Estimates of norms of the error in the Preconditioned Conjugate Gradient

Gérard MEURANT

CEA/DIF

July 14, 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 Gauss quadrature estimates by Golub & Meurant and give results for PCG

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

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

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 $y = Q^T u$

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

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 for $I[f]$ 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,

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

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

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

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

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

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 equivalently

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

$d > 0$

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

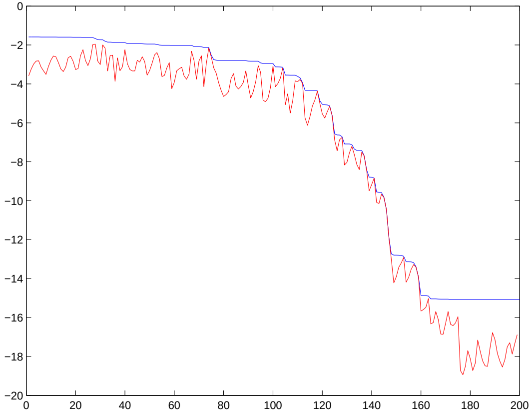
$$\|\epsilon^k\|^2 = \|r^0\|^2[(e^1, T_n^{-2}e^1) - (e^1, T_k^{-2}e^1)] - 2\frac{(e^k, T_k^{-2}e^1)}{(e^k, T_k^{-1}e^1)}\|\epsilon^k\|_A^2$$

This can be used to obtain estimates

All these formulas are also valid in finite precision arithmetic up to terms proportional to the roundoff unit (Strakoš and Tichý, Meurant)

Numerical example 1

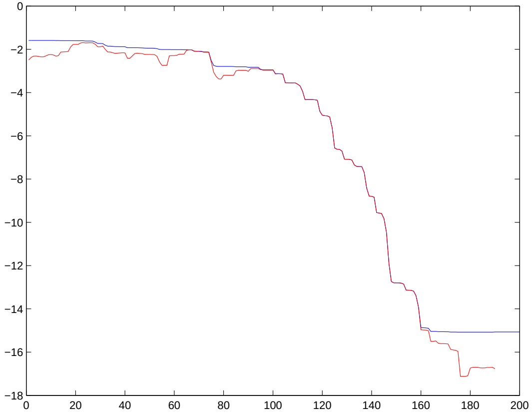
Matrix Bcsstk01, $n = 48$



A norm (blue), Gauss quad $d = 1$ (red)

Numerical example 2

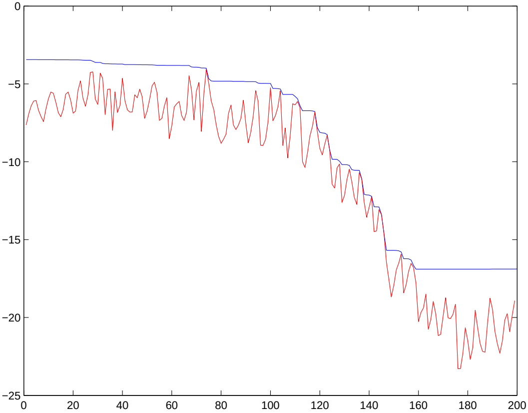
Matrix Bcsstk01, $n = 48$



A norm (blue), Gauss quad $d = 10$ (red)

Numerical example 3

Matrix Bcsstk01, $n = 48$



l_2 norm (blue), Gauss quad $d = 1$ (red)

Preconditioning

M symmetric positive definite

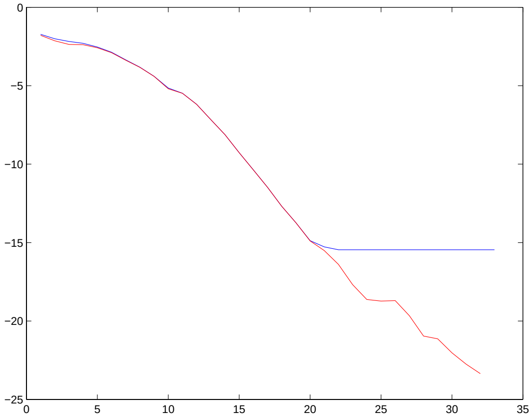
$$Mz^k = r^k$$

Gauss quadrature lower bound

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

Numerical examples with PCG 1

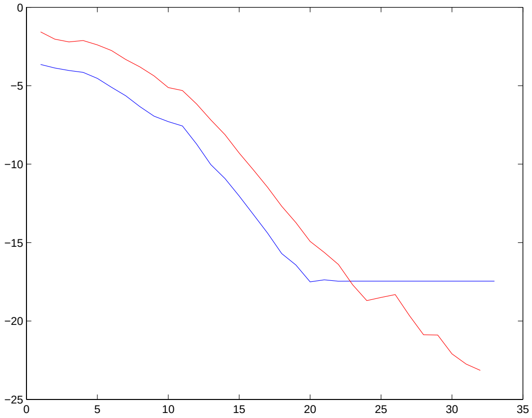
Matrix Bcsstk01, $n = 48$, IC(0)



A norm (blue), Gauss quad $d = 1$ (red)

Numerical examples with PCG 2

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



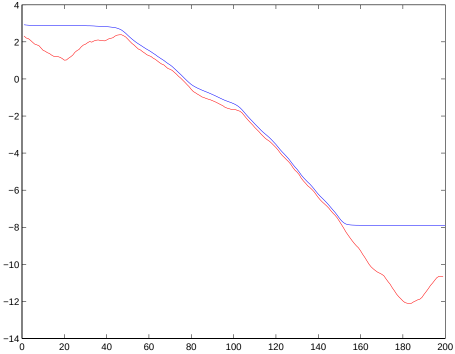
l_2 norm (blue), Gauss quad $d = 1$ (red)

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)

Conclusions

- All these estimates give good results

Conclusions

- All these estimates give good results
- Estimates for the A-norm are usually better than for the l_2 norm

Conclusions

- All these estimates give good results
- Estimates for the A-norm are usually better than for the l_2 norm
- The A-norm bound is easy to extend to PCG

Conclusions

- All these estimates give good results
- Estimates for the A-norm are usually better than for the l_2 norm
- The A-norm bound is easy to extend to PCG
- They work well in finite precision arithmetic and can be used to stop the iterations

Conclusions

- All these estimates give good results
- Estimates for the A-norm are usually better than for the l_2 norm
- The A-norm bound is easy to extend to PCG
- They work well in finite precision arithmetic and can be used to stop the iterations
- This gives a reliable stopping criterion for PCG