

[Q1] (10 pts) (a) For a generic Krylov subspace method that takes the initial approximation x_0 , gets the initial residual $r_0 = b - Ax_0$, develops the sequence of Krylov subspaces $\mathcal{K}_k(A, r_0)$ and constructs the approximate solution $x_k = x_0 + z_k$ where $z_k \in \mathcal{K}_k(A, r_0)$, the residual $r_k = b - Ax_k$ can be written as $r_k = p_k(A)r_0$, where p_k is a polynomial of degree no greater than k with $p_k(0) = 1$.

Consider $r_k = r_0 - AU_k y_k$
 Since $\mathcal{K}_k(A, r_0) = \mathcal{K}_k(A, r_0)$, we can write $U_k y_k$ as $q_{k-1}(A)r_0$

Thus, $r_k = (I - Aq_{k-1}(A))r_0$

Note $p_k(0) = 1$ since $I \rightarrow 1$ \square

(b) Let A be SPD, and x_0 and $r_0 = b - Ax_0$ be the initial approximation and residual, respectively. Consider the Lanczos relation $AU_k = U_k T_k + \beta_k u_{k+1} e_k^T$ (Arnoldi's method applied to a symmetric A), where $u_1 = \frac{r_0}{\|r_0\|_2}$. Show that the k -th iterate of CG can be written as $x_k = x_0 + U_k y_k$, where y_k satisfies $T_k y_k = \|r_0\|_2 e_1$. (Hint: use the fact that $r_k = b - Ax_k = r_0 - AU_k y_k \perp \mathcal{K}_k(A, r_0) = \text{col}(U_k)$)

$$r_k = b - Ax_k = r_0 - AU_k y_k$$

$$r_k \perp \mathcal{K}_k(A, r_0) = \text{col}(U_k)$$

$$r_0 = b - Ax_0$$

$$x_k = x_0 + U_k y_k$$

$$r_k = b - Ax_k$$

$$= b - A(x_0 + U_k y_k)$$

$$= b - Ax_0 - AU_k y_k$$

$$= r_0 - AU_k y_k$$

$$r_k = r_0 - (U_k T_k + \beta_k u_{k+1} e_k^T) y_k$$

$$U_k^T r_k = U_k^T r_0 - T_k y_k$$

$$U_k^T (b - Ax_k) = U_k^T (b - Ax_0) - T_k y_k$$

$$U_k^T A x_0 = U_k^T A x_k - T_k y_k$$

$$U_k^T A x_0 + T_k y_k = U_k^T A x_k$$

$$U_k^T A U_k = T_k$$

$$U_k^T A (x_0 + U_k y_k) = U_k^T A x_k$$

$$x_k = x_0 + U_k y_k$$

(c) Show that the k -th residual of GMRES $r_k = b - Ax_k$ satisfies $r_k \in \mathcal{K}_{k+1}(A, r_0)$, $r_k \perp \mathcal{A}\mathcal{K}_k(A, r_0)$, $(r_k, r_k) = (r_j, r_k)$ for all $0 \leq j \leq k-1$, and therefore $\|r_k\|_2 \leq \|r_j\|_2$.

$$\text{span}(\{A^n r_0\}_{n=1}^k) \quad r_1 = b - Ax_0$$

$$r_k = b - AQ_k y \quad r_1 \in \mathcal{K}_1(A, r_0)$$

Assume k

$$r_k \in \mathcal{K}_k(A, r_0) \quad r_k = b - Ax_k \in \text{span}\{A^n r_0\}_{n=0}^{k-1}$$

$$r_{k+1} = b - Ax_{k+1} \quad x_{k+1} = Q_k y$$

$$x_{k+1} \in \text{Col}(Q_k) = \mathcal{K}_k(A, r_0)$$

$$\Rightarrow b - Ax_{k+1} \in \mathcal{K}_{k+1}(A, r_0)$$

$$\therefore r_{k+1} \in \mathcal{K}_{k+1}(A, r_0)$$

$$\langle r_k, v \rangle = \langle r_0 - Ax_k, v \rangle$$

$$= v^T r_0 - v^T A x_k$$

$$= 0 + 0$$

$$\therefore \|r_k\| \leq \|r_j\| \quad \forall k \geq j$$

Let $v \in \mathcal{A}\mathcal{K}_k(A, r_0)$

Since by Arnoldi process they are orthogonal

35.2. (a) Let $S \subseteq \mathbb{C}$ be a set whose convex hull contains 0 in its interior. That is, S is contained in no half-plane disjoint from the origin. Show that there is no $p \in P_1$ (i.e., no polynomial p of degree 1 with $p(0) = 1$) such that $\|p\|_S < 1$.

(b) Let A be a matrix, not necessarily normal, whose spectrum $\Lambda(A)$ has the property (a). Show that there is no $p \in P_1$ such that $\|p(A)\| < 1$.

(c) Though the convergence in Figure 35.5 is slow, it is clear that $\|r_1\| < \|r_0\|$. Explain why this does not contradict the result of (b). Describe what kind of polynomial $p_1 \in P_1$ GMRES has probably found to achieve $\|r_1\| < \|r_0\|$.

$$a) \|p\|_S = \sup_{z \in S} |p(z)|$$

Assume $\exists p \in P_1$ s.t.

$$\begin{aligned} p(0) &= 1 \\ p(z) &= a + bz \\ p(0) &= a = 1 \end{aligned}$$

$$\Rightarrow p(z) = 1 + bz$$

WLOG, assume S is a circle, if not take $S \subseteq S'$ that is a circle and what is shown for S' will hold for S

$$\text{Let } S = B(0, r) \text{ so } \|p\|_S = \sup_{z \in S} |1 + bz| = |1 + br| > 1$$

(b) Let A be a matrix, not necessarily normal, whose spectrum $\Lambda(A)$ has the property (a). Show that there is no $p \in P_1$ such that $\|p(A)\| < 1$.

$$b) \|p(0)\| = 1 \Rightarrow p(A) = I + BA$$

Consider an eigenpair (λ, v) of A .

Let v be normalized. Choose direction of v s.t. $\operatorname{Re} \lambda \geq 0$

$$\begin{aligned} \|I + BA\| &= \sup_{\|x\|=1} \|(I + BA)x\| \\ &\geq \|v + BA v\| \\ &= \|v + B\lambda v\| \\ &= \|(1 + B\lambda)v\| \\ &= |1 + B\lambda| \\ &> 1 \end{aligned}$$

(c) Though the convergence in Figure 35.5 is slow, it is clear that $\|r_1\| < \|r_0\|$. Explain why this does not contradict the result of (b). Describe what kind of polynomial $p_1 \in P_1$ GMRES has probably found to achieve $\|r_1\| < \|r_0\|$.

$$r_0 = b \quad r_1 = b - Ax_0$$

It does not contradict since the residual decreases by looking at the minimized residual. So $\|r_{n+1}\| \leq \|r_n\| \quad \forall n \in \mathbb{N}$ (problem 1)

$$p(A) = I + bA$$

$$r_1 = p(A)r_0$$

$$\|r_1\| = \|p(A)r_0\| \leq \|p(A)\| \|r_0\|$$

If $\|p(A)\| = 1$ then $\|r_1\| \leq \|r_0\|$

(b) Let $A \in \mathbb{R}^{n \times n}$ be nonsymmetric and diagonalizable. Assume that all eigenvalues of A lie in the disk centered at $c \in \mathbb{C} \setminus \{0\}$ with radius $r < |c|$. Consider using GMRES to solve the linear system $Ax = b$ iteratively. Show that the k -th relative residual satisfies $\frac{\|r_k\|_2}{\|r_0\|_2} \leq C \left(\frac{r}{|c|}\right)^k$ for some constant C independent of k . What if A has a small number, say, $m \ll n$ eigenvalues outside such a disk?

$$\|r_k\| = \min_p \|p_k(A)r_0\|$$

$$\frac{\|r_k\|}{\|r_0\|} \leq \min_p \|p_k(A)\|$$

$$A = V D V^{-1}$$

$$\begin{aligned} \frac{\|r_k\|}{\|r_0\|} &\leq \min_p \kappa(V) \|p_k(A)\| \\ &= \max \kappa(V) \|H_k - \lambda I\| \quad \text{from class min} \\ &= \kappa(V) \lambda_1^k \quad \text{is characteristic} \\ &\leq \kappa(V) \left(\frac{r}{|c|}\right)^k \quad \text{polynomial} \\ &\quad \text{of } H \\ &\quad \text{max from eigen} \end{aligned}$$

$$\text{Since } \Lambda(A) \subseteq B(|c|, r)$$

I would depend on the magnitude of the eigenvalues outside of the disk.

c) (c) If A is an SPD matrix with the smallest eigenvalue λ_1 and the largest eigenvalue λ_n , what is the convergence factor obtained in part (b)? Compare this factor with that of CG we learned in class. Which one is better?

$$\frac{\lambda_1}{\lambda_n} \quad \text{which would be larger}$$

then CG's $\sqrt{\frac{\lambda_1}{\lambda_n}}$ rate

Note $\Lambda(A) = B(1, \lambda_n) \setminus B(1, \lambda_1)$ in this case

3. [Q3] (10 pts) Let x^* be the true solution of $Ax = b$ with SPD A , x_k be the k -th iterate of CG, and $\varphi(x) = \frac{1}{2}x^T Ax - b^T x$ for CG minimization.

(a) Note that $r_k \perp r_j$ for $0 \leq j \leq k-1$, and hence $r_k \perp U_k = \text{span}\{p_0, p_1, \dots, p_{k-1}\}$. Also note that $r_k = -\nabla \varphi(x_k)$, and any vector $x \in W_k = x_0 + U_k$. Explain from the optimization point of view, why $x_k = \text{argmin}_{x \in W_k} \varphi(x)$.

Hint: one possible (and easier) solution is to show that W_k is a convex set, and $\varphi(x)$ is a convex function defined on W_k ; then local minimizer of $\varphi(x)$ is necessarily a global minimizer. Please do a little search on convex set/functions yourselves. The condition $r_k \perp U_k$ is crucial to show the optimality here.

(b) Show directly that $x_k = \text{argmin}_{x \in W_k} \|x - x^*\|_A$, without referring to the connection between $\varphi(x)$ and $\|e_k\|_A$. (Hint: consider a different $\tilde{x}_k \in W_k$, with $d_k = \tilde{x}_k - x_k \neq 0$. Show that $\|\tilde{x}_k - x^*\|_A = \|d_k + x_k - x^*\|_A \geq \|x_k - x^*\|_A$)

~) Since $W_k = x_0 + U_k$ where U_k is a vector space, it follows U_k is convex $\Rightarrow W_k$ is convex.

$$\text{Let } x_1 = \sum_{i=1}^n \alpha_i p_i \quad x_2 = \sum_{i=1}^n \beta_i p_i$$

$$\begin{aligned} \text{WTS } \varphi(\lambda x_1 + (1-\lambda)x_2) &\leq \lambda \varphi(x_1) + (1-\lambda)\varphi(x_2) \\ \frac{1}{2}(\lambda x_1 + (1-\lambda)x_2)^T A(\lambda x_1 + (1-\lambda)x_2) - b^T(\lambda x_1 + (1-\lambda)x_2) & \leftarrow A \text{ is SPD } \in \mathbb{R}^{n \times n} \\ &= \frac{1}{2} \lambda^2 x_1^T A x_1 + \frac{1}{2} (1-\lambda)^2 x_2^T A x_2 + \lambda(1-\lambda) x_1^T A x_2 - \lambda b^T x_1 + (1-\lambda)b^T x_2 \end{aligned}$$

$$\begin{aligned} &= \lambda^2 \varphi(x_1) + (1-\lambda)^2 \varphi(x_2) + \lambda(1-\lambda) x_1^T A x_2 \\ &\leq \lambda \varphi(x_1) + (1-\lambda) \varphi(x_2) + \lambda(1-\lambda) x_1^T A x_2 \end{aligned}$$

Bc A is SPD $\lambda(1-\lambda) x_1^T A x_2$ is convex

Thus by the sum of two convex functions being convex, we see $\varphi(x)$ is convex

Therefore the local minimizer $x_k \in W_k$ for $\varphi(x)$ is the global minimizer.

(b) Show directly that $x_k = \text{argmin}_{x \in W_k} \|x - x^*\|_A$, without referring to the connection

(b) Show directly that $x_k = \operatorname{argmin}_{x \in W_k} \|x - x^*\|_A$, without referring to the connection between $\varphi(x)$ and $\|e_k\|_A$. (Hint: consider a different $\tilde{x}_k \in W_k$, with $d_k = \tilde{x}_k - x_k \neq 0$. Show that $\|\tilde{x}_k - x^*\|_A = \|d_k + x_k - x^*\|_A \geq \|x_k - x^*\|_A$)

$$d_k = \tilde{x}_k - x_k \neq 0 \Rightarrow \tilde{x}_k = d_k + x_k$$

$$\|\tilde{x}_k - x^*\|_A = \|d_k + x_k - x^*\|_A \geq \|\|x_k - x^*\|_A - \|d_k\|_A\|_A \geq \|x_k - x^*\|_A$$

Since $\|d_k\|_A \geq 0$

4a) B has pos & neg eig vals ✓
 \tilde{B} has pos eig vals ✓

Both symmetric

A & \hat{A} are not symmetric

The eigvals are close to the origin

There does not seem to be an obvious correlation to condition number as \tilde{B} was largest

```
>> HW9_linsolvecomp
Loading problem data completed.
SPD matrix A: preconditioner setup time 4.762 secs.
P-CG converges in 70 iters (12.203 secs) to relative tol 1.000e-08.
MATLAB's sparse Cholesky direct solver takes 369.842 secs.
nnz of A, incomplete and complete Cholesky factors: 6141493 27725520
1140363654.

Symmetric indefinite C: indefinite preconditioner constructed already.
P-SQMR converges in 67 iters (2.059 secs) to relative tol 1.000e-08.

Transforming the indefinite preconditioner to an SPD preconditioner.
Transformation to SPD preconditioner takes 0.186 secs.
P-MINRES converges in 210 iters (2.701 secs) to relative tol 1.000e-08.
MATLAB's sparse LDL direct solver takes 156.538 secs.
nnz of C, incomplete and complete LDL factors: 738598 617410 82593444.

Unsymmetric matrix A: preconditioner setup time 2.858 secs.
P-BICGSTAB(2) solve: 93 mvps (10.342 secs) to tolerance 1.000e-08.
P-IDR(4) solve: 86 mvps (10.190 secs) to tolerance 1.000e-08.
P-GMRES(60) solve: 90 mvps (24.967 secs) to tolerance 1.000e-08.
MATLAB's sparse LU direct solution takes 1129.186 secs.
nnz of B, incomplete and complete LU factors: 3604989 5659826 856629224.
```

It was more manageable and faster

```
rng('default'); n = 1024; A = randn(n,n); [A,R] = qr(A);
Ahat = A+1.2*eye(n); E = randn(n,n); E = E+E';
B = (A+A')/2; B = B+1e-4*E; Bhat = B+1.01*eye(n);

D=eig(B);
D=sort(D);

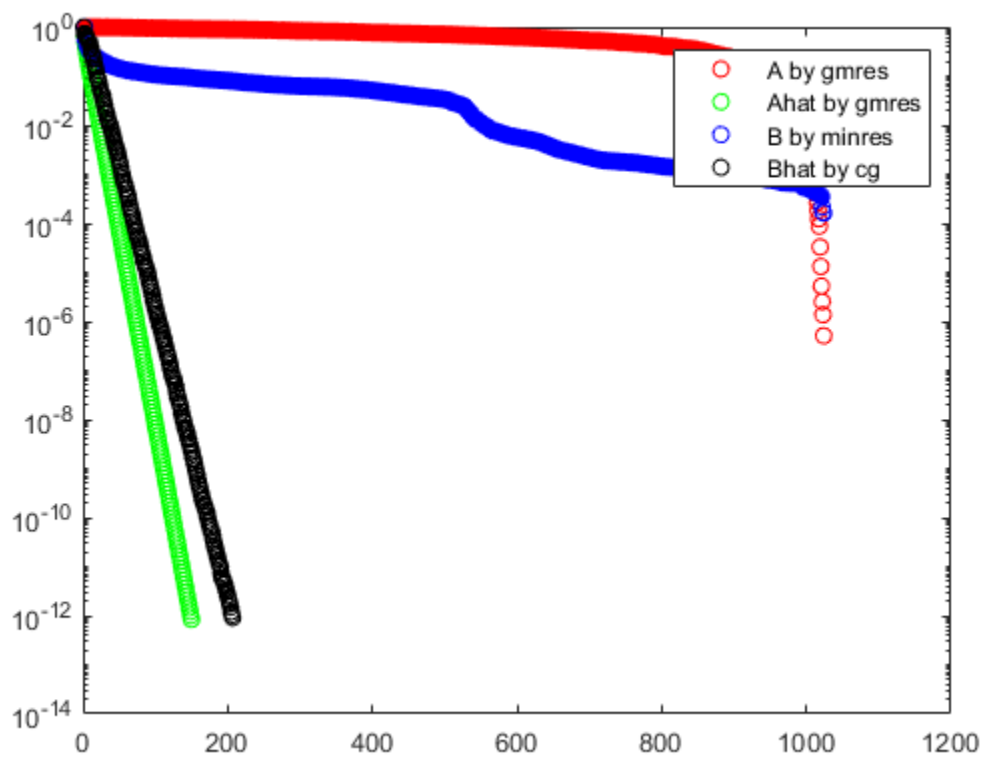
F=eig(Bhat);
F=sort(F);

ca=cond(A);
cahat=cond(Ahat);
cb=cond(B);
cbhat=cond(Bhat);

L=eig(A);
M=eig(Ahat);

hold on
plot(real(M),imag(M),'o')
plot(real(L),imag(L),'o')
plot(real(D),imag(D),'o')
plot(real(F),imag(F),'o')
hold off

f = ones(n,1); m = n-1; restart = 1; tol = 1e-12;
[x1,flag1,relres1,iter1,resvec1] = gmres(A,f,m,tol,1);
semilogy(resvec1/norm(f),'ro'); hold on;
[x2,flag2,relres2,iter2,resvec2] = gmres(Ahat,f,m,tol,1);
semilogy(resvec2/norm(f),'go'); hold on;
[x3,flag3,relres3,iter3,resvec3] = minres(B,f,tol,m);
semilogy(resvec3/norm(f),'bo'); hold on;
[x4,flag4,relres4,iter4,resvec4] = pcg(Bhat,f,tol,m);
semilogy(resvec4/norm(f),'ko'); hold on;
legend('A by gmres','Ahat by gmres','B by minres','Bhat by cg');
```

Published with MATLAB® R2023b