Math221 Homework # 2 Solutions
Instructor: James W. Demmel

1.7

Proof:

$$\begin{aligned}
\|xy^H\|_F^2 &= \sum_{i=1}^n \sum_{j=1}^n |x_i \bar{y_j}|^2 \\
&= \sum_{i=1}^n |x_i|^2 \sum_{j=1}^n |\bar{y_j}|^2 \\
&= \|x\|_2^2 \|y\|_2^2
\end{aligned}$$

First proof that $\|xy^H\|_2 = \|x\|_2\|y\|_2$:

$$\begin{aligned}
\|xy^H\|_2^2 &= \lambda_{\max}((xy^H)^H(xy^H)) \text{ by Part 7 of Lemma 1.7} \\
&= \lambda_{\max}(yx^H xy^H) \\
&= (x^H x)\lambda_{\max}(yy^H)
\end{aligned}$$

The rank-1 matrix yy^H has one (eigenvalue,eigenvector) pair equal to $(y^H y, y)$; plug in to confirm this. All the other eigenvalues are equal to 0, with eigenvectors equal to any vector orthogonal to y. Thus $\lambda_{\max}(yy^H) = y^H y$ and $\|xy^H\|_2^2 = \|x\|_2^2\|y\|_2^2$.

Second proof that $\|xy^H\|_2 = \|x\|_2\|y\|_2$:

$$\begin{aligned}
\|xy^H\|_2 &= \max_{z \neq 0} \|xy^H z\|_2 / \|z\|_2 \\
&= \max_{z \neq 0} |y^H z| \|x\|_2 / \|z\|_2
\end{aligned}$$

The dot product is known to satisfy $y^H z = \|y\|_2\|z\|_2 \cos\theta$, where θ is the angle between the vectors y and z. Thus $|y^H z| \leq \|y\|_2\|z\|_2$, and the upper bound is attained when $\theta = 0$, i.e. z is a multiple of y. Thus

$$\begin{aligned}
\|xy^H\|_2 &= \max_{z \neq 0} |y^H z| \|x\|_2 / \|z\|_2 \\
&= \|y\|_2\|y\|_2\|x\|_2 / \|y\|_2 = \|y\|_2\|x\|_2
\end{aligned}$$

1.13.

Solution: Given an inner product $< .,. >$, we construct a Hermitian positive definite (h.p.d.) matrix A such that $< x, y > = y^H Ax$ as follows. If e_i are the unit vectors in C^n then define $A_{ij} = < e_i, e_j >$. A is Hermitian ($A_{ij} = A_{ji}^H$, by a property of the inner product) and

$$< x, y > = < \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j > = \sum_{i=1}^n \sum_{j=1}^n x_i \bar{y_j} < e_i, e_j > = y^H Ax.$$

Also $x^H Ax = < x, x > \geq 0$, and $A^H Ax = < x, x > = 0$ if and only if $x = 0$, so A is h.p.d.

Conversely, if A is h.p.d. then
1) $y^H Ax = x^H Ay$
2) $(y+z)^H Ax = y^H Ax + z^H Ax$
3) $y^H A(ax) = ay^H Ax$
4) $x^H Ax > 0$ iff $x \neq 0, x^H Ax = 0$ iff $x = 0$.
So $y^H Ax$ is an inner product.

1.14

We prove Lemma 1.5:

$\|x\|_2^2 = \sum_i |x_i|^2 \leq (\sum_i |x_i|)^2 = \|x\|_1^2$ since $(\sum_i |x_i|)^2$ includes all terms $|x_i|^2$ and more, so $\|x\|_2 \leq \|x\|_1$.

$\|x\|_1 = \sum_{i=1}^n 1 \cdot |x_i| \leq \sqrt{\sum_{i=1}^n 1^2} \cdot \sqrt{\sum_{i=1}^n |x_i|^2} = \sqrt{n} \cdot \|x\|_2$ by the Cauchy-Schwartz inequality.

$\|x\|_\infty = \max_i |x_i| = \sqrt{\max_i |x_i|^2} \leq \sqrt{\sum_i |x_i|^2} = \|x\|_2$.

$\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} \leq \sqrt{n \cdot \max_i |x_i|^2} = \sqrt{n} \cdot \max_i |x_i| = \sqrt{n} \cdot \|x\|_\infty$.

$\|x\|_\infty = \max_i |x_i| \leq \sum_i |x_i| = \|x\|_1$.

$\|x\|_1 = \sum_{i=1}^n |x_i| \leq n \cdot \max_i |x_i| = n \cdot \|x\|_\infty$.

1.15

To show that an operator norm $\|A\|_{mn} = \max_{x \neq 0} \|Ax\|_m / \|x\|_n$ is a matrix norm, we need to check that it satisfies all 3 parts of Definition 1.7.

$\|A\|_{mn} \geq 0$ because it is defined as the quotient of nonnegative quantities. It can only equal 0 if $\|Ax\|_m = 0$ for all $x \neq 0$, which can happen if and only if $Ax = 0$ for all $x \neq 0$ (since $\|\cdot\|_m$ is a vector norm), which is true if and only $A = 0$.

$\|\alpha \cdot A\|_{mn} = \max_{x \neq 0} \|\alpha \cdot Ax\|_m / \|x\|_n = \max_{x \neq 0} |\alpha| \cdot \|Ax\|_m / \|x\|_n$ (again since $\|\cdot\|_m$ is a vector norm), and this in turn equals $|\alpha| \cdot \max_{x \neq 0} \|Ax\|_m / \|x\|_n = |\alpha| \cdot \|A\|_{mn}$.

$\|A+B\|_{mn} = \max_{x \neq 0} \|(A+B)x\|_m / \|x\|_n = \max_{x \neq 0} \|Ax + Bx\|_m / \|x\|_n \leq \max_{x \neq 0} (\|Ax\|_m + \|Bx\|_m) / \|x\|_n$ since $\|\cdot\|_m$ satisfies the triangle inequality. This last expression is in turn at most $\max_{x \neq 0} \|Ax\|_m / \|x\|_n + \max_{x \neq 0} \|Bx\|_m / \|x\|_n$ since maximizing over the two parts separately can only make it larger. This last expressions equals $\|A\|_{mn} + \|B\|_{mn}$.

1.16.1.

$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$ so $\|A\| \cdot \|x\| \geq \|Ax\|$ for $x \neq 0$ and is also true for $x = 0$. For the Frobenius norm from Cauchy-Schwartz we have:

$$\left(\sum_{j=1}^n a_{ij} x_j \right)^2 \leq \left(\sum_{j=1}^n a_{ij}^2 \right) \left(\sum_{j=1} x_j^2 \right)$$

for $1 \leq i \leq n$, so

$$\sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} x_j \right)^2 \leq \sum_{i=1}^n \left[\left(\sum_{j=1}^n a_{ij}^2 \right) \left(\sum_{j=1} x_j^2 \right) \right] = \left(\sum_{i,j=1}^n a_{ij}^2 \right) \left(\sum_{j=1} x_j^2 \right)$$

i.e. $\|A\|_F \|x\|_F \geq \|Ax\|_F$.

1.16.2. For an operator norm, let x be such a vector of norm 1 that $\|AB\| = \|ABx\|$. Then $\|AB\| = \|A(Bx)\| \leq \|A\|\|Bx\| \leq \|A\|\|B\|\|x\| = \|A\|\|B\|$. For the Frobenius norm we have

$$\left(\sum_{k=1}^{n} a_{ik} b_{kj}\right)^2 \leq \left(\sum_{k=1}^{n} a_{ik}^2\right)\left(\sum_{k=1}^{n} b_{kj}^2\right), \quad \text{by Cauchy Schwartz, so}$$

$$\sum_{i,j=1}^{n}\left(\sum_{k=1}^{n} a_{ik} b_{kj}\right)^2 \leq \sum_{i,j=1}^{n}\left[\left(\sum_{k=1}^{n} a_{ik}^2\right)\left(\sum_{k=1}^{n} b_{kj}^2\right)\right] = \left(\sum_{i,k=1}^{n} a_{ik}^2\right)\left(\sum_{j,k=1}^{n} b_{kj}^2\right),$$

i.e. $\|AB\|_F \leq \|A\|_F \|B\|_F$.

1.16.4.

If Q is a unitary matrix, then $\|Qx\|_2 = \|x\|_2$ for any x, by the Pythagorean Theorem (rotating a vector does change its length). Thus

$$\|QAZ\|_2 = \max_{x \neq 0} \frac{\|QAZx\|_2}{\|x\|_2} = \max_{x \neq 0} \frac{\|AZx\|_2}{\|x\|_2} = \max_{x \neq 0} \frac{\|AZx\|_2}{\|Zx\|_2} = \max_{y \neq 0} \frac{\|Ay\|_2}{\|y\|_2} = \|A\|_2$$

Next, again using $\|Qx\|_2 = \|x\|_2$,

$$\|QA\|_F^2 = \sum_i \|QA_{*i}\|_2^2 = \sum_i \|A_{*i}\|_2^2 = \|A\|_F^2$$

where A_{*i} is the i-th column of A. Now apply the same result to Z^H and $(QA)^H$, and use $\|X^H\|_F = \|X\|_F$.

1.16.5.

$$\|A\|_\infty = \max_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{\|x\|_\infty = 1} \|Ax\|_\infty \leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|.$$

We can pick $x_j = \text{sign}(a_{ij})$ for the row that maximizes the above sum.

1.16.6.

$$\|A\|_1 = \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{\|x\|_1 = 1} \|Ax\|_1 = \max_{\|x\|_1 = 1} \sum_{i=1}^{n} \left|\sum_{j=1}^{n} a_{ij} x_j\right|$$

$$\leq \max_{\|x\|_1 = 1} \sum_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}|\right) |x_j| \leq \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|$$

Both inequalities become equalities if we choose $x_j = 1$, where j maximizes the above sum, and other $x_i = 0$.

1.16.8.

First assume A is square. $\|A\|_2 = \sqrt{\lambda_{\max}(A^H A)} = \sqrt{\lambda_{\max}(AA^H)} = \|A^H\|_2$ since the nonzero eigenvalues of XY and YX are identical, for any square X and Y. This is enough for full credit.

We reduce the case of nonsquare A to the square case as follows. Suppose that A is m-by-n with $m > n$. Let $\hat{A}$ be a square matrix obtained by attaching $m-n$ zero columns to the right of A: $\hat{A} = [A, 0]$. It is easy to see from the definition that $\max_{x \neq 0} \|Ax\|_2/\|x\|_2 = \max_{\hat{x} \neq 0} \|\hat{A}\hat{x}\|_2/\|\hat{x}\|_2$, and the maximizing $\hat{x}$ for $\hat{A}$ is gotten by adding $m-n$ zeros to the end of the maximizing x for A. The 2-norms of A^H and $\hat{A}^H$ can be similarly seen to be equal.

1.16.10.

Using Lemma 1.5 we have $\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$. Then for all $x \neq 0$

$$\frac{\|Ax\|_2}{\sqrt{n}\|x\|_2} \leq \frac{\|Ax\|_1}{\|x\|_1} \leq \frac{\sqrt{n}\|Ax\|_2}{\|x\|_2}$$

$$\max_{x \neq 0} \frac{\|Ax\|_2}{\sqrt{n}\|x\|_2} \leq \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} \leq \max_{x \neq 0} \frac{\sqrt{n}\|Ax\|_2}{\|x\|_2}$$

$$n^{-1/2}\|A\|_2 \leq \|A\|_1 \leq n^{1/2}\|A\|_2$$

1.20.2

Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_0$. Consider the perturbed polynomial $p^H(x) = a_n^H x^n + a_{n-1}^H x^{n-1} + ... + a_0^H$, where $a_i^H = a_i(1+e_i), |e_i| \leq e \ll 1$. Let $p(r) = 0$ and $p^H(r(1+\alpha)) = 0$.

Then

$$0 = \sum_{i=0}^{n} a_i(1+e_i)r^i(1+\alpha)^i$$

Assuming $|\alpha| \ll 1$ and linearizing (i.e. taking the first 2 terms of the Taylor expansion)

$$0 = \sum_{i=0}^{n} a_i r^i (1+e_i)(1+i\alpha) = \sum_{i=0}^{n} a_i r^i (1+e_i+i\alpha) \text{ plus smaller terms that we ignore}$$

$$= p(r) + \alpha r p'(r) + \sum_{i=0}^{n} a_i r^i e_i = \alpha r p'(r) + \sum_{i=0}^{n} a_i r^i e_i$$

Therefore

$$|\alpha| = \left| \frac{\sum_{i=0}^{n} a_i r^i e_i}{r p'(r)} \right| \leq \frac{\sum_{i=0}^{n} |a_i r^i|}{|r p'(r)|} e \equiv c(r) \cdot e$$

So we may call $c(r)$ the condition number of a root r. The root (nearest r) of the perturbed polynomial will lie in a disk centered at r with radius $c(r) \cdot |r| \cdot e$.

Here is the code to be added to the end of Polyplot.m in order to plot those circles:

```
radius=e*polyval(abs(p),abs(r))./abs(polyval(polyder(p),r));
z=[0:0.001:2*pi];
for i=1:size(r,2)
   plot(real(r(i))+radius(i).*cos(z),imag(r(i))+radius(i).*sin(z));
end
```

1.20.3

If r is a multiple root we have $p'(r) = 0$, so $c(r) = \infty$. Let $p(x) = q(x)(x-r)^m$. For a slightly perturbed polynomial $p(x) - q(x)\epsilon = q(x)[(x-r)^m - \epsilon]$. If $r + \delta r$ is a root of this polynomial, then $(\delta r)^m = \epsilon$ and $|\delta r| = |\epsilon|^{1/m}$.

If ϵ is about the order of machine epsilon the number of guaranteed correct decimal digits is

$$-\log_{10}\left|\frac{\delta r}{r}\right| = -\log_{10}\frac{|\epsilon|^{1/m}}{|r|}$$

Thus we expect the roots with multiplicity m to be computed with about $(1/m)$-th of the maximum number of significant digits. For example if machine epsilon is about 10^{-16}, so that the maximum number of significant digits is 16, we expect double roots to have at most about 8 digits correct, quadruple roots to have 2 digits correct, and so on.