The University Of Sheffield.

SCHOOL OF MATHEMATICS AND STATISTICS — Autumn Semester 2008-2009

NUMERICAL LINEAR ALGEBRA — Two hours

Marks will be awarded for your best FOUR answers

1 (i) (a) Write down the conditions that must be satisfied by any matrix norm, $\|A\|$, and define the *subordinate matrix norm.* ***(5 marks)***

(b) Compute $\|A\|_1$ and $\|A\|_\infty$ for the matrix

$$A = \begin{bmatrix} 1 & -3 & 7 & -8 \\ -4 & 2 & 9 & 14 \\ 6 & 5 & -8 & 0 \\ 13 & -6 & 13 & -2 \end{bmatrix}.$$

(4 marks)

(ii) We wish to solve $A\mathbf{x} = \mathbf{b}$ where $\mathbf{b}$ is known exactly and A is subject to an uncertainty δA. Thus, in effect we necessarily solve

$$(A + \delta A)(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b}.$$

Define the condition number, $\mathcal{K}(A)$, and hence, assuming that $\|A^{-1}\|\|\delta A\| < 1$, prove that the relative error in $\mathbf{x}$ satisfies

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\mathcal{K}(A)}{\left\{1 - \mathcal{K}(A)\frac{\|\delta A\|}{\|A\|}\right\}} \frac{\|\delta A\|}{\|A\|}.$$

(9 marks)

(iii) We wish to solve $A\mathbf{x} = \mathbf{b}$ where

$$A \approx \begin{bmatrix} 1.0000 & 0.2500 & 0.1111 & 0.0625 \\ 0.2500 & 0.1111 & 0.0625 & 0.0400 \\ 0.1111 & 0.0625 & 0.0400 & 0.0278 \\ 0.0625 & 0.0400 & 0.0278 & 0.0204 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3.14159 \\ 3.14159 \\ 3.14159 \\ 3.14159 \end{bmatrix}.$$

Given that A is subject to an uncertainty $\|\delta A\| \approx 2 \times 10^{-5}$ and $\mathbf{b}$ is subject to an uncertainty $\|\delta\mathbf{b}\| \approx 3 \times 10^{-6}$ and that, in this case, the relative error in $\mathbf{x}$ satisfies

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\mathcal{K}(A)}{\left\{1 - \mathcal{K}(A)\frac{\|\delta A\|}{\|A\|}\right\}} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}\right),$$

then:

(a) write ***scilab*** code to determine bounds on $\|\delta\mathbf{x}\|_1/\|\mathbf{x}\|_1$; ***(5 marks)***

(b) comment on the usefulness of solutions computed for this system given this calculation yields $\|\delta\mathbf{x}\|_1/\|\mathbf{x}\|_1 \leq 1.42$, . ***(2 marks)***

2 (i) Given $m+1$ data points $(x_j,\ f_j)$, $j = 0, 1, \ldots, m$, where the x_j values are all distinct, and a set of suitably chosen basis functions, $\phi_0(x)$, $\phi_1(x)$, $\phi_2(x)$, then derive the normal equations

$$\begin{aligned} a_0 \sum_{j=0}^{m} \phi_0(x_j)\phi_k(x_j) + a_1 \sum_{j=0}^{m} \phi_1(x_j)\phi_k(x_j) + a_2 \sum_{j=0}^{m} \phi_2(x_j)\phi_k(x_j) &= \sum_{j=0}^{m} f_j \phi_k(x_j), \\ k &= 0, 1, 2 \end{aligned}$$

for determining the best least square fit to the data of the function

$$Y(x) = a_0\phi_0(x) + a_1\phi_1(x) + a_2\phi_2(x)$$

without weights.

(8 marks)

(ii) Hence, by choosing the basis functions $\phi_0(x) \equiv 1$, $\phi_1(x) \equiv x$ and $\phi_2(x) \equiv x^2$, write down the normal equations for the best quadratic fit, and write ***scilab*** code to compute this fit for the data

x_j	1.000	1.500	2.000	2.500	3.000	3.500	4.000	4.500
f_j	8.00	11.13	15.00	18.88	22.00	23.63	23.00	19.38

(10 marks)

(iii) Given that the best quadratic fit of the previous part is given by

$$P_2(x) = -2.25\,x^2 + 16.38\,x - 7.44$$

then, by computing the residuals on the sub-range $1.5 \leq x \leq 4.0$ only, comment briefly on whether or not you think the degree of the polynomial should be increased. ***(7 marks)***

3 (i) (a) Verify, by direct expansion, that the matrix equation

$$A^T A\alpha = A^T \mathbf{f} \tag{1}$$

where

$$A = \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^n \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_m \end{pmatrix}, \quad \alpha = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}$$

represents the normal equations arising from a least squares fit of the n-th order polynomial, $P_n(x)$, to the data $(x_j,\ f_j)$, $j = 0..m$.

(3 marks)

(b) Verify that, if P is an orthogonal matrix, then equations (1) remain as the normal equations when the transformed residuals

$$\hat{\mathbf{r}} \equiv P\mathbf{r} = P(A\alpha - \mathbf{f})$$

are minimized ***(3 marks)***

(ii) For any $(m+1) \times 1$ vector $\mathbf{w}$ satisfying $\|\mathbf{w}\| \neq 0$, an orthogonal matrix P can be defined according to

$$P = \left(I - 2\frac{\mathbf{w}\mathbf{w}^T}{\mathbf{w}^T\mathbf{w}} \right).$$

Suppose that $\mathbf{w}$ is defined according to

$$\mathbf{w} = \mathbf{x} + sign(x_0)\,||\mathbf{x}||_2\,\mathbf{e}_0,$$

for some arbitrary $\mathbf{x} = (x_0, x_1, \ldots, x_m)^T$ and where $\mathbf{e}_0$ is the first column of the $(m+1) \times (m+1)$ identity matrix. Show that $\hat{\mathbf{x}}$, defined according to

$$\hat{\mathbf{x}} = P\mathbf{x},$$

satisfies

$$\hat{\mathbf{x}} = -sign(x_0)||\mathbf{x}||_2\,\mathbf{e}_0.$$

(7 marks)

 Question 3 continued on next page

3 (continued)

(iii) It is required that $y = a_0 + a_1 x + a_2 x^2$ is fitted to a certain set of data in the least-squares sense. The residuals arising from this fit are given by

$$\mathbf{r} \equiv \begin{pmatrix} 1 & 0.0 & 0.00 \\ 1 & 0.2 & 0.04 \\ 1 & 0.4 & 0.16 \\ \vdots & \vdots & \vdots \\ 1 & 2.0 & 4.00 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} - \begin{pmatrix} 0.4000 \\ 0.2046 \\ 0.0329 \\ \vdots \\ 0.5972 \end{pmatrix}$$

where (a_0, a_1, a_2) are chosen such that $\|\mathbf{r}\|_2$ is minimised. A series of orthogonal transformations $P \equiv P_4 P_3 P_2 P_1$, based on the Housholder Reflection matrix, is applied to $\mathbf{r}$ to obtain $\hat{\mathbf{r}} \equiv P\mathbf{r}$ where

$$\hat{\mathbf{r}} \equiv \begin{pmatrix} -3.3166 & -3.3166 & -4.6433 \\ 0.0000 & 2.0977 & 4.1952 \\ 0.0000 & 0.0000 & 1.1717 \\ 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 \\ \vdots & \vdots & \vdots \\ 0.0000 & 0.0000 & 0.0000 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} - \begin{pmatrix} -0.2517 \\ 0.1432 \\ 0.8349 \\ -0.0662 \\ 0.0000 \\ \vdots \\ 0.0000 \end{pmatrix}.$$

Determine the values of (a_0, a_1, a_2) which minimizes $\|\hat{\mathbf{r}}\|_2$ and state the value of this minimized $\|\hat{\mathbf{r}}\|_2$. Work correct to four decimal places.

(7 marks)

(iv) We wish to apply a single Householder reflection to the matrix

$$A = \begin{bmatrix} 1 & 2.2 \\ 1 & 2.4 \\ 1 & 2.6 \\ 1 & 2.8 \\ 1 & 3.0 \end{bmatrix}.$$

Write ***scilab*** code to accomplish this task.

(5 marks)

4 (i) (a) The real symmetric matrix A has eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ satisfying

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq ... \geq |\lambda_n| > 0$$

with corresponding linearly independent eigenvectors $\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n$ which can be supposed normalized so that the largest element of each one is unity. The iteration

$$\begin{aligned} \mathbf{y}_k &= A\mathbf{z}_{k-1} \\ \mathbf{z}_k &= \frac{\mathbf{y}_k}{\mu_k}, \quad k = 1, 2, ..., \end{aligned}$$

where $\mathbf{z}_0$ is normalized so that its largest element is unity, and where μ_k is chosen so that the largest element of $\mathbf{z}_k$ is likewise unity, is called the *Power Method.* Prove that, for this iteration, $(\mu_k, \mathbf{z}_k) \to (\lambda_1, \mathrm{x}_1)$ as $k \to \infty$. ***(10 marks)***

(b) Supposing that A and $\mathbf{z}_0$ are already defined, then write ***scilab*** code which impliments ten iterations of the Power Method to produce an estimate of $(\lambda_1, \mathrm{x}_1)$. ***(5 marks)***

(ii) (a) Suppose now that A is a real symmetric matrix and that an estimate of $(\lambda_2, \mathrm{x}_2)$ is also required. The method of *Hotelling's Deflation,*

$$B = A - \lambda_1 \mathrm{x}_1 \mathrm{x}_1^T ,$$

where the eigenvectors are supposed normalized so that $\mathrm{x}_k^T \mathrm{x}_k = 1$, is useful for this purpose since B has eigenvalues $0, \lambda_2, \lambda_3, ..., \lambda_n$. Prove this latter statement. ***(5 marks)***

(b) Write additional ***scilab*** code which takes the results of your Power Method code (above) and uses Hotelling's Deflation to produce the matrix B, and then applies ten iterations of the Power Method to the matrix B to obtain an estimate of $(\lambda_2, \mathrm{x}_2)$. ***(5 marks)***

5 (i) The linear system

$$A\mathsf{x} = \mathbf{b},$$

where A is an $n \times n$ matrix of known coefficients, $\mathbf{b}$ is an $n \times 1$ column vector of known values can be rearranged, for the solution vector $\mathbf{x}$, in arbitrarily many ways in the form

$$\mathsf{x} = H\mathsf{x} + \mathbf{d}$$

which can subsequently be used to define the iteration

$$\mathsf{x}^{(k+1)} = H\mathsf{x}^{(k)} + \mathbf{d} \tag{2}$$

where H is some $n \times n$ matrix and $\mathbf{d}$ is a column vector of known values.

(a) Derive a sufficient condition, written in terms of $\|H\|$, which will guarantee that the iteration (2) above will give a sequence of iterates convergent to x, the solution of $A\mathsf{x} = \mathbf{b}$. ***(5 marks)***

(b) Starting from $A\mathsf{x} = \mathbf{b}$, write down the Jacobi iteration, and hence prove that **strict diagonal dominance** of the matrix A is sufficient to guarantee the convergence of the method. ***(8 marks)***

(ii) (a) Use the Jacobi iterative method to obtain two successive approximations to the solution of the system $A\mathsf{x} = \mathbf{b}$ where

$$A = \begin{pmatrix} 12 & -5 & -2 \\ -5 & 12 & -3 \\ -2 & -3 & 12 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

using $\mathsf{x}^{(0)} = (0, 0, 0)^T$ as the starting vector and working correct to four significant figures. ***(5 marks)***

(b) Write ***scilab*** code with implements the Jacobi method to generate an approximate solution to the system $A\mathsf{x} = \mathbf{b}$ defined above that will iterate until $\|A\mathsf{x} - \mathbf{b}\|_\infty \leq 10^{-5}$. ***(7 marks)***

End of Question Paper