On error estimation in the conjugate gradient method

Petr Tichý

Institute of Computer Science, Czech Academy of Sciences, and
Faculty of Mathematics and Physics, Charles University in Prague

joint work with
Gérard Meurant and Zdeněk Strakoš

March 17, 2016, Podbanské, Slovakia
ALGORITMY 2016

The conjugate gradient method

$\mathbf{A}x = b$, $\mathbf{A}$ is real, symmetric, positive definite, $\|b\| = 1$, $x_0 = 0$.

input $\mathbf{A}$, b

$r_0 = b$, $p_0 = r_0$

for $k = 1, 2, \ldots$ **do**

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

end for

$\mathbf{R}_k$

$$\begin{bmatrix} \frac{1}{\sqrt{\gamma_0}} & \sqrt{\frac{\delta_1}{\gamma_0}} & & \\ & \ddots & \ddots & \\ & & \ddots & \sqrt{\frac{\delta_{k-1}}{\gamma_{k-2}}} \\ & & & \frac{1}{\sqrt{\gamma_{k-1}}} \end{bmatrix}$$

Mathematical properties of CG

- CG generates an **orthogonal basis** $r_0, \dots, r_{k-1}$ of

$$\mathcal{K}_k(\mathbf{A}, b) \equiv \operatorname{span}\{b, \mathbf{A}b, \dots, \mathbf{A}^{k-1}b\}\,.$$

  using a coupled two-term recurrence.

- The CG approximation x_k is **optimal**,

$$\|x - x_k\|_{\mathbf{A}} = \min_{y \in \mathcal{K}_k} \|x - y\|_{\mathbf{A}}\,.$$

- orthogonal vectors $\rightarrow$
  - **orthogonal polynomials**
  - **three-term recurrence**, $\mathbf{T}_k = \mathbf{R}_k^T \mathbf{R}_k$
  - **Gauss quadrature** connection

How to measure quality of an approximation?

- **relative residual norm,**

$$\frac{\|r_k\|}{\|b\|}, \qquad \mathbf{A}x_k = b + \Delta b.$$

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

- **normwise backward error**

The normwise backward error is, as a base for stopping criteria, frequently recommended in the numerical analysis literature.

- **error estimates**

Estimates of the $\mathbf{A}$-norm of the error play an important role in stopping criteria in many problems [Deuflhard 1994], [Arioli 2004], [Jiránek, Strakoš, Vohralík 2010]

The normwise backward error

Given x_k, what are the norms of the smallest perturbations $\Delta\mathbf{A}$ of $\mathbf{A}$ and Δb of b (in the relative sense) such that

$$(\mathbf{A} + \Delta\mathbf{A})\, x_k = b + \Delta b\,?$$

We are interested in the quantity

$$\min\left\{\varepsilon : \; (\mathbf{A} + \Delta\mathbf{A})\, x_k = b + \Delta b, \; \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} \leq \varepsilon, \; \frac{\|\Delta b\|}{\|b\|} \leq \varepsilon\right\}$$

called the **normwise backward error**. It is given by

$$\frac{\|r_k\|}{\|\mathbf{A}\|\|x_k\| + \|b\|}\,.$$

[Rigal, Gaches 1967]

The normwise backward error

How to efficiently approximate $\|\mathbf{A}\|$?

$$\frac{\|r_k\|}{\|\mathbf{A}\|\|x_k\| + \|b\|}$$

- Using **Frobenius norm** of $\mathbf{A}$, $\frac{1}{\sqrt{n}}\|\mathbf{A}\|_F \leq \|\mathbf{A}\| \leq \|\mathbf{A}\|_F$,
  - \+ very simple
  - – for large n, not accurate enough
- Using the **largest eigenvalue** of $\mathbf{T}_k$, i.e., $\|\mathbf{T}_k\|$
  - \+ accurate
  - – solving eigenvalue problems, which k?
- Based on **incremental estimation** of $\|\mathbf{T}_k\|$
  - \+ accurate enough
  - \+ very cheap

How to incrementally estimate $\|\mathbf{T}_k\|$?

Since $\mathbf{T}_k = \mathbf{R}_k^T \mathbf{R}_k$, it holds that

$$\|\mathbf{T}_k\| = \|\mathbf{R}_k\|^2 .$$

- $\mathbf{R}_k$ is **upper triangular**, $\mathbf{R}_{k+1}$ arises from $\mathbf{R}_k$ by adding one column and one row $\Rightarrow$ one can use the ideas of [Bischof 1990], [Duff, Vömmel 2002], [Duintjer Tebbens, Tůma 2014].
- The idea is to **incrementally improve** an approximation of the maximum right singular vector.

The idea of incremental norm estimation

Given $\mathbf{R} \in \mathbb{R}^{k\times k}$ **upper triangular**, and z, $\|z\| = 1$, $\|\mathbf{R}z\| \approx \|\mathbf{R}\|$.
Let

$$\hat{\mathbf{R}} = \begin{bmatrix} \mathbf{R} & v \\ & \mu \end{bmatrix}, \qquad v \in \mathbb{R}^k, \quad \mu \in \mathbb{R}.$$

Consider new approximate max. right singular vector in the form

$$\hat{z} = \begin{bmatrix} sz \\ c \end{bmatrix},$$

where $s^2 + c^2 = 1$ are chosen such that $\|\hat{\mathbf{R}}\hat{z}\|$ is maximum,

$$\|\hat{\mathbf{R}}\hat{z}\|^2 = \begin{bmatrix} s \\ c \end{bmatrix}^T \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \begin{bmatrix} s \\ c \end{bmatrix}$$

$$\alpha = \|\mathbf{R}z\|^2, \qquad \beta = v^T\mathbf{R}z, \qquad \gamma = v^Tv + \mu^2.$$

Determine the **maximum eigenvalue** of the 2×2 matrix, and the corresponding **eigenvector**.

Specialization to upper bidiagonal matrices from CG

$$\mathbf{R}_{k+1} = \begin{bmatrix} \mathbf{R}_k & v \\ & \mu \end{bmatrix}, \qquad \mathbf{R}_k = \begin{bmatrix} \frac{1}{\sqrt{\gamma_0}} & \sqrt{\frac{\delta_1}{\gamma_0}} & & \\ & \ddots & \ddots & \\ & & \ddots & \sqrt{\frac{\delta_{k-1}}{\gamma_{k-2}}} \\ & & & \frac{1}{\sqrt{\gamma_{k-1}}} \end{bmatrix}$$

$$v = \sqrt{\frac{\delta_k}{\gamma_{k-1}}} e_k, \qquad \mu = \frac{1}{\sqrt{\gamma_k}}$$

Using some algebra → an updating formula for $\Delta_k \equiv \|\mathbf{R}_k z_k\|^2$,

$$\Delta_{k+1} = \Delta_k + \omega_k c_{k+1}^2,$$

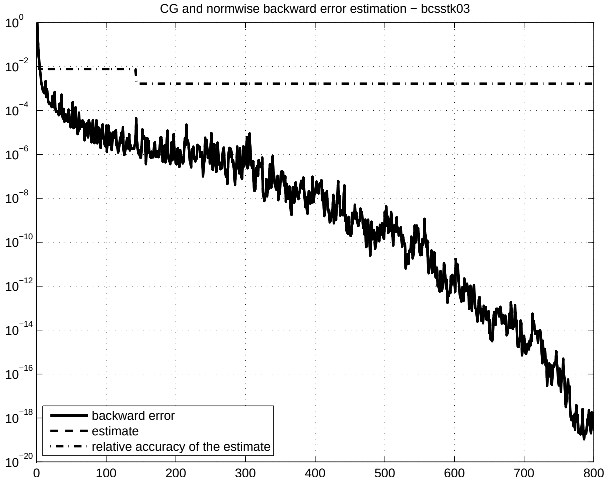
we don't **need to store** z_k!

CG with incremental norm estimation

input $\mathbf{A}$, b, x_0
$r_0 = b - \mathbf{A}x_0$, $p_0 = r_0$
$\delta_0 = 0$, $\gamma_{-1} = 0$
for $k = 1, \ldots,$ **do**
  CG iteration(k) $\rightarrow \gamma_{k-1}$, x_k, r_k, δ_k, p_k
  $\alpha_k = \frac{1}{\gamma_{k-1}} + \frac{\delta_{k-1}}{\gamma_{k-2}}$
  $\beta_k^2 = \frac{\delta_k}{\gamma_{k-1}^2}$
  if $k = 1$ **then**
    $c_1^2 = 1$, $\Delta_1 = \alpha_1$
  else
    $\omega_{k-1} = \sqrt{(\Delta_{k-1} - \alpha_k)^2 + 4\beta_{k-1}^2 c_{k-1}^2}$
    $c_k^2 = \frac{1}{2}\left(1 - \frac{\Delta_{k-1} - \alpha_k}{\omega_{k-1}}\right)$
    $\Delta_k = \Delta_{k-1} + \omega_{k-1} c_k^2$
  end if
end for

Accuracy of the upper bound

$$\Delta_k \le \|\mathbf{A}\| \quad \Rightarrow \quad \frac{\|r_k\|}{\|\mathbf{A}\|\|x_k\| + \|b\|} \le \frac{\|r_k\|}{\Delta_k\|x_k\| + \|b\|}$$



Estimating the $\mathbf{A}$-norm of the error

A brief history

- "The function $(x - x_k, \mathbf{A}(x - x_k))$ can be used as a measure of the "goodness" of x_k as an estimate of x." [Hestenes, Stiefel 1952]
- **Gene Golub** and collaborators: [Dahlquist, Golub, Nash 1978] relate error bounds to **Gauss quadrature**; see also [Golub, Meurant 1994].
- Idea of estimating errors in CG, behavior in finite precision arithmetic [Golub, Strakoš 1994], the CGQL algorithm [Golub, Meurant, 1997], intensively studied in many later papers.
- Numerical stability of the estimates based on Gauss quadrature [Strakoš, T. 2002], [Strakoš, T. 2005].
- Summary in the book [Golub, Meurant 2010].
- Improvement [Meurant, T. 2013] $\rightarrow$ the CGQ algorithm.

CG and Gauss quadrature

The lower bound on $\|x - x_k\|_{\mathbf{A}}$

At any iteration step k, CG determines (through $\mathbf{T}_k$) **weights** and **nodes** of the k-point **Gauss quadrature**. For $f(\lambda) = \lambda^{-1}$:

$$\|x\|_{\mathbf{A}}^2 = \sum_{j=0}^{k-1} \gamma_j \|r_j\|^2 + \|x - x_k\|_{\mathbf{A}}^2 \,.$$

Considering the same formula for some $k + d$ we obtain

$$\|x - x_k\|_{\mathbf{A}}^2 = \sum_{j=k}^{k+d-1} \gamma_j \|r_j\|^2 + \|x - x_{k+d}\|_{\mathbf{A}}^2$$

We have a **lower bound**. E.g., if $\frac{\|x - x_{k+d}\|_{\mathbf{A}}}{\|x - x_k\|_{\mathbf{A}}} < 0.8$, then

$$\sqrt{\sum_{j=k}^{k+d-1} \gamma_j \|r_j\|^2} < \|x - x_k\|_{\mathbf{A}} < 2\sqrt{\sum_{j=k}^{k+d-1} \gamma_j \|r_j\|^2} \,.$$

The lower bound

Practically relevant questions

$$\|x - x_k\|_{\mathbf{A}}^2 \;=\; \sum_{j=k}^{k+d-1} \gamma_j \|r_j\|^2 \;+\; \|x - x_{k+d}\|_{\mathbf{A}}^2 .$$

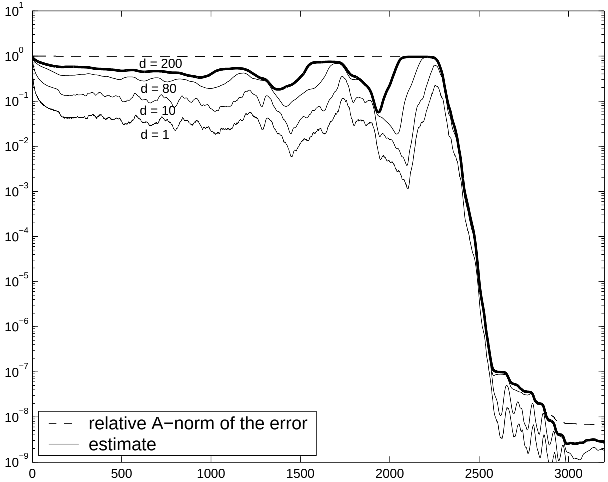
Practically relevant questions:

+ We can compute it almost **for free**.
+ It is **numerically stable** [Strakoš, T. 2002, 2005]:
- How to **control the quality** of the bound?
  - How to choose d adaptively?
  - How to detect a decrease of the $\mathbf{A}$-norm of the error?

The choice of d

R. Kouhia: Cylindrical shell (Matrix Market), matrix s3dkt3m2

PCG, $\kappa(\mathbf{A}) = 3.62e + 11$, $n = 90499$, cholinc$(\mathbf{A}, 0)$.



CG and Gauss-Radau quadrature

- Modification of Gauss quadrature rule.
- **Prescribe** μ such that $0 < \mu < \lambda_{\min}$.

Gauss-Radau quadrature rule can be written algebraically as

$$\|x\|_{\mathbf{A}}^2 = \sum_{j=0}^{k-2} \gamma_j \|r_j\|^2 + \gamma_{k-1}^{(\mu)} \|r_{k-1}\|^2 + \mathcal{R}_k^{(R)}$$

where $\mathcal{R}_k^{(R)} < 0$, $\gamma_0^{(\mu)} = \mu^{-1}$, and

$$\gamma_{j+1}^{(\mu)} = \frac{\left(\gamma_j^{(\mu)} - \gamma_j\right)}{\mu \left(\gamma_j^{(\mu)} - \gamma_j\right) + \delta_{j+1}} .$$

How to construct an **upper bound** on $\|x - x_k\|_{\mathbf{A}}$?

CG and Gauss-Radau quadrature

Upper bound

Using **Gauss** rule for k and **Gauss-Radau** rule for $k+d$,

$$\|x - x_k\|_{\mathbf{A}}^2 \quad = \quad \sum_{j=k}^{k+d-2} \gamma_j \|r_j\|^2 + \gamma_{k+d-1}^{(\mu)} \|r_{k+d-1}\|^2 + \mathcal{R}_{k+d}^{(R)}$$

and $\mathcal{R}_{k+d}^{(R)} < 0$.

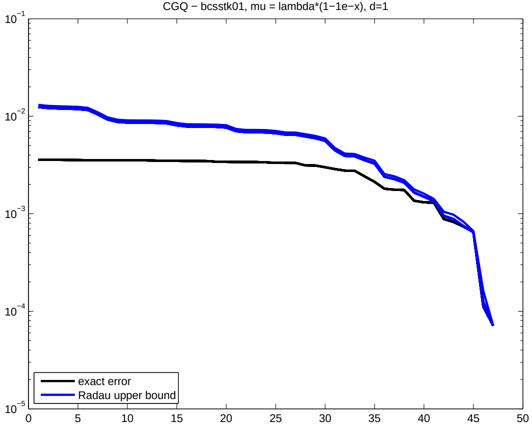
Practically relevant questions:

+ We can compute it almost **for free**.
- How to get μ?
- **Sensitivity** of the bound on μ?
- Numerical **behavior**?
- **Quality** of the bound?

Gauss-Radau upper bound, exact arithmetic

various values of μ, bcsstk01 matrix, $n = 48$

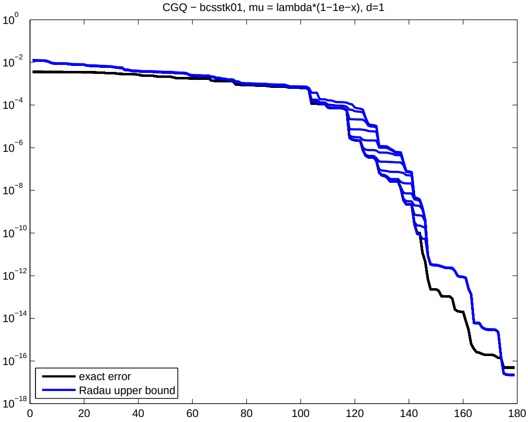
$$\mu = \lambda_{\min}(1 - 10^{-m}), \quad m = 1, \ldots, 14$$



Gauss-Radau upper bound, finite precision arithmetic

various values of μ, bcsstk01 matrix, $n = 48$

$$\mu = \lambda_{\min}(1 - 10^{-m}), \quad m = 1, \dots, 14$$



An upper bound on the upper bound

Find an envelope for all curves that corresponds to various μ's.

Multiply the updating formula for $\gamma_{j+1}^{(\mu)}$ by μ

$$\mu\gamma_{j+1}^{(\mu)} = \frac{\mu\left(\gamma_j^{(\mu)} - \gamma_j\right)}{\mu\left(\gamma_j^{(\mu)} - \gamma_j\right) + \delta_{j+1}}.$$

We can show that $\mu\gamma_{j+1}^{(\mu)} < f_{j+1}$, and f_{j+1} does not depend on μ,

$$f_{j+1} = \frac{f_j}{f_j + \delta_{j+1}}, \qquad f_0 = 1.$$

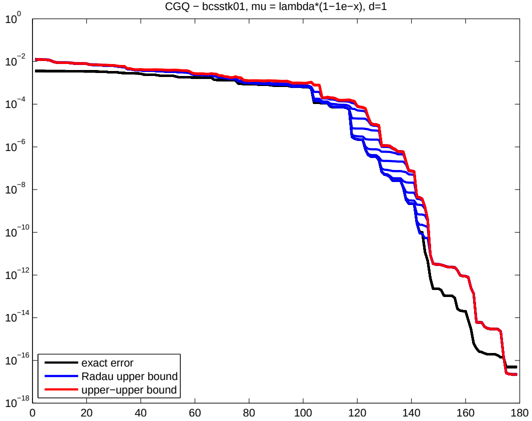
In fact, we can show that

$$f_j = \frac{\|r_j\|^2}{\|p_j\|^2}.$$

An upper bound on the upper bound

In summary, if $\mu < \lambda_{\min}$, then

$$\|x - x_k\|_{\mathbf{A}}^2 \leq \gamma_k^{(\mu)} \|r_k\|^2 < \frac{f_k}{\mu} \|r_k\|^2.$$



CG and Gauss-Radau quadrature

Upper bound on the upper bound

If $\mu < \lambda_{\min}$, then

$$\|x - x_k\|_{\mathbf{A}}^2 \quad < \quad \sum_{j=k}^{k+d-1} \gamma_j \|r_j\|^2 + \frac{f_{k+d}}{\mu} \|r_{k+d}\|^2$$

Practically relevant questions:

+ We can compute it almost **for free**.
+ No too much **sensitive** to the choice of μ.
+ Heuristics $\rightarrow$ we can use it even if $\mu > \lambda_{\min}$ (e.g., use the estimate of the smallest Ritz value).
+ Numerical **behavior**? Looks OK, so far no analysis.
– **Quality** of the bound. Not so tight.
+ We hope, it can be used for the **adaptive choice of d** in the (tight) lower bound.

Conclusions and questions

- $\mathbf{T}_k = \mathbf{R}_k^T \mathbf{R}_k$ represents an important **source of information** that can be used for estimating the CG convergence.
- The **normwise backward error** can be approximated almost for free using incremental norm estimation.
- The approximation Δ_k of $\|\mathbf{A}\|$ can also be used to estimate the **level of maximal attainable accuracy**.
- The estimation of the **$\mathbf{A}$-norm of the error** should be based on the numerical stable **lower bound**.
- How to **detect a decrease** of the $\mathbf{A}$-norm of the error? (How to choose d adaptively?).
- We found an **upper bound** on the Gauss-Radau upper bound that is insensitive to the choice of μ, and hope to be useful in the adaptive choice of d for the lower bound.

Related papers

- G. Meurant and P. Tichý, [On computing quadrature-based bounds for the $\mathbf{A}$-norm of the error in conjugate gradients, Numer. Algorithms, 62 (2013), pp. 163-191]
- G. H. Golub and G. Meurant, [Matrices, moments and quadrature with applications, Princeton University Press, USA, 2010.]
- Z. Strakoš and P. Tichý, [On error estimation in the conjugate gradient method and why it works in finite precision computations, Electron. Trans. Numer. Anal., 13 (2002), pp. 56–80.]
- G. H. Golub and G. Meurant, [Matrices, moments and quadrature. II. BIT, 37 (1997), pp. 687–705.]
- G. H. Golub and Z. Strakoš, [Estimates in quadratic formulas, Numer. Algorithms, 8 (1994), pp. 241–268.]

Thank you for your attention!