Estimates of norms of error in PCG

Gérard MEURANT

CEA/DIF

June 10, 2006

Introduction

Solve

$$Ax = b$$

Let $r^k = b - Ax^k$ be the residual vector
Since $\|r^k\|$ is often misleading for stopping iterative methods, it is of interest to obtain bounds or estimates of norms of the error

$$\epsilon^k = x - x^k$$

We have

$$A\epsilon^k = r^k$$

We shall review proposals by Brezinski (extrapolation) and Golub & Meurant (Gauss quadrature) and give results for PCG

Bresinski's estimates

C.B. (SISC v21 1999) considered the moments of $r = r^k$ (or $\epsilon = \epsilon^k$)

$$c_0 = (r, r) = \|r\|^2 = (A\epsilon, A\epsilon)$$

$$c_1 = (r, Ar) = (A\epsilon, A^2\epsilon)$$

$$c_2 = (Ar, Ar) = (A^2\epsilon, A^2\epsilon)$$

$$c_{-1} = (r, A^{-1}r) = (A\epsilon, \epsilon) = \|\epsilon\|_A^2$$

$$c_{-2} = (A^{-1}r, A^{-1}r) = (\epsilon, \epsilon) = \|\epsilon\|^2$$

c_0, c_1 and c_2 are computable

We would like to have estimates of c_{-1} and/or c_{-2}

Bresinski's estimates 2

Using the first terms in SVD expansions, Brezinski obtained the following estimates e_i^2 for c_{-2}

$$
\begin{aligned}
e_1^2 &= c_1^4 / c_2^3 \\
e_2^2 &= c_0 c_1^2 / c_2^2 \\
e_3^2 &= c_0^2 / c_2 \\
e_4^2 &= c_0^3 / c_1^2 \\
e_5^2 &= c_0^4 c_2 / c_1^4
\end{aligned}
$$

C.B. proved that

$$e_1 \leq e_2 \leq e_3 \leq e_4 \leq e_5$$

Bresinski's estimates 3

e_3 is usually the most appropriate estimate of $\|\epsilon\|$
Moreover, it can be derived by other techniques and is valid for any consistent norm
Therefore, we shall consider:

$$\|\epsilon\|^2 \simeq \frac{(r,r)^2}{(Ar,Ar)}$$

$$\|\epsilon\|_A^2 \simeq \frac{(r,Ar)}{(A^2r,Ar)}$$

Gauss quadrature bounds

Golub & Meurant (1993)

$$A\epsilon^k = r^k$$

Therefore,

$$\|\epsilon^k\|_A^2 = (A\epsilon^k, \epsilon^k) = (A^{-1}r^k, r^k) = (r^k)^T A^{-1} r^k$$

$$\|\epsilon^k\|^2 = (r^k)^T A^{-2} r^k$$

Suppose A is symmetric positive definite

$$A = Q\Lambda Q^T$$

Consider

$$u^T f(A) u, \quad f(A) = Q f(\Lambda) Q^T$$

Gauss quadrature bounds 2

$$I[f] = u^T f(A) u = \int_a^b f(\lambda)\, d\alpha(\lambda)$$

where the measure α is piecewise constant if $y = Q^T u$

$$\alpha(\lambda) = \begin{cases} 0 & \text{if } \lambda < a = \lambda_1, \\ \sum_{j=1}^{i} y_j^2 & \text{if } \lambda_i \leq \lambda < \lambda_{i+1}, \\ \sum_{j=1}^{n} y_j^2 & \text{if } b = \lambda_n \leq \lambda \end{cases}$$

We can use Gauss, Gauss–Radau and Gauss–Lobatto to obtain bounds for $I[f]$

Gauss quadrature bounds 3

This is closely linked to the Lanczos algorithm starting from $v^1 = u/\|u\|$

$$AV_k = V_k T_k + G_k, \quad G_k = \begin{pmatrix} 0 & \eta_{k+1} v^{k+1} \end{pmatrix}$$

T_k is tridiagonal

Gauss quadrature bounds 4

The Gauss lower bound at Lanczos iteration k is

$$(e^1)^T f(T_k) e^1$$

So, we compute

$$(e^1)^T T_k^{-1} e^1$$

$$(e^1)^T T_k^{-2} e^1$$

For Gauss–Radau and Gauss–Lobatto some elements of T_k have to be modified to obtain the prescribed nodes (the smallest and largest eigenvalues of A)

CG algorithm

x^0 given and $r^0 = b - Ax^0$:
for $k = 0, 1, \ldots$ until convergence do,

$$\beta_k = \frac{(r^k, r^k)}{(r^{k-1}, r^{k-1})}, \; \beta_0 = 0$$

$$p^k = r^k + \beta_k p^{k-1}$$

$$\gamma_k = \frac{(r^k, r^k)}{(Ap^k, p^k)}$$

$$x^{k+1} = x^k + \gamma_k p^k$$

$$r^{k+1} = r^k - \gamma_k Ap^k$$

Gauss quadrature estimates can be obtained by

$$\|\epsilon^{k-d}\|_A^2 \simeq \|r^0\|^2[(T_k^{-1}e^1, e^1) - (T_{k-d}^{-1}e^1, e^1)]$$

or

$$\|\epsilon^{k-d}\|_A^2 \simeq \sum_{j=k-d}^{k-1} \gamma_j \|r^j\|^2$$

$d > 0$

T_k is the Lanczos matrix obtained from CG coefficients

The first formula can be used by modifying T_k to obtain upper bounds (if we have an estimate of the smallest eigenvalue of A)

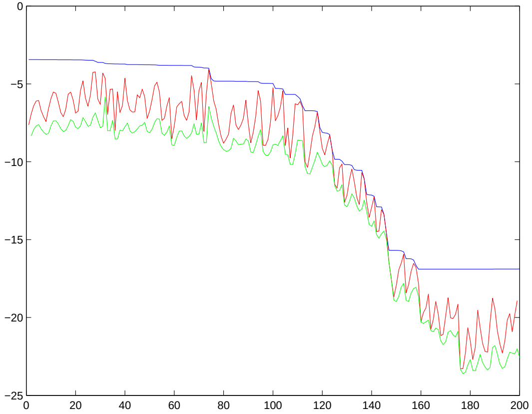
For the l_2 norm of the error we have

$$\|\epsilon^k\|^2 = \|r^0\|^2[(e^1, T_n^{-2}e^1) - (e^1, T_k^{-2}e^1)] - 2\frac{(e^k, T_k^{-2}e^1)}{(e^k, T_k^{-1}e^1)}\|\epsilon^k\|_A^2$$

This can be used to obtain estimates

Numerical example

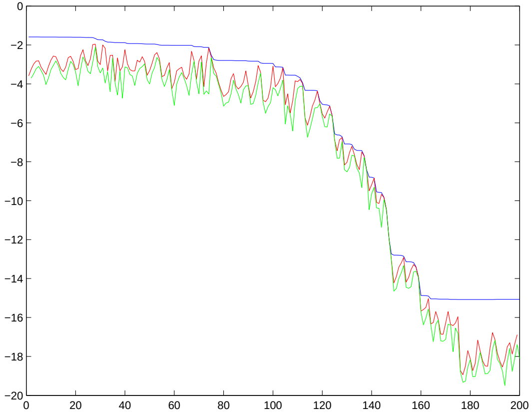
Matrix Bcsstk01, $n = 48$



l_2 norm (blue), Gauss quad $d = 1$ (red), C.B. (green)

Numerical example 2

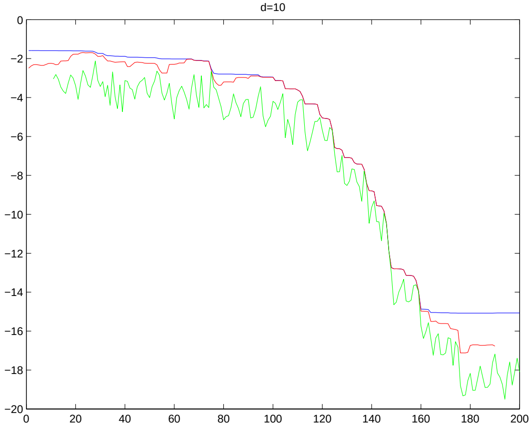
Matrix Bcsstk01, $n = 48$



A norm (blue), Gauss quad $d = 1$ (red), C.B. (green)

Numerical example 3

Matrix Bcsstk01, $n = 48$



A norm (blue), Gauss quad $d = 10$ (red), C.B. (green)

Preconditioning

M symmetric positive definite

$$Mz^k = r^k$$

Brezinski's estimates:

$$My^k = Az^k$$

l_2 norm

$$\|\epsilon^k\|^2 \simeq \frac{(z^k, r^k)^2}{(y^k, Az^k)}$$

Preconditioning 2

A-norm

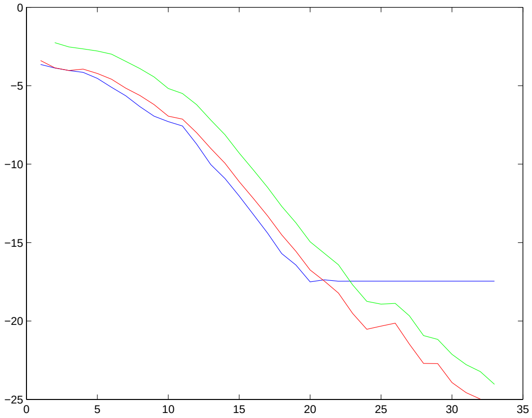
$$\|\epsilon^k\|_A^2 \simeq \frac{(y^k, r^k)^2}{(y^k, Ay^k)}$$

Gauss quadrature lower bound

$$\|\epsilon^{k-d}\|_A^2 \simeq \sum_{j=k-d}^{k-1} \gamma_j (r^k, z^k)$$

Numerical examples with PCG

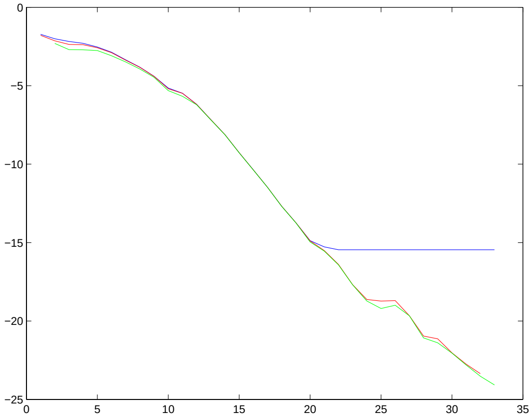
Matrix Bcsstk01, $n = 48$, IC(0)



l_2 norm (blue), Gauss quad $d = 1$ (red), C.B. (green)

Numerical examples with PCG 2

Matrix Bcsstk01, $n = 48$, IC(0)



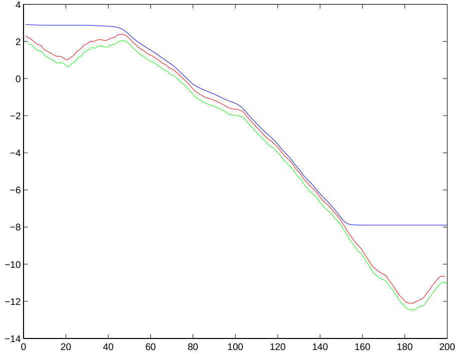
A norm (blue), Gauss quad $d = 1$ (red), C.B. (green)

Numerical examples with PCG 3

Pb sinus–sinus, diff. coeff.=

$$\frac{1}{(2 + p \sin \frac{x}{\eta})(2 + p \sin \frac{y}{\eta})}$$

$p = 1.99$ and $\eta = 0.01$, mesh 100×100, IC(0)



A norm (blue), Gauss quad $d = 1$ (red), C.B. (green)

Conclusions

- All these estimates give good results

Conclusions

- All these estimates give good results
- Gauss quadrature is better suited to PCG

Conclusions

- All these estimates give good results
- Gauss quadrature is better suited to PCG
- Brezinski's estimates are more general

Conclusions

- All these estimates give good results
- Gauss quadrature is better suited to PCG
- Brezinski's estimates are more general
- AND NOW ...

FROM

TO (Luc sur mer, Normandy)