d)} = Ay$ for some derivative d, then

$$\mathbf{r}_m(t) = A\mathbf{y}_m - \mathbf{y}_m^{(d)} = AV_m f(tH_m)\mathbf{e}_1 - \mathbf{y}_m^{(d)} = \ldots = \mathbf{v}_{m+1} h_{m+1,m}\mathbf{e}_m^T f(tH_m)\mathbf{e}_1$$

Evaluation of a matrix function. The inexact context.

$$AV_m = V_{m+1}\underline{H}_m + F_m, \qquad F_m = \mathcal{E}_m V_m$$

$$\begin{aligned}
\mathbf{r}_m &= A\mathbf{y}_m - \mathbf{y}_m^{(d)} = AV_m f(H_m)\mathbf{e}_1 - \mathbf{y}_m^{(d)} \\
&= -\mathcal{E}_m V_m f(H_m)\mathbf{e}_1 + V_m H_m f(H_m)\mathbf{e}_1 - \mathbf{y}_m^{(d)} + \mathbf{v}_{m+1} h_{m+1,m}\mathbf{e}_m^T f(H_m)\mathbf{e}_1 \\
&= -F_m f(H_m)\mathbf{e}_1 + \mathbf{v}_{m+1} h_{m+1,m}\mathbf{e}_m^T f(H_m)\mathbf{e}_1.
\end{aligned}$$

♣ The quantity $\|\mathbf{r}_m\|$ is not available! (A is not known), whereas $r(t) = |h_{m+1,m}\mathbf{e}_m^T e^{-tH_m}\mathbf{e}_1|$ computable

Distance between exact and computable residuals: for $F_m = [\mathbf{f}_1, \ldots, \mathbf{f}_m]$,

$$|\|\mathbf{r}_m\| - r_m| \le \|[\mathbf{f}_1, \ldots, \mathbf{f}_m] f(H_m)\mathbf{e}_1\| \le \sum_{j=1}^{m} \|\mathbf{f}_j\| \, |\mathbf{e}_j^T f(H_m)\mathbf{e}_1|$$

Proof of element-wise decay of $f(H_m)\mathbf{e}_1$ in Pozza-Simoncini, BIT '19

Evaluation of a matrix function. The inexact context.

$$AV_m = V_{m+1}\underline{H}_m + F_m, \qquad F_m = \mathcal{E}_m V_m$$

$$\begin{aligned}
\mathbf{r}_m &= A\mathbf{y}_m - \mathbf{y}_m^{(d)} = AV_m f(H_m)\mathbf{e}_1 - \mathbf{y}_m^{(d)} \\
&= -\mathcal{E}_m V_m f(H_m)\mathbf{e}_1 + V_m H_m f(H_m)\mathbf{e}_1 - \mathbf{y}_m^{(d)} + \mathbf{v}_{m+1} h_{m+1,m}\mathbf{e}_m^T f(H_m)\mathbf{e}_1 \\
&= -F_m f(H_m)\mathbf{e}_1 + \mathbf{v}_{m+1} h_{m+1,m}\mathbf{e}_m^T f(H_m)\mathbf{e}_1.
\end{aligned}$$

♣ The quantity $\|\mathbf{r}_m\|$ is not available! (A is not known), whereas $r(t) = |h_{m+1,m}\mathbf{e}_m^T e^{-tH_m}\mathbf{e}_1|$ computable

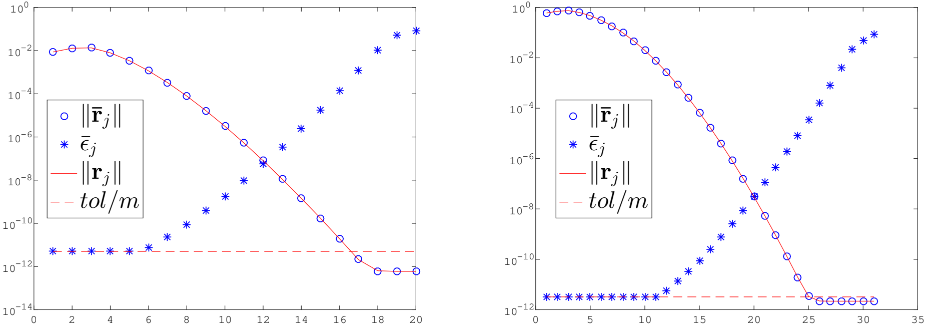
Distance between exact and computable residuals: for $F_m = [\mathbf{f}_1, \ldots, \mathbf{f}_m]$,

$$|\|\mathbf{r}_m\| - r_m| \le \|[\mathbf{f}_1, \ldots, \mathbf{f}_m] f(H_m)\mathbf{e}_1\| \le \sum_{j=1}^{m} \|\mathbf{f}_j\| \, |\mathbf{e}_j^T f(H_m)\mathbf{e}_1|$$

Proof of element-wise decay of $f(H_m)\mathbf{e}_1$ in Pozza-Simoncini, BIT '19

An example

Approximation of $e^{-A}\mathbf{v}$ with $\mathbf{v} = 1$ (normalized)



* Residual norm $\|\mathbf{r}_j\|$ with constant accuracy $\epsilon_j = tol/m$,
* residual norm $\|\bar{\mathbf{r}}_j\|$ with a variable strategy for the perturbation $\bar{\epsilon}_j$ as the inexact Arnoldi method proceeds

Left: For $A = \mathrm{Toeplitz}(1, \underline{2}, 0.1, -1)$

Right: For matrix pde225 from the Matrix Market repository

Lyapunov equation (and Sylvester equation)

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

Projection-type methods

Given a 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:

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

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

Lyapunov equation (and Sylvester equation)

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

Projection-type methods

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

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

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

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

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

Projected Lyapunov equation:

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

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

Lyapunov equation (and Sylvester equation)

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

Projection-type methods

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

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

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

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

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

Projected Lyapunov equation:

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

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

Residual and solution decay

$$\begin{aligned}
\|R\| &= \|AV_mYV_m^\top + V_mYV_m^\top A^\top - V_me_1e_1^\top V_m^\top\| \\
&= \|V_mT_mYV_m^\top + V_mYT^\top V_m^\top - V_me_1e_1^\top V_m^\top \\
&\qquad + v_{m+1}t_{m+1}e_m^\top YV_m^\top + V_mYe_mt_{m+1}v_{m+1}^\top\| \\
&= \|V_{m+1}\begin{bmatrix} T_mY + V_mYT^\top - e_1\|b\|^2e_1^\top V_m^\top & t_{m+1}e_m^\top Y \\ Ye_mt_{m+1} & 0 \end{bmatrix} V_{m+1}^\top\| \\
&= \|\begin{bmatrix} 0 & t_{m+1}e_m^\top Y \\ Ye_mt_{m+1} & 0 \end{bmatrix}\| \qquad B = b, \|b\| = 1
\end{aligned}$$

It is sufficient to show that $Y_{i,j} \to 0$ as i, j grow.

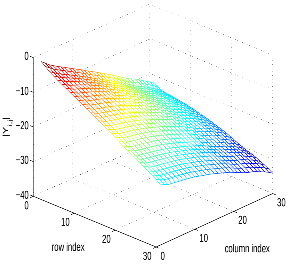
Residual and solution decay

$$
\begin{aligned}
\|R\| &= \|AV_mYV_m^\top + V_mYV_m^\top A^\top - V_m e_1 e_1^\top V_m^\top\| \\
&= \|V_mT_mYV_m^\top + V_mYT^\top V_m^\top - V_m e_1 e_1^\top V_m^\top \\
&\qquad + v_{m+1}t_{m+1}e_m^\top YV_m^\top + V_mYe_mt_{m+1}v_{m+1}^\top\| \\
&= \|V_{m+1}\begin{bmatrix} T_mY + V_mYT^\top - e_1\|b\|^2 e_1^\top V_m^\top & t_{m+1}e_m^\top Y \\ Ye_mt_{m+1} & 0 \end{bmatrix} V_{m+1}^\top\| \\
&= \|\begin{bmatrix} 0 & t_{m+1}e_m^\top Y \\ Ye_mt_{m+1} & 0 \end{bmatrix}\| \qquad B = b, \|b\| = 1
\end{aligned}
$$

It is sufficient to show that $Y_{i,j} \to 0$ as i, j grow.

Inexact computations

Typical decay pattern of Y:



$$AV_m = V_{m+1}\underline{H}_m + \underbrace{F_m}_{[f_1, f_2, \ldots, f_m]} \qquad F_m \text{ error matrix, } \|f_j\| = O(\epsilon_j)$$

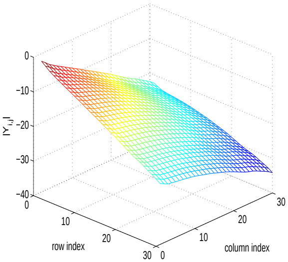
$$\begin{aligned}\|R\| &= \|AV_mYV_m^\top + V_mYV_m^\top A^\top - V_me_1e_1^\top V_m^\top + F_mYV_m^\top + V_mYF_m^\top\| \\ &= \left\|V_{m+1}\begin{bmatrix} 0 & t_{m+1}e_m^\top Y \\ Ye_mt_{m+1} & 0\end{bmatrix}V_{m+1}^\top + F_mYV_m^\top + V_mYF_m^\top\right\|\end{aligned}$$

Proofs of element-wise decay in Y:

- Standard Krylov (Simoncini '15)
- Rational Krylov (Pozza-Simoncini '19, see also Freitag-Kürschner '20)

Inexact computations

Typical decay pattern of Y:



$$AV_m = V_{m+1}\underline{H}_m + \underbrace{F_m}_{[f_1,f_2,\ldots,f_m]} \qquad F_m \text{ error matrix, } \|f_j\| = O(\epsilon_j)$$

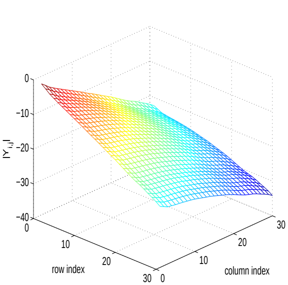
$$\begin{aligned} \|R\| &= \|AV_mYV_m^\top + V_mYV_m^\top A^\top - V_me_1e_1^\top V_m^\top + F_mYV_m^\top + V_mYF_m^\top\| \\ &= \left\|V_{m+1}\begin{bmatrix} 0 & t_{m+1}e_m^\top Y \\ Ye_mt_{m+1} & 0 \end{bmatrix} V_{m+1}^\top + F_mYV_m^\top + V_mYF_m^\top\right\| \end{aligned}$$

Proofs of element-wise decay in Y:

- Standard Krylov (Simoncini '15)
- Rational Krylov (Pozza-Simoncini '19, see also Freitag-Kürschner '20)

Inexact computations

Typical decay pattern of Y:



$$AV_m = V_{m+1}\underline{H}_m + \underbrace{F_m}_{[f_1, f_2, \ldots, f_m]} \qquad F_m \text{ error matrix, } \|f_j\| = O(\epsilon_j)$$

$$\begin{aligned} \|R\| &= \|AV_mYV_m^\top + V_mYV_m^\top A^\top - V_me_1e_1^\top V_m^\top + F_mYV_m^\top + V_mYF_m^\top\| \\ &= \left\|V_{m+1}\begin{bmatrix} 0 & t_{m+1}e_m^\top Y \\ Ye_mt_{m+1} & 0 \end{bmatrix} V_{m+1}^\top + F_mYV_m^\top + V_mYF_m^\top\right\| \end{aligned}$$

Proofs of element-wise decay in Y:

- Standard Krylov (Simoncini '15)
- Rational Krylov (Pozza-Simoncini '19, see also Freitag-Kürschner '20)

Multiterm linear matrix equation

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

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

Possibly large dimensions, structured 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 available for the solution of such problems.*
>
> *Peter Lancaster, SIAM Rev. 1970*

Multiterm linear matrix equation. Classical device

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

Kronecker formulation $\boxed{\left(B_1^\top \otimes A_1 + \ldots + B_\ell^\top \otimes A_\ell\right) \boldsymbol{x} = c \;\Leftrightarrow\; \mathcal{A}\boldsymbol{x} = c}$

Iterative methods: matrix-matrix multiplications and rank truncation

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

$$\text{Kronecker product}: \quad M \otimes P = \begin{bmatrix} m_{11}P & \ldots & m_{1n}P \\ \vdots & \ddots & \vdots \\ m_{n1}P & \ldots & m_{nn}P \end{bmatrix} \text{ and } \mathrm{vec}(AXB) = (B^\top \otimes A)\mathrm{vec}(X)$$

Alternatives to Kronecker form:

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

Current very active area of research

Multiterm linear matrix equation. Classical device

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

Kronecker formulation $$\boxed{\left(B_1^\top \otimes A_1 + \ldots + B_\ell^\top \otimes A_\ell\right) \boldsymbol{x} = c \;\Leftrightarrow\; \mathcal{A}\boldsymbol{x} = c}$$

Iterative methods: matrix-matrix multiplications and rank truncation

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

$$\text{Kronecker product}: \quad M \otimes P = \begin{bmatrix} m_{11}P & \ldots & m_{1n}P \\ \vdots & \ddots & \vdots \\ m_{n1}P & \ldots & m_{nn}P \end{bmatrix} \text{ and } \mathrm{vec}(AXB) = (B^\top \otimes A)\mathrm{vec}(X)$$

Alternatives to Kronecker form:

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

Current very active area of research

Truncated matrix-oriented CG (TCG) for Kronecker form

Input: $\mathcal{A}(\boldsymbol{X}) = A_1 \boldsymbol{X} B_1 + A_2 \boldsymbol{X} B_2 + \ldots + A_\ell \boldsymbol{X} B_\ell$, right-hand side $C \in \mathbb{R}^{n\times n}$ in low-rank format. Truncation operator $\mathcal{T}$.
Output: Matrix $X \in \mathbb{R}^{n\times n}$ in low-rank format s.t. $||\mathcal{A}(X) - C||_F/||C||_F \le tol$

$X_0 = 0,\ R_0 = C,\ P_0 = R_0,\ Q_0 = \mathcal{A}(P_0)$
$\xi_0 = \langle P_0, Q_0\rangle,\ k = 0$ $\qquad \langle X, Y\rangle = \mathrm{tr}(X^\top Y)$
While $||R_k||_F > tol$
  $\omega_k = \langle R_k, P_k\rangle/\xi_k$
  $X_{k+1} = X_k + \omega_k P_k,$ $\qquad X_{k+1} \leftarrow \mathcal{T}(X_{k+1})$
  $R_{k+1} = C - \mathcal{A}(X_{k+1}),$ $\qquad$ Optionally: $R_{k+1} \leftarrow \mathcal{T}(R_{k+1})$
  $\beta_k = -\langle R_{k+1}, Q_k\rangle/\xi_k$
  $P_{k+1} = R_{k+1} + \beta_k P_k,$ $\qquad P_{k+1} \leftarrow \mathcal{T}(P_{k+1})$
  $Q_{k+1} = \mathcal{A}(P_{k+1}),$ $\qquad$ Optionally: $Q_{k+1} \leftarrow \mathcal{T}(Q_{k+1})$
  $\xi_{k+1} = \langle P_{k+1}, Q_{k+1}\rangle$
  $k = k+1$
end while

♣ Iterates kept in factored form! $\qquad$ Kressner and Tobler, 2011
$\mathcal{T}(X_{k+1})$ acts on the SVD of X_{k+1}:
If X_k and P_k in factored form, then SVD on the augmented factor

Truncated matrix-oriented CG (TCG) for Kronecker form

Input: $\mathcal{A}(\boldsymbol{X}) = A_1\boldsymbol{X}B_1 + A_2\boldsymbol{X}B_2 + \ldots + A_\ell\boldsymbol{X}B_\ell$, right-hand side $C \in \mathbb{R}^{n\times n}$ in low-rank format.
Truncation operator $\mathcal{T}$.
Output: Matrix $X \in \mathbb{R}^{n\times n}$ in low-rank format s.t. $||\mathcal{A}(X) - C||_F/||C||_F \leq tol$

$X_0 = 0,\ R_0 = C,\ P_0 = R_0,\ Q_0 = \mathcal{A}(P_0)$

$\xi_0 = \langle P_0, Q_0\rangle,\ k = 0$ $\qquad\qquad \langle X, Y\rangle = \mathrm{tr}(X^\top Y)$

While $||R_k||_F > tol$

  $\omega_k = \langle R_k, P_k\rangle/\xi_k$

  $X_{k+1} = X_k + \omega_k P_k,$ $\qquad X_{k+1} \leftarrow \mathcal{T}(X_{k+1})$

  $R_{k+1} = C - \mathcal{A}(X_{k+1}),$ $\qquad$ Optionally: $R_{k+1} \leftarrow \mathcal{T}(R_{k+1})$

  $\beta_k = -\langle R_{k+1}, Q_k\rangle/\xi_k$

  $P_{k+1} = R_{k+1} + \beta_k P_k,$ $\qquad P_{k+1} \leftarrow \mathcal{T}(P_{k+1})$

  $Q_{k+1} = \mathcal{A}(P_{k+1}),$ $\qquad$ Optionally: $Q_{k+1} \leftarrow \mathcal{T}(Q_{k+1})$

  $\xi_{k+1} = \langle P_{k+1}, Q_{k+1}\rangle$

  $k = k + 1$

end while

♣ Iterates kept in factored form! $\qquad$ Kressner and Tobler, 2011
$\mathcal{T}(X_{k+1})$ acts on the SVD of X_{k+1}:
If X_k and P_k in factored form, then SVD on the augmented factor

Effect of truncation

Let $x_k = \mathrm{vec}(X_k)$ (and similarly for the other variables). Truncation can be written as

$$x^{(k+1)} = x_{ex}^{(k+1)} + \boldsymbol{e}_X^{(k+1)}, \qquad p^{(k+1)} = p_{ex}^{(k+1)} + \boldsymbol{e}_P^{(k+1)}$$

($\boldsymbol{e}_X^{(k+1)}, \boldsymbol{e}_P^{(k+1)}$ local truncation errors)

TH: Let $\Delta_k = \max\{\|\boldsymbol{e}_P^{(k)}\|, \|\boldsymbol{e}_X^{(k)}\|, \|\boldsymbol{e}_P^{(k+1)}\|, \|\boldsymbol{e}_X^{(k+1)}\|\}$ and also $\delta_k = \min\{\|\boldsymbol{e}_P^{(k)}\|, \|\boldsymbol{e}_X^{(k)}\|, \|\boldsymbol{e}_P^{(k+1)}\|, \|\boldsymbol{e}_X^{(k+1)}\|\}$. Then there exists $\eta \in [0,1]$ such that

$$\eta \frac{1}{\|\mathcal{A}^{-1}\|} \frac{\delta_k}{\|r^{(k+1)}\|} \le \frac{|r^{(k+1)})^\top p^{(k)}|}{\|r^{(k+1)}\|\|p^{(k)}\|} \le \|\mathcal{A}\| \frac{\Delta_k}{\|r^{(k+1)}\|},$$

and

$$\beta_k = -\frac{(r_{ex}^{(k+1)})^\top \mathcal{A} p^{(k)} - (\mathcal{A}\boldsymbol{e}_X^{(k+1)})^\top \mathcal{A} p^{(k)}}{(p^{(k)})^\top \mathcal{A} p^{(k)}}.$$

Moreover,

$$\frac{|r^{(k+1)})^\top r^{(k)}|}{\|r^{(k+1)}\|\|r^{(k)}\|} \le \gamma \frac{\Delta_k}{\|r^{(k+1)}\|} \qquad \gamma = \|\mathcal{A}p^{(k)}\| + (2|\beta_{k-1}| + |\beta_{k-1}\alpha_k|)\|\mathcal{A}p^{(k-1)}\| + \|r^{(k+1)}\|$$

Effect of truncation

Let $x_k = \mathrm{vec}(X_k)$ (and similarly for the other variables). Truncation can be written as

$$x^{(k+1)} = x_{ex}^{(k+1)} + \boldsymbol{e}_X^{(k+1)}, \qquad p^{(k+1)} = p_{ex}^{(k+1)} + \boldsymbol{e}_P^{(k+1)}$$

($\boldsymbol{e}_X^{(k+1)}, \boldsymbol{e}_P^{(k+1)}$ local truncation errors)

TH: Let $\Delta_k = \max\{\|\boldsymbol{e}_P^{(k)}\|, \|\boldsymbol{e}_X^{(k)}\|, \|\boldsymbol{e}_P^{(k+1)}\|, \|\boldsymbol{e}_X^{(k+1)}\|\}$ and also $\delta_k = \min\{\|\boldsymbol{e}_P^{(k)}\|, \|\boldsymbol{e}_X^{(k)}\|, \|\boldsymbol{e}_P^{(k+1)}\|, \|\boldsymbol{e}_X^{(k+1)}\|\}$. Then there exists $\eta \in [0, 1]$ such that

$$\eta \frac{1}{\|\mathcal{A}^{-1}\|} \frac{\delta_k}{\|r^{(k+1)}\|} \le \frac{|(r^{(k+1)})^\top p^{(k)}|}{\|r^{(k+1)}\|\|p^{(k)}\|} \le \|\mathcal{A}\| \frac{\Delta_k}{\|r^{(k+1)}\|},$$

and

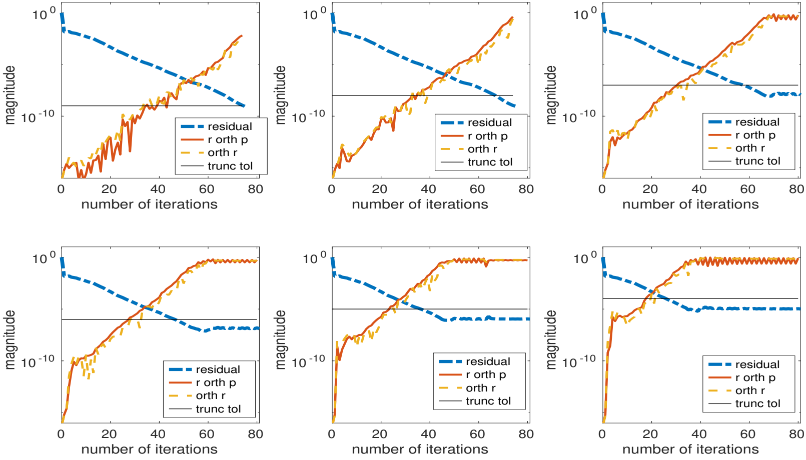
$$\beta_k = -\frac{(r_{ex}^{(k+1)})^\top \mathcal{A} p^{(k)} - (\mathcal{A}\boldsymbol{e}_X^{(k+1)})^\top \mathcal{A} p^{(k)}}{(p^{(k)})^\top \mathcal{A} p^{(k)}}.$$

Moreover,

$$\frac{|(r^{(k+1)})^\top r^{(k)}|}{\|r^{(k+1)}\|\|r^{(k)}\|} \le \gamma \frac{\Delta_k}{\|r^{(k+1)}\|} \qquad \gamma = \|\mathcal{A} p^{(k)}\| + (2|\beta_{k-1}| + |\beta_{k-1}\alpha_k|)\|\mathcal{A} p^{(k-1)}\| + \|r^{(k+1)}\|)/$$

An example: $AX + XA + MXM = c_1 c_1^\top$

A: 2D Laplace operator, M =pentadiag$(-0.5, -1, 3.2, -1, -0.5)$, c_1 random entries
Truncated CG residual norm (blue line) for different truncation values



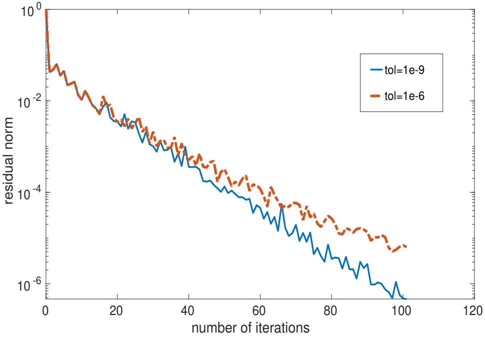
Also reported: Loss of orthogonality (cosine of the angles) between consecutive residuals and residual and directions

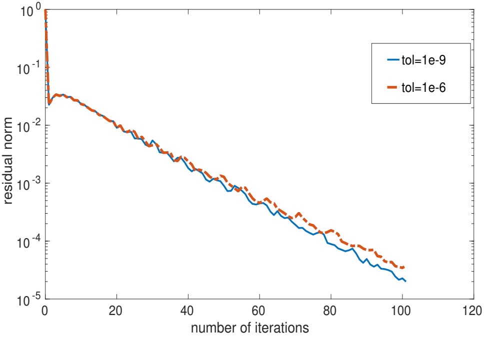
Another example

$A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_i = \lambda_1 + \frac{(i-1)}{(n-1)}(\lambda_n - \lambda_1)\rho^{n-i}$, $\lambda_1 = 0.1, \quad \lambda_n = 100$

M: diagonal matrix with elements logarithmically distributed in $[10^{-2}, 10^{0}]$

Convergence history of TCG for two truncation tolerances:





Left: $\rho = 0.4$

Right: $\rho = 0.8$

Conclusions

- Krylov-based approaches are very flexible
- Relaxation properties are usually not problem dependent
- Relaxation properties arise in disguise
- Extremely useful for practical purposes

Visit: www.dm.unibo.it/~simoncin

Email address: valeria.simoncini@unibo.it