

# Math Problem Set 3

Matthew Brown  
OSM Boot Camp 2018

July 8, 2018

**Problem 4.2.** Recall from the last homework that

$$D = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

This is an upper triangular matrix, so it has eigenvalues along the diagonal, so the only eigenvalue is 0. The eigenspace for this eigenvalue is the set of constant functions, which has dimension 1, so the geometric multiplicity is 1. The algebraic multiplicity, on the other hand, is 3, since 0 appears 3 times along the diagonal.

**Problem 4.4.** I'll do both parts at once.

*Proof.* If  $A = A^H$ , then the diagonal elements of  $A$  must be real (since they don't change when we take their complex conjugate). We know from exercise 4.3 that

$$p(\lambda) = \lambda^2 - \operatorname{tr} A \lambda + \det(A)$$

By the quadratic formula, we can look at the "discriminant"  $(\operatorname{tr} A)^2 - 4 \det(A)$  to see if the roots are real. Now

$$(\operatorname{tr} A)^2 - 4 \det(A) = (A_{11} + A_{22})^2 - 4(A_{11}A_{22} - A_{12}A_{21}) = (A_{11} - A_{22})^2 + 2A_{12}A_{21}$$

If  $A$  is Hermitian,

$$A_{12} = \overline{A_{21}} \implies A_{12}A_{21} \geq 0$$

which implies the discriminant is positive and the characteristic polynomial has real roots.

If  $A$  is Skew-Hermitian,

$$A_{12} = -\overline{A_{21}} \implies A_{12}A_{21} \leq 0$$

which implies the discriminant is negative and the characteristic polynomial has complex roots.  $\square$

**Problem 4.6.** *Proof.* Consider an upper triangular matrix  $A$ , with the elements  $\lambda$  on the diagonal like so:

$$A = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_k \end{bmatrix}$$

(Note of course that the  $\lambda$ s don't have to be unique.) A number  $c$  is an eigenvalue, iff the space  $\text{Ker}(cI - A)$  has some nonzero elements. We see that

$$cI - A = \begin{bmatrix} c - \lambda_1 & & & \\ & c - \lambda_2 & & \\ & & \dots & \\ & & & c - \lambda_k \end{bmatrix}$$

For  $c \neq \lambda_i$ , the matrix is invertible and  $\text{Ker}(cI - A) = \{0\} \implies c$  is not an eigenvalue.

For  $c = \lambda_i$ , some column of  $cI - A$  will have zero on the diagonal. Choose the first such column, call it column  $(cI - A)_r$ . I claim that column  $(cI - A)_r$  must be in the span of all the previous columns  $(cI - A)_1, \dots, (cI - A)_{r-1}$ .

To see this, think of all the columns as  $r - 1$ -dimensional vectors. I can do this because the coordinates in position  $k > r$  are all 0, and thus they "don't matter" (technical term!). Since I'm thinking now of vectors in  $\mathbb{F}^{r-1}$  and I notice that  $(cI - A)_1, \dots, (cI - A)_{r-1}$  are linearly independent, I can say that these vectors form a basis for  $\mathbb{F}^{r-1}$ . Therefore  $(cI - A)_{r-n} \in \mathbb{F}^{r-1}$  is in the span of  $\{(cI - A)_1, \dots, (cI - A)_{r-1}\}$ ... but there is no difficulty in passing to the whole column of  $(cI - A)$  because the other coordinates are 0!

That was long and there's probably a shorter way to do it (indeed I'm pretty sure the above isn't rigorous), but the point is that I've shown that the columns of  $(cI - A)$  are not linearly independent. Thus  $(cI - A)$  does not have full rank, so it's not injective, so its kernel must have nonzero elements, so  $c$  is an eigenvalue.  $\square$

**Problem 4.8.** Three Parts:

- *Proof.*  $S$  spans  $V$  by definition. It remains to show that I can't write any of the elements as a linear combination of any of the others. I feel like should *probably* be invoking that fourier stuff that Jan taught us right about now... but I'll wriggle my way out of that by doing a trick. Let's define for each function a 4-dimensional vector which evaluates the function at the points:  $0, \frac{\pi}{4}, \frac{\pi}{3}$  and  $\frac{\pi}{2}$ . The vectors are:

$$v_{\sin(x)} = \begin{bmatrix} \sin(0) = 0 \\ \sin(\frac{\pi}{4}) = .707 \\ \sin(\frac{\pi}{3}) = .866 \\ \sin(\frac{\pi}{2}) = 1 \end{bmatrix} \quad v_{\cos(x)} = \begin{bmatrix} 1 \\ .707 \\ .5 \\ 0 \end{bmatrix} \quad v_{\sin(2x)} = \begin{bmatrix} 0 \\ 1 \\ .866 \\ 0 \end{bmatrix} \quad v_{\cos(2x)} = \begin{bmatrix} 1 \\ 0 \\ -.5 \\ -1 \end{bmatrix}$$

I want to check to see if these vectors are linearly independent. One check is to see if a matrix with these vectors as columns row-reduces to the identity matrix, which it

does. And if I can't express any of the functions as a linear combination of the others at these 4 points, than I sure as heck know that I can't express any of them as a linear combination of the others generally, so they are linearly independent.  $\square$

- (ii)

$$D = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$$

- (iii) The spaces  $E_1 = \text{span}\{\sin(x), \cos(x)\}$  and  $E_2 = \text{span}\{\sin(2x), \cos(2x)\}$  are complementary and  $D$ -invariant.

**Problem 4.13.** First step is to find the basis of eigenvectors. Eigenvalues are 1, .4 with eigenvectors  $(2, 1)$  and  $(1, -1)$  So I want to change to this basis, and I'll use

$$P = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$$

as my transition matrix.

$$P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & .4 \end{bmatrix}$$

**Problem 4.15.** *Proof.* Let  $\mathcal{B} = \{b_1, \dots, b_n\}$  be an eigenbasis for  $A$ , with each  $b_i$  corresponding to an eigenvalue  $\lambda_i$ . Consider the action of  $f(A)$  on the vector  $f(b_i)$ . First I show that  $Af(b_i) = \lambda_i f(b_i)$

Proof of the above:

$$\begin{aligned} Af(b_i) &= A(a_0 + a_1b_i + a_2b_i^2 + \dots + a_nb_i^n) = a_0A + a_1Ab_i + a_2Ab_i^2 + \dots + a_nAb_i^n \\ &= a_0A + a_1\lambda_ib_i + a_2\lambda_i^2b_i + \dots + a_n\lambda_i^n b_i = f(\lambda_i)b_i \end{aligned}$$

After we've shown this, we can see that

$$\begin{aligned} f(A)f(b_i) &= (a_0 + a_1A + a_2A^2 + \dots + a_nA^n)f(b_i) \\ &= a_0f(b_i) + a_1\lambda_if(b_i) + a_2\lambda_i^2f(b_i) + \dots + \lambda_i^n f(b_i) \\ &= f(\lambda_i)f(b_i) \end{aligned}$$

which concludes the proof.  $\square$

**Problem 4.16.** Three Parts

- (i) Let  $D = P^{-1}AP$  as in problem 4.13. Then by proposition 4.3.10 for any  $k \in \mathbb{N}$ ,

$$D = P^{-1}AP \implies A = PDP^{-1} \implies A^k = PD^kP^{-1}$$

And I can apply a limit as  $k \rightarrow \infty$  to both sides. It's clear that  $D^k \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ , so

$$A^k \rightarrow P \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} P^{-1} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

- (ii) The answer for the  $\infty$ -norm will not change. Recall that we showed in week 2 that the  $\infty$ -norm and the 1-norm were *topologically equivalent*. Two norms that are topologically equivalent have the same convergent sequences. Indeed, the answer does not change when we use the Frobenius norm either, since by theorem 5.8.7, all norms on a finite-dimensional vector space are topologically equivalent, and we apply this fact to  $M_{2 \times 2}(\mathbb{F})$  to see that convergence is the same over all norms.
- (iii) For a polynomial  $p$ , The eigenvalues of  $p(A)$  are just  $p(\lambda_i)$  where  $\lambda - I$  are the eigenvalues of  $A$ .

$$p(1) = 3 + 5(1) + 1^3 = 8, \quad p(.4) = 3 + 5(.4) + .4^3 = 5.096$$

so our eigenvalues are 8 and 5.96.

**Problem 4.18.** *Proof.* First, look at remark 4.3.15 in the book to see that

$$x^T A = \lambda x^T \iff A^T x = \lambda x$$

In other words, we're looking to see if  $\lambda$  is an eigenvalue of  $A^T$ . But we know that

$$0 = \det(A - \lambda I) = \det(A^T - \lambda I) = 0$$

so this is indeed the case. □

**Problem 4.20.** *Proof.* Suppose  $A$  is Hermetian and there exist orthonormal matrices  $Q, Q^H$  such that  $Q A Q^H = B$ . Then

$$B^H = (Q A Q^H)^H = Q A Q^H = B$$

□

**Problem 4.24.** First of all, I am not concerned about the denominator since the norm will always be real. So examine  $\langle x, Ax \rangle$ . Recall that the adjoint with the normal inner product on  $\mathbb{F}$  is the hermitian. Then for  $A$  hermitian,

$$\langle x, Ax \rangle = \langle A^H x, x \rangle = \langle Ax, x \rangle = \overline{\langle x, Ax \rangle}$$

Since  $\langle x, Ax \rangle$  equals its own conjugate, this must be a real number. On the other hand, if  $A$  is skew-hermitian,

$$\langle x, Ax \rangle = \langle A^H x, x \rangle = \langle -Ax, x \rangle = \overline{\langle x, -Ax \rangle} = -\overline{\langle x, Ax \rangle}$$

and now the inner product must be imaginary.

**Problem 4.25.** Two parts

- (i)

*Proof.*  $A$  is a normal matrix, so we know that its orthonormal eigenvectors  $\{x_1, \dots, x_n\}$  form a basis for  $\mathbb{C}^n$ . If I show that a transformation fixes each of these basis vectors I have shown that it is the identity. And by orthonormality, for any basis vector  $x_j$ ,

$$(x_1 x_1^T + \dots + x_n x_n^T) x_j = 0 + \dots + x_j 1 + \dots + 0 = x_j$$

□

- (ii)

*Proof.* Very similar to the above: If I show that two transformations are equal on a basis, then they're equal everywhere. Let  $x_j$  be an arbitrary basis vector, and see that

$$*\lambda_1 x_1 x_1^T + \dots + x_n x_n^T) x_j = 0 + \dots + \lambda_j x_j 1 + \dots + 0 = \lambda_j x_j = A x_j$$

□

**Problem 4.27.** *Proof.* Let  $A \in M_n(\mathbb{F})$  be a positive definite matrix. I'll argue by contradiction. Suppose that  $A_{ii}$  is a diagonal entry in  $A$  that is not a positive real number. Then

$$A e_i = [a_{1i} \dots a_{ni}]$$

and

$$\langle e_i, A e_i \rangle = A_{ii} \notin \mathbb{R}^+$$

so we see that  $A$  is not positive semi definite and we have our contradiction. □

**Problem 4.28.** *Proof.* First I'll show the facts about trace.

Recall that the trace is the sum of the eigenvalues, so  $AB$  positive semidefinite  $\implies \text{tr}(AB) \geq 0$ . I must show therefore that  $AB$  is positive semidefinite. (This part of the proof is from Jaehyung!)

By the definition of a positive definite matrix  $\langle x, Ax \rangle, \langle x, Bx \rangle \geq 0$ , so their product is also  $\geq 0$ . See that

$$\begin{aligned} \langle x, Ax \rangle \langle x, Bx \rangle &= (x^T A x)(x^T B x) \\ &= (x x^T)(x^T A B x) \\ &= (x x^T) \langle x, A B x \rangle \end{aligned}$$

But for all nonzero  $x$ , we have  $x x^T$  is a positive scalar, and

$$\langle x, Ax \rangle \langle x, Bx \rangle \geq 0 \implies \langle x, A B x \rangle \geq 0$$

Which means that  $AB$  is positive semidefinite.

For the second inequality, recall that any positive semidefinite matrix is orthonormally similar to a diagonal matrix. Let  $A_D, B_D$  be the diagonal representations of  $A, B$ . Trace is invariant

over change of basis, so  $\text{tr}(AB) = \text{tr}(A_D B_D)$ . Let  $\lambda_1, \dots, \lambda_m$  be the positive eigenvalues of  $A$ , and  $\phi_1, \dots, \phi_n$  be the positive eigenvalues of  $B$ . Let  $k = \min\{m, n\}$ . Then

$$(A_D B_D) = \text{diag}(\lambda_1 \phi_1, \dots, \lambda_k \phi_k, 0, \dots, 0) \implies \text{tr}(A_D B_D) = \sum_{i=1}^k \lambda_i \phi_i$$

On the other hand,

$$\begin{aligned} \text{tr}(A) \text{tr}(B) &= (\lambda_1 + \dots + \lambda_m)(\phi_1 + \dots + \phi_n) \\ &= \sum_{i=1}^m \sum_{j=1}^n \lambda_i \phi_j \\ &\geq \sum_{i=1}^k \lambda_i \phi_i = \text{tr}(AB) \end{aligned}$$

Which is the second inequality.

Finally,

$$\|AB\|_F^2 = \text{tr}(AA^H BB^H) \leq \text{tr}(AA^H) \text{tr}(BB^H) = \|A\|_F \|B\|_F^2$$

Which is the property that makes  $\|\cdot\|_F$  a matrix norm. □

**Problem 4.31.** 4 parts

- (i)

*Proof.* Recall the definition that

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

Now in the SVD we decompose  $A$  into  $U\Sigma V^H$  with  $U$  and  $V^H$  orthonormal. Now, because  $U$  and  $V^H$  are orthonormal, they preserve norms, and the only way that  $A$  can change the norm of a vector is through the diagonal matrix  $\Sigma$ . I appeal to intuition when I say that the most "effective" way that  $\Sigma$  can do this is if it acts on the eigenvector of its highest eigenvalue  $\sigma_1$ . (The spectral norm of a diagonal matrix is its highest eigenvalue). And we see that

$$\|A\|_2 = \|\Sigma\|_2 = \sigma_1$$

□

- (ii)

*Proof.* Let  $U\Sigma V^H$  be the SVD of  $A$ . Then

$$A^{-1} = V\Sigma^{-1}U^H$$

Since  $\sigma_n$  is the smallest eigenvalue of  $\Sigma$ ,  $\sigma_n^{-1}$  is the largest eigenvalue of  $\Sigma^{-1}$  and it follows that it is the spectral norm by part (i). □

- (iii)

*Proof.* It's become clear by now that the spectral norm of  $A$  depends only on its singular values, and because singular values are positive real numbers,

$$\Sigma^H = \Sigma^T = \Sigma \implies \|A^H\|_2^2 = \|A^T\|_2^2 = \|A\|_2^2$$

For the last piece of the puzzle, see that

$$A^H A = (V \Sigma U)^H (V \Sigma U) = (U^H \Sigma^H V^H)(V \Sigma U) = U^H \Sigma^2 U$$

which is a singular value decomposition of  $A$ , and the largest eigenvalue of  $\Sigma^2$  is  $\sigma_1^2$  so we see that  $\|A^H A\|_2 = \|A\|_2^2$   $\square$

- (iv)

*Proof.* If  $A$  has an SVD  $U_1 \Sigma V_1^H = A$ , then

$$U A V = (U U_1) \Sigma (V_1^H V^H)$$

which is also a singular decomposition for  $U A V$ , so these two transformations have the same singular values and thus the same norms.  $\square$

**Problem 4.32. NOTE:** I'm doing (ii) first and then (i)

(ii)

*Proof.* If  $U \Sigma V^H$  is an SVD for  $A$ ,

$$\|A\|_F = \sqrt{\text{tr}(A^H A)} = \sqrt{\text{tr}((U^H \Sigma^H V^H)(V \Sigma U))} = \sqrt{\text{tr}(\Sigma^H \Sigma)} = \sqrt{\sum_{i=1}^n \sigma_i^2}$$

$\square$

(i)

*Proof.* We saw in 4.31 part (iv) that  $U A V$  and  $A$  have the same singular values. And since by (ii) the norm depends entirely on the singular values,  $U A V$  and  $A$  have the same norms.  $\square$

**Problem 4.33. Proof.** Note that  $y^H A x$  will be a field element. If we think of this as a linear map,  $y^H A x : \mathbb{F} \rightarrow \mathbb{F}$ , then we see that the spectral norm of this map is:

$$\|y^H A x\|_2 = \sup_{f \in \mathbb{F}} \frac{\|(y^H A x) f\|_2}{\|f\|_2} = |y^H A x|$$

(Where the first norm is the spectral norm and the norm in the fraction is the standard 2-norm.)  $\square$

**Problem 4.36.** Well, singular values are always strictly positive, so

$$-I = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$$

works since both its eigenvalues are negative (There are many examples).

**Problem 4.38.** Lots of parts. Recall the definition that for  $U\Sigma V^H$  an SVD of  $A$ ,

$$A^\dagger = V\Sigma^{-1}U^H$$

- (i)

*Proof.*

$$AA^\dagger A = (U\Sigma V^H)(V\Sigma^{-1}U^H)(U\Sigma V^H) = U\Sigma V^H = A$$

□

- (ii)

*Proof.*

$$A^\dagger AA^\dagger = (V\Sigma^{-1}U^H)(U\Sigma V^H)(V\Sigma^{-1}U^H) = (V\Sigma^{-1}U^H) = A^\dagger$$

□

- (iii)

*Proof.*

$$(AA^\dagger)^H = ((U\Sigma V^H)(V\Sigma^{-1}U^H))^H = U\Sigma^{-1}V^H V\Sigma U^H = UU^H = AA^\dagger$$

□

- (iv)

*Proof.*

$$(A^\dagger A)^H = ((V\Sigma^{-1}U^H)(U\Sigma V^H))^H = V\Sigma U^H U\Sigma^{-1}V^H = VV^H = A^\dagger A$$

□

- (v)

*Proof.* To show that a matrix is an orthogonal projection, I must show that it is hermitian ( $M^H = M$ ) and idempotent ( $M^2 = M$ ). By (iii)  $AA^\dagger$  is hermitian. By (i)  $AA^\dagger AA^\dagger = AA^\dagger \implies$  idempotence.

It remains to show  $\mathcal{R}(AA^\dagger) = \mathcal{R}(A)$ .

Trivially  $\mathcal{R}(AA^\dagger) \subset \mathcal{R}(A)$ . And (i)  $\implies \mathcal{R}(A) \subset \mathcal{R}(AA^\dagger)$

□



- (vi)

*Proof.* By (iv)  $A^\dagger A$  is hermitian. By (ii)  $A^\dagger A A^\dagger A = A^\dagger A \implies$  idempotence. It remains to show  $\mathcal{R}(A^\dagger A) = \mathcal{R}(A^H)$ . By (iv),

$$A A^\dagger = (A A^\dagger)^H = A^H (A^\dagger)^H \implies \mathcal{R}(A^\dagger A) \subset \mathcal{R}(A^H)$$

If we take the hermitian of both sides of (i), we see that

$$(A^\dagger A)^H A^H = (A^\dagger A)^H A^H A^H \implies \mathcal{R}(A^H) \subset \mathcal{R}(A^\dagger A)$$

(credit again to Jaehyung for these last two parts!)

□