The Austalian National University Semester 2, 2021
School of Computing Assignment 4 Theory Questions

COMP3670/6670: Introduction to Machine Learning

Question 1 **Properties of Eigenvalues** (5+5=10 credits)

Let $\mathbf{A}$ be an invertible matrix.

1. Prove that all the eigenvalues of $\mathbf{A}$ are non-zero.

   Solution. If $\lambda = 0$ is an eigenvalue of $\mathbf{A}$, then the equation $\mathbf{A}\mathbf{x} = \mathbf{0}$ has non-trivial solutions. But, multiplying by $\mathbf{A}^{-1}$, we obtain

   $$\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{A}^{-1}\mathbf{0}$$

   $$\mathbf{x} = \mathbf{0}$$

   a contradiction.

2. Prove that for any eigenvalue λ of $\mathbf{A}$, λ^{-1} is an eigenvalue of $\mathbf{A}^{-1}$.

   Solution. Let λ be an eigenvalue of $\mathbf{A}$. Then there exists $\mathbf{x} \neq \mathbf{0}$ such that

   $$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

   $$\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{A}^{-1}\lambda\mathbf{x}$$

   $$\mathbf{x} = \lambda\mathbf{A}^{-1}\mathbf{x}$$

   $$\frac{1}{\lambda}x = \mathbf{A}^{-1}\mathbf{x}$$

   $$\mathbf{A}^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}$$

   where the second to last step is justified by 1, as $\lambda \neq 0$.

Question 2 **Properties of Eigenvalues II** (10 credits)

Let $\mathbf{B}$ be a square matrix. Let $\mathbf{x}$ be an eigenvector of $\mathbf{B}$ with eigenvalue λ. Prove that for all integers $n \geq 1$, $\mathbf{x}$ is an eigenvector of $\mathbf{B}^n$ with eigenvalue λ^n.

Solution. We proceed by induction on n.
Base case, $n = 1$ is trivial, as we are already given that λ is an eigenvalue of $\mathbf{B}$. Step case. Assume that λ^n is an eigenvalue of $\mathbf{B}^n$. Let $\mathbf{x}$ be the corresponding eigenvector for λ^n. Then,

$$\mathbf{B}^{n+1}\mathbf{x} = \mathbf{B}\mathbf{B}^n\mathbf{x} = \mathbf{B}(\mathbf{B}^n x) = \mathbf{B}\lambda^n x = \lambda^n(\mathbf{B}x) = \lambda^n\lambda\mathbf{x} = \lambda^{n+1}x$$

and hence λ^n is an eigenvalue of $\mathbf{B}^n$ for all $n \geq 1$.

Question 3 **Distinct eigenvalues and linear independence** (20+5 credits)

Let $\mathbf{A}$ be a $n \times n$ matrix.

1. Suppose that $\mathbf{A}$ has n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, and corresponding non-zero eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Prove that $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is linearly independant.

   Hint: You may use without proof the following property: If $\{\mathbf{y}_1, \ldots, \mathbf{y}_m\}$ is linearly dependent then there exists some p such that $1 \leq p < m$, $\mathbf{y}_{p+1} \in \text{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_p\}$ and $\{\mathbf{y}_1, \ldots, \mathbf{y}_p\}$ is linearly independent.

Solution. Suppose for a contradiction that $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ is linearly dependent. Then, by the hint, there exists p such that $1 \le p < n$, $\mathbf{x}_{p+1} \in \operatorname{span}\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ and $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ is linearly independent. So,

$$\mathbf{x}_{p+1} = c_1\mathbf{x}_1 + \dots + c_p\mathbf{x}_p \tag{1}$$

for some collection of scalars $c_1, \dots, c_p$. Apply A to both sides of the above equation, (noting that $\mathbf{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i$) and obtain

$$\lambda_{p+1}\mathbf{x}_{p+1} = c_1\lambda_1\mathbf{x}_1 + \dots + c_p\lambda_p\mathbf{x}_p \tag{2}$$

We can also multiply both sides of the first equation by λ_{p+1}, and obtain

$$\lambda_{p+1}\mathbf{x}_{p+1} = c_1\lambda_{p+1}\mathbf{x}_1 + \dots + c_p\lambda_{p+1}\mathbf{x}_p \tag{3}$$

Equations (2) and (3) are equal to each other.

$$c_1\lambda_1\mathbf{x}_1 + \dots + c_p\lambda_p\mathbf{x}_p = c_1\lambda_{p+1}\mathbf{x}_1 + \dots + c_p\lambda_{p+1}\mathbf{x}_p$$
$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{x}_1 + \dots + c_p(\lambda_1 - \lambda_{p+1})\mathbf{x}_p = 0$$

Since each of the eigenvalues are distinct, $\lambda_i - \lambda_{p+1} \neq 0$ for all $1 \le i \le p$. Since $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ are linearly independent, the only way that $c_1(\lambda_1 - \lambda_{p+1})\mathbf{x}_1 + \dots + c_p(\lambda_1 - \lambda_{p+1})\mathbf{x}_p = 0$ could be true is if $c_1 = \dots = c_p = 0$. Therefore, by Equation (1), $\mathbf{x}_{p+1} = 0$, a contradiction.

2. Hence, or otherwise, prove that for any matrix $\mathbf{B} \in \mathbb{R}^{n\times n}$, there can be at most n distinct eigenvalues for $\mathbf{B}$.

   Solution. Suppose not. Then $\mathbf{B}$ has m distinct eigenvalues $\{\lambda_1, \dots, \lambda_m\}$ with $m > n$. Let $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ denote the corresponding eigenvectors. Then $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ is linearly independant by the previous theorem. But that would mean we have m many linearly independent vectors in $\mathbb{R}^n$ with $m > n$, a contradiction.

Question 4 **Properties of Determinants** (10+15=25 credits)

1. Prove $\det(A^T) = \det(A)$.

   Solution. For a given square matrix A of larger dimension, the **minor** M_{ij} of entry a_{ij} is the determinant of the submatrix obtained by deleting the i^{th} row and the j^{th} column from A. Let $C_{ij} := (-1)^{i+j}M_{ij}$ denote the cofactor of a_{ij}. By choosing any row i or column j of A, the determinant is defined recursively, by cofactor expansion along that row or column.

$$\det A = \sum_k a_{kj}C_{kj} = \sum_k a_{ik}C_{ik}$$

   By the definition of the determinant, we get the same answer whether we cofactor expand along a row or a column, and cofactor expanding along the first row of A is equivalent to cofactor expanding along the first column of A^T, so it must be the case that

$$\det(A) = \det(A^T)$$

2. Prove $\det(I_n) = 1$ where I_n is the $n \times n$ identity matrix.

   Solution. Proof by induction. Trivial for $n = 1$. Assume true for some n. Then by cofactor expanding along the first column of I_{n+1},

$$\det I_{n+1} = \sum_{k=1}^{n+1} I_{1k}C_{1k} = I_{11}C_{11} + \sum_{k=2}^{n+1} I_{1k}C_{1k}$$

   All off diagonals terms of I_{n+1} are zero, all diagonal terms are one. Therefore,

$$\det I_{n+1} = 1 \times C_{11} = (-1)^{1+1}M_{11} = \det I_n = 1$$

   Hence result follows by induction.

Question 5 **Eigenvalues of symmetric matrices** (15 credits)

1. Let $\mathbf{A}$ be a symmetric matrix. Let $\mathbf{v}_1$ be an eigenvector of $\mathbf{A}$ with eigenvalue λ_1, and let $\mathbf{v}_2$ be an eigenvector of $\mathbf{A}$ with eigenvalue λ_2. Assume that $\lambda_1 \neq \lambda_2$. Prove that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal. (Hint: Try proving $\lambda_1 \mathbf{v}_1^T \mathbf{v}_2 = \lambda_2 \mathbf{v}_1^T \mathbf{v}_2$. Recall the identity $\mathbf{a}^T\mathbf{b} = \mathbf{b}^T\mathbf{a}$.)

   Solution.

$$
\begin{aligned}
&\lambda_1 \mathbf{v}_1^T \mathbf{v}_2 \\
&= \lambda_1 \mathbf{v}_2^T \mathbf{v}_1 \\
&= \mathbf{v}_2^T (\lambda_1 \mathbf{v}_1) \\
&= \mathbf{v}_2^T (\mathbf{A}\mathbf{v}_1) \\
&= (\mathbf{A}\mathbf{v}_1)^T \mathbf{v}_2 \\
&= (\mathbf{v}_1^T \mathbf{A}^T) \mathbf{v}_2 \\
&= \mathbf{v}_1^T (\mathbf{A}^T \mathbf{v}_2) \\
&= \mathbf{v}_1^T (\mathbf{A}\mathbf{v}_2) \\
&= \mathbf{v}_1^T (\lambda_2 \mathbf{v}_2) \\
&= \lambda_2 \mathbf{v}_1^T \mathbf{v}_2
\end{aligned}
$$

   Hence $\lambda_1 \mathbf{v}_1^T \mathbf{v}_2 = \lambda_2 \mathbf{v}_1^T \mathbf{v}_2$, and therefore

$$(\lambda_1 - \lambda_2)\mathbf{v}_1^T \mathbf{v}_2 = 0$$

   Since, $\lambda_1 \neq \lambda_2$, $\lambda_1 - \lambda_2 \neq 0$, and thus $\mathbf{v}_1^T \mathbf{v}_2 = 0$. Hence, $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal.

Question 6 **Computations with Eigenvalues** (3+3+3+3+3=15 credits)

Let $\mathbf{A} = \begin{bmatrix} -1 & 2 \\ 3 & 4 \end{bmatrix}$.

1. Compute the eigenvalues of $\mathbf{A}$.

   Solution. We set up the characteristic polynomial, and solve.

$$
\begin{aligned}
&\det(\mathbf{A} - \lambda \mathbf{I}) \\
&= \det \begin{bmatrix} -1-\lambda & 2 \\ 3 & 4-\lambda \end{bmatrix} \\
&= (-1-\lambda)(4-\lambda) - 6 \\
&= \lambda^2 - 3\lambda - 10 \\
&= (\lambda + 2)(\lambda - 5)
\end{aligned}
$$

   So $\lambda = -2$ or $\lambda = +5$.

2. Find the eigenspace E_λ for each eigenvalue λ. Write your answer as the span of a collection of vectors.

   Solution. For the case of $\lambda = -2$, we form the equation $\mathbf{A}\mathbf{x} = -2\mathbf{x}$, or $(\mathbf{A} + 2\mathbf{I})\mathbf{x} = \mathbf{0}$, and solve using gaussian elimination.

$$\left[\begin{array}{cc|c} -1+2 & 2 & 0 \\ 3 & 4+2 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 6 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 6 & 0 \end{array}\right]$$

So $x_1 = -2x_2$ and x_2 is free, so the eigenspace E_{-2} is given by

$$E_{-2} = \text{span}\left\{\begin{bmatrix}-2\\1\end{bmatrix}\right\}$$

For the case of $\lambda = 5$, we form the equation $\mathbf{Ax} = 5\mathbf{x}$, or $(\mathbf{A}-5\mathbf{I})\mathbf{x} = \mathbf{0}$, and solve using gaussian elimination.

$$\left[\begin{array}{cc|c}-1-5 & 2 & 0\\3 & 4-5 & 0\end{array}\right] \sim \left[\begin{array}{cc|c}-6 & 2 & 0\\3 & -1 & 0\end{array}\right] \sim \left[\begin{array}{cc|c}-3 & 1 & 0\\0 & 0 & 0\end{array}\right]$$

So x_2 free, and $x_1 = x_2/3$, so the eigenspace E_5 is given by

$$E_5 = \text{span}[\begin{bmatrix}1/3\\1\end{bmatrix}] = \text{span}[\begin{bmatrix}1\\3\end{bmatrix}]$$

3. Verify the set of all eigenvectors of $\mathbf{A}$ spans $\mathbb{R}^2$.

   Solution. The set of all eigenvectors is given by

$$\text{span}[\begin{bmatrix}-2\\1\end{bmatrix}, \begin{bmatrix}1\\3\end{bmatrix}]$$

   Clearly, the two vectors are linearly independent, as one is not a multiple of the other. Since we have two linearly independant vectors, they must span $\mathbb{R}^2$.

4. Hence, find an invertable matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$.

   Solution. We form $\mathbf{P}$ from two linearly independant eigenvectors (we can use the vectors that span $\mathbb{R}^2$ in 3) and we form $\mathbf{D}$ by a diagonal matrix, with the eigenvalues of $\mathbf{A}$ on the diagonal, in the same order as the eigenvectors in $\mathbf{P}$.

$$\mathbf{P} = \begin{bmatrix}-2 & 1\\1 & 3\end{bmatrix} \quad \mathbf{D} = \begin{bmatrix}-2 & 0\\0 & 5\end{bmatrix}$$

   We compute the inverse using the closed form equation for 2x2 matricies.

$$\mathbf{P}^{-1} = \frac{1}{7}\begin{bmatrix}-3 & 1\\1 & 2\end{bmatrix}$$

   We can then verify $\mathbf{A} = \mathbf{PDP}^{-1}$ holds via matrix multiplication.

5. Hence, find a formula for efficiently [1] calculating $\mathbf{A}^n$ for any integer $n \geq 0$. Make your formula as simple as possible.

   Solution. From the week 8? tutorials, we proved that if $\mathbf{A} = \mathbf{PDP}^{-1}$, then $\mathbf{A}^n = \mathbf{PD}^n\mathbf{P}^{-1}$. Since $\mathbf{D}$ is a diagonal matrix, we have

$$\mathbf{D}^n = \begin{bmatrix}(-2)^n & 0\\0 & (5)^n\end{bmatrix}$$

   Hence,

$$\mathbf{A} = \begin{bmatrix}-2 & 1\\1 & 3\end{bmatrix}\begin{bmatrix}(-2)^n & 0\\0 & (5)^n\end{bmatrix}\frac{1}{7}\begin{bmatrix}-3 & 1\\1 & 2\end{bmatrix} = \frac{1}{7}\left(\begin{bmatrix}5^n - 3(-2)^{n+1} & (-2)^{n+1} + 2\times 5^n\\3\times 5^n - 3(-2)^n & (-2)^n + 6\times 5^n\end{bmatrix}\right)$$

[1] That is, a closed form formula for $\mathbf{A}^n$ as opposed to multiplying $\mathbf{A}$ by itself n times over.