MINRES Residual Norms of Diagonally Translated Linear Systems

Jurjen Duintjer Tebbens,

jointly with **Zdeněk Strakoš**

Institute of Computer Science

Academy of Sciences of the Czech Republic

MAT-TRIAD 2005, Będlewo, March 4, 2005

Outline

1. Motivation
2. An equality for MINRES residual vectors
3. An equality for MINRES residual vectors for diagonally translated systems
4. The equality in practice

1. Motivation

- Solution of large sparse symmetric indefinite systems remains a challenging task.
- Solution of large sparse symmetric positive definite systems can be done efficiently with the CG method

Rhetorical question: Can we transform solution of an indefinite system in solution of positive definite systems ?

Example: Diagonal translation. We consider a linear system

$$\mathbf{A}x = b$$

- $\mathbf{A}$ is nonsingular, symmetric and indefinite
- $\lambda_1 < 0$ is the smallest eigenvalue of $\mathbf{A}$
- Consider the positive definite system $(\mathbf{A} - \lambda\mathbf{I})z = b$, where $\mathbf{I}$ is the identity matrix and $\lambda < \lambda_1$
- Problem: If we have found the solution z of the translated system, the solution of the original system is $x = \left(\mathbf{I} + \lambda(\mathbf{A} - \lambda\mathbf{I})^{-1}\right)^{-1} z$

Computation of iterates from iterates of a translated system is not feasible. But we can formulate a concise relationship between the *residuals*.

2. An equality for MINRES residual vectors

The MINRES method [Paige, Saunders - 1975]: Starting with an initial guess x_0, iterates x_k minimize the residual norm $\|r_k\| = \|b - \mathbf{A}x_k\|$ according to

$$\|r_k\| = \min_{s \in \mathbf{A}\mathcal{K}_k(\mathbf{A}, r_0)} \|r_0 - s\|,$$

where $r_0 = b - \mathbf{A}x_0$ and

$$\mathcal{K}_k(\mathbf{A}, r_0) \equiv \mathrm{span}\{r_0, \mathbf{A}r_0, \ldots, \mathbf{A}^{k-1}r_0\}.$$

Hence r_0 is projected onto $\mathbf{A}\mathcal{K}_k(\mathbf{A}, r_0)$,

$$r_k \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}, r_0).$$

Theorem: *Let* $[r_0, \mathbf{A}r_0, \ldots, \mathbf{A}^k r_0]$ *have full column rank. Then*

$$r_k^T = \|r_k\|^2 \, e_1^T [r_0, \mathbf{A}r_0, \ldots, \mathbf{A}^k r_0]^+, \tag{1}$$

where $[\mathbf{X}]^+$ *denotes the Moore-Penrose pseudoinverse of a matrix* $\mathbf{X}$.

First proof of this equality by Stewart, later also by [Ipsen - 2000], [Zítko - 2000], [Liesen, Rozložník, Strakoš - 2002]. The elegant proof in the latter article exploits

- $\mathbf{X} \cdot \mathbf{X}^+ = (\mathbf{X} \cdot \mathbf{X}^+)^T$
- $\mathbf{X}$ has full column rank $\Rightarrow \mathbf{X}^+ \cdot \mathbf{X} = \mathbf{I}$

Proof: $r_k = r_0 - [\mathbf{A}r_0, \dots, \mathbf{A}^k r_0]\, y_k,$ for some $y_k \in \mathbb{R}^k$. Hence

$$r_k = [r_0, \dots, \mathbf{A}^k r_0] \begin{pmatrix} 1 \\ -y_k \end{pmatrix} = [r_0, \dots, \mathbf{A}^k r_0][r_0, \dots, \mathbf{A}^k r_0]^+ [r_0, \dots, \mathbf{A}^k r_0] \begin{pmatrix} 1 \\ -y_k \end{pmatrix}$$

$$= \left([r_0, \dots, \mathbf{A}^k r_0][r_0, \dots, \mathbf{A}^k r_0]^+\right)^T r_k = ([r_0, \dots, \mathbf{A}^k r_0]^+)^T [r_0, \dots, \mathbf{A}^k r_0]^T r_k.$$

- $r_k \perp [\mathbf{A}r_0, \dots, \mathbf{A}^k r_0]$
- $r_0^T r_k = (r_k + [\mathbf{A}r_0, \dots, \mathbf{A}^k r_0]\, y_k)^T r_k$
- $[r_0, \dots, \mathbf{A}^k r_0]^T r_k = \|r_k\|^2 e_1$

$$\Rightarrow \quad r_k = ([r_0, \dots, \mathbf{A}^k r_0]^+)^T \|r_k\|^2 e_1. \quad \square$$

3. An equality for MINRES residual vectors for diagonally translated systems

An equality for the special case of splitting of tridiagonal Toeplitz matrices is given in [Liesen, Strakoš - 2004]:

$$\mathbf{T} = \begin{pmatrix} \lambda & \gamma & & & \\ \gamma & \lambda & \gamma & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma & \lambda & \gamma \\ & & & \gamma & \lambda \end{pmatrix} = \gamma\mathbf{S} + \gamma\mathbf{S}^T + \lambda\mathbf{I},$$

where $\mathbf{S}$ is the downshift matrix $\mathbf{S} = (e_2, \ldots, e_n, 0)$.

Theorem: *Residual vectors generated by the MINRES method applied to a system* $\mathbf{T}x = b$ *with a given initial residual* $r_0 = b - \mathbf{A}x_0$ *such that* $[r_0, \mathbf{T}r_0, \ldots, \mathbf{T}^k r_0]$ *has full column rank, satisfy*

$$r_k^T = \|r_k\|^2 \, [1, -\lambda, \ldots, (-\lambda)^k] \cdot \left[r_0, (\gamma\mathbf{S} + \gamma\mathbf{S}^T) r_0, \ldots, (\gamma\mathbf{S} + \gamma\mathbf{S}^T)^k r_0 \right]^+ .$$

Generalization for arbitrary diagonally translated systems:

Theorem: *Let* $\lambda \in \mathbb{R}$ *and let* $\mathbf{B} = \mathbf{A} - \lambda \mathbf{I}$. *Then the residual vectors generated by MINRES applied to* $\mathbf{A}x = b$ *with initial residual* r_0 *such that* $[r_0, \mathbf{A}r_0, \ldots, \mathbf{A}^k r_0]$ *has full column rank, satisfy*

$$r_k^T = \|r_k\|^2 \, [1, -\lambda, \ldots, (-\lambda)^k] \cdot \left[r_0, \mathbf{B}r_0, \ldots, \mathbf{B}^k r_0 \right]^+,$$

and

$$\|r_k\| = \frac{1}{\| \, [1, -\lambda, \ldots, (-\lambda)^k] \, [r_0, \mathbf{B}r_0, \ldots, \mathbf{B}^k r_0]^+ \|} \ .$$

Proof:

$$r_k^T = \|r_k\|^2 e_1^T [r_0, \mathbf{A} r_0, \ldots, \mathbf{A}^k r_0]^+ \equiv \|r_k\|^2 g_k^T.$$

$$\Rightarrow \quad g_k^T [r_0, \mathbf{A} r_0, \ldots, \mathbf{A}^k r_0] = e_1^T \qquad (2)$$

We can prove $g_k^T [r_0, \mathbf{B} r_0, \ldots, \mathbf{B}^k r_0] = [1, -\lambda, \ldots, (-\lambda)^k]$: (3)

- Second entry: $0 = g_k^T \mathbf{A} r_0 = g_k^T (\mathbf{B} + \lambda \mathbf{I}) r_0 = g_k^T \mathbf{B} r_0 + \lambda \Rightarrow g_k^T \mathbf{B} r_0 = -\lambda$
- Next entries: $0 = g_k^T \mathbf{A}^j r_0 = g_k^T (\mathbf{B} + \lambda \mathbf{I})^j r_0 = g_k^T \left(\sum_{i=0}^{j} \binom{j}{i} \lambda^i \mathbf{B}^{j-i} \right) r_0 \ldots \Rightarrow g_k^T \mathbf{B}^j r_0 = (-\lambda)^j$

Multiply (3) from the right with $[r_0, \mathbf{B} r_0, \ldots, \mathbf{B}^k r_0]^+$:

$$g_k^T = [1, -\lambda, \ldots, (-\lambda)^k][r_0, \mathbf{B} r_0, \ldots, \mathbf{B}^k r_0]^+. \quad \square$$

Corollary: *Let* $\mathbf{B} = \mathbf{A} - \lambda\mathbf{I}$ *and, in addition to the assumptions of the preceding theorem,* $\mathbf{B}$ *be nonsingular. Let* $r_k^{\mathbf{A}}$ *denote the MINRES residuals for* $\mathbf{A}x = b$ *and the initial residual* r_0 *and let* $r_k^{\mathbf{B}}$ *denote the MINRES residuals for* $\mathbf{B}y = c$*, with the same initial residual* r_0*. Then*

$$\frac{(r_k^{\mathbf{A}})^T}{\|r_k^{\mathbf{A}}\|^2} = \frac{(r_k^{\mathbf{B}})^T}{\|r_k^{\mathbf{B}}\|^2} + [0, -\lambda, \ldots, (-\lambda)^k] \cdot \left[r_0, \mathbf{B}r_0, \ldots, \mathbf{B}^k r_0\right]^+.$$

Proof: Subtract

$$\frac{(r_k^{\mathbf{A}})^T}{\|r_k^{\mathbf{A}}\|^2} = [1, -\lambda, \ldots, (-\lambda)^k] \cdot \left[r_0, \mathbf{B}r_0, \ldots, \mathbf{B}^k r_0\right]^+$$

and

$$\frac{(r_k^{\mathbf{B}})^T}{\|r_k^{\mathbf{B}}\|^2} = [1,\ 0\ ,\ \ldots\ ,\ 0\ ] \cdot \left[r_0, \mathbf{B}r_0, \ldots, \mathbf{B}^k r_0\right]^+. \quad \square$$

4. The equality in practice

1. The relation

$$\frac{(r_k^{\mathbf{A}})^T}{\|r_k^{\mathbf{A}}\|^2} = \frac{(r_k^{\mathbf{B}})^T}{\|r_k^{\mathbf{B}}\|^2} + [0, -\lambda, \ldots, (-\lambda)^k] \cdot \left[r_0, \mathbf{B}r_0, \ldots, \mathbf{B}^k r_0\right]^+$$

shows that the residuals are connected in a rather complicated way, in spite of equality of the Krylov subspaces

$$\mathcal{K}_k(\mathbf{A}, r_0) = \text{span}\{r_0, \mathbf{A}r_0, \ldots, \mathbf{A}^{k-1}r_0\}$$

$$= \text{span}\{r_0, (\mathbf{A} - \lambda\mathbf{I})r_0, \ldots, (\mathbf{A} - \lambda\mathbf{I})^{k-1}r_0\} = \mathcal{K}_k(\mathbf{B}, r_0).$$

But

$$\mathbf{A}\mathcal{K}_k(\mathbf{A}, r_0) \neq (\mathbf{A} - \lambda\mathbf{I})\mathcal{K}_k(\mathbf{A}, r_0) = \mathbf{B}\,\mathcal{K}_k(\mathbf{B}, r_0),$$

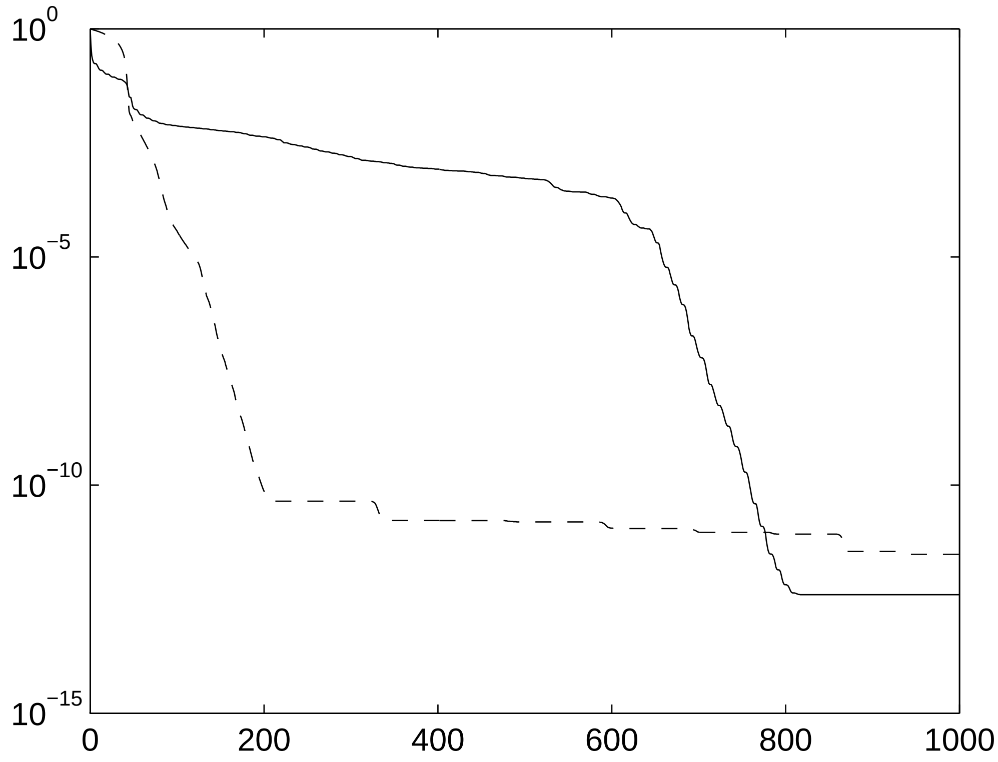
and the distance of r_0 to *these* spaces determines convergence speed.

Concerning norms we have

$$\|r_k^{\mathbf{A}}\| = \frac{\|r_k^{\mathbf{B}}\|^2}{\left\| r_k^{\mathbf{B}} + [0, -\lambda, \ldots, (-\lambda)^k] \cdot \left[r_0, \mathbf{B} r_0, \ldots, \mathbf{B}^k r_0 \right]^+ \right\|}.$$

Example: Shifted 2D Laplacian.
We choose $\mathbf{B} \equiv \mathbf{L}_{60}$, the 2D Laplacian of dimension 3600 from a 60×60 grid, and $\mathbf{A} \equiv \mathbf{L}_{60} - \frac{1}{10}\mathbf{I}$. Then $\lambda = -0.1$ and $[0, -\lambda, \ldots, (-\lambda)^k] \approx [0, 0.1, 0, \ldots, 0]$ but still the difference in convergence behavior is remarkable.

Laplacian and Shifted Laplacian (3600×3600)



Dashed line: $\frac{\|r_k^{\mathbf{L}_{60}}\|}{\|r_0\|}$ - Solid line: $\frac{\|r_k^{\mathbf{A}}\|}{\|r_0\|}$.

2. Derivation of residual bounds ? We have

$$\frac{\|r_k\|}{\|r_0\|} \leq \min_{p \in \pi_k} \max_{\lambda_i \in \sigma(\mathbf{A})} |p(\lambda_i)|,$$

where π_k denotes the polynomials of maximal degree k with the value 1 at the origin. In the positive definite case this gives the well-known bound

$$\frac{\|r_k\|}{\|r_0\|} \leq \min_{p \in \pi_k} \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |p(\lambda)| \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k, \quad \kappa = \frac{\lambda_{\max}}{\lambda_{\min}}.$$

In the indefinite case the interval $[\lambda_{\min}, \lambda_{\max}]$ contains zero and the bound becomes useless. Hence one has to divide the interval into

$$[\lambda_{\min}, \lambda_s] \cup [\lambda_{s+1}, \lambda_{\max}], \qquad \lambda_{\min} \leq \lambda_s < 0 < \lambda_{s+1} \leq \lambda_{\max}.$$

This makes it very difficult to handle the bound.

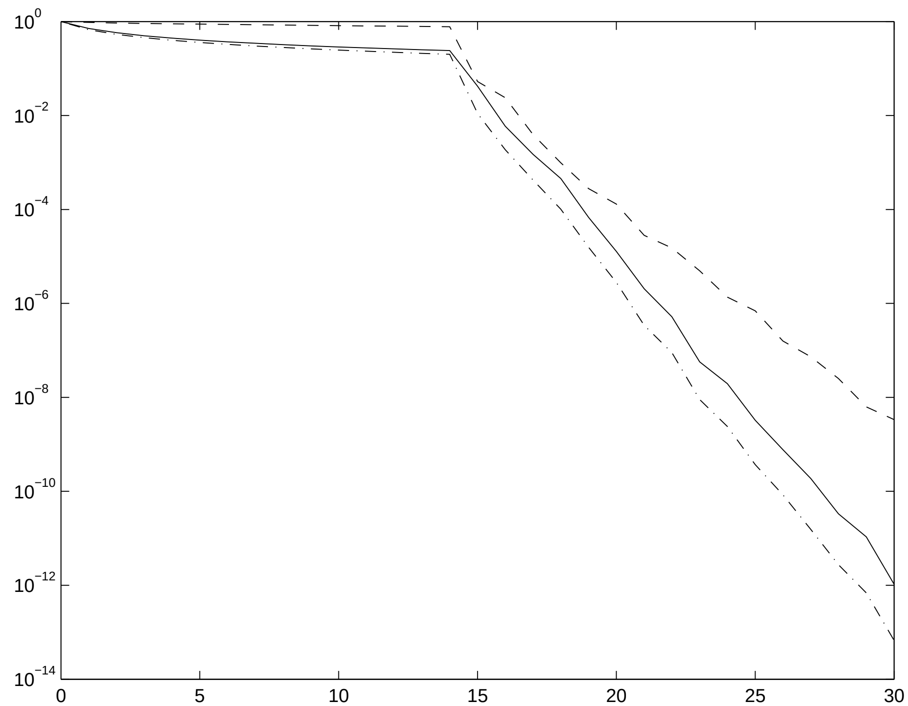
If we can choose $\mathbf{B} = \mathbf{A} - \lambda\mathbf{I}$ such that $\mathbf{B}$ is positive definite and $\|r_k^{\mathbf{B}} + [0, -\lambda, \ldots, (-\lambda)^k] \cdot \left[r_0, \mathbf{B}r_0, \ldots, \mathbf{B}^k r_0\right]^+\| \geq \|r_k^{\mathbf{B}}\|$, then

$$\|r_k^{\mathbf{A}}\| = \frac{\|r_k^{\mathbf{B}}\|^2}{\left\|r_k^{\mathbf{B}} + [0, -\lambda, \ldots, (-\lambda)^k] \cdot \left[r_0, \mathbf{B}r_0, \ldots, \mathbf{B}^k r_0\right]^+\right\|} \leq \|r_k^{\mathbf{B}}\|.$$

This should be worked out more in detail.

3. Our equality holds for any residual minimizing method. Originally: developed for the GMRES method to explain the following convergence behavior (solid line) for a convection diffusion model problem (see e.g. [Liesen, Strakoš - 2004]):

Discretized convection diffusion model problem



Solid line: GMRES residual norm reduction.

The system matrix **A** of dimension $15^2 \times 15^2$ is block tridiagonal with blocks of dimension 15×15:

$$\mathbf{A} = \begin{pmatrix} \lambda_1 \mathbf{I} & \mu_1 \mathbf{I} & & & \\ \gamma_1 \mathbf{I} & \lambda_2 \mathbf{I} & \mu_2 \mathbf{I} & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma_{14} \mathbf{I} & \lambda_{14} \mathbf{I} & \mu_{14} \mathbf{I} \\ & & & \gamma_{15} \mathbf{I} & \lambda_{15} \mathbf{I} \end{pmatrix},$$

where all $|\mu_i| << 1$ and λ_i are close.

We apply our theorem with $\mathbf{B} = \mathbf{A} - \lambda \mathbf{I}$ for a λ close to all λ_i. Then **B** is close to the matrix with only the blocks $\gamma_i \mathbf{I}$. Hence $\mathbf{B}^{15} r_0 \approx \mathbf{0}$ for any r_0. Now our equality

$$\|r_k\| = \frac{1}{\| [1, -\lambda, \ldots, (-\lambda)^k] \, [r_0, \mathbf{B} r_0, \ldots, \mathbf{B}^k r_0]^+ \|}$$

yields the bounds

$$\frac{\sigma_{\min}([r_0, \mathbf{B}r_0, \ldots, \mathbf{B}^k r_0])}{\sqrt{\sum_{j=0}^{k} |\lambda|^{2j}}} \le \|r_k\| \le \frac{\|\mathbf{B}^k r_0\|}{|\lambda|^k}$$

displayed above. For details we refer to a future article [DT, Greenbaum, Strakoš - 2005 ?].

4. From the proof it is clear that we can generalize as follows: **Theorem:** *Let us consider the residual vectors generated by MINRES applied to* $\mathbf{A}x = b$ *with an initial residual* r_0 *such that* $[r_0, \mathbf{A}r_0, \ldots, \mathbf{A}^k r_0]$ *has full column rank. If* $\mathbf{A} = \mathbf{B} + \mathbf{U}$ *with* $\mathbf{UB} = \mathbf{BU}$ *and* $\mathbf{U}r_0 = \lambda r_0$*, then*

$$r_k^T = \|r_k\|^2 \, [1, -\lambda, \ldots, (-\lambda)^k] \cdot \left[r_0, \mathbf{B}r_0, \ldots, \mathbf{B}^k r_0\right]^+ .$$

Thank you for your attention.

This work is supported by the Program Information Society under project 1ET400300415.