AMSC 667 Section 0101 Spring 2011

NUMERICAL ANALYSIS II

HOMEWORK #5 (Pbs 1-3 due May 3, Pbs 4-7 due May 10)

1 (10 pts). *Normal matrices*: A matrix A is *normal* if $A^H A = AA^H$. Show that A is nondefective. Hint: prove that an upper triangular normal matrix is diagonal, and use Schur decomposition theorem.

2 (15 pts) *Richardson Method*: Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and $\mathbf{b} \in \mathbb{R}^n$. Consider the following iteration for the solution of $A\mathbf{x} = \mathbf{b}$:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha(\mathbf{b} - A\mathbf{x}_k).$$

(a) Show that if all eigenvalues $\{\lambda_i\}_{i=1}^n$ of A lie in the right half plane, that is their real part is positive, then the iterates converge for any starting point $\mathbf{x}_0 \in \mathbb{R}^n$ provided

$$0 < \alpha < \frac{2Re(\lambda_i)}{|\lambda_i|^2}.$$

(b) Show that the optimal choice of α is

$$\alpha = \frac{2}{\lambda_1 + \lambda_n}$$

provided $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n > 0$; this is the case if A is SPD. Show that the rate of convergence is

$$\rho(A) = \frac{\kappa_2(A) - 1}{\kappa_2(A) + 1},$$

where $\kappa_2 = \|A\|_2\|A^{-1}\|_2$ is the condition number of A in the 2-norm.

(c) Show that if some eigenvalues of A have negative real part and some positive real part, then there is no α for which the iterations converge;

(d) Let $\sigma = \|I - \alpha A\| < 1$ for a matrix norm subordinate to a vector norm. Show that the error can be expressed in terms of the difference between consecutive iterates, namely

$$\|\mathbf{x} - \mathbf{x}_{k+1}\| \leq \frac{\sigma}{1-\sigma}\|\mathbf{x}_k - \mathbf{x}_{k+1}\|.$$

3 (15 pts). *Multigrid*: Consider either finite differences or finite elements over a uniform partition of $(0, 1)$ for the two-point boundary value problem $-u'' = f, u(0) = u(1) = 0$, which give rise to a SPD matrix A:

$$-U_{n-1} + 2U_n - U_{n+1} = F_n \quad 1 \leq n \leq N, \qquad U_0 = U_{N+1} = 0. \tag{1}$$

(a) Show that A has the eigenpairs $(\lambda_i(A), \mathbf{v}_i)_{i=1}^N$ where

$$\lambda_i(A) = 4\sin^2\Big(\frac{i\pi}{2(N+1)}\Big), \qquad v_{i,n} = \sin\Big(\frac{ni\pi}{N+1}\Big) \quad 1 \leq n \leq N,$$

and that the eigenvectors $\mathbf{v}_i$ are orthogonal. These are the so-called *Fourier modes*, and i is the *wavenumber*. Plot $\sin(i\pi x)$ on $[0, 1]$ for $i = 1, 3, 6$, which corresponds to taking $x = n/(N + 1)$. Note that small i corresponds to slowly varing $\mathbf{v}_i$ and large i to oscillatory $\mathbf{v}_i$.

(b) Let $0 < \omega \leq 1$ be a weighting factor. Consider the *weighted (or damped) Jacobi* iteration for (1):

$$v_n^{k+1} = (1-\omega)v_n^k + \frac{\omega}{2}(v_{n-1}^k + v_{n+1}^k), \quad 1 \leq n \leq N. \tag{2}$$

Note that for $\omega = 1$, (2) reduces to the *Jacobi* iteration. Consequently, v_n^{k+1} is an intermediate value between the previous iterate v_n^k and the Jacobi update. Determine the iteration matrix B and show that its eigenpairs $(\lambda_i(B), \mathbf{v}_i)_{i=1}^N$ satisfy $\lambda_i(B) = 1 - \frac{\omega}{2}\lambda_i$. Show that the iteration (2) converges.

(c) Let $\mathbf{e}_k$ be the error associated with (2). Show that if $\mathbf{e}_0 = \sum_{i=1}^N \alpha_i \mathbf{v}_i$ then

$$\mathbf{e}_k = \sum_{i=1}^N \alpha_i \lambda_i(B)^k \mathbf{v}_i.$$

This provides an eigenvector expansion for $\mathbf{e}_k$. We see that after k steps, the k-th mode of the initial error has been reduced by a factor $\lambda_i(B)^k$, and that weighted Jacobi does not mix modes.

(d) We now explore the choice of ω. Plot $\lambda_i(B)$ as a function of i for $\omega = 1/3, 1/2, 2/3, 1$. Show that for $\omega = 2/3$ we have $|\lambda_i(B)| \leq 1/3$ for all $(N+1)/2 \leq i \leq N$; these are the so-called high-frequency (or oscillatory) modes. Conclude that $\omega = 2/3$ is the most effective choice to reduce the high-frequency modes, and that no ω damps effectively the low-frequency modes.

Similar properties are true for all classical methods: they are effective in reducing high-frequency modes but rather ineffective for the low-frequency modes. This property is the basis of *multigrid*, the best method for sparse systems like (1): low-frequency modes for a partition are high-frequency modes for a partition of twice the meshsize (coarsened partition). We could then repeat (2) over a set of nested partitions and reduce selectively all frequencies of the error in order N operations.

4 (15 pts) *MATLAB*: Consider the two-point boundary value problem in $(0, 1)$:

$$-u'' + u = f, \qquad u(0) = u(1) = 0. \tag{3}$$

(a) Let $\{x_i\}_{i=0}^{N+1}$ be a uniform partition of $(0, 1)$ with meshsize $h = 1/(N+1)$. Derive the discrete equations resulting from applying centered differences to (3). Write the equations in matrix form $A\mathbf{U} = \mathbf{F}$, where $\mathbf{U} = (U_i)_{i=1}^N$ is the vector of nodal values and $\mathbf{F} = (f(x_i))_{i=1}^N$. Show that A is strictly diagonally dominant, symmetric and positive definite.

(b) Write a MATLAB function `[x,k] = gradient(A,b,x0,tol)` which implements the steepest descent (or gradient) method for an $N \times N$ symmetric and positive definite matrix `A`, right-hand side `b`, and starting value `x0`. The program should compute the 2-norm of the residual and stop when such a norm is less than a given tolerance `tol`, giving the current iterate vector `x` and number of iterations `k`.

(c) Write a MATLAB function `[x,k] = cg(A,b,x0,tol)` which implements the CG method. The arguments have the same meaning as in (b).

(d) Let f be the right-hand side of (3) corresponding to the exact solution $u(x) = \sin(\pi x) - \sin(3\pi x)$. Run the functions `gradient` and `cg` for $N = 10, 20, 40$, `x0=0` and `tol` $= 10^{-8}$. Plot the computed solutions together with $u(x)$. Plot also the log of the 2-norm of the residual in terms of the log of the number of steps, and draw conclusions.

5 (15 pts) *CG for Semi-Definite Matrices*: Let A be a symmetric positive semi-definite matrix. Let the linear system $A\mathbf{x} = \mathbf{b}$ be consistent, that is, $\mathbf{b}$ belongs to the range of A and thus solution exists. Prove that with initial guess $\mathbf{x}^{(0)} = 0$, the conjugate gradient method (CG) is guaranteed to produce a solution without component in the kernel of A. To this end proceed as follows.
(a) Show that the singular value decomposition of A can be written as $A = \mathbf{U}\Sigma\mathbf{U}^T$ with Σ diagonal, and study the properties of Σ.
(b) Derive a system equivalent to $A\mathbf{x} = \mathbf{b}$ for $\hat{\mathbf{x}} = \mathbf{U}\mathbf{x}$ and $\hat{\mathbf{b}} = \mathbf{U}\mathbf{b}$, and study the structure of $\hat{\mathbf{b}}$.
(c) Show that CG with $\mathbf{x}^{(0)} = 0$ is equivalent to CG for a symmetric positive definite matrix with zero initial guess.

6 (15 pts). *Rayleigh Quotient Method*. Whenever $A \in \mathbb{R}^{n\times n}$ is symmetric it is better to replace the power method by the following algorithm

$$\lambda_1^{m+1} = \frac{\mathbf{z}_m^T \cdot A\mathbf{z}_m}{\mathbf{z}_m^T \cdot \mathbf{z}_m} = \frac{\mathbf{z}_m^T \cdot \mathbf{w}_{m+1}}{\mathbf{z}_m^T \cdot \mathbf{z}_m}, \quad m \geq 0,$$

where

$$\mathbf{w}_{m+1} = A\mathbf{z}_m, \qquad \mathbf{z}_{m+1} = \mathbf{w}_{m+1}/\|\mathbf{w}_{m+1}\|_\infty.$$

Prove that

$$|\lambda_1^{m+1} - \lambda_1| = O\left((\lambda_2/\lambda_1)^{2m}\right).$$

So this method converges faster than the power method, namely with a rate $(\lambda_2/\lambda_1)^{2m}$ instead of $(\lambda_2/\lambda_1)^m$.

7 (15 pts). *Reduction to Hessenberg Form.* Let A be an upper Hessenberg matrix ($a_{ij} = 0$ if $i > j + 1$) and consider the QR factorization $A = QR$. Then $Q = P_1 \cdots P_{n-1}$ is a product of Householder reflections

$$P = I - 2\frac{\mathbf{v}_i \cdot \mathbf{v}_i^T}{\|\mathbf{v}_i\|_2^2}, \qquad 1 \le i \le n-1.$$

(a) Show that $\mathbf{v}_i = (v_{i1}, \cdots v_{in})^T$ has the special form

$$v_{ik} = 0 \qquad \text{if} \quad k < i \quad \text{and} \quad k > i + 1.$$

(b) Show that the form of P_i is such that Q will also be a Hessenberg matrix.
(c) Show that the product of an upper triangular matrix and a Hessenberg matrix is again Hessenberg.
(d) Examine how to convert a general matrix into a similar Hessenberg form via Householder reflections. Perform an operation count.
(e) Explain what happens if A is symmetric.

When combined, these results show that RQ is again a Hessenberg matrix, which in turn implies that next iterate $A_{m+1} = R_m Q_m$ in the QR method will possess the same Hessenberg form as $A_m = Q_m R_m$.