Problem Set #3

Reiko Laski

Exercise 2
Let $V = span(1, x, x^2)$ be a subspace of the inner product space $L^2([0,1];\mathbb{R})$ given by $D[p](x) = p'(x)$. Find the eigenvalues and eigenspaces of D and their algebraic and geometric multiplicities.

We start by evaluating D on the basis vectors: $D[1](x) = 0, D[x](x) = 1, D[x^2](x) = 2x$. The matrix representation of D with respect to this basis is

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

The eigenvalues of a triangular matrix are the elements on the diagonal $\implies \lambda = 0$ is the only eigenvalue.
To find the eigenspace spaces of λ we solve $\mathscr{N}(D - \lambda I)$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \implies \mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

The geometric multiplicity of λ is 1, while the algebraic multiplicity is 3.

Exercise 4
Recall that a matrix $A \in M_n(\mathbb{F})$ is Hermitian if $A^H = A$ and skew-Hermitian if $A^H = -A$. From Exercise 4.3, we know that the characteristic polynomial of any 2×2 matrix has the form

$$p(\lambda) = \lambda^2 - tr(A)\lambda + det(A).$$

Prove that
(i) a Hermitian 2×2 matrix has only real eigenvalues.
Proof:
The eigenvalues of the matrix are

$$\lambda = \frac{tr(A) \pm \sqrt{(tr(A))^2 - 4det(A)}}{2}.$$

To see if the matrix has only real eigenvalues, we need to check that

$$(tr(A))^2 - 4det(A) \geq 0 \tag{1}$$

Note that the diagonal elements of a Hermitian matrix must be real, and the off-diagonal elements must be conjugates. Consider the matrix A with $a, c \in \mathbb{R}, \beta \in \mathbb{C}$:

$$A = \begin{bmatrix} a & \beta \\ \bar{\beta} & c \end{bmatrix} \implies tr(A) = a + c \text{ and } det(A) = ac - \beta\bar{\beta} = ac - \|\beta\|^2$$

Substituting into the inequality in (1), we have that

$$\begin{aligned}(a+c)^2 - 4(ac - \|\beta\|^2) &= a^2 + c^2 + 2ac - 4ac + 4\|\beta\|^2 \\ &= a^2 + c^2 - 2ac + 4\|\beta\|^2 \\ &= (a-c)^2 + 4\|\beta\|^2 \geq 0.\end{aligned}$$

Therefore, the eigenvalues of the matrix must be real.

(ii) a skew-Hermitian 2×2 matrix has only imaginary eigenvalues.
Proof:
Similarly, to see if the matrix has only imaginary eigenvalues, we need to check that

$$(tr(A))^2 - 4det(A) < 0 \tag{2}$$

. Consider the matrix B with $a, c \in \mathbb{R}, \beta \in \mathbb{R}$

$$B = \begin{bmatrix} ai & \beta \\ -\bar{\beta} & ci \end{bmatrix} \implies tr(B) = ai + ci = (a+c)i \text{ and } det(B) = -ac + \beta\bar{\beta} = -ac + \|\beta\|^2$$

Substituting into the inequality in (2), we have that

$$\begin{aligned}((a+c)i)^2 - 4(-ac + \|\beta\|^2) &= -(a+c)^2 + 4ac - 4\|\beta\|^2 \\ &= -a^2 - c^2 - 2ac + 4ac - 4\|\beta\|^2 \\ &= -a^2 - c^2 + 2ac - 4\|\beta\|^2 \\ &= -(a-c)^2 - 4\|\beta\|^2 < 0.\end{aligned}$$

Therefore, the eigenvalues of the matrix must be imaginary.

Exercise 6
The diagonal entries of an upper-triangular (or lower-triangular) matrix are its eigenvalues.
Proof:
Let A be upper-triangular,

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{bmatrix}$$

Then the eigenvalues of A are given by

$$det(A - \lambda I) = \begin{vmatrix} a_{1,1} - \lambda & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} - \lambda & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} - \lambda \end{vmatrix} = (a_{1,1} - \lambda)(a_{2,2} - \lambda) \cdots (a_{n,n} - \lambda) = 0$$

$\implies \lambda_1 = a_{1,1}, \lambda_2 = a_{2,2}, ..., \lambda_n = a_{n,n}$
We can see that these are exactly the diagonal elements of A.

Exercise 8
Let V be the span of the set $S = \{\sin(x), \cos(x), \sin(2x), \cos(2x)\}$ in the vector space $C^\infty(\mathbb{R};\mathbb{R})$.
(i) S is a basis for V.
Proof:
By definition, we know that S spans the space V. To show that S is a basis for V, we only need to show that the elements of S are linearly independent. If this is true, then we know that there do not exist nonzero constants $a, b, c, d \in \mathbb{R}$ s.t.

$$a\sin(x) + b\cos(x) + c\sin(2x) + d\cos(2x) = 0.$$

Evaluating at $x = 0$ and $x = \pi$, we get

$$\begin{aligned} & b + d = 0 \\ & -b + d = 0 \\ & \implies b = d = 0. \end{aligned}$$

Similarly, if we evaulate at $x = \frac{\pi}{2}$, we get $a = 0$. Since we have already showed that $a, b, d = 0$, we are left with $c\sin(2x) = 0$. Clearly we must have that $c = 0$ for this equation to hold $\forall x$.

(ii) Let D be the derivative operator. The matrix representation of D in the basis S is

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

(iii) Two complementary D-invariant subspaces in V are $W_1 = span\{\sin(x), \cos(x)\}$ and $W_2 = span\{\sin(2x), \cos(2x)\}$.

Exercise 13
Let

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

First, we find the eigenvalues of A by solving

$$det(A - \lambda I) = \begin{vmatrix} 0.8 - \lambda & 0.4 \\ 0.2 & 0.6 - \lambda \end{vmatrix} = \lambda^2 - 1.4\lambda + 0.4 \implies \lambda_1 = 0.4, \lambda_2 = 1$$

Next, we find the eigenvectors corresponding to these eigenvalues.

$$\lambda_1 = 0.4 : \begin{bmatrix} 0.4 & 0.4 \\ 0.2 & 0.2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 1 : \begin{bmatrix} -0.2 & 0.4 \\ 0.2 & -0.4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\implies P = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} \implies P^{-1}AP = \begin{bmatrix} 0.4 & 0 \\ 0 & 1 \end{bmatrix}$$

Exercise 15
If $(\lambda_i)_{i=1}^n$ are the eigenvalues of a semisimple matrix $A \in M_n(\mathbb{F})$ and $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial, then $(f(\lambda_i))_{i=1}^n$ are the eigenvalues of $f(A) = a_0 I + a_1 A + \cdots + a_n A^n$.
Proof:
Since A is semisimple, we know that it is diagonalizable, i.e. there exist a nonsingular matrix P and a diagonalizable matrix D s.t. $A = P^{-1}DP$. We want to find the eigenvalues of

$$\begin{aligned} f(A) &= a_0 I + a_1 A + \cdots + a_n A^n \\ &= a_0 I + a_1 P^{-1}DP + \cdots + a_n P^{-1}D^n P \\ &= P^{-1}(a_0 I + a_1 D + \cdots + a_n D^n)P \\ &= P^{-1}f(D)P. \end{aligned}$$

Since $f(A)$ and $f(D)$ are similar matrices, we know that they have the same eigenvalues, and thus it suffices to show that the eigenvalues of $f(D)$ are $(f(\lambda_i))_{i=1}^n$. To prove this, we use the fact that $f(D)$ is diagonal, which means that its eigenvalues are its diagonal entries. Since we know that the eigenvalues of D^k are $(\lambda_i^k)_{i=1}^n$, we can see that the matrix

$$f(D) = a_0 I + a_1 D + \cdots + a_n D^n$$

has as its diagonal entries

$$a_0 I + a_1 \lambda_i + \cdots + a_n \lambda_i^n = f(\lambda_i).$$

These $(f(\lambda_i))_{i=1}^n$ are the eigenvalues of $f(D)$ and therefore of $f(A)$.

Exercise 16
Let

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

(i) We want to compute $\lim_{n\to\infty} A^n$ with respect to the 1-norm, i.e, find a matrix B s.t. for any $\epsilon > 0$, there exists an $N > 0$ with

$$\|A^k - B\|_1 < \epsilon \quad \text{whenever} \quad k > N \implies \|PD^kP^{-1} - B\|_1 < \epsilon$$

Using the diagonal matrix we found in Exercise 13,

$$D = \begin{bmatrix} 0.4 & 0 \\ 0 & 1 \end{bmatrix} \implies D^k = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{as} \quad k \to \infty$$

we see that

$$PD^kP^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix} = B$$

(ii) The 1-norm, the ∞-norm, and the Frobenius norm are topologically equivalent, so the answer does not depend on the choice of norm.

(iii) We want to find all the eigenvalues of the matrix $3I + 5A + A^3$. Since A is diagonalizable, we know that it is semisimple. Then we can apply the Semisimple Spectral Mapping Theorem. The eigenvalues of A are $\{0.4, 1\}$, so the eigenvalues of $f(A) = 3I+5A+A^3$ are $f(0.4) = 3+5(0.4)+0.4^3 = 5.064$ and $f(1) = 3+5(1)+1^3 = 9$.

Exercise 18

If λ is an eigenvalue of the matrix $A \in M_n(\mathbb{F})$, then there exists a nonzero row vector $\mathbf{x}^T$ s.t. $\mathbf{x}^T = \lambda \mathbf{x}^T$.

Proof:

Let $\mathbf{x}$ be the eigenvector of the matrix A^T corresponding to the eigenvalue λ. Then we know that $\mathbf{x}$ is nonzero and $A^T\mathbf{x} = \lambda\mathbf{x}$. Taking the transpose of both sides, we see that $\mathbf{x}^T A = \lambda \mathbf{x}^T$.

Exercise 20

If A is Hermitian and orthonormally similar to B, then B is also Hermitian.

Proof:

Since B is orthonormally similar to A, we know that there exists an orthonormal matrix U s.t. $B = U^H AU$. Then

$$B^H = (U^H AU)^H = U^H A^H U = U^H AU = B.$$

Exercise 24

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

$$\rho(\mathbf{x}) = \frac{\langle \mathbf{x}, A\mathbf{x} \rangle}{\|\mathbf{x}\|^2},$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{F}^n$. The Rayleigh quotient can only take on real values for Hermitian matrices and only imaginary values for skew-Hermitian matrices.

Proof:

The denominator of the Rayleigh quotient is real and strictly positive for $\mathbf{x} \neq \mathbf{0}$. Therefore, we want to analyze the numerator $\langle \mathbf{x}, A\mathbf{x} \rangle$. Taking the standard inner on $\mathbb{F}^n$, we have that

$$\begin{aligned} \langle \mathbf{x}, A\mathbf{x} \rangle &= \langle \mathbf{x}, \lambda \mathbf{x} \rangle \\ &= \sum_{i=1}^{n} \bar{x}_i \lambda x_i \\ &= \sum_{i=1}^{n} \|x_i\|^2 \lambda \end{aligned}$$

for the eigenvalue λ of A corresponding to $\mathbf{x}$. Since Hermitian matrices have only real eigenvalues and skew-Hermitian matrices have only imaginary eigenvalues, we know that the same holds for the Rayleigh quotient.

Exercise 25

Let $A \in M_n(\mathbb{C})$ be a normal matrix with eigenvalues $(\lambda_1, \cdots, \lambda_n)$ and corresponding orthonormal eigenvectors $[\mathbf{x}_1, \cdots, \mathbf{x}_n]$.

(i) The identity matrix can be written $I = \mathbf{x}_1\mathbf{x}_1^H + \cdots + \mathbf{x}_n\mathbf{x}_n^H$.

Proof:

Since the vectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are orthonormal, we know that $\|x_i^H x_j\| = 0$ for $i \neq j$. Then,

$$\begin{aligned}(\mathbf{x}_1\mathbf{x}_1^H + \cdots + \mathbf{x}_n\mathbf{x}_n^H)\mathbf{x}_j &= \mathbf{x}_1(\mathbf{x}_1^H\mathbf{x}_j) + \cdots + \mathbf{x}_j(\mathbf{x}_j^H\mathbf{x}_j) + \cdots + \mathbf{x}_n(\mathbf{x}_n^H\mathbf{x}_j)\\ &= 0 + \cdots + 1 \cdot \mathbf{x}_j + \cdots + 0\\ &= \mathbf{x}_j.\end{aligned}$$

Thus $\mathbf{x}_1\mathbf{x}_1^H + \cdots + \mathbf{x}_n\mathbf{x}_n^H$ must be equal to I.

(ii) A can be written as $A = \lambda_1\mathbf{x}_1\mathbf{x}_1^H + \cdots + \lambda_n\mathbf{x}_n\mathbf{x}_n^H$.

Proof:

As before, we can write

$$\begin{aligned}(\lambda_1\mathbf{x}_1\mathbf{x}_1^H + \cdots + \lambda_n\mathbf{x}_n\mathbf{x}_n^H)\mathbf{x}_j &= \lambda_1\mathbf{x}_1(\mathbf{x}_1^H\mathbf{x}_j) + \cdots + \lambda_j\mathbf{x}_j(\mathbf{x}_j^H\mathbf{x}_j) + \cdots + \mathbf{x}_n\lambda_n(\mathbf{x}_n^H\mathbf{x}_j)\\ &= 0 + \cdots + 1 \cdot \lambda_j\mathbf{x}_j + \cdots + 0\\ &= \lambda_j\mathbf{x}_j\\ &= A\mathbf{x}_j.\end{aligned}$$

Thus $\lambda_1\mathbf{x}_1\mathbf{x}_1^H + \cdots + \lambda_n\mathbf{x}_n\mathbf{x}_n^H$ must be equal to A.

Exercise 27

Assume $A \in M_n(\mathbb{F})$ is positive definite. Then all its diagonal entries are real and positive.

Proof:

Since A is positive definite, we know that $\langle \mathbf{x}, A\mathbf{x}\rangle > 0 \; \forall \mathbf{x} \neq \mathbf{0}$. Consider the unit vector e_i. We can access the diagonal entries of A as follows:

$$e_i^H A e_i = a_{i,i} > 0$$

This guarantees that the diagonal entries of A are strictly positive.

Exercise 28

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

$$0 \leq tr(AB) \leq tr(A)tr(B),$$

and $\|\cdot\|_F$ is a matrix norm.
Proof:
Since A, B are positive semidefinite, we know that there exist matrices S_A and S_B s.t. $A = S_A^H S_A$ and $B = S_B^H S_B$. Then

$$\begin{aligned}
tr(AB) &= tr(S_A^H S_A S_B^H S_B) \\
&= tr(S_B S_A^H S_A S_B^H) \\
&= tr((S_A S_B^H)^H (S_A S_B^H)) \\
&= \|S_A S_B^H\|_F^2 \geq 0
\end{aligned}$$

which proves the first inequality. We also know that if A, B are positive semidefinite, then A, B are orthonormally similar to some diagonal matrices D_A, D_B. Since trace is invariant with respect to changes of bases, we know that

$$\begin{aligned}
tr(A)tr(B) &= tr(D_A)tr(D_B) \\
&= \Big(\sum_{i=1}^{p} \lambda_i^A\Big)\Big(\sum_{i=1}^{1} \lambda_i^B\Big) \\
&\geq \sum_{i=1}^{k} \lambda_i^A \lambda_i^B \\
&= tr(D_A D_B) \\
&= tr(AB)
\end{aligned}$$

Exercise 33
Assume $A \in M_n(\mathbb{F})$. Then

$$\|A\|_2 = \sup_{\|\mathbf{x}\|_2=1, \|\mathbf{y}\|_2=1} |\mathbf{y}^H A\mathbf{x}|$$

Proof:

$$\begin{aligned}
\sup_{\|\mathbf{x}\|_2=1, \|\mathbf{y}\|_2=1} |\mathbf{y}^H A\mathbf{x}| &= \sup_{\|\mathbf{x}\|_2=1, \|\mathbf{y}\|_2=1} \langle \mathbf{y}, A\mathbf{x} \rangle \\
&= \sup_{\|\mathbf{x}\|_2=1, \|\mathbf{y}\|_2=1} \|\mathbf{y}\|_2 \|A\mathbf{x}\|_2 \\
&= \sup_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2
\end{aligned}$$

Exercise 36
Give an example of a 2×2 matrix whose determinant is nonzero and whose singular values are not equal to any of its eigenvalues:

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

The eigenvalue of the matrix is $\lambda = -1$, and the singular value is $\sigma = 1$.

Exercise 38
If $A \in M_{m\times n}(\mathbb{F})$, then the Moore-Penrose pseudoinverse of A satisfies the following:
(i) $AA^\dagger A = A$
Proof:

$$\begin{aligned}
AA^\dagger A &= (U_1\Sigma_1 V_1^H)(V_1\Sigma_1^{-1}U_1^H)(U_1\Sigma_1 V_1^H) \\
&= U_1\Sigma_1 V_1^H V_1\Sigma_1^{-1}U_1^H U_1\Sigma_1 V_1^H \\
&= U_1\Sigma_1\Sigma_1^{-1}\Sigma_1 V_1^H \\
&= U_1\Sigma_1 V_1^H \\
&= A
\end{aligned}$$

(ii) $A^\dagger AA^\dagger = A^\dagger$
Proof:

$$\begin{aligned}
A^\dagger AA^\dagger &= (V_1\Sigma_1^{-1}U_1^H)(U_1\Sigma_1 V_1^H)(V_1\Sigma_1^{-1}U_1^H) \\
&= V_1\Sigma_1^{-1}U_1^H U_1\Sigma_1 V_1^H V_1\Sigma_1^{-1}U_1^H \\
&= V_1\Sigma_1^{-1}\Sigma_1\Sigma_1^{-1}U_1^H \\
&= V_1\Sigma_1^{-1}U_1^H \\
&= A^\dagger
\end{aligned}$$

(iii) $(AA^\dagger)^H = AA^\dagger$
Proof:

$$\begin{aligned}
(AA^\dagger)^H &= ((U_1\Sigma_1 V_1^H)(V_1\Sigma_1^{-1}U_1^H))^H \\
&= (V_1\Sigma_1^{-1}U_1^H)^H(U_1\Sigma_1 V_1^H)^H \\
&= U_1(\Sigma_1^{-1})^H V_1^H V_1\Sigma_1^H U_1^H \\
&= U_1(\Sigma_1^{-1})^H\Sigma_1^H U_1^H \\
&= U_1(\Sigma_1^H)^{-1}\Sigma_1^H U_1^H \\
&= U_1 U_1^H \\
&= U_1\Sigma_1\Sigma_1^{-1}U_1^H \\
&= U_1\Sigma_1 V_1^H V_1\Sigma_1^{-1}U_1^H \\
&= (U_1\Sigma_1 V_1^H)(V_1\Sigma_1^{-1}U_1^H) \\
&= AA^\dagger
\end{aligned}$$

(iv) $(A^\dagger A)^H = A^\dagger A$

Proof:

$$\begin{aligned}(A^\dagger A)^H &= (V_1\Sigma_1^{-1}U_1^H)(U_1\Sigma_1V_1^H)\\ &= V_1\Sigma_1^{-1}U_1^HU_1\Sigma_1V_1^H\\ &= V_1V_1^H\\ &= V_1\Sigma_1^{-1}\Sigma_1V_1^H\\ &= V_1\Sigma_1^{-1}U_1^HU_1\Sigma_1V_1^H\\ &= (V_1\Sigma_1^{-1}U_1^H)(U_1\Sigma_1V_1^H)\\ &= A^\dagger A\end{aligned}$$

(v) $AA^\dagger = proj_{\mathscr{R}(A)}$ is the orthogonal projection onto $\mathscr{R}(A)$
Proof:
By part (i), we can see that $AA^\dagger AA^\dagger = AA^\dagger$, so $AA^\dagger$ is idempotent. From part(iii), we see that $AA^\dagger = U_1U_1^H$ s.t. the columns of U_1 form an orthonormal basis for $\mathscr{R}(A)$. Then for some vector $\mathbf{x}$, we have that

$$U_1U_1^H\mathbf{x} = U_1\left[\mathbf{u}_1^H\mathbf{x}\cdots\mathbf{u}_r^H\mathbf{x}\right]^H = \sum_{i=1}^{n}\mathbf{u}_i^H\mathbf{x}\mathbf{u}_i = \sum_{i=1}^{n}\langle\mathbf{u}_i^H\mathbf{x}\rangle\mathbf{u}_i = proj_{\mathscr{R}(A)}\mathbf{x}$$

(vi) $A^\dagger A = proj_{\mathscr{R}(A^H)}$ is the orthogonal projection onto $\mathscr{R}(A^H)$
Proof:
As with part (v), we note that $A^\dagger AA^\dagger = A^\dagger A$, so it is idempotent. Also, we can see from part (iv) that $V_1A^\dagger A = V_1^H$ s.t. the columns of V_1 form an orthonormal basis for $\mathscr{R}(A^H)$. Then for some vector $\mathbf{x}$, we have that

$$V_1V_1^H\mathbf{x} = V_1\left[\mathbf{v}_1^H\mathbf{x}\cdots\mathbf{v}_r^H\mathbf{x}\right]^H = \sum_{i=1}^{n}\mathbf{v}_i^H\mathbf{x}\mathbf{v}_i = \sum_{i=1}^{n}\langle\mathbf{v}_i^H\mathbf{x}\rangle\mathbf{v}_i = proj_{\mathscr{R}(A^H)}\mathbf{x}$$