

COMP3670/6670: Introduction to Machine Learning

Question 1

Matrix Properties

1. Uniqueness of inverses

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$. Assume \mathbf{A} is invertible. Prove that the inverse of \mathbf{A} is unique, (that is, there is only one matrix \mathbf{B} that satisfies $\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$)

Solution. Assume not for contradiction. Then at least two inverses of \mathbf{A} must exist (as \mathbf{A} is invertible.) Let \mathbf{X} and \mathbf{Y} denote distinct inverses of \mathbf{A} . (i.e that $\mathbf{X} \neq \mathbf{Y}$). Then by definition,

$$\mathbf{XA} = \mathbf{AX} = \mathbf{I}$$

$$\mathbf{YA} = \mathbf{AY} = \mathbf{I}$$

So then

$$\mathbf{AY} = \mathbf{AX}$$

Left multiplying by any inverse of \mathbf{A} (we choose \mathbf{X}).

$$\mathbf{X}(\mathbf{AY}) = \mathbf{X}(\mathbf{AX})$$

$$(\mathbf{XA})\mathbf{Y} = (\mathbf{XA})\mathbf{X}$$

$$\mathbf{IY} = \mathbf{IX}$$

$$\mathbf{Y} = \mathbf{X}$$

which is a contradiction. Hence inverses are unique.

2. Inverse of an inverse

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$. Assume \mathbf{A} is invertible. Prove that $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

Solution. We need to find a matrix \mathbf{X} such that

$$\mathbf{XA}^{-1} = \mathbf{A}^{-1}\mathbf{X} = \mathbf{I}$$

Choose $\mathbf{X} = \mathbf{A}$. Note from the definition of the inverse, we have that

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

Hence by definition, the inverse of \mathbf{A}^{-1} is \mathbf{A} , and

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$$

3. Distributing the transpose

For $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times n}$, prove that $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$

Solution. We check the i, j th element, and verify they both match.

$$\begin{aligned} & (\mathbf{A} + \mathbf{B})_{i,j}^T \\ &= (\mathbf{A} + \mathbf{B})_{j,i} \\ &= \mathbf{A}_{j,i} + \mathbf{B}_{j,i} \\ &= \mathbf{A}_{i,j}^T + \mathbf{B}_{i,j}^T \\ &= (\mathbf{A}^T + \mathbf{B}^T)_{i,j} \end{aligned}$$

The above proof works as addition is performed elementwise.

4. Matrix Cancellation

Let \mathbf{A} , \mathbf{B} , \mathbf{C} all be square matrices of the same dimension. Assume $\mathbf{AB} = \mathbf{AC}$. Does it always follow that $\mathbf{B} = \mathbf{C}$?

Solution. If \mathbf{A} is invertible, then yes, as

$$\begin{aligned}\mathbf{A}^{-1}\mathbf{AB} &= \mathbf{A}^{-1}\mathbf{AC} \\ \mathbf{B} &= \mathbf{C}\end{aligned}$$

If \mathbf{A} isn't invertible, then it might not hold. (E.g. If \mathbf{A} was the zero matrix, then the equation would hold for any \mathbf{B} and \mathbf{C} .)

Question 2

Moore-Penrose Inverse

Assuming \mathbf{A} is invertible, prove that the Moore-Penrose inverse $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ equals \mathbf{A}^{-1} .

How does this show that the Moore-Penrose inverse is more general than the inverse?

Give an example of a matrix that does not have a Moore-Penrose inverse.

Solution.

$$(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{A}^{-1}(\mathbf{A}^T)^{-1}\mathbf{A}^T = \mathbf{A}^{-1}\mathbf{I} = \mathbf{A}^{-1}$$

This is more general than the inverse, as the Moore-Penrose inverse can be defined for non-square matrices, e.g.

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

The square zero matrix of any dimension \mathbf{Z} has no Moore-Penrose inverse, as $\mathbf{Z}^T\mathbf{Z} = \mathbf{Z}$, and thus $(\mathbf{Z}^T\mathbf{Z})^{-1}$ is undefined.

Question 3

Linear Equations

Prove that a system of linear equations $\mathbf{Ax} = \mathbf{b}$ either has no solutions, a unique solution or infinitely many solutions.

(Hint: If there are at least two solutions \mathbf{p} and \mathbf{q} , consider the vector $\mathbf{v}_\lambda = \lambda\mathbf{p} + (1 - \lambda)\mathbf{q}$.)

Solution. If $\mathbf{Ax} = \mathbf{b}$ has no solutions or a unique solution, we are done. So assume not. So there exists at least two distinct solutions \mathbf{p} and \mathbf{q} . So we have $\mathbf{Ap} = \mathbf{b}$ and $\mathbf{Aq} = \mathbf{b}$. For some $\lambda \in \mathbb{R}$, let

$$\mathbf{v}_\lambda = \lambda\mathbf{p} + (1 - \lambda)\mathbf{q}$$

Then,

$$\begin{aligned}\mathbf{Av}_\lambda &= \mathbf{A}(\lambda\mathbf{p} + (1 - \lambda)\mathbf{q}) \\ &= \lambda\mathbf{Ap} + (1 - \lambda)\mathbf{Aq} \\ &= \lambda\mathbf{b} + (1 - \lambda)\mathbf{b} \\ &= \mathbf{b}\end{aligned}$$

Hence \mathbf{v}_λ is a solution for any $\lambda \in \mathbb{R}$, and we have infinitely many solutions.

Question 4

Linear Independence

Let $\mathbf{T} \in \mathbb{R}^{n \times m}$ be a matrix. Let $\{\mathbf{u}, \mathbf{v}\}$ be a set of linearly independent vectors in $\mathbb{R}^{m \times 1}$. Assume that $\{\mathbf{Tu}, \mathbf{Tv}\}$ are linearly dependent. Prove there exists non-zero $\mathbf{x} \in \mathbb{R}^{m \times 1}$ such that $\mathbf{T}\mathbf{x} = \mathbf{0}$.

Solution. Linear dependence means there exists scalars c_1 and c_2 , at least one of them non-zero, such that

$$c_1\mathbf{Tu} + c_2\mathbf{Tv} = \mathbf{0}$$

Using the fact that matrix multiplication distributes over scalar multiplication and vector addition,

$$\mathbf{T}(c_1\mathbf{u} + c_2\mathbf{v}) = \mathbf{0}$$

Now since \mathbf{u} and \mathbf{v} are linearly independent, and it is not the case that both c_1 and c_2 are zero, it follows that $(c_1\mathbf{u} + c_2\mathbf{v}) \neq \mathbf{0}$, hence we have a non-zero solution to $\mathbf{T}\mathbf{x} = \mathbf{0}$.

Question 5

Combining vector subspaces

Let V be a vector space. Let $A \subseteq V$ and $B \subseteq V$ be vector subspaces of V .

1. Prove that $A \cap B$ is a vector subspace of V .

Solution. We need to check the two properties.

Let \mathbf{x}, \mathbf{y} be in $A \cap B$. Then \mathbf{x} and \mathbf{y} are in A , and so $\mathbf{x} + \mathbf{y} \in A$, since A is a vector subspace, and is closed under addition. By a similar argument, $\mathbf{a} + \mathbf{b} \in B$. Hence, $\mathbf{a} + \mathbf{b} \in A \cap B$, and the set is closed under addition. Let $\lambda \in \mathbb{R}, \mathbf{x} \in A \cap B$. Then $\mathbf{x} \in A$ and $\mathbf{x} \in B$. Since both A and B are vector subspaces, $\lambda\mathbf{x} \in A, \lambda\mathbf{x} \in B$. Thus $\lambda\mathbf{x} \in A \cap B$, and the set is closed under scalar multiplication.

2. (**Tricky**) Prove that $A \cup B$ is a vector subspace of V if and only if A is contained in B , or B is contained in A .

(This proof is easy in one direction, and tricky the other direction. As a hint, if the sets are not contained in each other, then there must lie a vector in $A \setminus B$ and in $B \setminus A$. Consider the sum of these vectors.)

Solution. Assume that one of A or B is contained in the other. If $A \subseteq B$, then $A \cup B = B$, and the result immediately follows, as B is a vector subspace. Similar argument for $B \subseteq A$.

Assume A is not contained in B , and vice versa. Assume for contradiction that $A \cup B$ is a vector subspace. So there must exist an element $\mathbf{a} \in A \setminus B$ and an element $\mathbf{b} \in B \setminus A$. Consider $\mathbf{a} + \mathbf{b}$. Since by assumption $A \cup B$ is a vector subspace, it must be closed under vector addition. So $\mathbf{a} + \mathbf{b}$ lies in A or B (or both.) If $\mathbf{a} + \mathbf{b}$ is in A , then note that $(-1)\mathbf{a} = -\mathbf{a}$ is also in A (by closure under scalar multiplication), but $(\mathbf{a} + \mathbf{b}) + (-\mathbf{a}) = \mathbf{b}$ is not in A , violating the property of closure under addition.

We can make the same argument if $\mathbf{a} + \mathbf{b}$ is in B .

In either case we get a contradiction, and $A \cup B$ cannot be a vector subspace.