Behaviour of the Gauss-Radau upper bound

Petr Tichý

Charles University, Prague

joint work with

Gérard Meurant

SNA 2023, January 23-27, 2023, Ostrava

The conjugate gradient method

A is symmetric and positive definite, $Ax = b$

input A, b, x_0

$r_0 = b - Ax_0$, $p_0 = r_0$

for $k = 1, \ldots$ until convergence **do**

$$\begin{aligned} \gamma_{k-1} &= \frac{r_{k-1}^T r_{k-1}}{p_{k-1}^T A\, p_{k-1}} \\ x_k &= x_{k-1} + \gamma_{k-1} p_{k-1} \\ r_k &= r_{k-1} - \gamma_{k-1} A\, p_{k-1} \\ \delta_k &= \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}} \end{aligned}$$

$$p_k = r_k + \delta_k p_{k-1}$$

end for

$$\| x - x_k \|_A^2 = \min_{y \in x_0 + \mathcal{K}_k} \| x - y \|_A^2$$

How to measure quality of approximation?

... it depends on what problem we solve.

- **using residual information,**
  – normwise backward error,
  – relative residual norm.

  [Hestenes, Stiefel 1952]: "Using of the residual vector r_k as a measure of the "goodness" of the estimate x_k is not reliable"

- **using error estimates,**
  – estimate of the A-norm of the error,
  – estimate of the Euclidean norm of the error.

  [Hestenes, Stiefel 1952] : "The function $(x - x_k, A(x - x_k))$ can be used as a measure of the "goodness" of x_k as an estimate of x."

Estimating the A-norm of the error in CG

$$\|x - x_k\|_A^2$$

- An important role in **stopping criteria**:

  [Deuflhard 1994], [Arioli 2004],
  [Jiránek, Strakoš, Vohralík 2006], [Papež, Vohralík 2022]

- **Estimating errors** using quadrature approach:

  [Dahlquist, Golub, Nash 1978],
  [Golub, Meurant 1994, 1997], [Golub, Strakoš 1994],
  [Meurant 1997, 1999, 2005], [Calvetti, Morigi, Reichel, Sgallari, 2000, 2001],
  [Strakoš, T. 2002], [Meurant, T. 2013, 2019], [Meurant, Papež, T. 2021]

- Why it works in **finite precision** arithmetic?

  [Paige 1976, 1980, Greenbaum 1989],
  [Golub, Strakoš 1994], [Strakoš, T. 2002, 2005, 2011]

Quadrature bounds

Gauss quadrature (lower) bound

- It holds that

$$\gamma_k \|r_k\|^2 \; < \; \|x - x_k\|_A^2 \; \equiv \; \varepsilon_k \, .$$

- One can improve the lower bound at iteration $\ell \leq k$ using

$$\varepsilon_\ell \; = \; \underbrace{\sum_{j=\ell}^{k} \gamma_j \|r_j\|^2}_{\Delta_{\ell:k}} \; + \; \varepsilon_{k+1} \, .$$

[Golub, Strakoš 1994, Golub, Meurant 1997, Strakoš, T. 2002, 2005]

- How to choose $\ell \leq k$ such that

$$\frac{\varepsilon_\ell - \Delta_{\ell:k}}{\varepsilon_\ell} \; \leq \; \tau \, .$$

[Meurant, Papež, T. 2021]

Gauss-Radau (upper) bound

- Given $\mu \leq \lambda_{\min}$, it holds that

$$\|x - x_k\|_A^2 < \gamma_k^{(\mu)} \|r_k\|^2$$

where

$$\gamma_{k+1}^{(\mu)} = \frac{\left(\gamma_k^{(\mu)} - \gamma_k\right)}{\mu\left(\gamma_k^{(\mu)} - \gamma_k\right) + \delta_{k+1}}, \qquad \gamma_0^{(\mu)} = \frac{1}{\mu}.$$

[Meurant, T. 2013]

Practically relevant questions:

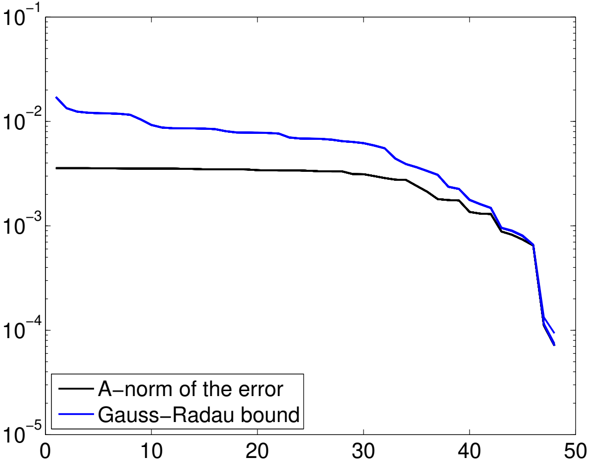
- How to get μ? [Gergelits, Mardal, Nielsen, Strakoš 2019, 2020, 2022] [Ladecký, Pultarová, Zeman 2021, 2021]
- **Quality** of the bound?
- Numerical **behavior**?

Behaviour of the upper bound

Upper bound in exact arithmetic

Gauss-Radau bound, bcsstk01 matrix, $n = 48$

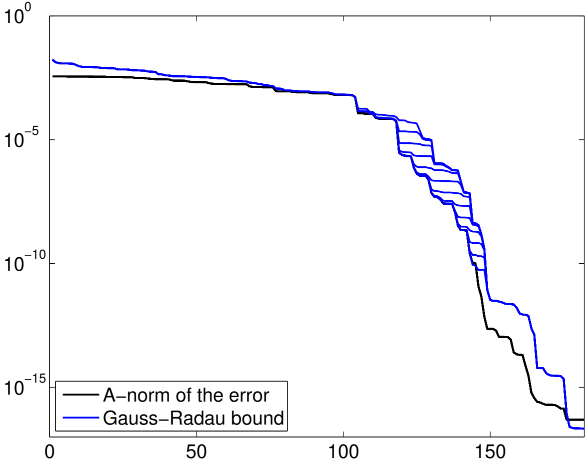
$$\mu = \frac{\lambda_{\min}}{1 + 10^{-m}}, \quad m = 2, \ldots, 14$$



Upper bound in finite precision arithmetic

Gauss-Radau bound, bcsstk01 matrix, $n = 48$

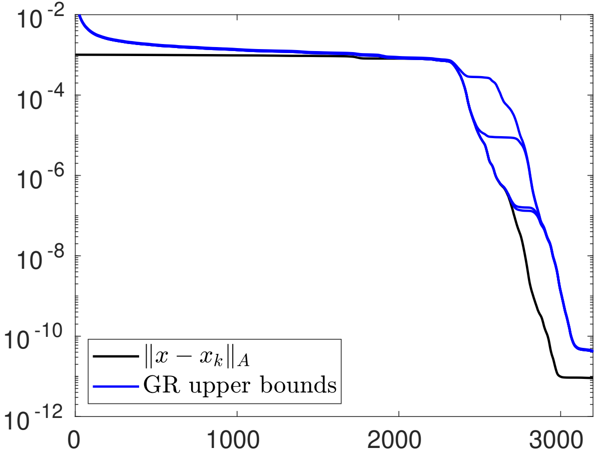
$$\mu = \frac{\lambda_{\min}}{1 + 10^{-m}}, \quad m = 2, \ldots, 14$$



Upper bound in finite precision arithmetic

Gauss-Radau bound, s3dkt3m2 matrix, $n = 90449$

$$\gamma_{k+1}^{(\mu)} = \frac{\left(\gamma_k^{(\mu)} - \gamma_k\right)}{\mu\left(\gamma_k^{(\mu)} - \gamma_k\right) + \delta_{k+1}}$$



Mathematical model

of finite precision CG computations

The results of **finite precision CG** can be interpreted (up to a small inaccuracy) as the results of **exact CG** applied to a larger problem with a matrix having **clustered eigenvalues** around λ_i's.

[Greenbaum 1989], [Greenbaum, Strakoš 1992], [Paige 1976, 1980]

$$A\,x \;=\; b \qquad\longleftrightarrow\qquad \hat{A}\,\hat{x} \;=\; \hat{b}$$

A model problem

[Meurant, T. 2023]

- **Consider**

$$A\,x \;=\; b\,,$$

$\|b\| = 1$, equal components, and $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$,

$$\lambda_i \;=\; \lambda_1 + \frac{i-1}{m-1}(\lambda_m - \lambda_1)\,\rho^{m-i},$$

see [Strakos 1991], **size** $m = 12$, $\rho = 0.8$, $\lambda_1 = 10^{-6}$, $\lambda_m = 1$.

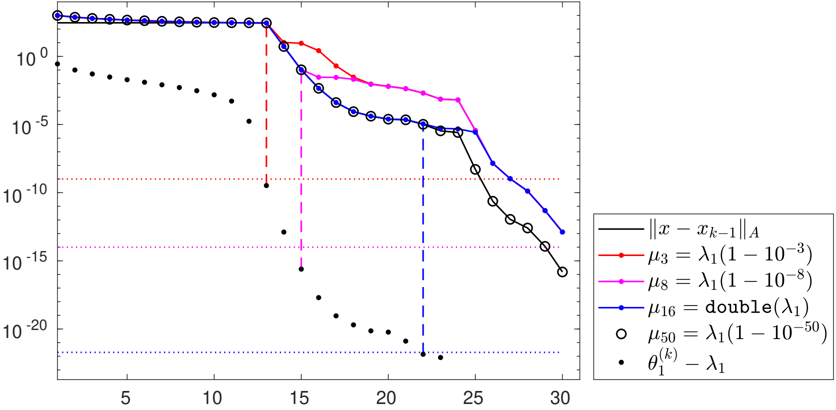
- **Blur** A and b resulting in $\hat{A}$ and $\hat{b}$ of **size** $N = 30$, solve

$$\hat{A}\,\hat{x} \;=\; \hat{b}$$

exactly $\rightarrow$ rename to $Ax = b$, A has eigenvalues λ_i.

Loss of accuracy of the Gauss-Radau upper bound

Current work [Meurant, T. 2023]



Analysis

Assumptions

1. λ_1 is **well separated** from λ_2
2. μ is a **tight underestimate** to λ_1

$$\lambda_1 - \mu \ll \lambda_2 - \lambda_1$$

3. $\theta_1^{(k)}$ **converges** to λ_1,

$$\theta_1^{(k)} - \lambda_1 \ll \lambda_1 - \mu$$

for some k.

Analysis [Meurant, T. 2023]

Based on modified tridiagonal matrices

$$T_{k+1}^{(\mu)} = \left[\begin{array}{cccc|c} \alpha_1 & \beta_1 & & & \\ \beta_1 & \ddots & \ddots & & \\ & \ddots & \ddots & \beta_{k-1} & \\ & & \beta_{k-1} & \alpha_k & \beta_k \\ \hline & & & \beta_k & \alpha_{k+1}^{(\mu)} \end{array}\right]$$

that have μ as an eigenvalue,

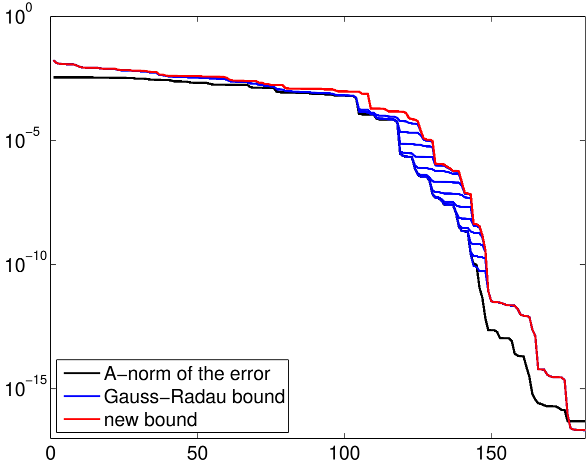
$$\alpha_{k+1}^{(\mu)} = \mu + \sum_{i=1}^{k} \eta_{i,k}^{(\mu)}, \qquad \eta_{i,k}^{(\mu)} \equiv \frac{\left(\beta_k s_{k,i}^{(k)}\right)^2}{\theta_i^{(k)} - \mu}.$$

Analysis of behaviour of the term $\eta_{1,k}^{(\mu)}$, other terms are not sensitive to small modifications of μ.

Simple upper bound for $\mu < \lambda_{\min}$

bcsstk01, $n = 48$, [Meurant, T. 2019]

$$\|x - x_k\|_A^2 \; < \; \gamma_k^{(\mu)} \|r_k\|^2 \; < \; \frac{\|r_k\|^2}{\mu \|p_k\|^2} \|r_k\|^2$$



Closeness of the Gauss-Radau upper bound

and the simple upper bound

We observe that if the upper bound is delayed, then

$$\gamma_k^{(\mu)} \approx \frac{\|r_k\|^2}{\mu\|p_k\|^2}.$$

Explanation based on the formula

$$\gamma_k^{(\mu)} = \left(\frac{\mu\|p_k\|^2}{\|r_k\|^2} + \sum_{i=1}^{k} \left(\frac{\mu}{\theta_i^{(k)}} \right)^2 \eta_{i,k}^{(\mu)} \right)^{-1}.$$

[Meurant, T. 2023]

Conclusions

The behaviour of the Gauss-Radau upper bound

- The Gauss-Radau upper bound is **delayed** if

$$\theta_1^{(k)} - \lambda_1 \; < \; \lambda_1 - \mu \, .$$

- This can happen, e.g, if A has **clustered** eigenvalues.

- If the Gauss-Radau upper bound is delayed, then

$$\gamma_k^{(\mu)} \; \approx \; \frac{\|r_k\|^2}{\mu \|p_k\|^2} \, .$$

Related papers

G. Meurant and P. Tichý, [The behaviour of the Gauss-Radau upper bound of the error norm in CG, submitted to Numer. Algorithms.]

- G. Meurant, J. Papež, and P. Tichý, [Accurate error estimation in CG, Numer. Algorithms, 88 (2021), pp. 1337-1359.]
- G. Meurant and P. Tichý, [Approximating the extreme Ritz values and upper bounds in CG, Numer. Algorithms, 82 (2019), pp. 937-968]
- G. Meurant and P. Tichý, [On computing quadrature-based bounds for the A-norm of the error in CG, Numer. Algorithms, 62 (2013), pp. 163-191]

Thank you for your attention!