Advanced Numerical Analysis Homework 9

Michael Nelson

Throughout this homework, $\|\cdot\|$ denotes the ℓ_2-norm. We also let $\langle\cdot,\cdot\rangle$ denote the standard Euclidean inner-product on $\mathbb{C}^m$ (thus

$$\langle x, y\rangle = \sum_{i=1}^{m} x_i \overline{y}_i$$

for all $x, y \in \mathbb{C}^m$).

1 Problem 1

Exercise 1. Solve the following:

1. 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 $r_k = p_{k+1}(A)r_0$, where p_{k+1} is a polynomial of degree no greater than $k+1$ with $p_{k+1}(0) = 1$.

2. Let A be SPD, and x_0 and $r_0 = b - Ax_0$ be the initial approximation and residual, respectively. Consider the Lanczos relation
$$AU_k = U_k T_k + \beta_k u_{k+1} e_k^\top$$
(Arnoldi's method applied to a symmetric A), where $u_1 = r_0/\|r_0\|$. u Show that the kth iterate of CG can be written as $x_k = x_0 + U_k y_k$, where y_k satisfies $T_k y_k = \|r_0\| e_1$ (Hint: use the fact that
$$\begin{aligned} r_k &= b - Ax_k \\ &= x_0 - AU_k y_k \\ &\perp \mathcal{K}_k(A, r_0) \\ &= \operatorname{col}(U_k). \end{aligned}$$

3. Show that the kth residual of GMRES $r_k = b - Ax_k$
$$r_k \in \mathcal{K}_{k+1}(A, r_0), \quad r_k \perp A\mathcal{K}_k(A, r_0), \quad \text{and} \quad (r_k, r_k) = (r_j, r_k)$$
for all $0 \le j \le k-1$, and therefore $\|r_k\| \le \|r_j\|$.

Solution 1.

2 Problem 2

Exercise 2. Solve the following:

1. Trefethen's book, Prob. 35.2.

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

3. 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 (2)? Compare this factor with that of CG we learned in class. Which one is better?

Solution 2.

3 Problem 3

Exercise 3. Let x^* be the true solution of $Ax = b$ with SPD A, x_k be the kth iterate of CG, and

$$\varphi(x) = \frac{1}{2}x^\top Ax - b^\top x = \frac{1}{2}\|x\|_A - \langle b, x\rangle$$

for CG minimization. (a)

1. Note that $r_k \perp r_j$ for $0 \le j \le k-1$, and hence $r_k \perp U_k = \text{span}\{p_0, p_1, \ldots, p_{k-1}\}$. 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.

2. 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 $\widetilde{x}_k \in W_k$ with $d_k = \widetilde{x}_k - x_k \neq 0$. Show that

$$\|\widetilde{x}_k - x^*\|_A = \|d_k + x_k - x^*\|_A \ge \|x_k - x^*\|_A.$$

Solution 3.

4 Problem 4

Exercise 4. Solve the following:

1. A common misconception is that Krylov subspace methods solving $Ax = b$ converge rapidly if the condition number, say, $\kappa_2(A)$ is small. This is largely true if A is SPD, but in general not true otherwise. To explore this point, construct three matrices as follows

```
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);
```

Check that A and $\widehat{A}$ are unsymmetric, B is symmetric and indefinite, and $\widehat{B}$ is SPD, and find $\kappa_2(A)$, $\kappa_2(\widehat{A})$, $\kappa_2(B)$, and $\kappa_2(\widehat{B})$. Are these condition numbers really large at all? Use eig to compute all eigenvalues of A, $\widehat{A}$, B, and $\widehat{B}$ and plot them on the complex plane. How are these eigenvalues distributed around the origin?

Solution 4.

Appendix

QR Post Process

```
function [Q,R] = QRpostprocess(Q,R)

[m,n] = size(Q);

for i = 1:n
   if R(i,i) < 0
```

```
        R(i,:) = -R(i,:);
        Q(:,i) = -Q(:,i);
    end;
end;
```

QR Algorithm

```
function [Qk,Rk,Ak] = QRalgorithm(A,k)

[m,n] = size(A); Ak = A; Rk = eye(n); Qk = eye(n);

for i = 1:k
    [Q,R] = qr(Ak);
    [Q,R] = QRpostprocess(Q,R);
    Ak = R*Q;
    Qk = Qk*Q;
    Rk = R*Rk;
end;
```

Simultaneous Iteration

```
function [Qk,Rk,Ak] = SimultaneousIteration(A,k)

[m,n] = size(A); Qk = eye(n); Rk = eye(n);

for i = 1:k
    Z = A*Qk;
    [Qk,R] = qr(Z);
    [Qk,R] = QRpostprocess(Qk,R);
    Ak = Qk'*A*Qk;
    Rk = R*Rk;
end;
```