STAT 309: MATHEMATICAL COMPUTATIONS I
FALL 2012
PROBLEM SET 2

1. Let $\mathbf{x} \in \mathbb{C}^n$ and $A \in \mathbb{C}^{m \times n}$. We write $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^*\mathbf{x}}$ and $\|A\|_2 = \sup_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2$ for the vector 2-norm and matrix 2-norm respectively.

(a) Show that there is no ambiguity in the notation, i.e. if $A \in \mathbb{C}^{n \times 1} = \mathbb{C}^n$, then $\|A\|_2$ is the same whether we regard it as the vector or matrix 2-norm. What if $A \in \mathbb{C}^{1 \times n}$?

(b) Show that the vector 2-norm is unitarily invariant, i.e.

$$\|U\mathbf{x}\|_2 = \|\mathbf{x}\|_2$$

for all unitary matrices $U \in \mathbb{C}^{n \times n}$. *Bonus*: Show that no other vector p-norm is unitarily invariant, $1 \le p \le \infty$, $p \ne 2$.

(c) Show that the matrix 2-norm is unitarily invariant, i.e.

$$\|UAV\|_2 = \|A\|_2$$

for all unitary matrices $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$.

(d) Show that the Frobenius norm is unitarily invariant, i.e.

$$\|UAV\|_F = \|A\|_F$$

for all unitary matrices $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$. (*Hint*: First show that $\|A\|_F^2 = \operatorname{tr}(A^*A) = \operatorname{tr}(AA^*)$).

(e) Let $U \in \mathbb{C}^{n \times n}$. Show that the following are equivalent statements:

- (i) $\|U\mathbf{x}\|_2 = \|\mathbf{x}\|_2$ for all $\mathbf{x} \in \mathbb{C}^n$;
- (ii) $(U\mathbf{x})^*U\mathbf{y} = \mathbf{x}^*\mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$;
- (iii) U is unitary.

2. Let $A \in \mathbb{C}^{n \times n}$. Let $\|\cdot\|$ be an operator norm of the form

$$\|A\| = \max_{\mathbf{0} \ne \mathbf{v} \in \mathbb{C}^n} \frac{\|A\mathbf{v}\|_\alpha}{\|\mathbf{v}\|_\alpha} \tag{2.1}$$

for some vector norm $\|\cdot\|_\alpha : \mathbb{C}^n \to [0, \infty)$. Show that if $\|A\| < 1$, then $I - A$ is nonsingular and furthermore,

$$\frac{1}{1 + \|A\|} \le \|(I - A)^{-1}\| \le \frac{1}{1 - \|A\|}.$$

3. Recall that in the lectures, we mentioned that (i) there are matrix norms that are not submultiplicative and an example is the Hölder ∞-norm; (ii) we may always construct a norm that approximates the spectral radius of a given matrix A as closely as we want.

(a) Show that if $\|\cdot\| : \mathbb{C}^{m \times n} \to \mathbb{R}$ is a norm, then there always exists a $c > 0$ such that the constant multiple $\|\cdot\|_c := c\|\cdot\|$ defines a submultiplicative norm, i.e.

$$\|AB\|_c \le \|A\|_c \|B\|_c$$

for any $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{n \times p}$ (even if $\|\cdot\|$ does not have this property). Find the constant c for the Hölder ∞-norm.

Date: October 24, 2012 (Version 1.4); due: November 1, 2012.

(b) Let $J \in \mathbb{C}^{n\times n}$ be in Jordan form, i.e.

$$J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{bmatrix}$$

where each block J_r, for $r = 1, \dots, k$, has the form

$$J_r = \begin{bmatrix} \lambda_r & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_r \end{bmatrix}.$$

Let $\varepsilon > 0$ and $D_\varepsilon = \operatorname{diag}(1, \varepsilon, \varepsilon^2, \dots, \varepsilon^{n-1})$. Verify that

$$D_\varepsilon^{-1} J D_\varepsilon = \begin{bmatrix} J_{1,\varepsilon} & & \\ & \ddots & \\ & & J_{k,\varepsilon} \end{bmatrix}$$

where $J_{r,\varepsilon}$ is the matrix you obtain by replacing the 1's on the superdiagonal of J_r by ε's,

$$J_{r,\varepsilon} = \begin{bmatrix} \lambda_r & \varepsilon & & \\ & \ddots & \ddots & \\ & & \ddots & \varepsilon \\ & & & \lambda_r \end{bmatrix}$$

(c) Show that

$$\|D_\varepsilon^{-1} J D_\varepsilon\|_\infty \le \rho(J) + \varepsilon.$$

(d) Hence, or otherwise, show that for any given $A \in \mathbb{C}^{n\times n}$ and $\varepsilon > 0$, there exists an operator norm $\|\cdot\|$ of the form (2.1) with the property that

$$\rho(A) \le \|A\| \le \rho(A) + \varepsilon.$$

(*Hint*: Transform A into Jordan form).

4. Let $\|\cdot\|$ be an operator norm of the form (2.1). When we proved that $\rho(A) \le \|A\|$ in the lectures, we used a very simple argument: if $\lambda \in \mathbb{C}$ is an eigenvalue of A with eigenvector $\mathbf{v} \in \mathbb{C}^n$ of unit norm, then $|\lambda| = \|\lambda\mathbf{v}\|_\alpha = \|A\mathbf{v}\|_\alpha \le \|A\|$. The trouble is that when $A \in \mathbb{R}^{n\times n}$, then $\|A\|$ is often defined as

$$\max_{\mathbf{0}\ne\mathbf{v}\in\mathbb{R}^n} \frac{\|A\mathbf{v}\|_\alpha}{\|\mathbf{v}\|_\alpha} \tag{4.2}$$

instead of (2.1), i.e. the maximum is taken over all real non-zero vectors instead of all complex non-zero vectors. Since an eigenvector is complex in general, the simple argument does not work when we use (4.2). Here we will show that $\rho(A) \le \|A\|$ is true for $A \in \mathbb{R}^{n\times n}$ even if we use (4.2) as the definition of $\|A\|$ [Thanks to Ridg Scott for this problem and the proof below].

(a) Let $A \in \mathbb{R}^{n\times n}$. Let $\lambda \in \mathbb{C}$ be the largest eigenvalue of A in magnitude, i.e. $|\lambda| = \rho(A) =: \rho$ and let $\mathbf{z} \in \mathbb{C}^n$ be a λ-eigenvector. Write λ in polar form

$$\lambda = \rho(\cos\theta + i\sin\theta)$$

for some $\theta \in [0, 2\pi]$ and write $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Show that

$$A\mathbf{x} = \rho(\cos\theta\mathbf{x} - \sin\theta\mathbf{y}), \quad A\mathbf{y} = \rho(\sin\theta\mathbf{x} + \cos\theta\mathbf{y})$$

and deduce that

$$A(\cos\phi\mathbf{x} + \sin\phi\mathbf{y}) = \rho[\cos(\phi - \theta)\mathbf{x} + \sin(\phi - \theta)\mathbf{y}]$$

for all $\phi \in [0, 2\pi]$.

(b) Define f on $[0, 2\pi]$ by

$$f(\phi) = \frac{\|A(\cos\phi\mathbf{x} + \sin\phi\mathbf{y})\|_\alpha}{\|\cos\phi\mathbf{x} + \sin\phi\mathbf{y}\|_\alpha}.$$

Show that f is continuous and there exist $\phi_1, \phi_2 \in [0, 2\pi]$ such that

$$f(\phi_1) \le \rho \le f(\phi_2).$$

Hence, by the intermediate value theorem, there exists ϕ_* such that

$$f(\phi_*) = \rho,$$

and thus

$$\max_{\mathbf{0}\neq\mathbf{v}\in\mathbb{R}^n} \frac{\|A\mathbf{v}\|_\alpha}{\|\mathbf{v}\|_\alpha} \ge \frac{\|A(\cos\phi_*\mathbf{x} + \sin\phi_*\mathbf{y})\|_\alpha}{\|\cos\phi_*\mathbf{x} + \sin\phi_*\mathbf{y}\|_\alpha} = \rho = \rho(A).$$

(c) Is it true that for $A \in \mathbb{R}^{n\times n}$ and $\|\cdot\|_\alpha$ any vector norm defined on $\mathbb{C}^n$, we will always have

$$\max_{\mathbf{0}\neq\mathbf{v}\in\mathbb{R}^n} \frac{\|A\mathbf{v}\|_\alpha}{\|\mathbf{v}\|_\alpha} = \max_{\mathbf{0}\neq\mathbf{v}\in\mathbb{C}^n} \frac{\|A\mathbf{v}\|_\alpha}{\|\mathbf{v}\|_\alpha}?$$

5. Let $A = [a_{ij}]$ be an $n \times n$ matrix with entries

$$a_{ij} = \begin{cases} n + 1 - \max(i, j) & i \le j + 1, \\ 0 & i > j + 1. \end{cases}$$

This is an example of an *upper Hessenberg* matrix: it is upper triangular except that the entries on the subdiagonal $a_{i,i+1}$ may also be non-zero. For $n = 12$ and $n = 25$, do the following:

(a) Compute $\|A\|_\infty$ and $\|A\|_1$.

(b) Compute $\rho(A)$ and $\|A\|_2$.

(c) Using Gerschgorin's theorem, describe the domain that contains all of the eigenvalues.

(d) Compute all of the eigenvalues and singular values of A. How many of the eigenvalues are real and how many are complex?

6. Let $A = [a_{ij}]$ be a 32×32 matrix defined by

$$a_{ij} := \begin{cases} 1 & \text{if } j = i, \\ i - 11 & \text{if } j = i + 1,\ i < 11, \\ i - 10 & \text{if } j = i + 1,\ i \ge 11, \\ 0 & \text{otherwise.} \end{cases} \tag{6.3}$$

In other words, A is a bidiagonal matrix with 1's on its diagonal, $-10, -9, \dots, -1, 1, \dots, 21$ on its superdiagonal, and 0's everywhere else.

(a) Construct the Gerschgorin disks for A.

(b) Let $\varepsilon > 0$. Construct a diagonal matrix D so that DAD^{-1} is bidiagonal with 1's on its diagonal, ε's on its superdiagonal, and 0's everywhere else, ie.

$$DAD^{-1} = \begin{bmatrix} 1 & \varepsilon & & & \\ & 1 & \varepsilon & & \\ & & 1 & \ddots & \\ & & & \ddots & \varepsilon \\ & & & & 1 \end{bmatrix}. \tag{6.4}$$

What are the Gerschgorin disks for DAD^{-1}?

(c) Give an algorithm to reduce A in (6.3) to the form in (6.4) with $\varepsilon = 10^{-4}$. How stable is your algorithm?