Numerical Analysis Screening Exam
Fall 2016

Problem 1. (Iterative Methods) Consider the block matrix

$$A = \begin{bmatrix} I_n & 0 & -M_1 \\ -M_2 & I_n & 0 \\ 0 & -M_3 & I_n \end{bmatrix}$$

where I_n is the $n \times n$ identity matrix and M_1, M_2, and M_3, are $n \times n$ matrices.

a) Find the Jacobi and Gauss-Seidel iteration matrices for solving the system of equations $Ax = b$ (for some fixed vector b).
b) Show the methods either both converge or both diverge and, in the case of the former, the Gauss-Seidel method converges three times as fast as the Jacobi method.
c) Now consider the Jacobi and Gauss-Seidel methods for the system of equations $A^T x = b$. Show they either both converge or both diverge and, in the case of the former, the Gauss-Seidel method converges one-and-a-half times as fast as the Jacobi method.

Problem 2. (Least Squares) Let $A \in C^{m \times n}$ and consider the following set of equations in the unknown matrix $X \in C^{n \times m}$ known as the Moore-Penrose equations (MP):

$$AXA = A$$

$$XAX = X$$

$$(AX)^* = AX$$

$$(XA)^* = XA$$

(a) Show that the system (MP) has at most one solution (Hint: Show that if both $X$and Y are solutions to (MP) then $X = XAY$ and X=YAX.)

(b) When $A = \text{zeros}(m, n)$, show that there exists a solution to the system (MP). (Hint: Find one!)

(c) Assume that A has full column rank and find the solution to the least squares problem given by $\min \|AX - I\|_F^2$ where I denotes the $m \times m$ identity matrix and $\|\cdot\|_F$ denotes the Frobenius norm on $C^{m \times m}$ ($\|B\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{m} b_{i,j}^2\right)^{\frac{1}{2}} = \left(\sum_{j=1}^{m} \sum_{i=1}^{m} b_{i,j}^2\right)^{\frac{1}{2}}$). (Hint: Note that the given least squares problem decouples into m separate standard least squares problems!)

(d) Assume that A has full column rank and use part (c) to show that the system (MP) has a solution. (Hint: Find one!).

(e) Assume $\text{rank}(A) = r$, $0 < r < n$, and show that there exists a permutation matrix $P \in C^{n \times n}$ such that $AP = [\hat{A} | \hat{A}R]$, or $A = [\hat{A} | \hat{A}R]P^T$ where $\hat{A} \in C^{m \times r}$ has full column rank and $R \in C^{r \times (n-r)}$. (A permutation matrix is a square matrix that has precisely a single 1 in every row and column and zeros everywhere else.)

(f) Assume $\text{rank}(A) = r$, $0 < r < n$, assume that a solution to the system (MP) has the form $X = P \begin{bmatrix} S \\ T \end{bmatrix}$, where $S \in C^{r \times m}$ and $T \in C^{(n-r) \times m}$ and use parts (c), (d), and (e) and the first equation in the system (MP) to determine the matrices $S \in C^{r \times m}$ and $T \in C^{(n-r) \times m}$ in terms of the matrix $\hat{A}$, and therefore show that the system (MP) has a solution.

Problem 3. (Direct Methods) Consider a vector norm $\|\cdot\|_V$ for $\mathbb{R}^n$. Define another norm $\|\cdot\|_{V^*}$ for $\mathbb{R}^n$ by $\|x\|_{V^*} = \max_{y\in\mathbb{R}^n, \|y\|_V \le 1} |x^T y|$.

It is known that for every vector $x \neq 0 \in \mathbb{R}^n$, there exists a vector $y \in \mathbb{R}^n$ such that $y^T x = \|y\|_{V^*}\|x\|_V = 1$. A vector y with this property is called a dual element of x.

a. Consider a nonsingular $n \times n$ matrix A. We define the distance between A and the set of singular matrices by

$\mathrm{dist}(A) = \min\{\|\delta A\|_V : \delta A \in \mathbb{R}^{n\times n}, \text{ where } A + \delta A \text{ is singular}\}$,

where $\|A\|_V$ is the operator norm induced by $\|\cdot\|_V$. Show that

$$\mathrm{dist}(A) \ge \frac{1}{\|A^{-1}\|_V}.$$

Hint: Suppose the matrix $A + \delta A$ is singular. Then there exists $x \neq 0$ such that $(A + \delta A)x = 0$.

b. Let x be a unit vector such that $\|A^{-1}x\|_V = \|A^{-1}\|_V$ and $y = A^{-1}x/\|A^{-1}\|_V$. Consider a dual element z of y and the matrix

$$\delta A = -\frac{xz^T}{\|A^{-1}\|_V}.$$

Show that $A + \delta A$ is singular.

c. Show that $\|\delta A\|_V = \|A^{-1}\|_V^{-1}$ and

$$\mathrm{dist}(A) = \frac{1}{\|A^{-1}\|_V}.$$

Problem 4. (Eigenvalue/Eigenvector Problems) Let A be a real symmetric matrix and q_1 a unit vector. Let

$$\mathcal{K}(A, q_1, j) = \text{Span}\{q_1, Aq_1, A^2q_1, \dots, A^{j-1}q_1\}$$

be the corresponding Krylov subspaces. Suppose $\{q_1, q_2, \dots, q_n\}$ is an orthonormal basis for $\mathbb{R}^n$ and $\{q_1, q_2, \dots, q_j\}$ is an orthonormal basis for $\mathcal{K}(A, q_1, j)$, $1 \le j \le n$. Let Q_j be the $n \times j$ matrix whose columns are $q_1, q_2, \dots, q_j$, i.e. $Q_j = [q_1, q_2, \dots, q_j]$. It is easy to see that $Q_j^T A Q_j = T_j$ where T_j has the form

$$T_j = \begin{pmatrix} \alpha_1 & \beta_1 & 0 & & & \\ \beta_1 & \alpha_2 & \beta_2 & \cdots & & 0 \\ 0 & \beta_2 & \alpha_3 & & & \\ & \vdots & & \ddots & & \vdots \\ & & & \cdots & \alpha_{j-1} & \beta_{j-1} \\ & 0 & & & \beta_{j-1} & \alpha_j \end{pmatrix}$$

a) Derive an algorithm to compute the vectors q_j and the numbers α_j and β_j iteratively from the fact that $AQ_n = Q_n T_n$. This algorithm is known as the Lanczos algorithm.
b) Let M_j be the largest eigenvalue of T_j. Show that M_j increases as j increases and that M_n is equal to the largest eigenvalue of A, λ_1.
Hint: Recall that $M_j = \max_{x \in \mathbb{R}^j} x^T T_j x$ and $\lambda_1 = \max_{x \in \mathbb{R}^n} x^T A x$.
c) A standard approach for finding λ_1 is to use Householder transformations to tridiagonalize A and then use the QR algorithm. Suppose A is large and sparse. How could you use the Lanczos algorithm to improve on this method? What are the advantages of this alternative approach?