Rules:

- Unless otherwise stated, all answers must be mathematically justified.
- Partial answers will be graded.
- Your submission has to be uploaded to Gradescope. In Gradescope, indicate the page on which each problem is written.
- You can work in groups but each student must write his/her/their own solution based on his/her/their own understanding of the problem. Please list on your submission the students you work with for the homework (this will not affect your grade).
- Problems with a ($\star$) are extra credit, they will not (directly) contribute to your score of this homework. However, for every 4 extra credit questions successfully answered your lowest homework score get replaced by a perfect score.
- If you have any questions, feel free to ask them on Ed Discussion (so that everyone can benefit from the answer) or stop at the office hours.

Problem 6.1 (2 points). **(a)** *Consider the matrix*

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Show that 3, −1, and 2 are eigenvalues of A and give an associated eigenvector in each case. What is the spectrum of A? Justify.

We compute the characteristic polynomial for matrix A, which is given by det$(A - \lambda I_3) = (3-\lambda)(-2-\lambda)(2-\lambda) = 0$, which shows that the spectrum of A are indeed $3, -1, 2$. Now we compute the eigenvectors under each eigenvalue.

(a) $\lambda = 3$, we have to compute the basis for $\mathcal{N}(A - 3I_3)$, which is given by $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

(b) $\lambda = -1$, we have to compute the basis for $\mathcal{N}(A - (-1)I_3)$, which is given by $\begin{pmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \\ 0 \end{pmatrix}$.

(c) $\lambda = 2$, we have to compute the basis for $\mathcal{N}(A - 2I_3)$, which is given by $\begin{pmatrix} -8/\sqrt{74} \\ 1/\sqrt{74} \\ 3/\sqrt{74} \end{pmatrix}$.

(b) *Let $A \in \mathbb{R}^{n \times n}$ be an upper-triangular matrix:*

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{pmatrix}.$$

For which values of $\lambda \in \mathbb{R}$ is $A - \lambda \mathrm{Id}$ invertible? What is the spectrum of A?

In order for the matrix A to be invertible, we have to make sure det$(A) \neq 0$, otherwise there exists two columns that are multiple of each other, resulting in linear dependence. Thus in order for $A - \lambda I_d$ to be invertible, we have to ensure det$(A - \lambda I_d) \neq 0$, which implies that det$(A - \lambda I_d) = p(\lambda) = (a_{1,1} - \lambda)(a_{2,2} - \lambda) \times \cdots \times (a_{n,n} - \lambda) \neq 0$, which means $\lambda \neq a_{i,i} \forall i = 1, 2, \cdots, n$.
The spectrum of A are just the solution to the characteristic polynomial given by $p(\lambda) = 0$, which are $a_{i,i}, \forall i = 1, 2, \cdots, n$.

Problem 6.2 (3 points). *Let $A \in \mathbb{R}^{n\times n}$ be a square matrix. Suppose that there exists a basis $(e_1, \dots, e_n)$ of $\mathbb{R}^n$ composed of eigenvectors of A, associated with eigenvalues $\lambda_1, \dots, \lambda_n$.*

(a) *Let $x \in \mathbb{R}^n$ with coordinates $(x_1, \dots, x_n)$ in the basis $(e_1, \dots, e_n)$. What are the coordinates of Ax in the basis $(e_1, \dots, e_n)$?*

We have $\vec{x} = \sum_{i=1}^n x_i\vec{e}_i$ by the problem statement. We then apply linear transformation A and get $A\vec{x} = \sum_{i=1}^n x_iA\vec{e}_i = \sum_{i=1}^n x_i\lambda_i\vec{e}_i$ due to the fact that $\vec{e}_i$'s are eigenvectors corresponding to the spectrum of A. Thus we have $\vec{x} = \sum_{i=1}^n (x_i\lambda_i)\vec{e}_i$, which implies that the coordinates are $(\lambda_1, \dots, \lambda_n)$.

(b) *Show that in the basis $(e_1, \dots, e_n)$, the matrix of the linear map A is diagonal.*

Denote the transformation matrix under eigen-basis by $\widetilde{A}$, then we know from the part (a) that in the eigen-basis, for each basis vector $\vec{e}_i$ we have $A(\vec{e}_i) = (0, \cdots, \lambda_i, \cdots 0)$, thus the i-th column of $\widetilde{A}$ by definition should be $A(\vec{e}_i)^\top$. This implies that $\widetilde{A} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{bmatrix}$, which is a diagonal matrix.

(c) *Suppose that A is invertible. Calculate the coordinates of $A^{-1}x$ in the basis $(e_1, \dots, e_n)$ and give the expression of the matrix of the linear map A^{-1} in this basis.*

We know that the eigenvalues for matrix A^{-1} should be λ_i^{-1} and the eigenvectors are the same as A. Using the same logic from part (a), we have $\forall \vec{x} = \sum_{i=1}^n x_i\vec{e}_i, A^{-1}\vec{x} = \sum_{i=1}^n x_iA^{-1}\vec{e}_i = \sum_{i=1}^n x_i\lambda_i^{-1}\vec{e}_i$, which implies that the coordinate should be $(\lambda_1^{-1}, \dots, \lambda_n^{-1})$. Using the same logic as part (b), denote the transformation matrix under eigen-basis by $\widetilde{A^{-1}}$, then we know from the part (a) that in the eigen-basis, for each basis vector $\vec{e}_i$ we have $A^{-1}(\vec{e}_i) = (0, \cdots, \lambda_i^{-1}, \cdots 0)$, thus the i-th column of $\widetilde{A^{-1}}$ by definition should be $A^{-1}(\vec{e}_i)^\top$. This implies that $\widetilde{A^{-1}} = \begin{bmatrix} \lambda_1^{-1} & 0 & \cdots & 0 \\ 0 & \lambda_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n^{-1} \end{bmatrix}$.

Problem 6.3 (3 points). *The trace of a square $A \in \mathbb{R}^{n\times n}$ matrix is defined as the sum of its diagonal elements*

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{i,i}.$$

(a) *Show that for any $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{m\times n}$, $\mathrm{Tr}(BC) = \mathrm{Tr}(CB)$.*

$$\begin{aligned}
Tr(BC) &= \sum_{i=1}^{n}\sum_{k=1}^{m} B_{ik}C_{ki} && \text{(By definition)}\\
&= \sum_{k=1}^{m}\sum_{i=1}^{n} C_{ki}B_{ik} && \text{(Change the summation sequence)}\\
&= Tr(CB) && \text{(By definition)}
\end{aligned}$$

(b) *Consider a matrix $A \in \mathbb{R}^{n\times n}$ that can be written $A = PDP^{-1}$ where $P \in \mathbb{R}^{n\times n}$ is invertible and $D \in \mathbb{R}^{n\times n}$ is diagonal. Calculate $\mathrm{Tr}(A)$ as a function of D.*

$$\begin{aligned}
Tr(PDP^{-1}) &= Tr(P^{-1}PD) && \text{(Using the result from part (a))}\\
&= Tr(D) && \text{(Property of matrix inverse)}
\end{aligned}$$

(c) *Show that $\langle B, C\rangle = \mathrm{Tr}(B^\top C)$ defines an inner product on the space of rectangular matrices $\mathbb{R}^{n\times m}$.*

We have to verify the three properties:

(a) Symmetry: $\langle B, C\rangle = \mathrm{Tr}(B^\top C) = Tr(C^\top B) = \langle C, B\rangle$(By the fact that $Tr(A^\top) = Tr(A) \forall A \in \mathbb{R}^{n\times n}$)

(b) Linearity:

$$\begin{aligned}
\langle \alpha A + B, C\rangle &= \mathrm{Tr}((\alpha A + B)^\top C) && (1)\\
&= \mathrm{Tr}(\alpha A^\top C + B^\top C). && \text{(By the linearity of trace)}\\
&= \alpha \mathrm{Tr}(A^\top C) + \mathrm{Tr}(B^\top C) && \text{(By the linearity of trace)}\\
&= \alpha\langle A, C\rangle + \langle B, C\rangle && \text{(By defintion)}
\end{aligned}$$

(c) Positive Definiteness:

$$\begin{aligned}
\langle A, A\rangle &= Tr(A^\top A) && \text{(By definition)}\\
&= \sum_{i=1}^{n}\sum_{k=1}^{n} A^\top_{i,k}A_{k,i} && \text{(By definition)}\\
&= \sum_{i=1}^{n}\sum_{k=1}^{n} A_{k,i}A_{k,i} && \text{(Transpose)}\\
&= \sum_{i=1}^{n}\sum_{k=1}^{n} A_{k,i}^2 && \geq 0
\end{aligned}$$

. When $\langle A, A\rangle = 0$, we have $\sum_{i=1}^{n}\sum_{k=1}^{n} A_{k,i}^2 = 0$, which implies $A_{i,k} = 0, \forall 1 \leq i, k \leq n$, which implies that $A = 0$.

Problem 6.4 (3 points). *Let $P \in \mathbb{R}^{n\times n}$ be a stochastic matrix.*

(a) *Let $x \in \Delta_n$ be a probability vector. Show that Px is a probability vector.*

By definition, we have $\sum_{i=1}^n \vec{x}_i = 1$, then $\sum_{i=1}^n P\vec{x} = \sum_{i=1}^n \sum_{k=1}^n P_{ik}\vec{x}_k = \sum_{k=1}^n (\sum_{i=1}^n P_{ik})\vec{x}_k = \sum_{k=1}^n 1 \times \vec{x}_k = 1$, which implies that $P\vec{x}$ is a probablity vector of $\mathbb{R}^n$.

(b) *Assume that there exists a basis $(e_1, \dots, e_n)$ of $\mathbb{R}^n$ composed of eigenvectors of P associated with eigenvalues $\lambda_1, \dots, \lambda_n$. Let $x = x_1e_1 + \cdots + x_ne_n$. Show that $P^k x = \lambda_1^k x_1 e_1 + \cdots \lambda_n^k x_n e_n$.*

We could prove it by induction, assume the inductive hypothesis is $M(k) : P^k\vec{x} = \lambda^k x_i\vec{e}_1 + \cdots + \lambda_n^k x_n \vec{e}_n$,

(a) Base Step $k = 0$: $P^0\vec{x} = \vec{x} = \sum_{i=1}^n x_i\vec{e}_i = \sum_{i=1}^n \lambda_i^0 x_i \vec{e}_i$, which means $P(0)$ holds.

(b) Inductive Step $k = m$, assume $P(m)$ holds, then $P^i\vec{x} = \sum_{i=1}^n \lambda_m^i x_i\vec{e}_i$. Apply the linear transformation matrix P again we have $P^{m+1}\vec{x} = P\sum_{i=1}^n \lambda_i^m x_i\vec{e}_i = \sum_{i=1}^n \lambda_i^m x_i P\vec{e}_i = \sum_{i=1}^n \lambda_i^{m+1} x_i \vec{e}_i$, which means $P(m+1)$ holds.

. Thus by mathematical induction, we have $P^k x = \lambda_1^k x_1 e_1 + \cdots \lambda_n^k x_n e_n$ holds for all $k \in \mathbb{N}$. But things get tricky when $k < 0$(this is indeed possible). But our stochastic matrix P are not guaranteed to have inverse, so the statement does't necessaily hold for $k < 0$.

(c) *Further assume that $\lambda_1 = 1$ and $|\lambda_k| < 1$ for $k \geq 2$. Show that $P^k x \to x_1 e_1$ when $k \to \infty$.*

From part (b) we know that $P^k x = \lambda_1^k x_1 e_1 + \cdots \lambda_n^k x_n e_n$. If $\lambda_1 = 1$ and $|\lambda_i| < 1$ for $i \geq 2$, then $|\lambda_i|^k \to 0$ when $k \to \infty$ for $i \geq 2$. Thus $\lim_{n\to\infty} P^k\vec{x} \to \lambda_1^k x_1\vec{e}_1 = x_1\vec{e}_1$.

(d) *What can you say about the long-term behavior of the Markov chain with transition matrix P started at some $x_0 \in \Delta_n$?*

We can write $\vec{x}_0 = \sum_{i=1}^n x_i\vec{e}_i, \forall \vec{x}_0 \in \Delta_n$, then by Perron-Frobenius Theorem, there exists an eigenvalue with value 1($\vec{u}$ is its eigenvector with geometric multiplicity 1)and that $P^k\vec{x}_0$ will converge to this $\vec{\mu}$ when $k \to \infty$.

Problem 6.5 ($\star$)**.** *We say that a symmetric matrix $M \in \mathbb{R}^n$ is positive semi-definite if, for any $x \in \mathbb{R}^n$,*

$$x^\top M x \geq 0.$$

(a) *Are the following matrices positive semi-definite?*

$$M_1 = \mathrm{Id}_n, \quad M_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

(a) M_1 is PSD since $\vec{x}^\top M_1 \vec{x} = x_1^2 + x_2^2 \geq 0, \forall x_1, x_2 \in \mathbb{R}$.

(b) M_2 is not PSD since $\vec{x}^\top M_2 \vec{x} = x_1^2 + 4x_1x_2 + x_2^2$, which is negative when $x_1 = 1, x_2 = -1$.

(c) M_3 is PSD. Since $\vec{x}^\top M_3 \vec{x} = x_1^2 - 2x_1x_2 + x_2^2 = (x_1 - x_2)^2 \geq 0, \forall x_1, x_2 \in \mathbb{R}$.

(b) *Show that if M is positive semi-definite, then all its eigenvalues are non-negative.*

By spectrum theorem (which can be proved by Schur Decomposition), any symmetric matrix M can be diagonalized(have n linearly independent eigenvectors) and thus can be written as $M = VDV^\top$ where the column of V is M's eigenvectors and the diagonal entries of D are M's spectrum. Then $\vec{x}^\top M \vec{x} = \vec{x}^\top VDV^\top \vec{x} = \vec{y}D\vec{y} = \sum_{i=1}^n D_{i,i}y_i^2 \geq 0, \forall y_i \in \mathbb{R}$, which implies that $D_{i,i}$ are non-negative, which proves the original statement.