Solution to Homework 3

MA502 Linear Algebra

Problem 1.
An $n \times n$ matrix A with entries from field $\mathbb{F}$ is diagonalizable over $\mathbb{F}$ if there exists an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

(a) Prove that A is diagonalizable over field $\mathbb{F}$ if and only if $V = \mathbb{F}^n$ admits a basis each vector of which is an eigenvector for A.

Proof.
($\Rightarrow$): Suppose A is diagonalizable, then there exists an invertible P and a diagonal D such that $A = PDP^{-1}$. Write $P = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \end{bmatrix}$, where $\mathbf{p}_i$ is i^{th} column vector of P. Meanwhile, we can write $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$. Then, we have $AP = PD$, which we may write

$$A \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \ldots & \mathbf{p}_n \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \ldots & \mathbf{p}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Extracting the i^{th} column of both sides of this equation, we find $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$ (where λ_i could equal λ_j for $i \neq j$.) This means that each $\mathbf{p}_i$ (obviously non-zero since P is invertible) is an eigenvector for A with corresponding eigenvalue λ_i. Meanwhile, since P is invertible, then $\text{rank}(P) = n$. Hence its set of columns $\{\mathbf{p}_i\}_{i=1}^n$ form a basis of eigenvectors for $\mathbb{F}^n$.

($\Leftarrow$): Suppose we have a basis $\mathcal{B} = \{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$ for $\mathbb{F}^n$ consisting entirely of eigenvectors for A. Then there exist scalars (eigenvalues) $\lambda_1, \ldots, \lambda_n$ in $\mathbb{F}$ such that $A\mathbf{p}_j = \lambda_j \mathbf{p}_j$ for $1 \leq j \leq n$. Stacking the n vectors $\mathbf{p}_j$ together as columns of a square matrix P, we get

$$\begin{aligned} A \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \end{bmatrix} &= \begin{bmatrix} \lambda_1 \mathbf{p}_1 & \lambda_2 \mathbf{p}_2 & \cdots & \lambda_n \mathbf{p}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \end{aligned}$$

Since $\{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$ is a basis, the $n \times n$ matrix $P = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \end{bmatrix}$ is invertible. Defining D to be the diagonal matrix with (j, j)-entry λ_j, we have $AP = PD$. Hence, $A = PDP^{-1}$, and A is diagonalizable. □

(b) Prove: If A is diagonalizable, then $\operatorname{tr}(A)$ is the sum of all eigenvalues of A, counting multiplicities.

Proof. Since A is diagonalizable, then there exists invertible P and diagonal D such that $A = PDP^{-1}$. We find

$$\operatorname{tr}(A) = \operatorname{tr}(PDP^{-1}) = \operatorname{tr}(P^{-1}PD) = \operatorname{tr}(D)$$

using the fact that $\operatorname{tr}(MN) = \operatorname{tr}(NM)$. Suppose $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$, where the scalars λ_j are the eigenvalues of A by the construction above, counting multiplicities. Hence, $\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$. □

Problem 2. We compute the spectral decomposition of the matrix $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$.

(a) Find and orthonormal basis for the eigenspace belonging to eigenvalue $\lambda = 1$.

Proof. When $\lambda = 1$, $A - \lambda I = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. This reduced row echelon form allows us to read off a basis for the eigenspace: $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$.

We apply Gram-Schmidt orthogonalization to find

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

Hence, $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthonormal basis for the eigenspace associated to $\lambda = 1$. □

(b) Find an orthogonal matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

Proof. Since A has trace zero, our remaining eigenvalue must be -2, and its eigenvector is easily found to be $\begin{bmatrix} -1 & 1 & 1 \end{bmatrix}^\top$. Since A is symmetric, this is already orthogonal to our previously-discovered eigenvectors. So we normalize to obtain

$$\mathbf{v}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Then we have $P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & 1 & -\sqrt{2} \\ 0 & 2 & \sqrt{2} \\ \sqrt{3} & -1 & \sqrt{2} \end{bmatrix}$, $D = \begin{bmatrix} 1 & & \\ & 1 & \\ & & -2 \end{bmatrix}$, and $A = PDP^{-1}$. □

(c) Compute the orthogonal projection matrix E_λ from $\mathbb{R}^3$ onto the eigenspace V_λ for each eigenvalue λ of A and compute $\sum_\lambda \lambda E_\lambda$. Explain.

Proof. Based on the construction above, we have $P^\top$ as orthogonal matrix, which means $P^{-1} = P^T$.

$$\begin{aligned}
A &= \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & & -2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ \mathbf{v}_3^\top \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ \mathbf{v}_3^\top \end{bmatrix} + \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 0 & & \\ & 0 & \\ & & -2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ \mathbf{v}_3^\top \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ 0 \end{bmatrix} + (-2) \cdot \begin{bmatrix} 0 & 0 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \mathbf{v}_3^\top \end{bmatrix} \\
&= 1 \cdot (\mathbf{v}_1\mathbf{v}_1^\top + \mathbf{v}_2\mathbf{v}_2^\top) + (-2) \cdot \mathbf{v}_3\mathbf{v}_3^\top
\end{aligned} \tag{1}$$

Since $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ are orthonormal basis, we have $\mathbf{v}_1\mathbf{v}_1^\top + \mathbf{v}_2\mathbf{v}_2^\top$ and $\mathbf{v}_3\mathbf{v}_3^\top$ are projection matrices. (The second one is obvious, and the first one is due to $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$.) For $\lambda = 1$ and $\mu = -2$, the projection matrices are

$$P_\lambda = \frac{1}{3} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}, \qquad P_\mu = \frac{1}{3} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$$

□

(d) Use part (b) to compute A^9.

Proof. Since $A = PDP^{-1}$, $A^n = PDP^{-1}PDP^{-1} \dots PDP^{-1} = PD^nP^{-1}$. Clearly the diagonal entries of D^9 are 1, 1, and $(-2)^9 = -512$. So

$$A^9 = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & & -2^9 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ \mathbf{v}_3^\top \end{bmatrix} = \begin{bmatrix} -170 & 171 & 171 \\ 171 & -170 & -171 \\ 171 & -171 & -170 \end{bmatrix}.$$

(Note that we can also obtain this using part (c): since $A = \lambda P_\lambda + \mu P_\mu$ and $P_\lambda P_\mu = 0$, we have $A^9 = 1^9 P_\lambda + (-2)^9 P_\mu$.) □

Problem 3.
Consider the vector space $V = \mathcal{C}([0,\pi])$ of real-valued continuous functions on the interval $[0,\pi]$. Any bivariate continuous function $A(x,y)$ determines a linear operator $\tau\colon V \to V$ given by

$$\tau f(x) = \int_0^{\pi/2} A(x,y) f(y) dy$$

In this exercise, we study $A(x,y) = \cos(x+y)$.

(a) Find all eigenfunctions $f(x)$ for τ that lie in the two-dimensional subspace

$$S = \{a \sin x + b \cos x \mid a, b \in \mathbb{R}\}.$$

Proof. For some $f(x) = a \sin x + b \cos x$ in S, we need to find a, b and λ such that $\tau f = \lambda f$. We need

$$\int_0^{\pi/2} \cos(x+y)(a \sin y + b \cos y) dy = \lambda(a \sin x + b \cos x).$$

We immediately see that, since

$$\cos(x+y) = \cos x \cos y - \sin x \sin y,$$

we can simplify the left-hand side to

$$\int_0^{\pi/2} \cos(x+y)(a \sin y + b \cos y) dy =$$

$$-\sin x \left[\int_0^{\pi/2} \sin(y)(a \sin y + b \cos y) dy \right] + \cos x \left[\int_0^{\pi/2} \cos(y)(a \sin y + b \cos y) dy \right].$$

(This is called a "separable kernel": $A(x,y)$ is expressible as a sum of products $p(x)q(y)$.) Since

$$\begin{aligned} \int_0^{\pi/2} \sin(y)(a \sin y + b \cos y) dy &= \frac{1}{4}(2b + \pi a) \\ \int_0^{\pi/2} \cos(y)(a \sin y + b \cos y) dy &= \frac{1}{4}(2a + \pi b) \end{aligned}$$

we must solve the system of equations

$$\left\{ \frac{1}{4}(2b + \pi a) = -\mu a, \quad \frac{1}{4}(2a + \pi b) = \mu b \right\}.$$

For $\mu = \pm\frac{1}{4}\sqrt{\pi^2 - 4}$ (sorry for the typo!), we have eigenvectors (or "eigenfunctions")

$$f(x) = -\left(\frac{b\mu}{2}\right) \sin x + \ b \ \cos x.$$

So we have two distinct eigenvalues and a one-dimensional eigenspace corresponding to each choice of μ. ☐

(b) Consider the subspace W of V spanned by $\{\cos(nx), \sin(nx)\}_{n=1}^{\infty}$. Is W a τ-invariant subspace? Explain.

Proof. We now know, from our trig identity that allowed us to separate $A(x, y)$, that any $f(x)$ which is continuous on $[0, \pi/2]$ gives us $\tau f \in \text{span}\{\sin x,\ \cos x\}$. So, yes, this space is τ-invariant: for an arbitrary linear combination of basis elements

$$f(x) = \sum_{k=1}^{M} a_k \sin(n_k x) + b_k \cos(n_k x)$$

we have

$$\tau f(x) = -\sin x \left[\int_0^{\pi/2} \sin(y) f(y) dy\right] + \cos x \left[\int_0^{\pi/2} \cos(y) f(y) dy\right]$$

where the coefficients of $\sin x$ and $\cos x$ are some real constants that we don't care about. So subspace W is τ-invariant (and much more!). □

Problem 4.
Consider the vector space V spanned by the following complex-valued functions on the interval $[-\pi, \pi]$: $\mathbf{v}_1 = e^z + \cos z$, $\mathbf{v}_2 = e^z + \sin z$, $\mathbf{v}_3 = \sin z$. This space V is invariant under the differential operator $\tau \in \mathcal{L}(V)$ given by $\tau f = \frac{df}{dz}$.

(a) Compute the matrix $[\tau]_{\mathcal{B}}$ of τ with respect to basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

Proof. Since

$$\begin{aligned} \tau\mathbf{v}_1 &= e^z - \sin z = \mathbf{v}_2 - 2\mathbf{v}_3 \\ \tau\mathbf{v}_2 &= e^z + \cos z = \mathbf{v}_1 \\ \tau\mathbf{v}_3 &= \cos z = \mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3 \end{aligned} \tag{2}$$

we have $A = [\tau]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ -2 & 0 & 1 \end{bmatrix}$. □

(b) Let us write $A = [\tau]_{\mathcal{B}}$. Using a computer, if necessary, find the eigenvalues of τ and determine a basis of eigenfunctions $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$. Write down the matrix $[\tau]_{\mathcal{C}}$.

Proof. Eigenvalues for A are $1, \pm i$.
When $\lambda_1 = 1$, we require $\frac{df}{dz} = f$, and hence we choose $\mathbf{w}_1 = e^z$.
For $\lambda_2 = i$, we seek $\frac{df}{dz} = if$, and hence we take $\mathbf{w}_2 = \cos(z) + i\sin(z)$.
For $\lambda_3 = -i$, we need $\frac{df}{dz} = -if$ so we choose $\mathbf{w}_3 = \cos(z) - i\sin(z)$.

So, for $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$, $[\tau]_{\mathcal{C}} = \begin{bmatrix} 1 & & \\ & i & \\ & & -i \end{bmatrix}$. □

(c) Can you choose an inner product on space V with respect to which operator τ is self-adjoint? Explain.

Proof. No, we cannot have such a inner product. If τ were self-adjoint, then all eigenvalues of τ would be real. However, in part (b), we already find that its eigenvalues include $\pm i$. □

Problem 5.

(a) For which $n \times n$ matrices is the minimal polynomial linear? Explain.

Proof. Claim: A should be in the form of λI ($\lambda \in \mathbb{C}$, $\lambda \neq 0$)

We know the minimal polynomial is a monic polynomial of lowest degree which, when evaluated at A, gives the zero matrix. If $m_A(t) = t - c_0$ is the minimal polynomial for A, then

$$A - c_0 I = 0$$

or $A = c_0 I$. So A is just a non-zero scalar multiple of the identity matrix. □

(b) Find a 3×3 matrix with real entries which is not diagonalizable yet whose minimal polynomial is quadratic.

Proof. From our knowledge of generalized eigenspaces, we see that A must have just one eigenvalue, if we ignore multiplicities. Let $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Then, we have

$$A^2 = \begin{bmatrix} 4 & 4 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} = 4A - 4I$$

and A has minimal polynomial $m_A(t) = (t-2)^2$. □

Problem 6.

Let S be a subspace of $\mathbb{R}^n$ with basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ and let U be the $n \times k$ matrix with j^{th} column $\mathbf{v}_j$. In class, we showed that, when $\mathcal{B}$ is an orthonormal basis, $P = UU^\top$ is the matrix representing orthogonal projection of $\mathbb{R}^n$ onto S with respect to the standard basis.

(a) No longer assume that $\mathcal{B}$ is orthonormal, but that it is just any basis for S. Prove that the projection operator $\rho_{S,S^\perp}$ is represented by the matrix $P = U(U^\top U)^{-1}U^\top$.

Proof. By direct computation, it is easy to see that $P^2 = P$. So P is idempotent. We need to prove that $\forall \mathbf{v} \in S$, $P\mathbf{v} = \mathbf{v}$, and $\forall \mathbf{v} \in S^\perp$, $P\mathbf{v} = 0$.

Given $\mathbf{v} \in S$, since $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a basis, we can write $\mathbf{v} = \sum_{k=1}^{n} a_k \mathbf{v}_k = U\mathbf{a}$, where $\mathbf{a} = \begin{bmatrix} a_1 & \ldots & a_k \end{bmatrix}^\top$. Then,

$$\begin{aligned} P\mathbf{v} &= U(U^\top U)^{-1}U^\top \mathbf{v} \\ &= U(U^\top U)^{-1}U^\top U\mathbf{a} \\ &= U\mathbf{a} = \mathbf{v} \end{aligned} \tag{3}$$

On the other hand, for $\mathbf{w} \in S^\perp$, we have $\mathbf{v}_i^\top \mathbf{w} = 0$ for $1 \leq j \leq k$ so $U^\top \mathbf{w} = 0$. Then

$$P\mathbf{w} = U(U^\top U)^{-1}U^\top \mathbf{w} = 0$$

This proves that P is projection operator $\rho_{S,S^\perp}$. □

[NOTE: We just assumed that $U^\top U$ is invertible. If not, then there would be a nonzero vector $\mathbf{a}$ with $(U^\top U)\mathbf{a} = 0$. But, as we've just seen $U\mathbf{a}$ is a vector in subspace S and the only vector $\mathbf{s}$ in S which can be orthogonal to every vector in a basis is the zero vector: $U^\top \mathbf{s} = 0$ implies $\mathbf{s} = 0$.]

(b) Does the matrix $P = UU^\top$ represent $\rho_{S,T}$ for some T ? Explain.

Proof. We show that P is a projection matrix if and only if $U^\top U = I_k$. Indeed, if $\mathbf{s} \in S$, then there is a vector $\mathbf{a}$ of length k with $\mathbf{s} = U\mathbf{a}$. If P is the matrix representing $\rho_{S,T}$, then $P\mathbf{s} = \mathbf{s}$ So $UU^\top U\mathbf{a} = U\mathbf{a}$ for every possible vector $\mathbf{a}$ of length k. We multiply both sides on the left by $U^\top$ to find

$$(U^\top U)(U^\top U)\mathbf{a} = (U^\top U)\mathbf{a} \ .$$

We just showed that $U^\top U$ is invertible. So we may multiply both sides on the left by its inverse to see that

$$(U^\top U)\mathbf{a} = I\mathbf{a}$$

for every $\mathbf{a} \in \mathbb{R}^k$. Since the two matrices have the same action, they are equal: $U^\top U = I$. □

(c) Now let V be the real vector space of polynomial functions $c_0 + c_1 x + \cdots + c_n x_n$ on the interval $[0, 1]$ with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$ and let S be the space spanned by $\{1, x\}$. Find a simple expression for the orthogonal projection of V onto S with respect to this inner product.

Proof. First, we need to make the basis orthonormal. Hence, by Gram-Schmidt orthogonalization, we have $f_1 = 1$, and

$$\bar{f}_2 = x - \langle 1, x\rangle \cdot 1 = x - \int_0^1 x\,dx = x - \frac{1}{2}$$

Then normalize $\bar{f}_2$ to get $f_2 = \sqrt{12}(x - \frac{1}{2})$, and then, for any $g \in V$, our projection is given by

$$P(g) = \langle g, 1\rangle \cdot 1 + \langle g, f_2\rangle \cdot f_2$$

□