On prescribing CG convergence

Gérard MEURANT

HHXXI, June 2022

Conjugate Gradient (CG)

input A symmetric positive definite, b, x_0
$r_0 = b - Ax_0$
$p_0 = r_0$
for $k = 1, \ldots$ until convergence **do**

$$\gamma_{k-1} = \frac{r_{k-1}^T r_{k-1}}{p_{k-1}^T A p_{k-1}}$$

$$x_k = x_{k-1} + \gamma_{k-1} p_{k-1}$$

$$r_k = r_{k-1} - \gamma_{k-1} A p_{k-1}$$

$$\delta_k = \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}}$$

$$p_k = r_k + \delta_k p_{k-1}$$

end for

CG minimizes the A-norm of the error

$$\|\varepsilon_k\|_A = (A(x - x_k), x - x_k)^{1/2}$$

which is decreasing

We would like to construct matrices A and right-hand sides b such that the residual norms $\|r_k\|$ have prescribed positive values

The answer to this question was (almost) given by Hestenes and Stiefel page 432

M.R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Nat. Bur. Standards, v 49 n 6 (1952), pp. 409–436

Theorem 18.3

There is no restriction whatever on the positive constants a_i, b_i (*our* γ_k *and* δ_k) in the cg-process, that is, given two sequences of positive numbers $a_0, \ldots, a_{n-1}$ and $b_0, \ldots, b_{n-1}$, there is a symmetric positive definite matrix A and a vector r_0 such that the cg-algorithm applied to A, r_0 yield the given numbers ...

Furthermore, the formula

$$b_i = \frac{\|r_{i+1}\|^2}{\|r_i\|^2}$$

shows that there is no restriction at all on the behavior of the length of the residual vector during the cg-process

Hence, it seems that we are done!

However, we will show that we can construct matrices and right-hand sides such that we can prescribe the residual norms and also the A-norms of the error

CG is related to the Lanczos algorithm. It implicitly constructs a Cholesky factorization of the Lanczos tridiagonal matrix

We will construct symmetric tridiagonal matrices T such that CG (with $x_0 = 0$) yields prescribed relative residual norms with the right-hand side e_1
Then, the linear system $Ax = b$ is obtained from $A = VTV^T$ and $b = Ve_1$ where V is any orthonormal matrix

Strangely enough, let us start with the nonsymmetric case...

The nonsymmetric case

We consider the Full Orthogonalization Method (FOM) of Y. Saad for solving nonsymmetric linear systems

FOM $\equiv$ Lanczos $\equiv$ CG when A is symmetric

Prescribing the residual norms for FOM was studied some years ago

We can construct linear systems for which the matrix has prescribed eigenvalues and such that FOM delivers prescribed residual norms and gives also prescribed (harmonic) Ritz values at all iterations, see

J. Duintjer Tebbens and G. Meurant, *Any Ritz value behavior is possible for Arnoldi and for GMRES*, SIAM J. Matrix Anal. Appl., v 33, n 3 (2012), pp. 958–978

Assume FOM does not stop early, then we get

$$AV = VH$$

V is the orthonormal matrix whose columns are the Arnoldi basis vectors and H is an unreduced upper Hessenberg matrix with positive subdiagonal entries

$$H = UCU^{-1}$$

where U is upper triangular with $u_{1,1} = 1$ and C is the companion matrix corresponding to the eigenvalues of A (and H)

$$U = \begin{pmatrix} e_1 & He_1 & \cdots & H^{n-1}e_1 \end{pmatrix}$$

It was proved (see JDT-GM) that the inverses of the absolute values of the entries of the first row of U^{-1} are equal to the FOM relative residual norms $\|r_k^F\|/\|r_0\|$

Note that the other entries of U^{-1} can be chosen to prescribe the Ritz values but we won't use that in the symmetric case

We use results in

A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, in Recent advances in iterative methods, G.H. Golub, A. Greenbaum and M. Luskin Eds., Springer (1994), pp. 95–118

Theorem

Let w_i, $i = 1, \ldots, k$ be an orthonormal basis for $A\mathcal{K}_k(A, v)$ with $k \leq n$, W be the matrix whose columns are the vectors w_i, $i = 1, \ldots, n$ and H the upper Hessenberg matrix such that $AW = WH$. Then, $A\mathcal{K}_k(A, v)$ and $B\mathcal{K}_k(B, v)$ are the same for $k = 1, \ldots, n$ if and only if

$$B = WRHW^*$$

where R is any nonsingular upper triangular matrix

Let $k = n$ and $Y = RH$

Then $R^{-1} = HY^{-1}$ and the matrix HY^{-1} must be upper triangular

From G and S, HX is upper triangular if and only if

$$\mathcal{L}(X) = \mathcal{L}(H^{-1})D$$

where $\mathcal{L}(F)$ is the lower triangular matrix whose lower triangular part is the same as for the matrix F, D is a diagonal matrix and $\mathcal{L}(H^{-1})$ has no zero column

It means that the lower triangular part of $X = Y^{-1}$ is the lower triangular part of H^{-1} with any column scaling

It is sufficient to construct a lower triangular matrix X whose lower triangular part is the same as that of H^{-1}

Let $H = UCU^{-1}$ with

$$U^{-1} = \begin{pmatrix} 1 & \hat{\nu}^T \\ 0 & \hat{U}^{-1} \end{pmatrix} \Rightarrow U = \begin{pmatrix} 1 & -\hat{\nu}^T \hat{U} \\ 0 & \hat{U} \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & \dots & 0 & -\alpha_1 \\ 0 & 1 & \dots & 0 & -\alpha_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & -\alpha_{n-1} \end{pmatrix}$$

Let $\hat{\alpha} = \begin{pmatrix}\alpha_1 & \cdots & \alpha_{n-1}\end{pmatrix}^T$

Since $\alpha_0 \neq 0$, C is nonsingular and its inverse is

$$C^{-1} = \begin{pmatrix} -\hat{\alpha}/\alpha_0 & I_{n-1} \\ -1/\alpha_0 & 0 \end{pmatrix}$$

We denote the entries of the first column as

$$\beta_1 = -\alpha_1/\alpha_0, \quad \hat{\beta}^T = \begin{pmatrix} -\alpha_2/\alpha_0 & \cdots & -\alpha_{n-1}/\alpha_0 & -1/\alpha_0 \end{pmatrix}$$

Let E be the matrix of order $n-1$ which is zero except for the first upper diagonal whose entries are equal to 1,

$$C^{-1} = \begin{pmatrix} \beta_1 & e_1^T \\ \hat{\beta} & E \end{pmatrix}$$

$$H^{-1} = UC^{-1}U^{-1}$$

Theorem

Using the previous notation the inverse of H is

$$H^{-1} = \begin{pmatrix} \beta_1 - \hat{\nu}^T \hat{U}\hat{\beta} & (\beta_1 - \hat{\nu}^T \hat{U}\hat{\beta})\hat{\nu}^T + (e_1^T - \hat{\nu}^T \hat{U} E)\hat{U}^{-1} \\ \hat{U}\hat{\beta} & \hat{U}\hat{\beta}\hat{\nu}^T + \hat{U} E \hat{U}^{-1} \end{pmatrix}$$

It shows that the lower triangular part of H^{-1} is the same as the lower triangular part of a rank-one matrix

See Y. Ikebe, *On inverses of Hessenberg matrices*, Linear Algebra Appl., v 24 (1979), pp. 93–97

We construct a lower triangular matrix X such that $\mathcal{L}(X) = \mathcal{L}(H^{-1})$

To simplify the notation, let $\tilde{\beta} = \beta_1 - \hat{\nu}^T \hat{U} \hat{\beta}$

$$X = \begin{pmatrix} \tilde{\beta} & 0 \\ \hat{U}\hat{\beta} & \tilde{U} \end{pmatrix} \Rightarrow X^{-1} = \begin{pmatrix} \frac{1}{\tilde{\beta}} & 0 \\ -\frac{1}{\tilde{\beta}} \tilde{U}^{-1} \hat{U} \hat{\beta} & \tilde{U}^{-1} \end{pmatrix}$$

with $\tilde{U} = \mathcal{L}(\hat{U}\hat{\beta}\hat{\nu}^T)$

The matrix X^{-1} is, in fact, lower bidiagonal

FOM applied to (X^{-1}, e_1) yields residual norms such that $\|r_0\|/\|r_k^F\| = \nu_{k+1}$

The symmetric case

We follow G-S and define

$$X_s = \mathcal{L}(H^{-1}) + \hat{\mathcal{L}}(H^{-1})^T$$

where $\hat{\mathcal{L}}$ gives the strict lower triangular part of the matrix

Let $\sigma = \hat{U}\hat{\beta}$

$$X_s = \begin{pmatrix} \tilde{\beta} & (\hat{U}\hat{\beta})^T \\ \hat{U}\hat{\beta} & \mathcal{L}(\hat{U}\hat{\beta}\hat{\nu}^T) + \hat{\mathcal{L}}(\hat{U}\hat{\beta}\hat{\nu}^T)^T \end{pmatrix} = \begin{pmatrix} \sigma_0 & \sigma_1 & \cdots & \sigma_{n-1} \\ \sigma_1 & \sigma_1\nu_2 & \cdots & \sigma_{n-1}\nu_2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n-1} & \sigma_{n-1}\nu_2 & \cdots & \sigma_{n-1}\nu_n \end{pmatrix}$$

From the structure of X_s we know that X_s^{-1} is a symmetric tridiagonal matrix whose nonzero entries can be computed from the entries of X_s

Let $\nu = \begin{pmatrix} 1 & \nu_2 & \cdots & \nu_n \end{pmatrix}^T$ and μ_i, $i = 1, \ldots, n$, η_i, $i = 1, \ldots, n-1$ be the diagonal and subdiagonal entries of the tridiagonal matrix X_s^{-1}

$$\mu_1 = -\frac{\nu_2}{\sigma_1 - \nu_2\sigma_0}, \quad \eta_1 = \frac{1}{\sigma_1 - \nu_2\sigma_0} = -\frac{\mu_1}{\nu_2},$$

and for $i = 2, \ldots, n-1$,

$$d_i = \nu_i(\nu_i\sigma_i - \nu_{i+1}\sigma_{i-1})$$

$$\mu_i = -\frac{\nu_{i+1}}{d_i} - \eta_{i-1}\frac{\nu_{i-1}}{\nu_i}, \quad \eta_i = \frac{\nu_i}{d_i},$$

the last diagonal entry being equal to

$$\mu_n = \frac{1}{\nu_n\gamma_{n-1}} - \eta_{n-1}\frac{\nu_{n-1}}{\nu_n}$$

We have to assume that $d_i \neq 0$

FOM ($\equiv$ CG) applied to (X_s^{-1}, e_1) gives relative residual norms equal to the inverses of the absolute values of the components of ν

However for CG we need the tridiagonal matrix to be positive definite

We can choose the σ_i's to obtain a positive definite matrix

From the Cholesky-like factorization $X_s^{-1} = L\Omega^{-1}L^T$ with L lower bidiagonal and Ω diagonal, we see that if we choose the ν_i's and σ_i's strictly positive, the condition is

$$\frac{\sigma_{i-1}}{\sigma_i} > \frac{\nu_i}{\nu_{i+1}}, \; i = 1, \ldots, n-1$$

Prescribing the A-norm of the error

Let $\varepsilon_k = x - x_k$ be the error vector, $A = VTV^T$. It is known that

$$\|\varepsilon_k\|_A^2 = \|r_0\|^2[(T^{-1})_{1,1} - (T_k^{-1})_{1,1}]$$

where T_k is the principal submatrix of the Lanczos matrix T of order k

Theorem

$$(T^{-1})_{1,1} - (T_k^{-1})_{1,1} = \frac{\sigma_k}{\nu_{k+1}}, \quad k = 1, \ldots, n-1$$

Corollary

The square of the A-norm of the error is given by

$$\|\varepsilon_k\|_A^2 = \|r_0\|^2 \frac{\sigma_k}{\nu_{k+1}} = |\sigma_k| \, \|r_0\| \, \|r_k\|, \;\; k = 0, \ldots, n-1$$

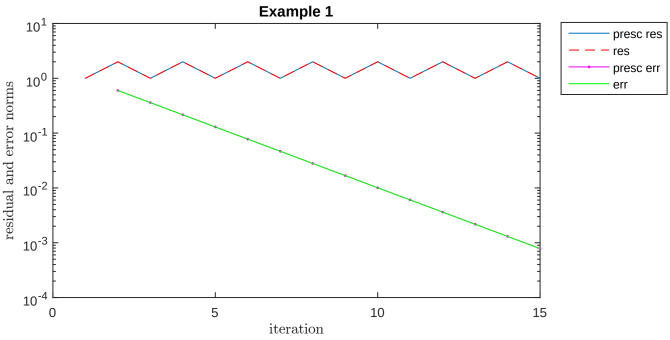
If we prescribe decreasing values for $\|\varepsilon_k\|_A$ we obtain the values of the positive σ_k's

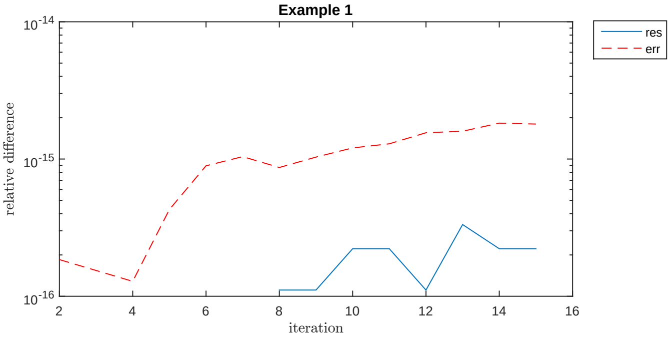
Prescribing decreasing values corresponds to the condition to have T positive definite

Mathematically, any residual norm convergence curve is possible for CG with any decreasing A-norm convergence curve

A numerical experiment

$n = 15$, res $= [1\ 2\ 1\ 2\ 1\ 2\ 1\ 2\ 1\ 2\ 1\ 2\ 1\ 2\ 1]$, err:
$g_0 = 1, \quad g_i = 0.6\, g_{i-1},\ i = 1, \ldots, n$, $\mathrm{cond}(T) = 6.37\ 10^6$




Example 1
res
err
relative difference
iteration
10^-14
10^-15
10^-16
2
4
6
8
10
12
14
16

Conclusion

We have shown how to construct symmetric positive definite linear systems with a prescribed CG residual norm convergence curve as well as a prescribed A-norm of the error convergence curve

For details, see
G.M., *On prescribing the convergence behavior of the conjugate gradient algorithm*, Numerical Algorithms, v 84 n 4 (2020), pp. 1353-1380

There are Matlab functions to construct these matrices on my Web site

https://gerard-meurant.pagesperso-orange.fr/

$\rightarrow$ Software $\rightarrow$ GM Toolbox $\rightarrow$ Linear systems $\rightarrow$ Symmetric

Functions gm_presc_conv_CG and gm_presc_conv_CG_err

A commercial

Claude Brezinski, Gérard Meurant and Michela Redivo-Zaglia

A Journey through the History of Numerical Linear Algebra

to be published by SIAM