Jiaming Lai
A53314574

Assignment #4

Problem 1: Moore–Penrose Pseudoinverse

Solution

(a) Suppose A^+ is not unique and there is A_1^+ and A_2^+. According to the difinition, for any matrix $A \in \mathbb{R}^{m\times n}$,

$$AA_1^+A = A$$
$$AA_2^+A = A$$
$$\Rightarrow (AA_1^+ - AA_2^+)A = A(A_1^+A - A_2^+A) = 0$$

Because the above equation is valid for any matrix $A \in \mathbb{R}^{m\times n}$, hence there must be $AA_1^+ = AA_2^+$ and $A_1^+A = A_2^+A$. So

$$A_1^+ = A_1^+AA_1^+ = A_1^+AA_2^+ = A_2^+AA_2^+ = A_2^+$$

Hence A^+ must be unique.

(b) Denote $(A^TA)^{-1}A^T$ as A^+. Because matrix A is tall, hence $n \le m$.

$$A^+A = (A^TA)^{-1}A^TA = (A^TA)^{-1}(A^TA) = I$$

So $(A^TA)^{-1}A^T$ a left inverse of matrix A.

$$AA^+A = A(A^+A) = AI = A$$
$$A^+AA^+ = (A^+A)A^+ = IA^+ = A^+$$

$A^+A = I$, so A^+A is symmetric. Meanwhile,

$$(AA^+)^T = (A^+)^TA^T = [(A^TA)^{-1}A^T]^TA^T = A(A^TA)^{-1}A^T = AA^+$$

So AA^+ is symmetric. In conclusion, $(A^TA)^{-1}A^T$ is the pseudoinverse and a left inverse of matrix A.

(c) Denote $A^T(AA^T)^{-1}$ as A^+.

$$AA^+ = AA^T(AA^T)^{-1} = (AA^T)(AA^T)^{-1} = I$$

So $A^T(AA^T)^{-1}$ is a right inverse of matrix A.

$$AA^+A = (AA^+)A = IA = A$$
$$A^+AA^+ = A^+(AA^+) = A^+I = A^+$$

$AA^+ = I$, so AA^+ is symmetric. Meanwhile,

$$(A^+A)^T = A^T(A^+)^T = A^T[A^T(AA^T)^{-1}]^T = A^T(AA^T)^{-1}A = A^+A$$

so A^+A is symmetric. In conclusion, $A^T(AA^T)^{-1}$ is the pseudoinverse and a right inverse of matrix A.

(d)

$$AA^{-1}A = IA = A \; and \; A^{-1}AA^{-1} = IA^{-1} = A^{-1}$$

Also $AA^{-1} = A^{-1}A = I$ is symmetric. So in conclusion, A^{-1} is the pseudoinverse of a full-rank square matrix A.

(e) For a projection matrix A, $A^2 = A$ and $A^T = A$. Hence

$$AAA = AA = A$$

Because $A^T = A$, so AA is symmetric. In conclusion, A is the pseudoinverse of itself for a projection matrix A.

(f)

$$A^T(A^+)^T A^T = [AA^+A]^T = A^T$$
$$(A^+)^T A^T (A^+)^T = [A^+AA^+]^T = (A^+)^T$$

Meanwhile,

$$[A^T(A^+)^T]^T = A^+A \;\Rightarrow\; symmetric$$
$$[(A^+)^T A^T]^T = AA^+ \;\Rightarrow\; symmetric$$

So in conclusion, $(A^T)^+ = (A^+)^T$.

(g) i.

$$AA^T[(A^+)^T A^+]AA^T = A(A^+A)^T A^+AA^T = A(A^+A)A^+AA^T = AA^+AA^T = AA^T$$

Hence $(AA^T)^+ = (A^+)^T A^+$.

ii.

$$A^TA[A^+(A^+)^T]A^TA = A^TAA^+(AA^+)^TA = A^TAA^+AA^+A = A^TA$$

Hence $(A^TA)^+ = A^+(A^+)^T$.

(h) i. For any $y \in R(A^+)$, there exit a vector x. S.T.

$$y = A^+x = A^+AA^+x = (A^+A)^T A^+x = A^T(A^+)^T A^+x$$

Hence $y \in R(A^T)$ and $R(A^+) \subset R(A^T)$. For any $y \in R(A^T)$, there exit a vector x. S.T.

$$y = A^Tx = A^T(A^T)^+A^Tx = A^T(A^+)^T A^Tx = (A^+A)^T A^Tx = A^+AA^Tx$$

Hence $y \in R(A^+)$ and $R(A^T) \subset R(A^+)$. In conclusion, $R(A^T) = R(A^+)$.

ii. For any $x \in N(A^T)$, $A^T x = 0$, hence

$$A^+x = A^+AA^+x = A^+(AA^+)^T x = A^+(A^+)^T A^T x = 0$$

So $N(A^T) \subset R(A^+)$. For any $x \in N(A^+)$, $A^+x = 0$, hence

$$A^T x = A^T(A^+)^T A^T x = A^T(AA^+)^T x = A^T AA^+ x = 0$$

So $N(A^+) \subset R(A^T)$. In conclusion, $N(A^T) = N(A^+)$.

(i) First, both AA^+ and A^+A are symmetric. Second,

$$P^2 = AA^+AA^+ = AA^+$$

$$Q^2 = A^+AA^+A = A^+A$$

Hence P and Q are projection matrix.

(j) Recall the result of problem 5 in Homework 3, $y = Px$ is the projection of x onto $R(P)$. For $\forall y \in R(P)$, there must exists $x \in \mathbb{R}^m$ S.T.

$$y = AA^+x \;\Rightarrow\; y = A(A^+x)$$

Hence for $\forall y \in R(P)$, there must exists $z = A^+x \in \mathbb{R}^n$, S.T. $y = Az$. Hence $R(P) = R(A)$. So $y = Px$ is the projection of x onto $R(A)$.

Similarly, $y = Qx$ is the projection of x onto $R(Q)$. For $\forall y \in R(Q)$, there must exists $x \in \mathbb{R}^n$ S.T.

$$y = A^+Ax \;\Rightarrow\; y = A^+(Ax)$$

Hence for $\forall y \in R(Q)$, there must exists $z = Ax \in \mathbb{R}^m$, S.T. $y = A^+z$. Hence $R(Q) = R(A^+) = R(A^T)$. So $y = Qx$ is the projection of x onto $R(A^T)$.

(k) Recall the result in problem (j), the projection matrix onto $R(A)$ is $P = AA^+$,

$$Ax^* = AA^+b$$

Hence Ax^* is the orthogonal projection of b onto $R(A)$. Hence $x^* = A^+b$ is a least-squares solution.

(l) It is clear that $x^* = A^+b = A^+Ax$ is the orthogonal projection of x onto $R(A)$. Hence $(x - x^*) \perp x^*$, hence $x^* = A^+b$ is the least norm solution.

Problem 2: Eigenvalues

Solution

(a) The characteristic polynomial of A is

$$p(\lambda) = det(\lambda I - A) = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n)$$

Hence

$$p(\lambda = 0) = det(-A) = (-1)^n det(A) = (-1)^n \lambda_1 \lambda_2 \ldots \lambda_n$$

So $det(A) = \lambda_1 \lambda_2 \ldots \lambda_n$.

(b) Because

$$\lambda I - A^T = (\lambda I - A)^T$$

and

$$det((\lambda I - A)^T) = det(\lambda I - A)$$

So $det(\lambda I - A) = det(\lambda I - A^T)$. A^T and A have the same characteristic polynomial. Hence the eigenvalues of A^T and A are the same.

(c) Give the fact that $Av = \lambda_i v$, where $i = 1, 2, \ldots, n$,

$$A^k v = A^{k-1}\lambda_i v = A^{k-2}\lambda_i^2 v = \ldots = \lambda_i^k v$$

Hence $\lambda_i^k,\ i = 1, 2, \ldots, n$ are eigenvalues of matrix A^k.

(d) If matrix A is invertible, then suppose matrix A has a zero eigenvalue λ. There must be a vector $v \neq 0$, S.T.

$$Av = \lambda v = 0$$

Obviously $v \in N(A)$. Because A is invertible, so $dim(N(A) = 0) \Rightarrow N(A) = 0$. Hence $v = 0$. But this contradicts the fact that $v \neq 0$. So if A is invertible, it does not have a zero eigenvalue.

If matrix A does not have a zero eigenvalue, then there is no a vector $v \neq 0$ S.T.

$$Av = \lambda v,\ \lambda = 0 \ \Rightarrow \ Av = 0$$

The only solution to $Av = 0$ is $v = 0$, which means $dim[N(A)] = 0$. Hence matrix A is full-rank and invertible.

In conclusion, A is invertible if and only if it does not have a zero eigenvalue.

(e) Accoding to the definition,

$$Av = \lambda_i v \;\Rightarrow\; A^{-1}Av = \lambda_i A^{-1}v \;\Rightarrow\; v = \lambda_i A^{-1}v \;\Rightarrow\; \lambda_i^{-1}v = A^{-1}v$$

Hence $\lambda_i^{-1},\ i = 1, 2, \ldots, n$ are eigenvalues of A^{-1}.

(f) The characteristic polynomial of $T^{-1}AT$ is

$$\begin{aligned} det(T^{-1}AT - \lambda I) &= det(T^{-1}AT - \lambda T^{-1}IT) \\ &= det[T^{-1}(A - \lambda I)T] \\ &= det(T^{-1})det(A - \lambda I)det(T) \end{aligned}$$

Because $det(T^{-1})det(T) = det(T^{-1}T) = 1$. Hence

$$det(T^{-1}AT - \lambda I) = det(A - \lambda I)$$

So A and $T^{-1}AT$ have the same eigenvalues.

Problem 3: Trace

Solution

(a) The characteristic polynomial of A is

$$p(\lambda) = det(\lambda I - A) = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n)$$

Then the coefficients of λ^{n-1} is the sum of eigenvalues.

Consider the computation process of $det(\lambda I - A)$, the only term that contains λ^{n-1} is

$$\sigma(1, 2, 3, \ldots, n)(\lambda I - A)_{11}(\lambda I - A)_{22}\ldots(\lambda I - A)_{nn}$$

If we expand the equation above, it is easy to find that coefficients of λ^{n-1} is the sum of diagonal entries. Hence

$$tr(A) = \sum_{i=1} \lambda_i$$

(b) Using the result from problem 2(c), $\lambda_i^k,\ i = 1, 2, \ldots, n$ are eigenvalues of matrix A^k. Hence

$$tr(A^k) = \sum_{i=1}^{n} \lambda_i^k$$

Problem 4: More on Eigenvalues

Solution

Use the Schwarz Triangularization Theorem, for the square matrix A, A could be trangularized by an unitary matrix:

$$A = UTU^H \ , where\ UU^H = U^H U = I$$

$$\Rightarrow ||A||_F = ||UTU^H||_F$$

Hence $||A||_F = \sum_{i=1}^n \sum_{j=1}^n |A_{ij}|^2 = ||UTU^H||_F$.

As for the $||UTU^H||_F$, $||UTU^H||_F = tr[(UTU^H)^H UTU^H] = tr(UT^H TU^H)$,

$$tr(UT^H TU^H) = \sum_i \sum_j \sum_k U_{ij} (T^H T)_{jk} U^H_{ki} = \sum_j \sum_k \sum_i (T^H T)_{jk} U^H_{ki} U_{ij} = tr(T^H TU^H U) = tr(T^H T)$$

$$tr(T^H T) = \sum_i \sum_j T^H_{ij} T_{ji} \geq \sum_i T^H_{ii} T_{ii}$$

Because the diagonal elements of T is the eigenvalues of A, hence $\sum_i T^H_{ii} T_{ii} = \sum_{i=1}^n |\lambda_i|^2$. In conclusion, $\sum_{i=1}^n \sum_{j=1}^n |A_{ij}|^2 = ||A||_F = ||UTU^H||_F \geq \sum_{i=1}^n |\lambda_i|^2$.

Problem 5: Limit

Solution

$$det(\lambda I - A) = \lambda^2 - 1.2\lambda + 0.2 = (\lambda - 1)(\lambda - 0.2)$$

Hence the eigenvalues of matrix A is $\lambda_1 = 0.2$ and $\lambda_2 = 1$.

$$A - \lambda_1 I = \begin{bmatrix} 4 & -1.6 \\ 12 & -4.8 \end{bmatrix} \Rightarrow N(A - \lambda_1 I) = span \left\{ \begin{bmatrix} 0.4 \\ 1 \end{bmatrix} \right\}$$

$$A - \lambda_0.2I = \begin{bmatrix} 4.8 & -1.6 \\ 12 & -4 \end{bmatrix} \Rightarrow N(A - \lambda_1 I) = span \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$$

$$\Rightarrow A = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P^{-1}$$

where

$$P = \begin{bmatrix} 0.4 & 1 \\ 1 & 3 \end{bmatrix}$$

Hence

$$\lim_{n \to \infty} A^n = P \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} P^{-1} = P \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = \begin{bmatrix} -5 & 2 \\ -15 & 6 \end{bmatrix}$$