Advanced Numerical Linear Algebra Homework 1
AMSC/CMSC 763 Due September 18, 2019

1. Let $\langle f, g\rangle$ denote an inner product for functions defined on a real interval $[a, b]$. Typical examples have the form

$$\langle f, g\rangle \equiv \int_a^b f(x)g(x)w(x)dx$$

where $w(x)$ is a positive weight function on (a, b). For example, the choice $w(x) \equiv 1$ gives the ℓ_2 inner product and $w(x) = 1/\sqrt{1-x^2}$ on $[-1, 1]$ gives the Chebyshev inner product.

a. Show that polynomials $\{\psi_j \,|\, j = 1, 2, \ldots\}$ orthogonal with respect to such an inner product can be defined via a three-term recurrence

$$\gamma_{j+1}\psi_{j+1}(x) = x\psi_j(x) - \delta_j\psi_j(x) - \gamma_j\psi_{j-1}(x) \tag{1}$$

such that ψ_j has degree $j-1$, $\|\psi_j\| = 1$ where $\|f\| \equiv \langle f, f\rangle^{1/2}$, and $\psi_0 = 0$.

b. Show that for $k+1 \le n$, the roots of ψ_{k+1} are the eigenvalues of the tridiagonal matrix T_k determined by the recurrence (1).
Hint: Consider the characteristic polynomial of T_k.

c. The expression (1) resembles the recurrence that defines the Lanczos algorithm for a symmetric matrix A of order n,

$$\gamma_{j+1}v_{j+1} = Av_j - \delta_j v_j - \gamma_j v_{j-1}, \tag{2}$$

which produces orthogonal vectors $\{v_j\}$. Derive a variant of (2) of the form

$$\gamma_{j+1}w_{j+1} = \Lambda w_j - \delta_j w_j - \gamma_j w_{j-1},$$

where $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues of A.

d. Show that

$$\langle p, q\rangle \equiv \sum_{i=1}^n p(\lambda_i)q(\lambda_i)[w_1]_i^2$$

is an inner product on the space of polynomials defined on $\mathbb{R}$. What does this tell you about the polynomials from (1) defined using this inner product?

2. Let $Ax = b$ be a linear system of equations where A is symmetric and positive-definite. Let x_k be the kth iterate generated by the conjugate gradient method (CG). Show that if $x_k \ne x$, then the vectors generated by CG satisfy

$$\begin{aligned}
&\text{(i)} \quad \langle r_k, p_j\rangle = \langle r_k, r_j\rangle = 0, \qquad j < k,\\
&\text{(ii)} \quad \langle Ap_k, p_j\rangle = 0, \qquad j < k,\\
&\text{(iii)} \quad \text{span}\{r_0, r_1, \ldots, r_{k-1}\} = \text{span}\{p_0, p_1, \ldots, p_{k-1}\}\\
&\qquad = \mathcal{K}(A, r_0) \equiv \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}.
\end{aligned}$$

Hint: Prove (i) and (ii) simultaneously by induction on k, and use a dimensionality argument for (iii).

3. Let $Ax = b$ be as in Problem 2. Starting from an arbitrary initial iterate x_0, the steepest descent method generates a sequence of iterates $x_1, x_2, \ldots$ by the computation

$$x_{k+1} = x_k + \alpha_k r_k \,,$$

where r_k is the residual $b - Ax_k$ and α_k is a scalar chosen so that the norm $\|x - x_{k+1}\|_A$ is minimal.

a. Explain the name "steepest descent method."

b. Show that the error $e_k = x - x_k$ satisfies

$$\|e_k\|_A \leq \left(\frac{\kappa - 1}{\kappa + 1}\right)^k \|e_0\|_A \,.$$

where $\kappa = \Lambda/\lambda$ is the *condition number* of A, that is, the ratio of the largest eigenvalue of A to its smallest eigenvalue.

4. A demo (soon to be) given in class shows the effect damped Jacobi smoothing had on the discrete one-dimensional diffusion equation.

a. Implement this demo yourself. That is, show that a few steps of damped Jacobi smoothing makes the error smooth. You can generate the matrix and right-hand side using the code

```
e1 = ones(n,1);
h = 1/(n+1);
A = spdiags([-e1 2*e1 -e1], [-1,0,1], n, n)/h;
f = h*e1;
```

Reasonable choices for n are 31 or 63, but feel free to play with anything you like. To make the case, start with a random initial value and then plot the error in one or two figures.

b. Continue this experiment by implementing the *two-grid* algorithm. This will require construction of the coarse-grid matrix A_{2h} and the prolongation and restriction operators, P and R. You can then take one step of the two-grid algorithm to consist of two smoothing steps, followed by restriction, coarse-grid correction, and prolongation. Show that this algorithm displays "textbook" multigrid behavior, that is, the number of steps needed for the error to be smaller than a given tolerance is independent of the discretization mesh size.