Math Homework #3

Eric C. Miller

July 10, 2017

Question #4.2

Let $V = span(\{1, x, x^2\})$ be a subspace of the inner product space $L^2([0,1], \mathbb{R})$. Let D be the derivative operator $D : V \to V$ given by $D[p(x)] = p'(x)$.

Question: Find the eigenvalues and eigenspaces of D, along with their geometric and algebraic multiplicities.
We have that:

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

We proceed to calculate the characterizing polynomial:

$$det(I\lambda - D) = \left(\lambda - \frac{2}{x}\right)\left(\lambda - \frac{1}{x}\right)\lambda = (\lambda x - 2)(\lambda x - 1) \tag{1}$$

Hence, we have that $\lambda = \frac{1}{x}, \frac{2}{x}$, with corresponding eigenvectors $v_1 = 2x^+ x,\ v_2 = x^2 + \frac{x}{2}$. Their eigenspaces are the corresponding span sets, and the algebraic and geometric multiplicities are both 1.

Question #4.4

Recall that a matrix $A \in M_n(\mathbb{F})$ is Hermitian if $A^H = A$ and skew-Hermitian if $A^H = -A$.

Claim: A Hermitian 2x2 matrix has only real eigenvalues. A skew-Hermitian matrix has only complex eigenvalues.
Proof: Let our matricies be:

$$A = A^H = \begin{bmatrix} a_1 & a_2 \\ \overline{a_2} & a_3 \end{bmatrix}, \text{ and } B = -B^H = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$$

From this, we see that $det(I\lambda - A) = \lambda(\lambda - a_1 - a_3) + det(A)$. As A is a Hermitian matrix, $det(A), a_1, a_3 \in \mathbb{R}$. Hence, we have that the eigenvalues of A must also be real.

For the skew-Hermitian case, observe that $det(I\lambda - B) = \lambda(\lambda - b_1 - b_4) + det(B)$, where $b_1, b_4 \in \mathbb{I}$ and thus $det(B) \in \mathbb{I}$. Hence, we have that all eigenvalues must be purely imaginary.

Question #4.6

Claim: The diagonal entries of an upper-triangular or lower-triangular matrix are its eigenvalues.
Proof: We can show the full statement by showing the property for just the upper-triangular case without loss of generality. We seek to show by induction that this property holds. For square matrix $n = 1$, this is trivially true. Let A_n be an upper-triangular matrix with diagonals equal to its eigenvectors. We then see that $det(I_{n+1}\lambda - A_{n+1}) = (\lambda - a_{1,1})det(I_n\lambda - A_n)$. Hence, we have that this property holds for all upper triangular matricies.

Question #8

Let $V = span\{S\},\ S = \{sin(x), cos(x), sin(2x), cos(2x)\} \in C^\infty(\mathbb{R}, \mathbb{R})$.

Claim: S is a basis for V.
Proof: We seek to show that S is linearly independent, and $V = span\{S\}$. With the second true by definition, linear independence is shown by noticing that no elements of S are scalar multiples of one another. Hence, we have that S is a basis for V. We can write a matrix D (where D is the derivative operator identified in 4.2) such that it is in the basis S:

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

In addition, we can choose $W_1 = span\{sin(x), cos(x)\},\ W_2 = span\{sin(2x), cos(2x)\}$ and show that they are D-invariant:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ -0 & 0 & -2 & 0 \end{bmatrix} \begin{bmatrix} sin(x) \\ cos(x) \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} cos(x) \\ -sin(x) \\ 0 \\ 0 \end{bmatrix} \in W_1$$

and

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ -0 & 0 & -2 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ sin(2x) \\ cos(2x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ cos(2x) \\ -sin(2x) \end{bmatrix} \in W_2$$

Question #13

Let

$$A = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$$

We compute the transition matrix P such that $P^{-1}AP$ is diagonal. We see that $\sigma(A) = \{1, 0.4\}$, as $det(I\lambda - A) = (\lambda - 0.8)(\lambda - 0.6) - 0.1 = \lambda^2 - 1.4\lambda + 0.4 = (\lambda - 1)(\lambda - 0.4)$. The corresponding eigenvectors are then $v_1 = [2, 1]^T$ and $v_2 = [1, -1]^T$ respectively (computationally achieved). We then have:

$$P = [v_1^T, v_2^T] = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$$

Question #15

Claim: If $\{\lambda_i\}_{i=1}^n$ are the eigenvalues of a semisimple matrix $A \in M_n(\mathbb{F})$ and $f(x) = a_0 + a_1 x + ... + a_n x^n$ is a polynomial, then $\{f(\lambda_i)\}_{i=1}^n$ are the eigenvalues of $f(A) = a_0 I + a_1 A + ... + a_n A^n$.
Proof: We have that A is semisimple $\iff$ it is diagonalizable. Hence, let $P \in M_n(\mathbb{R})$ such that $D = P^{-1}AP$, where D is a diagonal matrix similar to A, with the eigenvalues of A along the diagonal of D. To find the eigenvalues for the matrix A, we must simply conduct this operation. Now:

$$\begin{aligned} f(D) &= a_1 I + a_2 D + a_3 D^2 + ... + a_n D^n \\ &= a_1 I + a_2 P^{-1}AP + a_3 P^{-1}A^2 P + ... + a_n P^{-1}A^n P \\ &= P^{-1}[a_1 I + a_2 A + a_3 A^2 + ... + a_n A^n]P \\ &= P^{-1} f(A) P \end{aligned} \tag{2}$$

Hence, we see that $P(D)$, which are the eigenvalues $\{f(\lambda_i)\}_{i=1}^n$, are the eigenvalues for $f(A)$.

Question #16

Let

$$A = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}, \quad P = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 0.4 \end{bmatrix}$$

by the previous calculations. As at least one eigenvalue is as least one, we have that $lim_{k\to 0} A^k \neq 0$. Also, we have $A^k = PD^kP^{-1}$. We have that:

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

Hence, we compute the limit for the 1-norm by letting $B = PD^\infty P^{-1}$, where

$$B = PD^\infty P^{-1} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

This process can be repeated for any matrix norm, as B is the true limit, and thus we will always have the limit property.

We then can compute the eigenvalues of $3I+5A+A^3$ using our result from #15: namely that the eigenvalues of $3I + 5A + A^3 = 3 + 5\lambda_i + \lambda_i^3, \ \forall i = 9, 5.064$.

Question #18

Claim: if λ is an eigenvalue of the matrix $A \in M_n(\mathbb{F})$, then there exists a nonzero row vector x^T such that $x^T A = \lambda x^T$.
Proof: We have by hypothesis $Ay = \lambda y$. We have further that $Ay = \lambda y \ \Rightarrow \ (Ay)^T = \lambda y^T \ \Rightarrow \ y^T A^T = \lambda y^T$. However, the eigenspace of A and A^T have the same dimension; hence, we have that there exists a nonzero row vector x^T such that $x^T A = \lambda x^T$.

Question #20

Claim: If A is Hermitian and orthonormally similar to B, then B is also Hermitian.
Proof: We have A is Hermitian, and $A = U^H BU$, for some orthonormal matrix U. However, as the product of two orthonormal matricies is also orthonormal, we extend this notion to show that B is also orthonormal.

Question #24

Given $A \in M_n(\mathbb{C})$, define the *Rayleigh quotient* as:

$$\rho(x) = \frac{\langle x, Ax\rangle}{\|x\|^2} \tag{3}$$

Claim: The Rayleigh quotient can only take on real values for Hermitian matricies and only imaginary values for skew-Hermitian matricies.
Proof: For Hermitian matricies:

$$\frac{\langle x, Ax\rangle}{\|x\|^2} = \frac{\langle A^H x, x\rangle}{\|x\|^2} = \frac{\langle Ax, x\rangle}{\|x\|^2} = \frac{\overline{\langle x, Ax\rangle}}{\|x\|^2} \tag{4}$$

Notice that as $\|x\|^2$ is also real-valued, the Rayleigh quotient can only take on real values for a Hermitian matrix.

For skew-Hermitian matricies, we compute the inner product the same way: $\langle x, Ax\rangle = x^H Ax$. Here, we see that as the real components of all complex non-diagonal values cancel (from the fact that the real components of non-diagonal elements being negative reflections of each other in a skew-Hermitian matrix across the diagonal). Hence, the elements remaining are the imaginary diagonal elements and remaining imaginary components of the complex elements. Hence, we have that the entire numerator of the quotient is imaginary, and thus the quotient itself.

Question #27

Assume $A \in M_n(\mathbb{F})$ is positive definite.

All of A's diagonal entries are real and positive.
Proof: A is positive definite, and is hence Hermitian; thus, we have that its values are real. We then have that for the eigenvalues $\{\lambda_i\}_{i=1}^n$ of A, the inner product defining positive definite properties $\langle x, Ax\rangle = \langle x, \lambda_i x\rangle = \lambda\langle x, x\rangle > 0$ where $x \neq 0$. From this, we see that as A is diagonalizable, its diagonal decomposition D has all positive diagonal entries, and hence, A has all positive diagonal entries.

Question #28

Assume $A, B \in M_n(\mathbb{F})$ are positive semi definite.

Claim: $0 \leq tr(AB) \leq tr(A)tr(B)$**.**
Proof: We begin by noting that the trace of a matrix is the sum of its diagonal elements. As a positive semi-definite matrix has all non-negative diagonal components, its trace is positive unless all diagonal components are 0. Thus, we have $0 \leq tr(A)tr(B)$. For the remaining inequality:

$$\begin{aligned} &0 \leq \|AB\| \leq \|A\|\|B\| \\ &0 \leq \sqrt{\|AB\|} \leq \sqrt{\|A\|}\sqrt{\|B\|} \\ &0 \leq tr(AB) \leq tr(A)tr(B) \end{aligned} \tag{5}$$

by the standard Frobenius norm.

Question #31

Assume $A \in M_{mxn}(\mathbb{F})$ and A is not identically zero.

Claim: $\|A\|_2 = \sigma_1$, **where** σ_1 **is the largest singular value of A.**
Proof:

$$\|A\|_2 = sup_{x\neq 0}\frac{\|Ax\|}{\|x\|} = sup_{x\neq 0}\frac{\|U\Sigma V^H x\|}{\|x\|} = sup_{x\neq 0}\frac{\|\Sigma V^H x\|}{\|x\|} = sup_{x\neq 0}\frac{\|\Sigma V^H x\|}{\|Vx\|} = \sigma_{max} \tag{6}$$

Claim: If A is invertible, then $\|A^{-1}\|_2 = \sigma_n^{-1}$.
Proof:

$$\begin{aligned} \|A^{-1}\|_2 = sup_{x\neq 0}\frac{\|A^{-1}x\|}{\|x\|} = sup_{x\neq 0}\frac{\|(U\Sigma V^H)^{-1}x\|}{\|x\|} &= sup_{x\neq 0}\frac{\|V\Sigma^{-1}U^H x\|}{\|x\|} \\ &= sup_{x\neq 0}\frac{\|\Sigma^{-1}U^H x\|}{\|Ux\|} \\ &= sup_{x\neq 0}\frac{\|\Sigma^{-1}x\|}{\|x\|} \\ &= \sigma_n^{-1} \end{aligned} \tag{7}$$

Claim: $\|A^H\|_2^2 = \|A^T\|_2^2 = \|A^HA\|_2^2 = \|A\|_2^2$
Proof:

$$\begin{aligned}
\|A\|_2 &= sup_{x\neq 0}\frac{\|U\Sigma V^Hx\|}{\|x\|} = \sigma_{max} = sup_{x\neq 0}\frac{\|V\Sigma^HU^Hx\|}{\|x\|} = \|A^H\|_2 \\
\|A^H\|_2 &= sup_{x\neq 0}\frac{\|V\Sigma^HU^Hx\|}{\|x\|} = \sigma_{max} = sup_{x\neq 0}\frac{\|V\Sigma^TU^Hx\|}{\|x\|} = \|A^T\|_2 \\
\|A^HA\|_2 &= sup_{x\neq 0}\frac{\|A^HAx\|}{\|x\|} = sup_{x\neq 0}\frac{\|V\Sigma^H\Sigma V^Hx\|}{\|x\|} = \sigma_{max}^2 = \|A\|_2^2
\end{aligned} \tag{8}$$

Claim: If $U \in M_n(\mathbb{F})$ and $V \in M_n(\mathbb{F})$ are orthonormal, then $\|UAV\|_2 = \|A\|_2$
Proof:

$$\|UAV\|_2 = sup_{x\neq 0}\frac{\|UAVx\|}{\|x\|} = sup_{x\neq 0}\frac{\|UU^H\Sigma V^HVx\|}{\|x\|} = sup_{x\neq 0}\frac{\|\Sigma x\|}{\|x\|} = \sigma_{max} = \|A\|_2 \tag{9}$$

Question #32

Assume $U \in M_{mxn}(\mathbb{F})$ is of rank r.

Claim: If $U \in M_m(\mathbb{F})$ and $V \in M_n(\mathbb{F})$ are orthonormal, then $\|UAV\|_F = \|A\|_F$.
Proof:

$$\begin{aligned}
\|UAV\|_F &= \sqrt{tr((UAV)^HUAV)} = \sqrt{tr(V^HA^HU^HUAV)} = \sqrt{tr(V^HA^HAV)} \\
&= \sqrt{tr(A^HAVV^H)} = \sqrt{tr(A^HA)} = \|A\|_F
\end{aligned} \tag{10}$$

Claim: $\|A\|_F = (\sigma_1^2 + \sigma_2^2 + ... + \sigma_r^2)^{\frac{1}{2}}$**.**
Proof:

$$\begin{aligned}
\|A\|_F &= \sqrt{tr(A^HA)} = \sqrt{tr((U\Sigma V^H)^HU\Sigma V^H)} = \sqrt{tr(V\Sigma U^HU\Sigma V^H)} \\
&= \sqrt{tr(V\Sigma\Sigma V^H)} = \sqrt{tr(\Sigma\Sigma V^HV)} \\
&= \sqrt{tr(\Sigma^2)} \\
&= \sqrt{\sigma_1^2 + \sigma_2^2 + ... + \sigma_r^2}
\end{aligned} \tag{11}$$

Question #33

Assume $A \in M_n(\mathbb{F})$.

Claim: $\|A\|_2 = sup\{|y^HAx| : \|x\|_2 = \|y\|_2 = 1\}$
Proof:

$$\|A\|_2 = sup\frac{\|Ax\|}{\|x\|} = sup\frac{\|y^TAx\|_2}{\|x\|\|y\|} = sup_{\|x\|=\|y\|=1}\|y^TAx\| = sup|y^TAx| \tag{12}$$

Question #36

Give an example of a 2x2 matrix whose determinant is nonzero and whos singular values are not equal to any of its eigenvalues.

Solution: We seek to find a matrix that is non-singular, and whose eigenvalues are irrational. The following matrix satisfies this property:

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$$

Question #38

Let $A \in M_{mxn}(\mathbb{F})$ and $A^{\dagger} =$ the Moore-Penrose pseudoinverse of A.

Claim: $AA^{\dagger}A = A$
Proof:

$$AA^{\dagger}A = AV\Sigma^{-1}U^{H}U\Sigma V^{H} = AV\Sigma^{-1}\Sigma V^{H} = AVV^{H} = A \tag{13}$$

Claim: $A^{\dagger}AA^{\dagger} = A^{\dagger}$
Proof:

$$A^{\dagger}AA^{\dagger} = A^{\dagger}U\Sigma V^{H}V\Sigma^{-1}U^{H} = A^{\dagger}U\Sigma\Sigma^{-1}U^{H} = A^{\dagger}UU^{H} = A^{\dagger} \tag{14}$$

Claim: $(AA^{\dagger})^{H} = A^{\dagger}$
Proof:

$$(AA^{\dagger})^{H} = (U\Sigma V^{H}V\Sigma^{-1}U^{H})^{H} = I^{H} = I = U\Sigma V^{H}V\Sigma^{-1}U^{H} = AA^{\dagger} \tag{15}$$

Claim: $(A^{\dagger}A)^{H} = A^{\dagger}A$
Proof:

$$(A^{\dagger}A)^{H} = (V\Sigma^{-1}U^{H}U\Sigma V^{H})^{H} = I^{H} = I = V\Sigma^{-1}U^{H}U\Sigma V^{H} = A^{\dagger}A \tag{16}$$

Claim: $AA^{\dagger} = proj_{\mathscr{R}(A)}$ **is the orthogonal projection onto** $\mathscr{R}(A)$**.**
Proof: We have that $(AA^{\dagger})^{2} = AA^{\dagger}AA^{\dagger} = AA^{\dagger}$ and that $(AA^{\dagger})^{H} = A^{\dagger}A$. Thus, $AA^{\dagger}$ is an orthogonal projection. It falls onto $\mathscr{R}(A)$ because $AA^{\dagger}$ is the least-squares operator, which maps to the range of A.

Claim: $A^{\dagger}A = proj_{\mathscr{R}(A^{H})}$ **is the orthogonal projection onto** $\mathscr{R}(A^{H})$**.**
Proof: We have that $(A^{\dagger}A)^{2} = A^{\dagger}AA^{\dagger}A = A^{\dagger}A$ and that $(A^{\dagger}A)^{H} = A^{\dagger}A$. Thus, $A^{\dagger}A$ is an orthogonal projection. It falls onto $\mathscr{R}(A^{T})$ because $A^{\dagger}Ab = A^{\dagger}(\hat{x}) \Rightarrow Ab = \hat{x}$.