MATH-6800, CLA: Problem Set 1, 9-17-15

Name: Michael Hennessey

1. Let B be a 4×4 matrix to which we apply the following operations:

   (a) Double column 1

   (b) Halve row 3

   (c) Add row 3 to row 1

   (d) Interchange columns 1 and 4

   (e) Subtract row 2 from each of the other rows

   (f) Replace column 4 by column 3

   (g) Delete column 1

   (a) Write the result as a product of 8 matrices

$$e = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}, c = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, b = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$a = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, d = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, f = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, g = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$ecbBadfg$$

   (b) Write it again as a product ABC

$$\begin{bmatrix} 1 & -1 & 1/2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1/2 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} B \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

2. The Pythagorean theorem asserts that for a set of n orthogonal vectors $\{x_i\}$

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2$$

(a) Prove this is the case for $n = 2$ by an explicit computation of $\|x_1 + x_2\|^2$.

$$\|x_1 + x_2\|^2 = (x_1 + x_2)^*(x_1 + x_2) = x_1^* x_1 + x_1^* x_2 + x_2^* x_1 + x_2^* x_2$$

But $x_1 \perp x_2$, so $x_1^* x_2 = x_2^* x_1 = 0$. Therefore,

$$\|x_1 + x_2\|^2 = x_1^* x_1 + x_2^* x_2 = \|x_1\|^2 + \|x_2\|^2$$

(b) Show that this computation establishes the general case by induction.
Assume the theorem is true up to k terms. Then,

$$\left\| \sum_{i=1}^{k} x_i \right\|^2 = (x_1 + ... + x_k)^*(x_1 + ... + x_k) = \sum_{i=1}^{k} \|x_i\|^2$$

We check if the theorem holds for $k + 1$ terms:

$$\left\| \sum_{i=1}^{k+1} x_i \right\|^2 = (x_1 + ... + x_k + x_{k+1})^*(x_1 + ... + x_k + x_{k+1})$$

By induction and orthogonality,

$$= x_1^* x_1 + ... + x_k^* x_k + x_{k+1}^*(x_1 + ... + x_k + x_{k+1}) = \|x_1\|^2 + ... + \|x_k\|^2 + x_{k+1}^* x_{k+1} = \sum_{i=1}^{k+1} \|x_i\|^2$$

3. Let $A \in \mathbb{C}^{m \times m}$ be hermitian ($A^* = A$)
(a) Prove that all eigenvalue are real
Proof. Let $Ax = \lambda x$. Premultiply by x^* to obtain $x^* A x = \lambda x^* x$. We then take the Hermitian of this equation:

$$(x^* A x = \lambda x^* x)^* \implies x^* A^* x = \lambda x^* x$$

Since A is hermitian, we then have

$$x^* A x = \bar{\lambda} x^* x$$

where $\bar{\lambda}$ is the complex conjugate of λ. We can then substitute $Ax = \lambda x$ into the equation to get

$$\lambda x^* x = \bar{\lambda} x^* x \implies \lambda = \bar{\lambda}$$

Therefore, λ must be real-valued. ■
(b) Prove that if x and y are eigenvectors corresponding to distinct eigenvalues, then $x^* y = 0$.

Proof. If x and y are eigenvectors corresponding to distinct eigenvalues denoted λ, μ respectively, then we have two equations:

$$Ax = \lambda x \text{ and } Ay = \mu y.$$

Premultiply the first equation by y^* to get

$$y^* A x = \lambda y^* x$$

We then take the hermitian of the equation

$$(y^* A x = \lambda y^* x)^* \implies x^* A^* y = \bar{\lambda} x^* y$$

Since A is hermitian, we can rewrite the equation, then substitute:

$$x^* A^* y = x^* A y = \bar{\lambda} x^* y \implies x^* \mu y = \bar{\lambda} x^* y$$

However, we know that all eigenvalues are real-valued ($\bar{\lambda} = \lambda$) and therefore,

$$\mu x^* y = \lambda x^* y \implies (\mu - \lambda) x^* y = 0$$

Since we know that $\mu \neq \lambda$, $x^* y = 0$. ■

4. If u and v are m-vectors, the matrix $A = I + uv^*$ is known as a rank-one perturbation of the identity. Show that if A is nonsingular, then its inverse has the form $A^{-1} = I + \alpha uv^*$ for some scalar α, and give an expression for α. For what u and v is A singular? If it is singular, what is $null(A)$?
We check if $AA^{-1} = I$:

$$I = AA^{-1} = (I + uv^*)(I + \alpha uv^*) = I + \alpha uv^* + uv^* + \alpha uv^* uv^*$$

Since $v^* u$ is a constant, we move it to the front of the last term, subtract I from both sides, and factor uv^*:

$$0 = (\alpha + 1 + \alpha v^* u) uv^*$$

If $uv^* = 0$, $A = I \implies A^{-1} = I$. If $uv^* \neq 0$, we get

$$0 = \alpha + 1 + \alpha v^* u \implies 0 = \alpha(1 + v^* u) + 1 \implies \alpha = \frac{-1}{1 + v^* u}$$

Therefore, with $\alpha = \dfrac{-1}{1 + v^* u}$ we see that $A^{-1} = I + \alpha uv^* = I - \dfrac{1}{1 + v^* u} uv^*$.
If A is singular, we know that A^{-1} is undefined. Mathematically speaking, as we have found an expression for A^{-1} with $\alpha = \dfrac{-1}{1 + v^* u}$, if $v^* u = -1$, α becomes undefined. Since A^{-1} is unique for each u and v, we choose u and v such that $v^* u = -1$. Then A is singular and

$$Au = (I + uv^*) u = u + uv^* u = u - u = 0$$

Therefore, $null(A) = span\{u\}$.

5. Prove that if W is an arbitrary nonsingular matrix, the function $\|\cdot\|_W$ defined by $\|x\|_W = \|Wx\|$ is a vector norm.
Proof. To prove that $\|\cdot\|_W$ is a vector norm, we must show that it holds the three properties:

   (a) $\|x\| \geq 0$, and $\|x\| = 0$ only if $x = 0$

   (b) $\|x + y\| \leq \|x\| + \|y\|$

   (c) $\|\alpha x\| = |\alpha| \, \|x\|$

   (a) Since the matrix-vector product Wx gives a new vector z, we know

$$\|x\|_W = \|Wx\| = \|z\| = (\sum_{i=1}^{m} |z_i|^p)^{1/p} \geq 0$$

   because each $|z_i| \geq 0$. As W is nonsingular, we also know that $Wx = 0 \iff x = 0$, then

$$\|x\|_W = \|Wx\| = \|z\| = 0 \iff z = 0 \iff Wx = 0 \iff x = 0.$$

   (b) Following the same logic as (a), we have

$$\|x + y\|_W = \|W(x + y)\| = \|Wx + Wy\| = \|z + u\|$$

   where $Wx = z$ and $Wy = u$.We then examine two cases: Case (1) - Both z and u have components z_i and u_i that have the same sign for each i. Then,

$$\|z + u\| = \|z\| + \|u\| .$$

   Case (2) - Without loss of generality, choose z to contain at least one negatively valued component in a position where u has a positively valued component. Then,

$$\|z + u\| \leq \|u\| + \|z\| .$$

   (c)

$$\|\alpha x\|_W = \|W(\alpha x)\| = \|\alpha Wx\| = \|\alpha z\| = |\alpha| \, \|z\|$$

   ■

6. Vector and matrix p-norms are related by various inequalities, often involving the dimensions m or n. For each of the following, verify the inequality and give an example of a nonzero vector or matrix (for general m, n) for which equality is achieved. $x \in \mathbb{C}^m, A \in \mathbb{C}^{m \times n}$

   (a) $\|x\|_\infty \leq \|x\|_2$
   We can express the infinite norm in a similar manner to the 2-norm: $\|x\|_\infty = \max x_i = x_j = \sqrt{x_j}^2$. Then,

$$\|x\|_2 = \sqrt{x_1^2 + ... + x_j^2 + ... + x_m^2} \geq \sqrt{x_j^2} = \|x\|_\infty$$

A nonzero vector such that equality is achieved is the vector $x = [\ 0 \quad 0 \quad ... \quad 0 \quad k \quad 0 \quad ... \quad 0\ ]^T$ where k is an arbitrary constant.

$$\|x\|_\infty = k, \|x\|_2 = \sqrt{k^2} = k$$

(b) $\|x\|_2 \leq \sqrt{m}\,\|x\|_\infty$
This inequality holds when we expand the 2-norm then divide by $\sqrt{m}$:

$$\|x\|_2 = \sqrt{x_1^2 + ... + x_m^2}$$

$$\|x\|_\infty = \max x_i \geq \sqrt{\frac{x_1^2 + ... + x_m^2}{m}}$$

In words, the maximum component is larger than or equal to the root mean square (quadratic mean) of the vector. An example vector such that equality holds is the vector with equally sized components: $x = [\ k \quad ... \quad k\ ]^T$.

(c) $\|A\|_\infty \leq \sqrt{n}\,\|A\|_2$
Since the infinite norm of A is the maximum row sum of A, we can express $\|A\|_\infty$ as $\|Ax\|_\infty$ where x is the 1-vector in $\mathbb{C}^n$. If we use the same x in the calculation of the 2 norm of A, we get

$$\|Ax\|_2 = \|A\|_2\,\|x\|_2 = \sqrt{n}\,\|A\|_2$$

By the definition of the induced matrix norm, we can write the infinite norm thus:

$$\frac{\|Ax\|_\infty}{\|x\|_\infty} \leq C \implies \|Ax\|_\infty \leq C\,\|x\|_\infty \leq C\,\|x\|_2$$

We know that the maximum row sum ($\|A\|_\infty$) is going to be less than or equal to the root-square of the row sums ($\|A\|_2$). However, when we divide the $\|A\|_2$ by $\sqrt{m}$, we get the quadratic mean of the row sums. Since the maximum row sum must be greater than or equal to the quadratic mean of the row sums, we know:

$$\|A\|_\infty \leq \sqrt{n}\,\|A\|_2\,.$$

An example matrix such that equality holds is the matrix with one nonzero row with equal entries.

(d) $\|A\|_2 \leq \sqrt{m}\,\|A\|_\infty$
We show that this inequality holds by expanding both the infinite norm and the 2 norm as row sums:

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{k=1}^{n} a_{ik}$$

$$\|A\|_2 = \max_{\|x\|_2=1} \sqrt{\sum_{i=1}^{m} |\sum_{k=1}^{n} a_{ik}x_k|^2} \leq \sqrt{\sum_{i=1}^{m} |\sum_{k=1}^{n} a_{ik}|^2}.$$

If we divide the last term in the line above by $\sqrt{m}$ we see that the quadratic mean of the row sums is greater than or equal to $\|A\|_2$. Since the maximum row sum is always larger than or equal to the quadratic mean of row sums, the inequality holds. An example matrix such that equality holds is the matrix with the same entry in each position.

7. Let $A \in \mathbb{C}^{m\times n}$ with columns a_i and $B \in \mathbb{C}^{p\times n}$ with columns b_i. Show that

$$AB^* = a_1b_1^* + ... + a_nb_n^*$$

in two ways: first using the component-wise definition for the elements of the product of two matrices and secondly using block-matrix multiplication.

Where a_i are the columns of A and b_i^* are the rows of B^*, we can see that the ij-th entry of AB^* comes from the component-wise definition of the product of two matrices:

$$(AB^*)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}^*.$$

We can also see, that for each $a_kb_k^*$, the ij-th element is defined similarly:

$$(a_kb_k^*)_{ij} = a_{ik}b_{kj}^*$$

As we sum each rank one matrix, we can clearly see equality:

$$\sum_{k=1}^{n}(a_kb_k^*)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}^* = (AB^*)_{ij}$$

We can also use block multiplication to come to the same result:
We begin by expanding the right side of the original equation into a block matrix.

$$\sum_{i=1}^{n} a_ib_i^* = \sum \left[\begin{array}{c|c} a_ib_i^* & a_ib_i^* \\ \hline a_ib_i^* & a_ib_i^* \end{array}\right] = \left[\begin{array}{c|c} \sum_{i=1}^{n} a_ib_i^* & \sum_{i=1}^{n} a_ib_i^* \\ \hline \sum_{i=1}^{n} a_ib_i^* & \sum_{i=1}^{n} a_ib_i^* \end{array}\right]$$

Now we expand the product AB^* using block multiplication:

$$AB^* = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \left[\begin{array}{c|c} B_{11}^* & B_{12}^* \\ \hline B_{21}^* & B_{22}^* \end{array}\right] = \left[\begin{array}{c|c} A_{11}B_{11}^* + A_{12}B_{21}^* & A_{11}B_{12}^* + A_{12}B_{22}^* \\ \hline A_{21}B_{11}^* + A_{22}B_{21}^* & A_{21}B_{12}^* + A_{22}B_{22}^* \end{array}\right]$$

If we denote the uv-th block matrix of AB^* as AB_{uv}, we see that for each block matrix in the product, we get a sum:

$$AB_{11}^* = A_{11}B_{11}^* + A_{12}B_{21}^* = \sum_{i=1}^{n_1} a_ib_i^* + \sum_{i=1}^{n_2} a_ib_i = \sum_{i=1}^{n} a_ib_i^* \text{ with } a_i \in \mathbb{C}^{m_1}, b_i^* \in \mathbb{C}^{p_1}$$

$$AB_{12}^* = A_{11}B_{12}^* + A_{12}B_{22}^* = \sum_{i=1}^{n_1} a_i b_i^* + \sum_{i=1}^{n_2} a_i b_i^* = \sum_{i=1}^{n} a_i b_i^* \text{ with } a_i \in \mathbb{C}^{m_1}, b_i^* \in \mathbb{C}^{p_2}$$

$$AB_{21}^* = A_{21}B_{11}^* + A_{22}B_{21}^* = \sum_{i=1}^{n_1} a_i b_i^* + \sum_{i=1}^{n_2} a_i b_i = \sum_{i=1}^{n} a_i b_i^* \text{ with } a_i \in \mathbb{C}^{m_2}, b_i^* \in \mathbb{C}^{p_1}$$

$$AB_{22}^* = A_{21}B_{12}^* + A_{22}B_{22}^* = \sum_{i=1}^{n_1} a_i b_i^* + \sum_{i=1}^{n_2} a_i b_i = \sum_{i=1}^{n} a_i b_i^* \text{ with } a_i \in \mathbb{C}^{m_2}, b_i^* \in \mathbb{C}^{p_2}$$

As the dimensions agree, we can clearly see that equality holds:

$$AB^* = \left[\begin{array}{c|c} A_{11}B_{11}^* + A_{12}B_{21}^* & A_{11}B_{12}^* + A_{12}B_{22}^* \\ \hline A_{21}B_{11}^* + A_{22}B_{21}^* & A_{21}B_{12}^* + A_{22}B_{22}^* \end{array}\right] = \left[\begin{array}{c|c} \sum_{i=1}^{n} a_i b_i^* & \sum_{i=1}^{n} a_i b_i^* \\ \hline \sum_{i=1}^{n} a_i b_i^* & \sum_{i=1}^{n} a_i b_i^* \end{array}\right] = \sum_{i=1}^{n} a_i b_i^*$$