Math128B
Feb. 1, 2005
Jonathan Dorfman
Homework 2 Solutions

Problem 2.1

First, a quick review of some facts about complex numbers:

1. for $z = x + iy = r\cos\theta + ir\sin\theta = re^{i\theta} \in \mathbb{C}$ we have (magnitude of complex scalar)
$$|z| \stackrel{\text{def}}{=} \sqrt{x^2 + y^2} = r$$
2. if we think of $z = x + iy$ as a point in the "complex plane", then
$$|z| = \|(x, y)\|_2 = \text{Euclidean length of } (x, y) \in \mathbb{R}^2$$
3. since the triangle inequality for the Euclidean norm states:
$$\|(x, y) + (X, Y)\|_2 \leq \|(x, y)\|_2 + \|(X, Y)\|_2$$
we also know $|z + w| \leq |z| + |w|$
4. if $\alpha = a + ib = re^{i\theta}$ and $z = x + iy = Re^{i\Theta}$ then
$$|\alpha z| = |re^{i\theta} \cdot Re^{i\Theta}| = |rRe^{i(\theta+\Theta)}| = Rr = |\alpha| \cdot |z|$$

Now for $v = (z_1, z_2, \ldots, z_n) = (x_1 + iy_1, x_2 + iy_2, \ldots, x_n + iy_n) \in \mathbb{C}^n$ we define
$$\|v\|_1 \stackrel{\text{def}}{=} |z_1| + |z_2| + \cdots + |z_n| = \|(x_1, y_1)\|_2 + \|(x_2, y_2)\|_2 + \cdots + \|(x_n, y_n)\|_2$$
We must verify that the defining properties of a norm are satisfied.

1. $\|v\|_1 \geq 0$ and $= 0 \iff v = 0$ Proof: This follows easily from the fact that each of the terms $\|(x_k, y_k)\|_2 \geq 0$ and $= 0 \stackrel{(2)}{\iff} x_k = y_k = 0$
2. $\|\alpha \cdot v\|_1 = |\alpha|\|v\|_1 (\forall \alpha \in \mathbb{C})$ Proof: This follows easily from the definitions:
$$\|\alpha \cdot v\|_1 \stackrel{\text{def}}{=} \|(\alpha z_1, \ldots, \alpha z_n)\|_1 \stackrel{\text{def}}{=} |\alpha z_1| + |\alpha z_2| + \cdots + |\alpha z_n| \stackrel{(4)}{=} |\alpha||z_1| + \cdots + |\alpha||z_n| = |\alpha|\big(|z_1| + \cdots + |z_n|\big) \stackrel{\text{def}}{=} |\alpha|\|v\|_1$$
3. $\|v + w\|_1 \leq \|v\|_1 + \|w\|_1 (\forall v, w \in \mathbb{C}^n)$ Proof: Again this follows from the corresponding property of the Euclidean-norm in $\mathbb{R}^2$ applied to each of the terms:
$$\|v + w\|_1 = \|(z_1, \ldots, z_n) + (w_1, \ldots, w_n)\|_1 = \|(z_1 + w_1, \ldots, z_n + w_n)\|_1$$
$$= |z_1 + w_1| + \cdots + |z_n + w_n| \stackrel{(3)}{\leq} |z_1| + |w_1| + \cdots + |z_n| + |w_n| = \|v\|_1 + \|w\|_1 \quad \blacksquare$$

Finally, $\|v\|_1 \geq \|v\|_2$ follows easily from the inequality
$$\|v\|_1^2 = \sum_{k=1}^{n} |z_k|^2 + 2\sum_{i=1}^{n}\sum_{j>i}^{n} |z_i||z_j| \geq \sum_{k=1}^{n} |z_k|^2 = \|v\|_2^2$$

Problem 2.2 & 2.3

file	function
Hmwk2Main.m	receives matrix A as input, calls EigenInfo, DisplayEigenInfo, FormatPoly
EigenInfo.m	calls Matlab functions to compute eigenvalues, eigenvectors, norms
DisplayEigenInfo.m	displays eigenvalues and eigenvectors in formatted text output
FormatPoly.m	displays polynomial in formatted output

The main functionality is contained in the EigenInfo function:

```
%EigenInfo.m
%Math 128B Spring 2005
%Hmwk2 - Solutions (Feb. 1, 2005)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [V, eigValues, radius, charPoly] = EigenInfo(A)

%use Matlab's poly to return characteristic polynomial
charPoly = poly(A);

%use Matlab's eig to return matrix of eigenvalues and matrix of eigenvectors
%V(:,i)=i'th eigenvector corresponding to eigenvalue D(i,i)
[V, D] = eig(A);

%eigenvalues are diag(D)
eigValues = diag(D);

%spectral radius is max of absolute values (= complex norm) of eigenvalues
radius = norm(eigValues, inf);

return
```

A diary of the output follows:

```
A =

     1     2
     3     2

Hmwk2Main(A)
spectral radius =  4

Characteristic Polynomial =  1x^2  -  3x   -  4
eigenvalues: -1,  4

eigenvector corresponding to eigVal=-1:
       [-0.70711     0.70711]
eigenvector corresponding to eigVal=4:
       [-0.5547    -0.83205]
```

```
Norms:
inf-norm of A = 5

2-norm of A = 4.1306
- - - - - - - - - - - - - - - - - - - - - - -
A =

    -2     3
     3    -2

Hmwk2Main(A)
spectral radius =  5

Characteristic Polynomial =  1x^2  +  4x   -  5
eigenvalues: -5,  1

eigenvector corresponding to eigVal=-5:
       [0.70711    -0.70711]
eigenvector corresponding to eigVal=1:
       [0.70711     0.70711]

Norms:
inf-norm of A = 5

2-norm of A = 5
- - - - - - - - - - - - - - - - - - - - - - -
A =

     1     1     1     1
     0     2     2     3
     0     0     3     2
     0     0     0     4

Hmwk2Main(A)
spectral radius =  4

Characteristic Polynomial =  1x^4  -  10x^3  +  35x^2  -  50x   +  24
eigenvalues: 1,  2,  3,  4

eigenvector corresponding to eigVal=1:
       [1  0  0  0]
eigenvector corresponding to eigVal=2:
       [0.70711     0.70711           0           0]
eigenvector corresponding to eigVal=3:
       [0.55709     0.74278     0.37139           0]
eigenvector corresponding to eigVal=4:
       [0.46252     0.74715     0.42694     0.21347]

Norms:
inf-norm of A = 7

2-norm of A = 6.2813
```

Problem 2.4

We wish to establish (for A real): $\qquad A = A^t \Longrightarrow \|A\|_2 = \rho(A)$

First, we know that *for any matrix A* $\;\rho(A) \leq \|A\|_2$, so we need only show that when A is symmetric the reverse inequality holds.
Since A is symmetric we know by the "Principal Axis Theorem" that there exists an orthonormal basis of eigenvectors $\mathbf{f_k}$ for which

$$A(\mathbf{f_k}) = \lambda_k \mathbf{f_k} \tag{1}$$

where the $\{\lambda_k\}$ are the (real) eigenvalues of A. Also, since the $\{\mathbf{f_k}\}$ are orthonormal we know that every x can be rewritten in terms of the $\{\mathbf{f_k}\}$:

$$x = \sum_{k=1}^{n} \langle x, \mathbf{f_k} \rangle \mathbf{f_k} = \sum_{k=1}^{n} x_k \mathbf{f_k} \tag{2}$$

and by the above formula for A

$$Ax \overset{(2)}{=} A\Big(\sum_{k=1}^{n} x_k \mathbf{f_k}\Big) = \sum_{k=1}^{n} A(x_k \mathbf{f_k}) = \sum_{k=1}^{n} x_k A\mathbf{f_k} \overset{(1)}{=} \sum_{k=1}^{n} x_k \lambda_k \mathbf{f_k} \tag{3}$$

Next, letting $M = \max\{\lambda_k^2\} = \rho(A)^2$, we compute Ax:

$$\begin{aligned}
\|Ax\|_2^2 &\overset{\text{def}}{=} \big\langle Ax, Ax \big\rangle \\
\text{(by (3) above)} &= \Big\langle \sum_{j=1}^{n} x_j \lambda_j \mathbf{f_j}, \sum_{k=1}^{n} x_k \lambda_k \mathbf{f_k} \Big\rangle \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} x_j \cdot x_k \cdot \lambda_j \cdot \lambda_k \cdot \big\langle \mathbf{f_j}, \mathbf{f_k} \big\rangle \\
\text{(since } \mathbf{f_j} \text{ are orthonormal)} &= \sum_{j=1}^{n} \sum_{k=1}^{n} x_j \cdot x_k \cdot \lambda_j \cdot \lambda_k \cdot \delta_{j,k} \\
\text{(only } j = k \text{ terms survive)} &= \sum_{k=1}^{n} x_k^2 \cdot \lambda_k^2 \\
&\leq \max\{\lambda_k^2\} \sum_{k=1}^{n} x_k^2 \\
\text{(definition of } M\text{)} &= M \cdot \|x\|_2^2
\end{aligned}$$

Dividing both sides by $\|x\|_2^2$ we get:

$$\frac{\|Ax\|_2^2}{\|x\|_2^2} \leq M$$

and taking the sup over $x \neq 0$ of the LHS we get $\|A\|_2^2 \leq M$, which implies $\|A\|_2 \leq \sqrt{M} = \rho(A)$ ∎

Remark: If you're not used to this type of abstract argument, a much more concrete presentation is given on the following pages. Also, if you use the formula $\|A\|_2 = \sqrt{\rho(A^*A)}$ (Theorem 7.15, part (i)) then the result is quite immediate since $\|A\|_2^2 = \rho(A^*A) = \rho(A^2) = \rho(A)^2$ - but this is not really valid (in spite of the author's comment on page 434) since its proof could (the one I know in fact does) logically depend on the result you're asked to prove in this homework assignment.

Problem 2.4 (concrete version)

We wish to establish (for A real): $\quad A = A^t \Longrightarrow \|A\|_2 = \rho(A)$

We break the proof up into several steps:

1. The result is true if A is a real diagonal matrix: $A = D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$

2. The LHS of the equality is unchanged if A is replaced by UAU^{-1} with U orthogonal:
$$\|UAU^{-1}\|_2 = \|A\|_2$$

3. The RHS of the equality is unchanged if A is replaced by UAU^{-1} with U orthogonal:
$$\rho(UAU^{-1}) = \rho(A)$$

4. Finally, for every real symmetric A there exists an orthogonal U such that $UAU^{-1} = D$ is (real) diagonal.

Proof of (1): If $A = D$ and $M = \max\{|\lambda_k|\}$ and $\|x\|_2 = 1$ then

$$\begin{aligned}
\|Ax\|_2^2 &= \big\langle Ax, Ax \big\rangle \\
&= \big\langle (\lambda_1 x_1, \dots, \lambda_n x_n), (\lambda_1 x_1, \dots, \lambda_n x_n) \big\rangle \\
&= (\lambda_1 x_1)^2 + \dots + (\lambda_n x_n)^2 \\
&= \lambda_1^2 x_1^2 + \dots + \lambda_n^2 x_n^2 \\
&\leq (M^2)(x_1^2 + \dots + x_n^2) \\
&= M^2 \|x\|_2^2 \\
&= M^2
\end{aligned}$$

Since the above inequality holds for each x with $\|x\|_2 = 1$, it will remain true after taking the sup of the $\big($ LHS $\big)$ over all such x. But $\sup\limits_{\|x\|_2=1} \big(\text{LHS}\big) \overset{\text{def}}{=} \|A\|_2^2, \quad$ so we've proved $\|A\|_2 \leq \max\{|\lambda_k|\}$
Also note that if k is the index for which $M = |\lambda_k|$ then letting $x = \mathbf{e_k}$ gives $Ax = \lambda_k \mathbf{e_k} = \lambda_k x$ which gives $\|Ax\|_2 = |\lambda_k| \|x\|_2 = M$, proving the LHS attains the value M. ∎

Remark: The matrix D takes the unit sphere (in the 2-norm) into an ellipsoid whose major axis (and all other principal axes) are along the coordinate axes. The norm of D and the spectral radius of D agree because they're both equal to the length of this major axis.

Proof of (2): We will prove a more general result, namely that $\|UAV\|_2 = \|A\|_2$ for any isometries U, V. Recall U is defined to be an isometry if it preserves lengths:

$$\|Ux\| = \|x\| \qquad \forall x$$

Now:

$$\begin{aligned}
\|UAV\|_2 &\stackrel{\text{def}}{=} \sup_{x\neq 0} \frac{\|UAVx\|}{\|x\|} \\
\text{(since } U \text{ is an isometry)} &= \sup_{x\neq 0} \frac{\|AVx\|}{\|x\|} \\
\text{(since } V \text{ is an isometry)} &= \sup_{x\neq 0} \frac{\|AVx\|}{\|Vx\|} \\
\text{(since every } y \text{ is } = Vx \text{ for some } x) &= \sup_{y\neq 0} \frac{\|Ay\|}{\|y\|} \\
&\stackrel{\text{def}}{=} \|A\|_2
\end{aligned}$$

In particular, if $V = U^{-1}$ then we have $\|UAU^{-1}\|_2 = \|A\|_2$ for any orthogonal U. (To see why a real matrix is orthogonal iff it's an isometry, see below (**)) ∎

Remark: Since the norm of A is the "diameter" of the image of the unit sphere, and since pre-multiplying A by an isometry leaves the unit sphere in the domain unchanged, and since post-multiplying A by any (possibly different) isometry leaves the diameter of a set in the range unchanged, this result follows.

Proof of (3): Again, we will prove a more general result:

$$\rho(UAU^{-1}) = \rho(A) \qquad \text{for any invertible } U$$

Since $\det(UAU^{-1} - \lambda Id) = \det(U(A - \lambda Id)U^{-1}) = \det(A - \lambda Id)$ we see similar matrices have the same characteristic polynomials hence the same set of eigenvalues - in particular, the maximum (in norm) of their eigenvalues must agree, and so their spectral radius must agree. Since orthogonal matrices are invertible ($\det = \pm 1$), (3) follows.

Proof of (4): This fundamental result called the "principal axis theorem" (or "spectral theorem for self-adjoint operators") is proved in Linear Algebra courses, and can be assumed without proof.

(**) Proof that U is an isometry iff U is orthogonal:

$$\begin{aligned}
\|Ux\| = \|x\| &\iff \|Ux\|^2 = \|x\|^2 \qquad \forall x \\
&\iff \big\langle Ux, Ux \big\rangle = \big\langle x, x \big\rangle \qquad \forall x \\
&\iff \big\langle U^*Ux, x \big\rangle = \big\langle x, x \big\rangle \qquad \forall x \\
\text{(this step is called "polarization")} &\iff \big\langle U^*U(y+z), (y+z) \big\rangle = \big\langle (y+z, y+z) \big\rangle \qquad \forall y, z \\
&\iff \big\langle U^*Uy, z) \big\rangle = \big\langle (y, z) \big\rangle \qquad \forall y, z \\
&\iff \big\langle (U^*U - Id)y, z) \big\rangle = 0 \qquad \forall y, z \\
&\iff (U^*U - Id)y = 0 \qquad \forall y \\
&\iff U^*U - Id = 0 \\
&\iff U^* = U^{-1} \\
&\iff U \text{ is orthogonal}
\end{aligned}$$