Indefinite Problems

CG Method

$r^0 = b - Ax^0;$
$\alpha_0 = (r^0)^T r^0;$
$d^1 = r^0;\ k = 1;$
while $\alpha_{k-1} > 0$
  $s^k = Ad^k;$
  $\tau_k = \alpha_{k-1}/(d^k)^T s^k;$
  $x^k = x^{k-1} + \tau_k d^k;$
  $r^k = r^{k-1} - \tau_k s^k;$
  $\beta_k = 1/\alpha_{k-1};$
  $\alpha_k = (r^k)^T r^k;$
  $\beta_k = \beta_k \alpha_k;$
  $d^{k+1} = r^k + \beta_k d^k;$
  $k = k + 1;$ end

For A symmetric, but not positive definite the CG method can break down with $(d^k)^T s^k = (d^k)^T Ad^k = 0$ for some $k < n$.

As in the positive definite case the search directions satisfy $(d^j)^T Ad^i = 0$ for $i \neq j$, and hence the d^j are linearly independent.

Thus, as long as it does not break down the CG method determines a basis of the Krylov space $\mathcal{K}_k(r^0, A)$, and the iterates $x^k = x^0 + \mathcal{K}_k(r^0, A)$ satisfy the Galerkin condition

$$(d^j)^T (Ax^k - b) = 0, \quad j = 1, \dots, k.$$

This method can become unstable if the matrix

$$A_k := ((d^i)^T Ad^j)_{i,j=1,\dots,k}$$

is nearly singular.

For $k \leq n$ let

$$\theta_1^{(k)} \leq \theta_2^{(k)} \leq \ldots \leq \theta_k^{(k)}$$

be the eigenvalues of the matrix A_k, the so called Ritz values of A with respect to $\mathcal{K}_k(r^0, A)$.

Since $\mathcal{K}_j(r^0, A) \subset \mathcal{K}_{j+1}(r^0, A)$, $j =, \ldots, n-1$, the minmax-characterization of the eigenvalues of a symmetric matrix yields, that for fixed j the sequence of the Ritz values decreases monotonely and is bounded below by the j smallest eigenvalue of the system matrix A.

$$\theta_j^{(j)} \geq \theta_j^{(j+1)} \geq \theta_j^{(j+2)} \geq \ldots \geq \lambda_j.$$

If the matrix A has exactly one negative eigenvalue (and if $(r^0)^T A r^0 > 0$) then there exists exactly one critical phase in the CG method, namely when the smallest Ritz value $\theta_1^{(k)}$ changes its sign. Every other sequence $\{\theta_j^{(k)}\}$ is bounded below by $\lambda_j > 0$.

Correspondingly there exist exactly m critical phases in the CG method if the matrix A has m negative und $p := n - m \geq m$ positive eigenvalues.

Consider

$$-\Delta u - 163.84u = 0 \text{ in } \Omega = (0,1)\times(0,1), \quad u = 0 \text{ on } \partial\Omega$$

Discretizing Δ with central differences with stepsize $h = 1/128$ yields a linear system

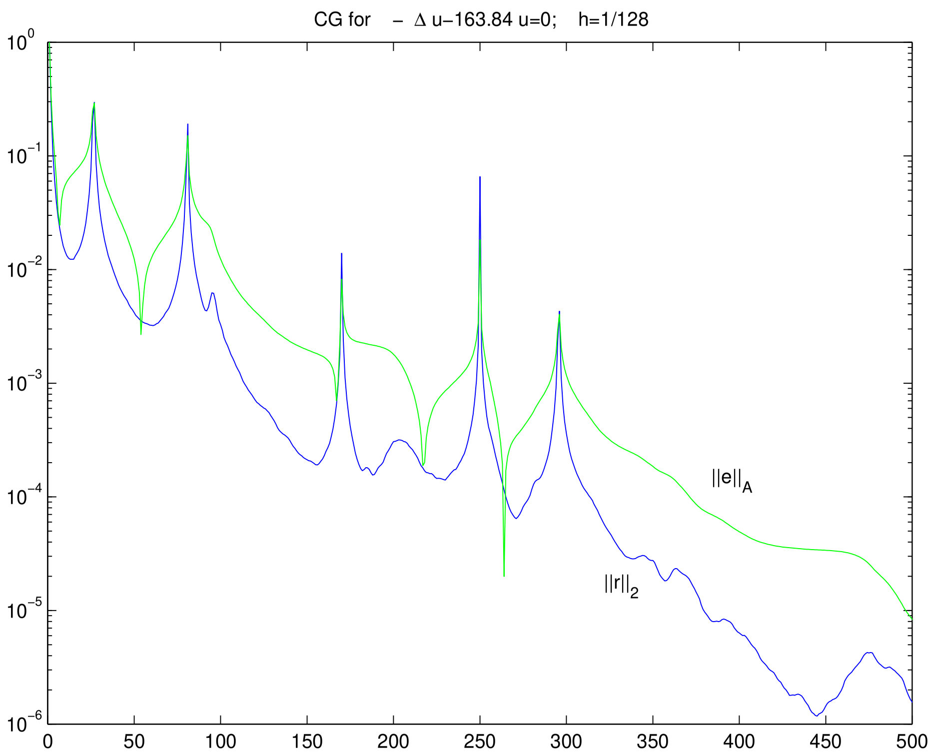
$$3.9U_{ij} - U_{i-1,j} - U_{i,j-1} - U_{i,j+1} - U_{i+1,j} = 0,\ i,j = 1,\dots,127,$$

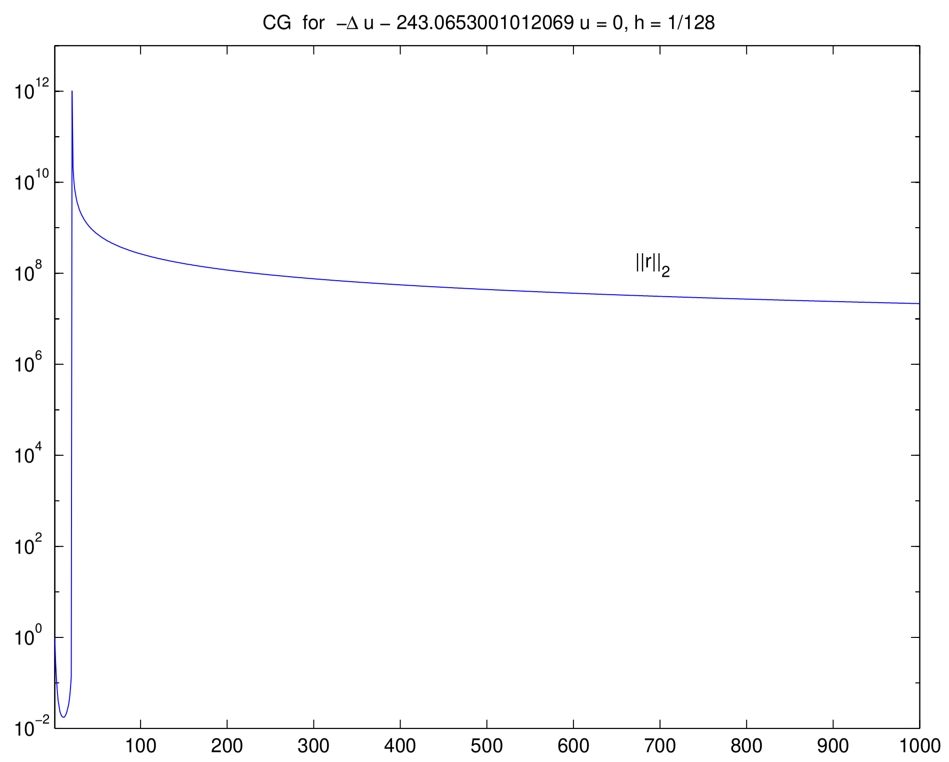
of dimension $n = 127^2 = 16129$.

The coefficient matrix has 8 negative eigenvalues

$$\begin{array}{ll} -8.795274784817014e-003 & -6.988549802753376e-003 \\ -6.988549802753343e-003 & -5.181824820689691e-003 \\ -3.978550749789028e-003 & -3.978550749789008e-003 \\ -2.171825767725394e-003 & -2.171825767725365e-003 \end{array}$$

5 of which are distinct.





Lanczos method

determines an orthonormal basis $\{q^1, \dots, q^k\}$ of the Krylov space

$$\mathcal{K}_k(r^0, A) := \text{span}\{r^0, Ar^0, \dots, A^{k-1}r^0\},\ \ k = 1, \dots, n,$$

such that

$$T_k := Q_k^T A Q_k, \quad Q_k := (q^1, \dots, q^k)$$

is tridiagonal.

Lanczos tridiagonalization

$q^0 = 0;\ k = 1;$
$\beta_0 = \text{norm}(r^0);$
while $r^{k-1} \neq 0$
  $q^k = r^{k-1}/\beta_{k-1};$
  $r^k = Aq^k;$
  $r^k = r^k - \beta_{k-1}q^{k-1};$
  $\alpha_k = (q^k)^T r^k;$
  $r^k = r^k - \alpha_k q^k;$
  $\beta_k = \text{norm}(r^k);$
  end

Then with $T_k = \text{tridiag}\{\beta_{j-1}, \alpha_j, \beta_j\}$

$$AQ_k = Q_k T_k + r^k (e^k)^T, \quad r^k = Aq^k - \alpha_k q^k - \beta_{k-1} q^{k-1}.$$

Consider the CG method for the linear system

$$Ax = b, \quad A \text{ symmetric and positive definite}$$

Let $x^0 = 0$ (else consider $Ay = b - Ax^0 =: \tilde{b}$)

The approximation x^k after k steps minimizes

$$\phi(x) := \frac{1}{2}x^T Ax - b^T x \quad \text{in } \mathcal{K}_k(r^0, A) = \mathcal{K}_k(b, A).$$

Restricting ϕ to $\mathcal{K}_k(b, A)$ yields

$$\phi_k(y) := \frac{1}{2} y^T Q_k^T A Q_k y - y^T Q_k^T b, \quad y \in \mathrm{IR}^k,$$

and the CG iterate x^k satisfies

$$T_k y^k = Q_k^T b, \quad x^k = Q_k y^k.$$

Disadvantage (at a first glance): All q^j, $j = 1, \ldots, k$, are needed to determine x^k.

T_k is SPD. Hence, it allows an LDL^T factorization $T_k = L_k D_k L_k^T$,

$$\begin{aligned} D_k &= \mathrm{diag}\{d_1, \ldots, d_k\},\ d_j > 0, \\ L_k &= \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ \mu_2 & 1 & 0 & \ldots & 0 & 0 \\ 0 & \mu_3 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \mu_k & 1 \end{pmatrix} \end{aligned}$$

Comparing the elements in $T_k = L_k D_k L_k^T$ yields

$d_1 = \alpha_1$;
for $j = 2 : k$
  $\mu_j = \beta_{j-1}/d_{j-1}$;
  $d_j = \alpha_j - \beta_{j-1}\mu_j$;
  end

With

$$z^k := L_k^T y^k, \quad C_k := Q_k L_k^{-T}$$

it holds

$$L_k D_k z^k = Q_k^T b, \quad x^k = C_k z^k,$$

and

$$C_k L_k^T = (c^1, \mu_2 c^1 + c^2, \dots, \mu_k c^{k-1} + c^k) = (q^1, \dots, q^k) = Q_k.$$

Since $L_k D_k$ is a lower triangular matrix and since $Q_k^T b = \|b\|_2 e^1$, one obtains the solution of

$$L_k D_k z^k = \|b\|_2 e^1$$

from the solution z^{k-1} of $L_{k-1} D_{k-1} z^{k-1} = \|b\|_2 e^1$ just adding the last component

$$\zeta_k = -\mu_k d_{k-1} \zeta_{k-1} / d_k.$$

CG method with Lanczos process

$r^0 = b - Ax^0;\ q^0 = 0;\ c^0 = 0;\ \beta_0 = \|r^0\|_2;$
for $k = 1, 2, \dots$ until convergence
  $q^k = r^{k-1}/\beta_{k-1};$
  $r^k = Aq^k - \beta_{k-1} q^{k-1};$
  $\alpha_k = (q^k)^T r^k;$
  $r^k = r^k - \alpha q^k;$
  $\beta_k = \|r^k\|_2;$
  if $k == 1$
    $d_1 = \alpha_1;\ \zeta_1 = \beta_0/d_1;\ c^1 = q^1;$
  else
    $\mu_k = \beta_{k-1}/d_{k-1};\ d_k = \alpha_k - \beta_{k-1}\mu_k;\ \zeta_k = -\mu_k d_{k-1}\zeta_{k-1}/d_k;$
    $c^k = q^k - \mu_k c^{k-1};$
    end
  $x^k = x^{k-1} + \zeta_k c^k;$ end

Conclusion: The CG approximations x^k can be obtained by the Lanczos method.

Cost of CG with Lanczos: 1 matrix-vector products
2 scalar products
5 _axpy

Storage requirements: 4 vectors

Let A be symmetric, but not necessarily positive definite

Lanczos' method does not need the definiteness of A, hence the projected system

$$T_k y^k = \|b\|_2 e^1$$

can be determined in a stable way.

The LDL^T factorization of T_k does not necessarily exist and can not be computed in a stable way.

Paige & Saunders (1975): Use the LQ factorization

$$T_k = \tilde{L}_k U_k, \quad U_k \text{ orthogonal}, \quad \tilde{L}_k \text{ lower triangular.}$$

Determine U_k as a product of Givens reflections

$$U_k = G_k \begin{pmatrix} G_{k-1} & 0 \\ 0^T & 1 \end{pmatrix} \cdot \ldots \cdot \begin{pmatrix} G_2 & O \\ O & I_{k-2} \end{pmatrix} =: G_k \begin{pmatrix} U_{k-1} & 0 \\ 0^T & 1 \end{pmatrix}$$

where

$$G_j = \begin{pmatrix} I_{j-2} & 0 & 0 \\ 0^T & c_j & s_j \\ 0^T & s_j & -c_j \end{pmatrix} \in \mathbb{R}^{(j,j)}, \quad c_j^2 + s_j^2 = 1.$$

$$\begin{aligned} T_k U_k^T &= \begin{pmatrix} T_{k-1} & \beta_{k-1} e^{k-1} \\ \beta_{k-1}(e^{k-1})^T & \alpha_k \end{pmatrix} \begin{pmatrix} U_{k-1}^T & 0 \\ 0^T & 1 \end{pmatrix} G_k \\ &= \begin{pmatrix} T_{k-1} U_{k-1}^T & \beta_{k-1} e^{k-1} \\ \beta_{k-1}(U_{k-1} e^{k-1})^T & \alpha_k \end{pmatrix} G_k \\ &= \begin{pmatrix} \tilde{L}_{k-1} & \beta_{k-1} e^{k-1} \\ \beta_{k-1}(e^{k-1})^T U_{k-1}^T & \alpha_k \end{pmatrix} G_k, \end{aligned}$$

where G_k is chosen such that β_{k-1} at the position $(k-1, k)$ is annihilated.

$$\begin{aligned} \beta_{k-1}(e^{k-1})^T U_{k-1}^T &= \beta_{k-1}(e^{k-1})^T \begin{pmatrix} U_{k-2}^T & 0 \\ 0^T & 1 \end{pmatrix} G_{k-1} \\ &= \beta_{k-1}(e^{k-1})^T G_{k-1} = (0, \ldots, 0, \beta_{k-1} s_{k-1}, -\beta_{k-1} c_{k-1}). \end{aligned}$$

Multiplying a matrix with G_k from the right only the last two columns are combined. Hence, in $\tilde{L}_k = T_k U_k^T$ only 3 diagonals are different from zero, i.e. $\tilde{L}_k$ has the form

$$\tilde{L}_k = \begin{pmatrix} \gamma_1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \delta_1 & \gamma_2 & 0 & 0 & \dots & 0 & 0 & 0 \\ \varepsilon_1 & \delta_2 & \gamma_3 & 0 & \dots & 0 & 0 & 0 \\ 0 & \varepsilon_2 & \delta_3 & \gamma_4 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \varepsilon_{k-2} & \delta_{k-1} & \tilde{\gamma}_k \end{pmatrix}$$

In the transition from $\tilde{L}_k$ to

$$\begin{aligned} \tilde{L}_{k+1} &= T_{k+1} \begin{pmatrix} U_k^T & 0 \\ 0^T & 1 \end{pmatrix} G_{k+1} \\ &= \begin{pmatrix} \tilde{L}_k & \beta_k e^k \\ \beta_k (e^k)^T U_k^T & \alpha_k \end{pmatrix} G_{k+1} \\ &= \begin{pmatrix} L_k & 0 \\ 0, \dots, 0, \varepsilon_{k-1}, \delta_k & \tilde{\gamma}_{k+1} \end{pmatrix} \end{aligned}$$

in the leading principal (k,k) matrix $\tilde{L}_k$ only the diagonal element $\tilde{\gamma}_k$ is changed.

With

$$\tilde{W}_k := (w^1, \dots, w^{k-1}, \tilde{w}^k) := Q_k U_k^T$$

and

$$\tilde{z}_k := (\zeta_1, \dots, \zeta_{k-1}, \tilde{\zeta}_k)^T := U_k y^k$$

the linear system

$$T_k y^k = \|b\|_2 e^1, \quad x^k = Q_k y^k$$

can be rewritten as

$$\tilde{L}_k \tilde{z}_k = \|b\|_2 e^1, \quad x^k = \tilde{W}_k \tilde{z}^k.$$

(As in the computation of the CG approximation with the Lanczos method for the SPD case) this decomposition can be used to solve the projected problem.

For $k = 2$ the solution $\tilde{z}^2$ is

$$\zeta_1 = \|b\|_2/\gamma_1, \quad \tilde{\zeta}_2 = -\delta_1\zeta_1/\tilde{\gamma}2.$$

If $\tilde{z}^{k-1}$ is already known the leading $k-2$ components of $\tilde{z}^k$ coincide with those of $\tilde{z}^{k-1}$, and it holds that

$$\begin{aligned} \zeta_{k-1} &= \tilde{\zeta}_{k-1}\tilde{\gamma}_{k-1}/\gamma_{k-1} = c_k\tilde{\zeta}_{k-1} \\ \tilde{\zeta}_k &= -(\varepsilon_{k-2}\zeta_{k-2} + \delta_{k-1}\zeta_{k-1})/\tilde{\gamma}_k. \end{aligned}$$

From

$$\begin{aligned}\tilde{W}_k &= (w^1, \dots, w^{k-1}, \tilde{w}^k) = (q^1, \dots, q^k)U_k^T \\ &= (q^1, \dots, q^k) \begin{pmatrix} U_{k-1}^T & 0 \\ 0^T & 1 \end{pmatrix} G_k = ((q^1, \dots, q^{k-1})U_{k-1}^T, q^k)G_k \\ &= (w^1, \dots, w^{k-2}, \tilde{w}^{k-1}, q^k) \begin{pmatrix} I_{k-2} & 0 & 0 \\ 0^T & c_k & s_k \\ 0^T & s_k & -c_k \end{pmatrix}\end{aligned}$$

we obtain w^{k-1} and $\tilde{w}^k$ in the k-th step from

$$(w^{k-1}, \tilde{w}^k) = (\tilde{w}^{k-1}, q^k) \begin{pmatrix} c_k & s_k \\ s_k & -c_k \end{pmatrix}, \quad \tilde{w}^1 = q^1.$$

Although the LQ factorization is determined in a stable way (even for indefinite matrices) the algorithm should not be implemented directly since intermediate matrices $\tilde{L}_k$ may become singular or $|\tilde{\gamma}_k|$ can become very small as a Ritz value passes the origin.

Since $\gamma_k = \sqrt{\beta_k^2 + \tilde{\gamma}_k^2} \neq 0$ as long as $\beta_k \neq 0$ (i.e. es long as the Lanczos process did not detect an invariant subspace of A the matrix L_k is nonsingular. Hence, the linear system

$$L_k z^k = \|b\|_2 e^1$$

for every k has a unique solution z^k, which can be determined in a more stable way than $\tilde{z}^k$.

Paige and Saunders propose to update the vectors

$$\hat{x}^k := W_k z^k := \hat{x}^{k-1} + \zeta_k w^k$$

instead of x^k from which we can get the CG approximation

$$x^{k+1} = \hat{x}^k + \tilde{\zeta}_{k+1} \tilde{w}^{k+1}$$

after the method has converged.

This method is called SYMMLQ.

SYMMLQ

$c_1 = -1;\ s_1 = 0;\ c_0 = 1;\ s_0 = 0;\ \zeta_0 = -1;\ \zeta_{-1} = 0;$
$\hat{x}^0 = x^0;\ \tilde{w}^0 = 0;\ q^0 = 0;\ \tilde{q} = b - Ax^0;\ \beta_0 = \|\tilde{q}\|_2;\ q^1 = \tilde{q}/\beta_0;$
for $k = 1, 2, \dots$ until convergence

$$\begin{aligned}
&\tilde{q} = Aq^k - \beta_{k-1} q^{k-1};\\
&\alpha_k = \tilde{q}^T q^k;\\
&\tilde{q} = \tilde{q} - \alpha_k q^k;\\
&\beta_k = \|\tilde{q}\|_2;\\
&\tilde{\gamma}_k = -c_k \alpha_k - c_{k-1} s_k \beta_{k-1};\ \gamma_k = \sqrt{\tilde{\gamma}_k^2 + \beta_k^2};\\
&\delta_{k-1} = s_k \alpha_k - c_k c_{k-1} \beta_{k-1};\ \varepsilon_{k-2} = s_{k-1} \beta_{k-1};\\
&c_{k+1} = \tilde{\gamma}_k / \gamma_k;\ s_{k+1} = \beta_k / \gamma_k;\\
&\tilde{w}^k = s_k \tilde{w}^{k-1} - c_k q^k;\\
&\zeta_k = -(\varepsilon_{k-2} \zeta_{k-2} + \delta_{k-1} \zeta_{k-1}) / \gamma_k;\\
&q^{k+1} = \tilde{q}/\beta_k;\ \hat{x}^k = \hat{x}^{k-1} + \zeta_k (c_{k+1} \tilde{w}^k + s_{k+1} q^{k+1});\ \text{end}
\end{aligned}$$

$x^{k+1} = \hat{x}^k + \zeta_k \tilde{w}^k / c_{k+1};$

Cost of SYMMLQ: 1 matrix-vector produkt
2 scalar products
7 _axpy

Storage requirements: 5 vectors

The iteration is stopped, if the norm of the residuum is small enough. Neither the residuum nor its norm is determined in the given algorithm. However, it can be shown that

$$r^k = -\frac{1}{c_{k+1}} s_2 s_3 \dots s_{k+1} \|b\|_2 q^{k+1}$$

from which we can update $\|r^k\|_2$.

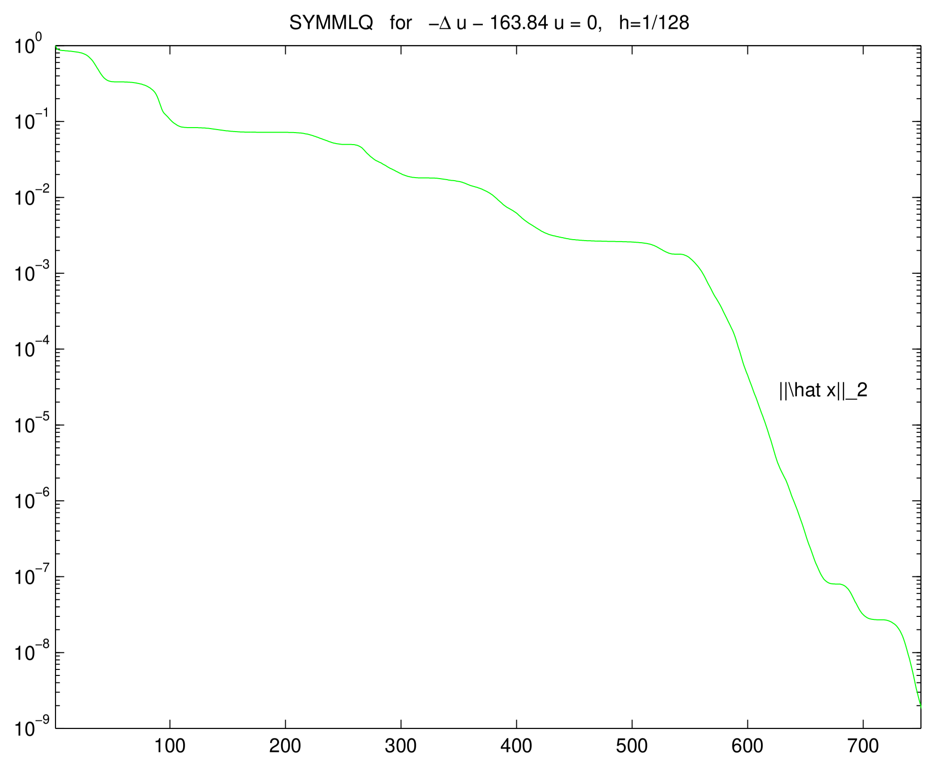
Theorem

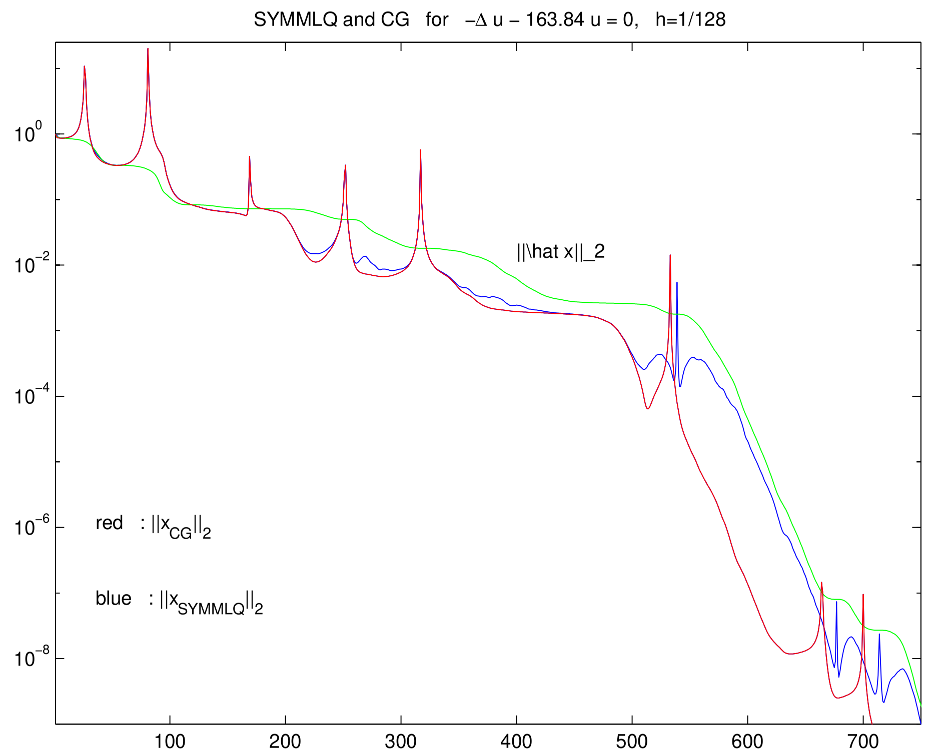
The approximation $\hat{x}^k$ is the unique solution of the minimization problem

$$\|e\|_2 = \text{min!},\ x \in A\mathcal{K}_k(b, A) = \text{span}\{Ab, A^2 b, \dots, A^k b\},$$

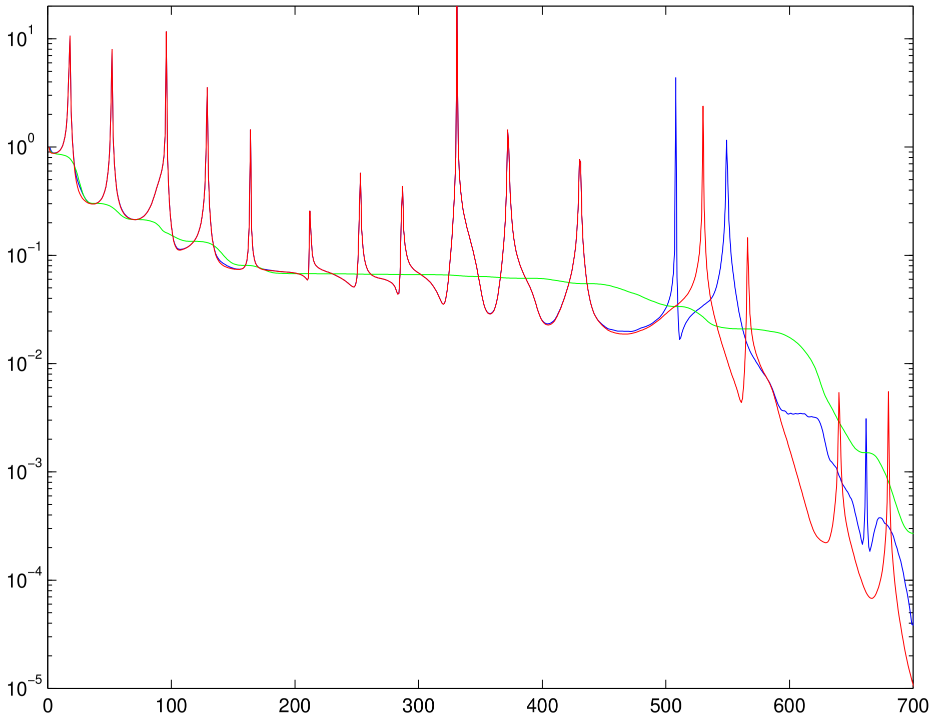
where $e := A^{-1}b - x$ denotes the error of x.

The CG iterates x^k do not minimize an error norm in the indefinite case. However, they are usually better approximations to the solution than the vectors $\hat{x}^k$.

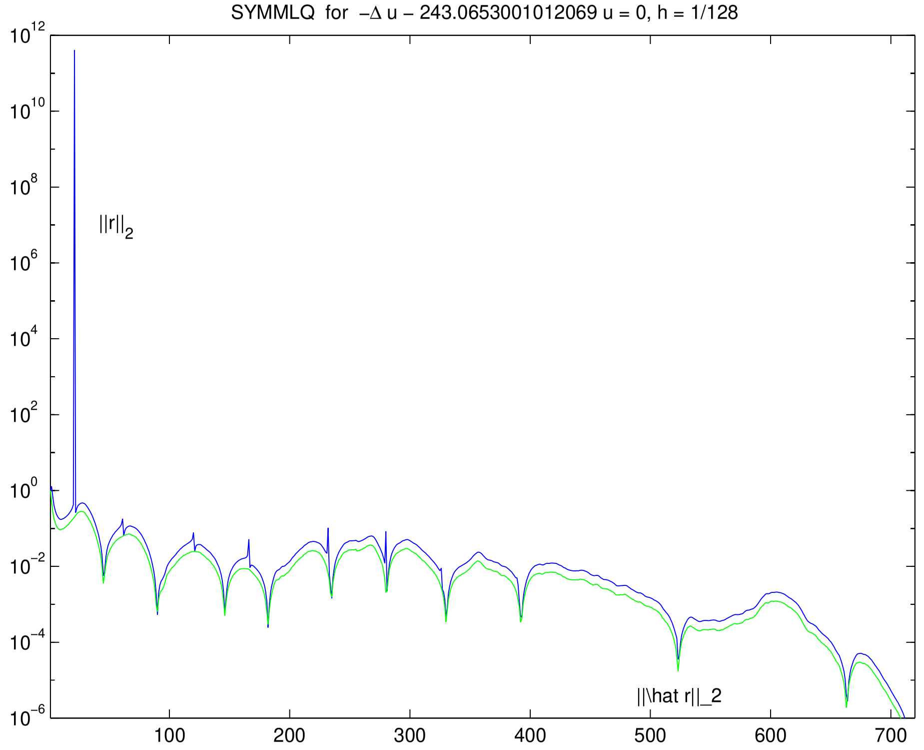

SYMMLQ for −Δ u − 163.84 u = 0, h=1/128
||\hat x||_2
10^0
10^−1
10^−2
10^−3
10^−4
10^−5
10^−6
10^−7
10^−8
10^−9
100
200
300
400
500
600
700

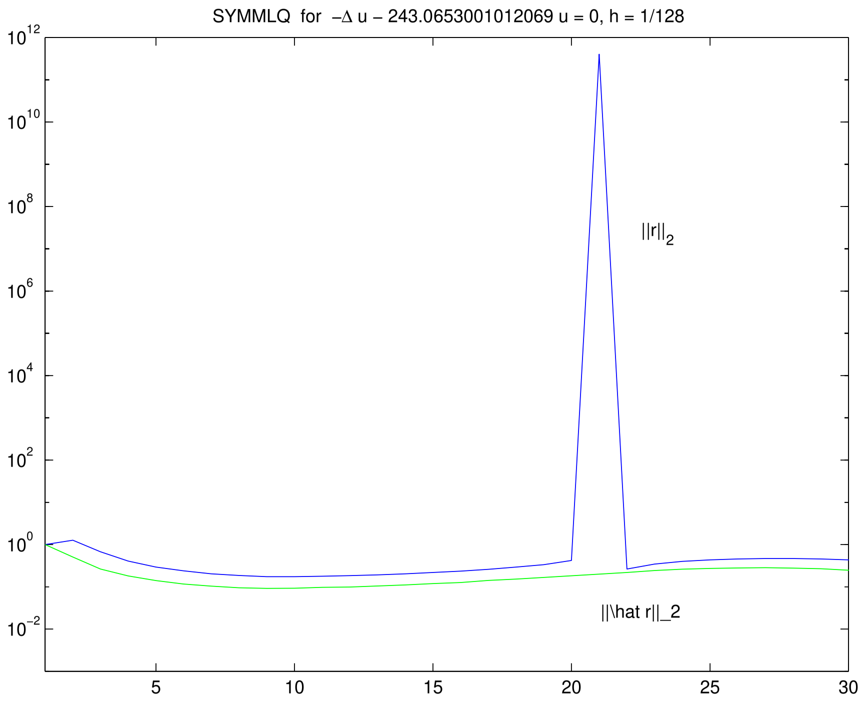

SYMMLQ and CG for −Δ u − 163.84 u = 0, h=1/128
||\hat x||_2
red : $||x_{CG}||_2$
blue : $||x_{SYMMLQ}||_2$
10^0
10^−2
10^−4
10^−6
10^−8
100
200
300
400
500
600
700

SYMMLQ for $-\Delta u - 327.68\, u = 0$, h=1/128



SYMMLQ for $-\Delta u - 243.0653001012069\, u = 0$, h = 1/128





A further stable method for solving indefinite systems is a modification of the

Method of conjugate residuals (CR method)

Let A be symmetric and positive definite. Determine $x^k \in x^0 + \mathcal{K}_k(r^0, A)$ such that the error with respect to the A^2 norm is minimal.

As for the CG method the approximations x_k can be obtained from a sequence of one dimensional line searches where the search directions satisfy a 3-term-recurrence.

CR method

$r^0 = b - Ax^0;$
$t^0 = Ar^0;$
$\alpha_0 = (t^0)^T r^0;$
$d^1 = r^0;\ s^1 = t^0;$
for $k = 1, 2, \ldots$ until convergence
$\tau_k = \alpha_{k-1}/(s^k)^T s^k;$
$x^k = x^{k-1} + \tau_k d^k;$
$r^k = r^{k-1} - \tau_k s^k;$
$\beta_k = 1/\alpha_{k-1};$
$t^k = Ar^k;$
$\alpha_k = (t^k)^T r^k;$
$\beta_k = \alpha_k \beta_k;$
$d^{k+1} = r^k + \beta_k d^k;$
$s^{k+1} = t^k + \beta_k s^k;$ end

Cost of CR: 1 matrix-vector product
2 scalar products
4 _axpy

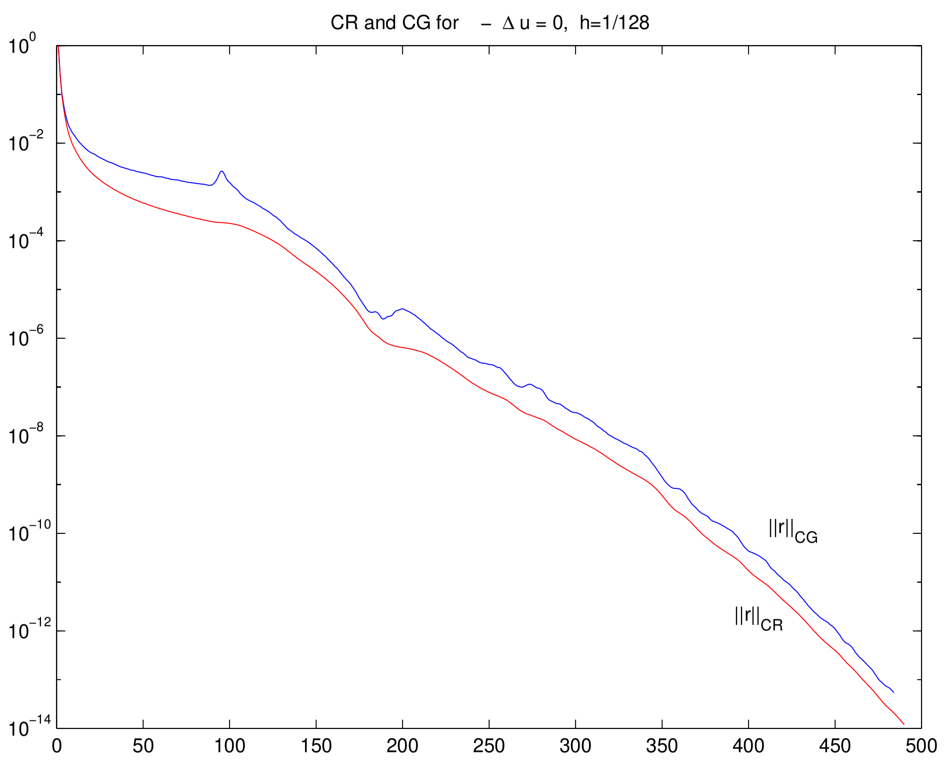
Storage requirements: 5 Vectors

For $x^* := A^{-1}b$ it holds

$$\|x - x^*\|_{A^2}^2 = \|A(x - x^*)\|_2^2 = \|Ax - b\|_2^2 = \|r\|_2^2$$

Hence, the CR method minimizes the Euclidean norm of the residuum in $x^0 + \mathcal{K}_k(r^0, A)$. Moreover, it can be shown that the residuals $r^k := b - Ax^k$ determined by the CR method are A-conjugate (this explains the name of the method).

If A is symmetric but indefinite the approach of the CR method (to minimize the residuum on $\mathcal{K}_k(r^0, A)$) is still reasonable. However, the CR algorithm can break down with $\alpha_k = (r^k)^T Ar^k = 0$.



Paige & Saunders (1975): MINRES

Use the Lanczos process to tridiagonalize A and the LQ factorization of T_k to solve the normal equations

$$Q_k^T A^2 Q_k y^k = Q_k^T A b, \quad x^k = Q_k y^k$$

of the least squares problem

$$\|AQ_k y - b\|_2^2 = \text{min!}$$

Again we assume without restricting generality $x^0 = 0$.

Since r^k is orthogonal to $q^1, \dots, q^k$ it holds

$$\begin{aligned} Q_k^T A^2 Q_k &= (Q_k T_k + r^k (e^k)^T)^T (Q_k T_k + r^k (e^k)^T) \\ &= T_k^2 + \beta_k^2 e^k (e^k)^T \\ Q_k^T A b &= \beta_0 Q_k^T A q^1 = \beta_0 Q_k^T A Q_k e^1 \\ &= \beta_0 Q_k^T (Q_k T_k + (r^k)^T e^k) e^1 = \beta_0 T_k e^1 \end{aligned}$$

From the LQ factorization of $T_k = \tilde{L}_k U_k$ which can be obtained in the same manner as in the SYMMLQ method currently along with the Lanczos process one gets the Cholesky factorization

$$T_k^2 + \beta_k e^k (e^k)^T = \tilde{L}_k \tilde{L}_k^T + \beta_k^2 e^k (e^k)^T = L_k L_k^T$$

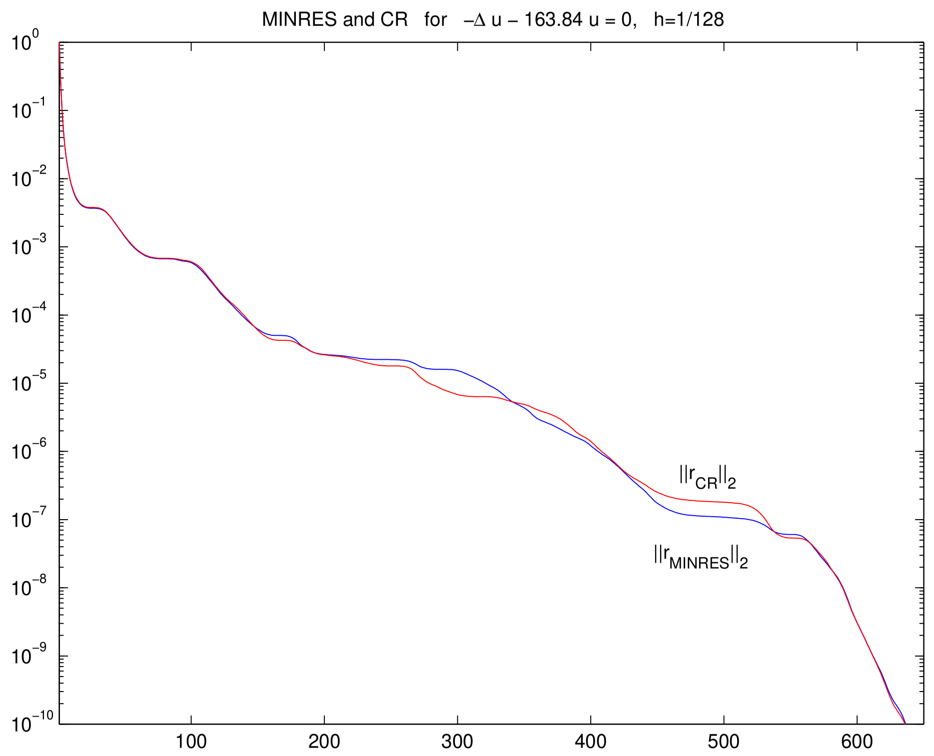
without computing T_k^2 explicitly.

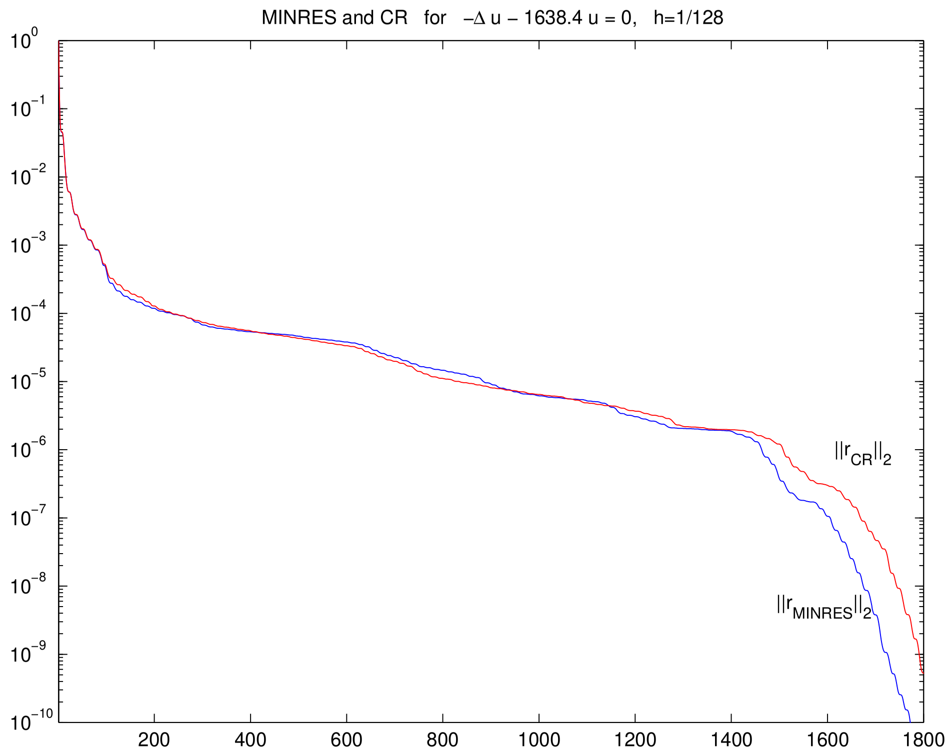
MINRES

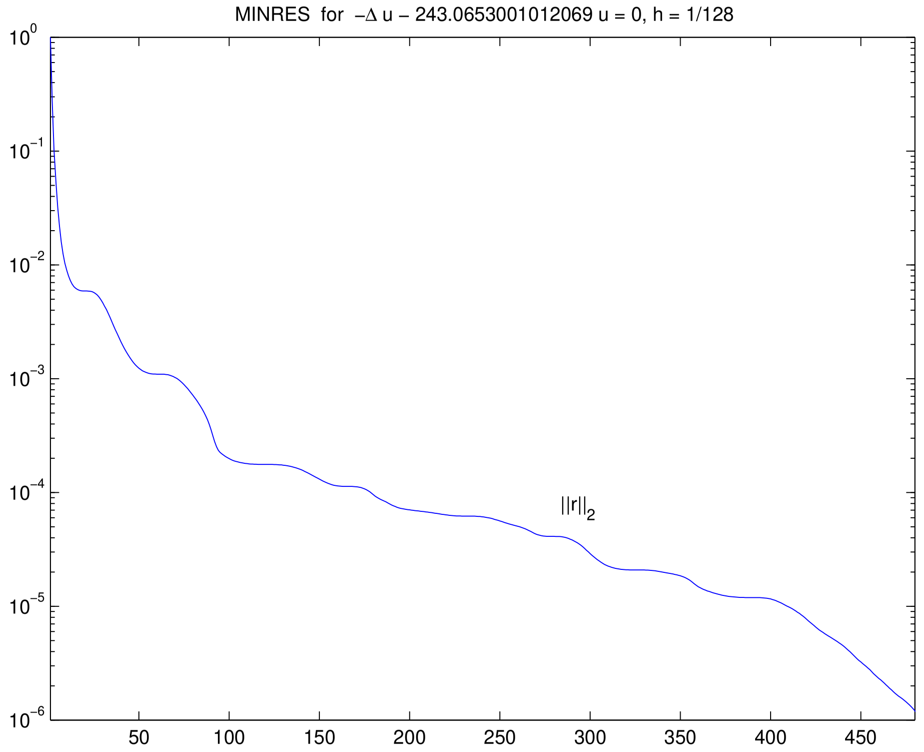
$c_1 = -1;\ s_1 = 0;\ c_0 = 1;\ s_0 = 0;\ q^0 = 0;\ v^0 = 0;\ v^{-1} = 0;$
$\tilde{q} = b - Ax^0;\ \beta_0 = \|\tilde{q}\|_2;\ q^1 = \tilde{q}/\beta_0;\ \eta_1 = \beta_0;$
for $k = 1, 2, \dots$ until convergence
$\quad \tilde{q} = Aq^k - \beta_{k-1} q^{k-1};$
$\quad \alpha_k = \tilde{q}^T q^k;$
$\quad \tilde{q} = \tilde{q} - \alpha_k q^k;$
$\quad \beta_k = \|\tilde{q}\|_2;$
$\quad \tilde{\gamma}_k = -c_k \alpha_k - c_{k-1} s_k \beta_{k-1};\ \gamma_k = \sqrt{\tilde{\gamma}_k^2 + \beta_k^2};$
$\quad \delta_{k-1} = s_k \alpha_k - c_k c_{k-1} \beta_{k-1};\ \varepsilon_{k-2} = s_{k-1} \beta_{k-1};$
$\quad c_{k+1} = \tilde{\gamma}_k / \gamma_k;\ s_{k+1} = \beta_k / \gamma_k;$
$\quad v^k = (q^k - \varepsilon_{k-2} v^{k-2} - \delta_{k-1} v^{k-1}) / \gamma_k;$
$\quad x^k = x^{k-1} + c_{k+1} \eta_k v^k;$
$\quad q^{k+1} = \tilde{q} / \beta_k;$
$\quad \eta_{k+1} = s_{k+1} \eta_k;$ end

Cost of MINRES: 1 matrix-vector product
2 scalar products
7 _axpy

Storage requirements: 6 Vectors




MINRES and CR for −Δ u − 1638.4 u = 0, h=1/128
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
10^{-10}
200
400
600
800
1000
1200
1400
1600
1800
$\|r_{CR}\|_2$
$\|r_{MINRES}\|_2$


MINRES for −Δ u − 243.0653001012069 u = 0, h = 1/128
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
50
100
150
200
250
300
350
400
450
$\|r\|_2$

Error bounds for MINRES

$x^k \in x^0 + \text{span}\{x^0, Ax^0, \dots, A^{k-1}x^0\}$ is determined such that $\|r^k\|_2$ is minimal in

$$r^0 + \text{span}\{Ar^0, \dots, A^k r^0\},$$

i.e. $r^k = p_k^M(A)r^0$ with $p_k^M \in \Pi_k$, $p_k^M(0) = 1$,

$$\|r^k\|_2 = \min_{p_k \in \Pi_k,\ p_k(0)=1} \|p_k(A)r^0\|_2.$$

For $A = U\Lambda U^T$, $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$, it follows

$$\|r^k\|_2 \le \min_{p_k \in \Pi_k,\ p_k(0)=1} \|p_k(\Lambda)\|_2 \|r^0\|_2,$$

i.e.

$$\frac{\|r^k\|_2}{\|r^0\|_2} \le \min_{p_k \in \Pi_k,\ p_k(0)=1} \max_{j=1,\dots,n} |p_k(\lambda_j)|.$$

If A is positive definite (CR method!) we obtain in the same way as for the CG method

$$\frac{\|r^k\|_2}{\|r^0\|_2} \le 2\left(\frac{\sqrt{\text{cond}_2(A)}-1}{\sqrt{\text{cond}_2(A)}+1}\right)^k.$$

For indefinite problems typically only very few ($\ell \ll n$) eigenvalues $\lambda_1 \le \dots \le \lambda_\ell < 0$ are negative, and most of the eigenvalues $0 < \lambda_{\ell+1} \le \dots \le \lambda_n$ are positive.

If

$$p_k(\lambda) = q_\ell(\lambda)\left[T_{k-\ell}\left(\frac{2\lambda-\lambda_{\ell+1}-\lambda_n}{\lambda_n-\lambda_{\ell+1}}\right) \Big/ T_{k-\ell}\left(\frac{-\lambda_{\ell+1}-\lambda_n}{\lambda_n-\lambda_{\ell+1}}\right)\right]$$

with $q_\ell(\lambda) = (-1)^\ell \prod_{j=1}^\ell (\lambda_j - \lambda) / \prod_{j=1}^\ell \lambda_j$ and Chebyshev polynomials $T_{k-\ell}$, one gets

$$\frac{\|r^k\|_2}{\|r^0\|_2} \le q_\ell(\lambda_n) 2 \left(\frac{\sqrt{\kappa_{n-\ell-1}}-1}{\sqrt{\kappa_{n-\ell-1}}+1}\right)^{k-\ell}, \quad \kappa_{n-\ell-1} := \frac{\lambda_n}{\lambda_{\ell+1}}.$$

The last error bound demonstrates that it is reasonable to use positive definite preconditioners even if the problem is indefinite.

To preserve symmetry of the problem we consider

$$C^{-1}AC^{-T}y = C^{-1}b, \; x := C^T y.$$

From

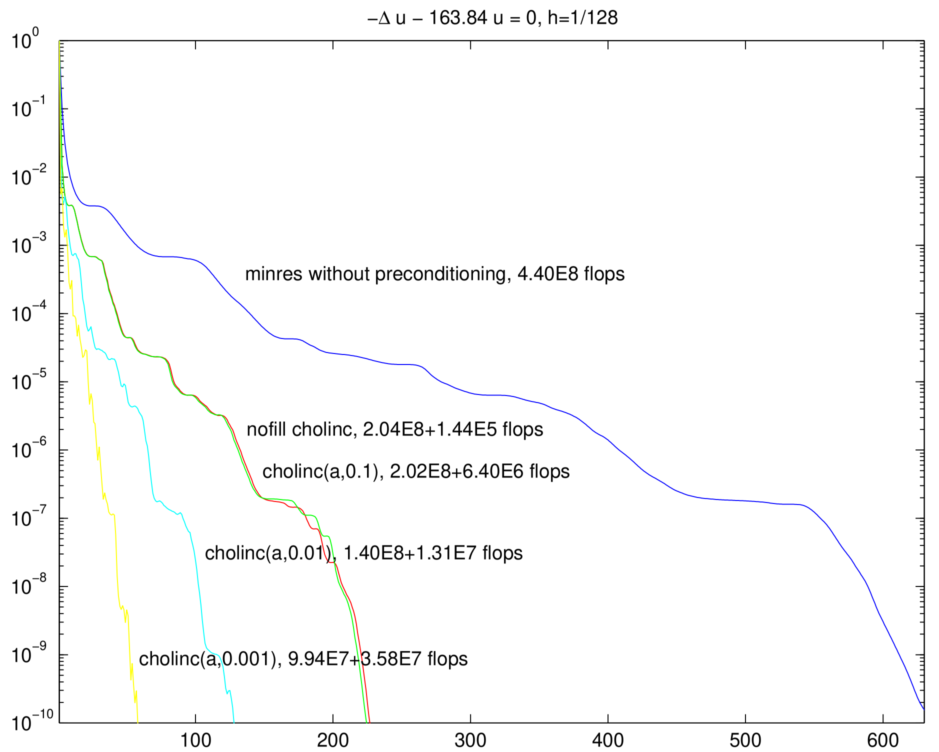
$$\begin{aligned} C^{-1}AC^{-T}z = \mu z \quad &\Longleftrightarrow \quad C^{-1}AY = \mu C^T y, \; y = C^{-T}z \\ &\Longleftrightarrow \quad C^{-T}C^{-1}Ay = (CC^T)^{-1}Ay = \mu y \end{aligned}$$

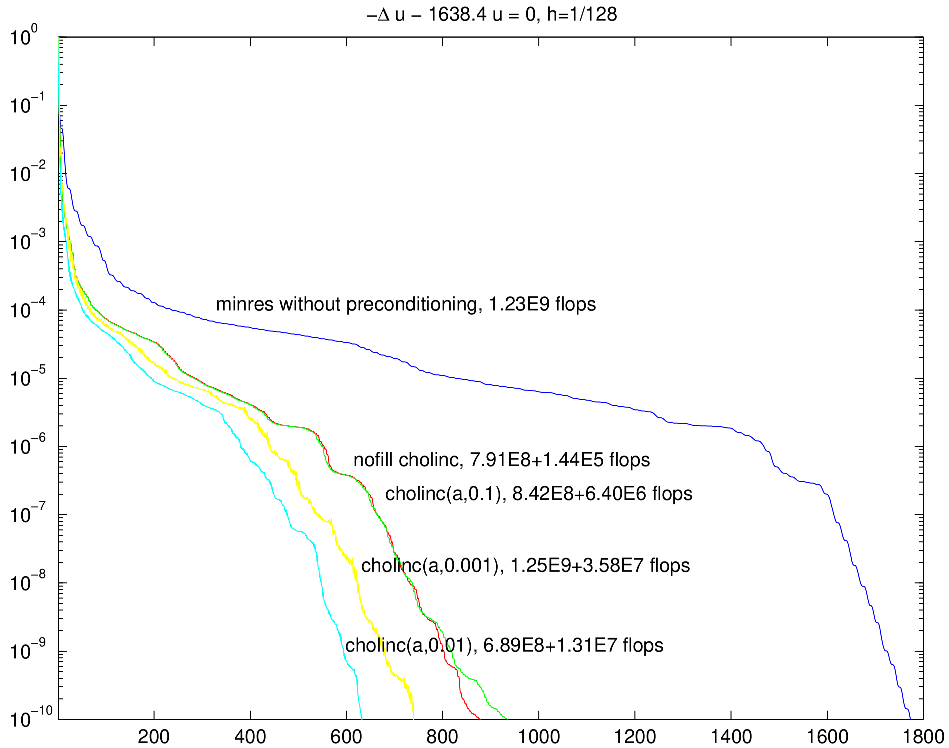
we obtain that the Cholesky factor C of the preconditioner $M := CC^T$ should be chosen such that:

$$\kappa_{n-\ell-1}(M^{-1}A) \ll \kappa_{n-\ell-1}(A)$$

MINRES with preconditioning

$c_1 = -1;\ s_1 = 0;\ c_0 = 1;\ s_0 = 0;\ q^0 = 0;\ v^0 = 0;\ v^{-1} = 0;$
$q = b - Ax;\ w = M^{-1}q;\ \beta_0 = \sqrt{q^T w};$
$q^1 = q/\beta_0;\ w^1 = w/\beta_0;\ \eta_1 = \beta_0;$
for $k = 1, 2, \dots$ until convergence
- $q = Aw^k - \beta_{k-1}q^{k-1};\ \alpha_k = q^T w^k;$
- $q = q - \alpha_k q^k;$
- $w^{k+1} = M^{-1}q;\ \beta_k = \sqrt{q^T w^{k+1}};$
- $\tilde{\gamma}_k = -c_k\alpha_k - c_{k-1}s_k\beta_{k-1};\ \gamma_k = \sqrt{\tilde{\gamma}_k^2 + \beta_k^2};$
- $\delta_{k-1} = s_k\alpha_k - c_k c_{k-1}\beta_{k-1};\ \varepsilon_{k-2} = s_{k-1}\beta_{k-1};$
- $c_{k+1} = \tilde{\gamma}_k/\gamma_k;\ s_{k+1} = \beta_k/\gamma_k;$
- $v^k = (w^k - \varepsilon_{k-2}v^{k-2} - \delta_{k-1}v^{k-1})/\gamma_k;$
- $x^k = x^{k-1} + c_{k+1}\eta_k v^k;$
- $q^{k+1} = q/\beta_k;\ w^{k+1} = w^{k+1}/\beta_k;$
- $\eta_{k+1} = s_{k+1}\eta_k;$ end


−Δ u − 163.84 u = 0, h=1/128
minres without preconditioning, 4.40E8 flops
nofill cholinc, 2.04E8+1.44E5 flops
cholinc(a,0.1), 2.02E8+6.40E6 flops
cholinc(a,0.01), 1.40E8+1.31E7 flops
cholinc(a,0.001), 9.94E7+3.58E7 flops
100
200
300
400
500
600


−Δ u − 1638.4 u = 0, h=1/128
minres without preconditioning, 1.23E9 flops
nofill cholinc, 7.91E8+1.44E5 flops
cholinc(a,0.1), 8.42E8+6.40E6 flops
cholinc(a,0.001), 1.25E9+3.58E7 flops
cholinc(a,0.01), 6.89E8+1.31E7 flops
200
400
600
800
1000
1200
1400
1600
1800