# How to control the accuracy of error estimates in CG?

Petr Tichý

Charles University, Prague

based on joint work with

Gérard Meurant, Jan Papež, Zdeněk Strakoš

Householder Symposium XXI

June 12-17, Selva di Fasano, Italy

# The conjugate gradient method

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

**input** $A$, $b$

$r_0 = p_0 = b$

**for** $k = 0, \ldots$ until convergence **do**

$$\left.\begin{aligned} \alpha_k &= \tfrac{r_k^T r_k}{p_k^T A p_k} \\ x_{k+1} &= x_k + \alpha_k p_k \\ r_{k+1} &= r_k - \alpha_k A p_k \\ \beta_{k+1} &= \tfrac{r_{k+1}^T r_{k+1}}{r_k^T r_k} \\ p_{k+1} &= r_{k+1} + \beta_{k+1} p_k \end{aligned}\right\} \quad \texttt{cgiter(k)}$$

**end for**

$$\varepsilon_k \;\equiv\; \|x - x_k\|_A^2 = \min_{y \in \mathcal{K}_k} \|x - y\|_A^2$$

# 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

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

- **Estimating errors** using quadrature approach:

  [Dahlquist, Golub, Nash 1978], [Golub, Meurant 1994], [Golub, Strakoš 1994], [Golub, Meurant, 1997], [Calvetti et al. 2000], [Strakoš, T. 2002], [Meurant, T. 2013, 2019]

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

  [Golub, Strakoš 1994], [Strakoš, T. 2002, 2005, 2011]

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

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

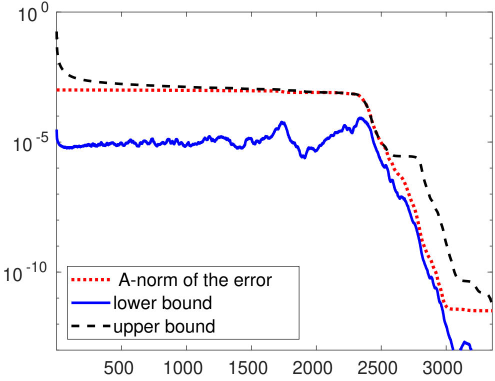
- How to **improve** and **control** the accuracy?

# Basic bounds

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

Given $\mu \leq \lambda_{\min}$,

$$\alpha_k \|r_k\|^2 < \varepsilon_k < \alpha_k^{(\mu)} \|r_k\|^2$$



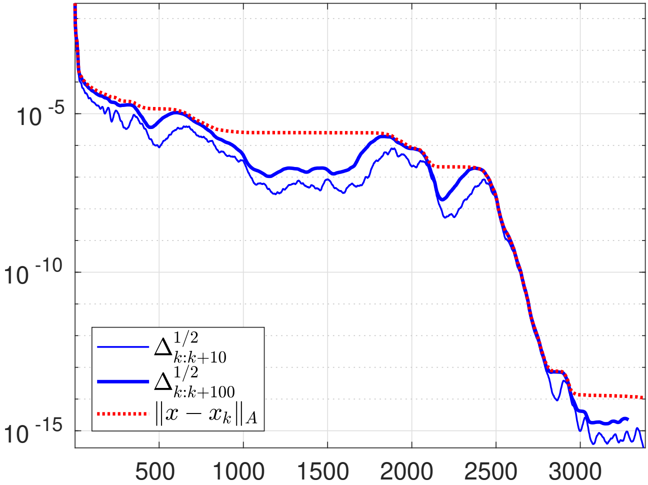
# How to improve the accuracy of the estimates?

$$\varepsilon_k \;=\; \underbrace{\sum_{j=k}^{\ell-1} \alpha_j \|r_j\|^2}_{\Delta_{k:\ell-1}} \;+\; \varepsilon_\ell$$

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

## $\ell = k + d$ with a constant $d$

s3dkq4m2, $n = 90449$, `ichol`




A need to **determine $d$ adaptively**.

# Prescribing the accuracy of the estimate

$$\varepsilon_k \; = \; \Delta_{k:\ell-1} \; + \; \varepsilon_\ell$$

Ideally, we would like to determine $\ell > k$ such that

$$\frac{\varepsilon_k - \Delta_{k:\ell-1}}{\varepsilon_k} \; = \; \frac{\varepsilon_\ell}{\varepsilon_k} \; \leq \; \tau \, ,$$

where $\tau \in (0, 1)$ is a given tolerance. Then

$$\Delta_{k:\ell-1} \; < \; \varepsilon_k \; \leq \; \frac{\Delta_{k:\ell-1}}{1 - \tau} \, .$$

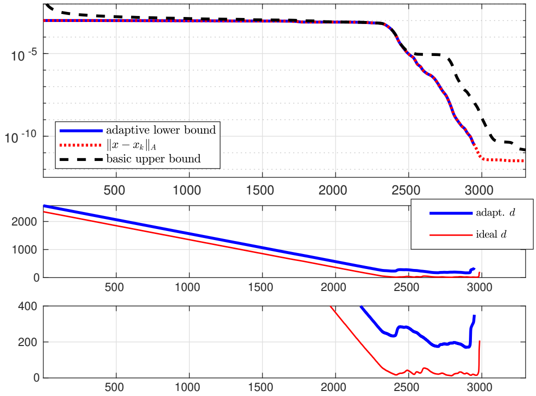
The delay $\ell - k$ should be as small as possible.

## Find $\ell$ such that

$$\frac{\varepsilon_k - \Delta_{k:\ell-1}}{\varepsilon_k} = \frac{\varepsilon_\ell}{\varepsilon_k} \leq \tau$$

# Using the upper bound

$$\frac{\varepsilon_\ell}{\varepsilon_k} \leq \frac{\alpha_\ell^{(\mu)} \|r_\ell\|^2}{\Delta_{k:\ell-1}} \leq \tau$$




Safe, but requires $\mu$ and, moreover, $\ell - k$ is far from being optimal!

# Heuristic strategy

## Learn from the history

It holds that [Meurant, Papež, T. 2021]

$$\Delta_j < \varepsilon_j < \kappa(A)\,\Delta_j\,.$$

Idea $\rightarrow$ find $S_\ell$ such that

$$\varepsilon_\ell \approx S_\ell\,\Delta_\ell.$$

Define

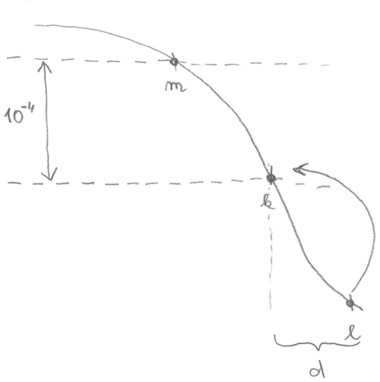
$$S_\ell \equiv \max_{m \leq j < \ell} \widetilde{S}_j, \quad \widetilde{S}_j \equiv \frac{\Delta_{j:\ell}}{\Delta_j} \approx \frac{\varepsilon_j}{\Delta_j}\,.$$

Approximate

$$\frac{\varepsilon_\ell}{\varepsilon_k} \approx \frac{S_\ell \Delta_\ell}{\Delta_{k:\ell-1}} \leq \tau\,.$$

# How far to go into history?

## Learn from the latest significant decrease




$$\varepsilon_k = \Delta_{k:\ell-1} + \varepsilon_\ell$$

- find $m$ such that

$$\frac{\varepsilon_k}{\varepsilon_m} \approx \frac{\Delta_{k:\ell}}{\Delta_{m:\ell}} \leq 10^{-4}$$

- define

$$S_\ell = \max_{m \leq j < \ell} \widetilde{S}_j$$

- test

$$\frac{S_\ell \Delta_\ell}{\Delta_{k:\ell-1}} \leq \tau$$

# Estimating $\varepsilon_k$ with a prescribed accuracy

[Meurant, Papež, T. 2021]

```
input A, b, τ
r_0 = p_0 = b, k = 0
cgiter(0)
for ℓ = 1, ..., do
    cgiter(ℓ)
    compute Δ_{k:ℓ-1} and Δ_ℓ
    determine S_ℓ
    while ℓ > k and (S_ℓ Δ_ℓ)/(Δ_{k:ℓ-1}) ≤ τ do
        accept Δ_{k:ℓ}
        k = k + 1
    end while
end for
```

Preconditioned CG $\rightarrow$ $\varepsilon_k = \sum_{j=k}^{\ell-1} \hat{\alpha}_j \, z_j^T r_j + \varepsilon_\ell$

```
function [x,estim,delay] = pcga(A,b,tau,maxit,L,x)
r = b - A * x;
z = L\r; z = L'\z; p = z;
rr = z' * r;
k = 1;

for ell = 1:maxit+1

    RR    = rr;                          % ... begin cgiter(ell)
    Ap    = A * p;
    alpha = RR/(p' * Ap);
    x     = x + alpha * p;
    r     = r - alpha * Ap;
    z     = L \ r; z = L' \ z;
    rr    = z' * r;
    beta  = rr / RR;
    p     = z + beta * p;                % ... end cgiter(ell)

    Delta(ell) = alpha * RR;
    history(ell) = 0; history = history + Delta(ell);

    if ell > 1                           % ... adaptive choice of the delay
        S = findS(history,Delta,k);
        num = S * Delta(ell);
        den = sum(Delta(k:ell-1));
        while (ell > k) && (num/den <= tau)
            delay(k) = ell-k;
            estim(k) = den;
            k = k + 1;
            den = sum(Delta(k:ell-1));
        end
    end
end
end % of function

function [S] = findS(history,Delta,k)
ind = find((history(k)./history) <= 1e-4, 1, 'last');
if isempty(ind), ind = 1; end
S = max(history(ind:end-1)./Delta(ind:end-1));
end
```

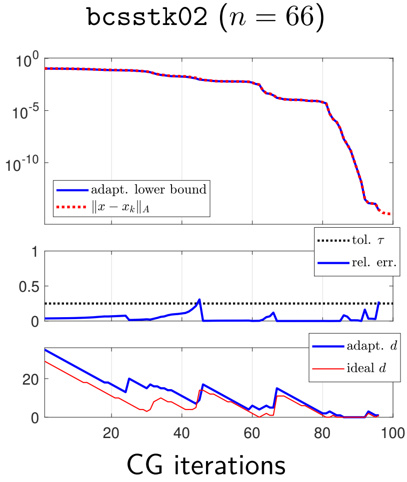
# Numerical experiments

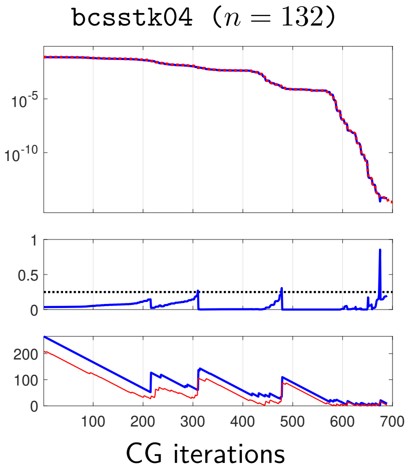
# Test problems

## SuiteSparse Matrix collection

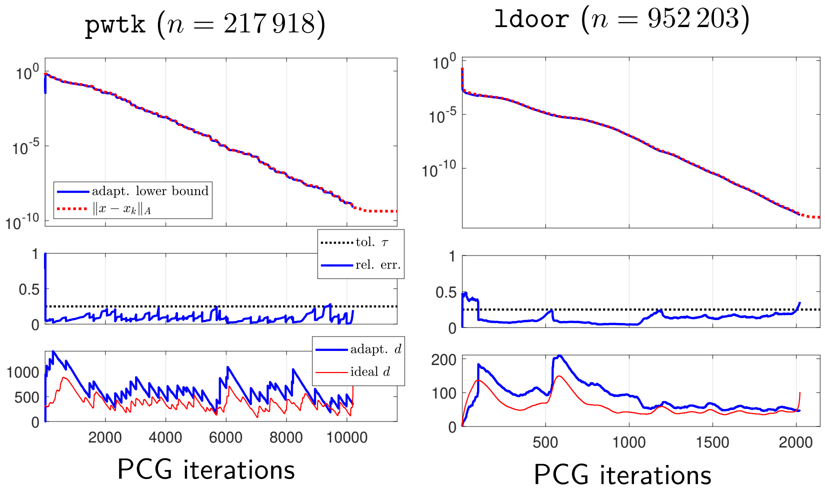
| name | size | rhs $b$ | $M = LL^T$ |
|---|---|---|---|
| bcsstk02 | 66 | equal components | — |
| bcsstk04 | 132 | equal components | — |
| bcsstk09 | 1083 | equal components | ict(1e-3, 1e-2) |
| s3dkt3m2 | 90 449 | comes with the matrix | ict(1e-5, 1e-2) |
| s3dkq4m2 | 90 449 | $\text{rand}(-1, 1)$ | ict(1e-5, 1e-2) |
| pwtk | 217 918 | $\text{rand}(-1, 1)$ | ict(1e-5, 1e-1) |
| af_shell3 | 504 855 | $\text{rand}(-1, 1)$ | zero-fill |
| tmt_sym | 726 713 | $\text{rand}(-1, 1)$ | zero-fill |
| ldoor | 952 203 | $\text{rand}(-1, 1)$ | zero-fill |

# Problems without preconditioning



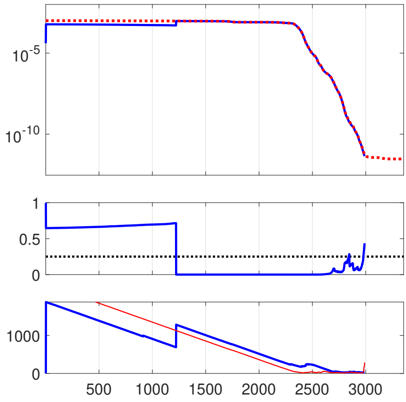



# Problems with preconditioning



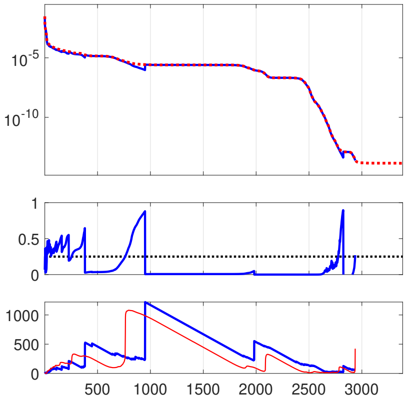
# Difficult problems

s3dkt3m2 ($n = 90\,449$)




PCG iterations

s3dkq4m2 ($n = 90\,449$)




PCG iterations

# Conclusions

- One can **improve** the accuracy of estimates of $\varepsilon_k$ using the information from the forthcoming CG iterations.
- We can **control** the accuracy of the estimates using:
  - Gauss-Radau **upper bound** $\rightarrow$ safe, far from being optimal.
  - A **heuristic strategy** $\rightarrow$ robust, reliable, often almost optimal.
- Generalization is possible for other CG-like methods.

# Related papers

G. Meurant, J. Papež, and P. Tichý,

[Accurate error estimation in CG, Numer. Algorithms, 88 (2021), pp. 1337-1359.]

- G. H. Golub and Z. Strakoš, [Estimates in quadratic formulas, Numer. Algorithms, 8 (1994), pp. 241–268.]
- 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]
- Z. Strakoš and P. Tichý, [On error estimation in CG and why it works in FP computations, Electron. Trans. Numer. Anal., 13 (2002), pp. 56–80.]

**Thank you for your attention!**

# Delay of the Gauss-Radau upper bound?

Work in progress ...