ELEC 405 - Homework 2

Gul Sena Altintas - 64284

November 22, 2020

Contents

1. Nearest Neighbor partitioning and halfspaces

$$\{x \mid \|x-a\| \leq \|x-b\|\} = \{x \mid c^T x \leq d\},\ c \in \mathcal{R}^n, d \in \mathcal{R}^n \tag{1}$$

Take the square of 1, as $\|.\| \geq 0$, then

$$\begin{aligned} \|x-a\|^2 &\leq \|x-b\|^2 \\ < x-a, x-a > &\leq < x-b, x-b > \\ \|x\|^2 - 2a^T x + \|a\|^2 &\leq \|x\|^2 - 2b^T x + \|b\|^2 \\ 2(b-a)^T x &\leq \|b\|^2 - \|a\|^2 \end{aligned} \tag{2}$$

Then $c = 2(b-a)$ and $d = \|b\|^2 - \|a\|^2$.

Let $a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$c = 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$

$d = \|b\|^2 - \|a\|^2 = 2 - 1 = 1$

the hyperplane = $\left\{ x \mid (0 \ \ 2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq 1 \right\}$

$0x_1 + 2x_2 \leq 1$

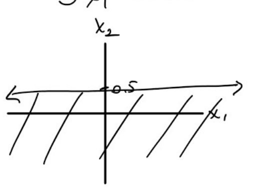

Figure 1: Hyperplane

2. Rank and matrix products

For the sake of this question let $A \in \mathcal{R}^{m\times n}$ and $B \in \mathcal{R}^{n\times p}$

1. If AB is full rank then A and B are full rank.
   Not necessarily. Let $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. $AB = \begin{bmatrix} 1 & 1 \end{bmatrix}$. AB is full rank but B is not.
2. If A and B are full rank then AB is full rank.
   Not necessarily. Let $A = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. $AB = \begin{bmatrix} 0 \end{bmatrix}$. A and B are full rank but AB is not.
3. If A and B have zero nullspace, then so does AB.
   Assume $\mathcal{N}(AB) \neq \{0\}$ and let $x \in \mathcal{R}^p$ such that $ABx = 0$ then $A(Bx) = 0$ so either $Bx \in \mathcal{R}^n$ is in null space of A but null space of A is trivial. So $Bx = 0$, then $x \in \mathcal{N}(B)$ but since null space of B is also trivial, x must be 0. Hence AB has zero nullspace.
4. If A and B are onto, then so is AB.
   We need to show that $\forall y \in \mathcal{R}^m \ \exists x \in \mathcal{R}^p$ such that $ABx = y$. Since A is onto, $\exists z \in \mathcal{R}^n$ such that $Az = y$. Similarly, since B is onto, $\exists x \in \mathcal{R}^p$ such that $Bx = z$. Hence $A(Bx) = Az = y$.

3. Fundamental Subspaces and Determinant

Let $A \in \mathcal{R}^{m\times n}$

1. If $\mathcal{R}(A) = R^m$, then AA^T is full rank.
   If $\mathcal{R}(A) = \mathcal{R}^m$ then A^T is full rank and its null space is trivial ($\{0\}$). Let $x \in \mathcal{R}^m$ such that $(AA^T)x = 0$, multiply each side with $x^T \Rightarrow x^TAA^Tx = (A^Tx)^T(A^Tx) =< A^Tx, A^Tx >= \|A^Tx\|^2$, since A^T is full rank we know that $x = 0$ which proves that the null space of AA^T is trivial and the matrix is full rank.
2. If $\mathcal{N}(A) = \{0\}$ then A^TA is full rank.
   For A^TA to be full rank, we need to show that $\dim(\mathcal{R}(A^TA)) = n$. In other words, if we show that the null space of A^TA is trivial then we prove that it is full rank. Let $x \in \mathcal{R}^n$, such that $(A^TA)x = 0$ is we multiply both sides with x^T then we get $x^TA^TAx = (Ax)^T(Ax) =< Ax, Ax >= \|Ax\|^2$ and we know that $\|y\| = 0$ if and only if $y = 0$ since A is full rank therefore A^TA is also full rank.

4. Characteristic Polynomial of a Square Matrix

Let $A \in \mathcal{R}^{n\times n}$ and $X(s) = det(sI - A)$ be the charateristic polynomial of A.

1. When we think about how we take the determinant, we can construct the result of the determinant as follows:
   First take the first row and column into account so that A' is A without the first row and column. Then
   $$det(sI - A) = (s - a_{11})detA' + C$$
   where C is a constant. Iteratively, this will yield a similar result so that we can write $det(sI - A)$ of the form:
   $$det(sI - A) = \Pi_{i=1}^{n}(s - a_{ii}) + C^* \tag{3}$$
   Therefore, the coefficient of s^n is 1.

2. Show that s^{n-1} coefficient of $X(s)$ is given by $-TrA$.
   Note that in (3) C is a polynomial of degree at most $n-1$ because A' is a $(n-1)\times(n-1)$ matrix. Similarly, A' is the only matrix with $n-1\times n-1$ elements, so the coefficient of s^{n-1} comes from $\Pi_{i=1}^{n}(s-a_{ii})$ which is $\sum_{i=1}^{n}-a_{ii}=TrA$.

3. Show that the constant coefficient of $X(s)$ is given by $det(-A)$.
   Like in any polynomial, we can take $X(0)$ to find the constant coefficient and find $det(0I-A)=det(-A)$.

4. Consider $n=1$, then $X(s)=det(sI-A)=s-\lambda_1$, and the relationship holds. Now assume that the relationship holds for some positive integer k such that $a_{k-1}=-\sum_{i=1}^{k}\lambda_i$ and $a_0 =_{i=1}^{n} -\lambda_i$. Multiply $X(s)$ with $(s-\lambda_{k+1})$ which will give us $sX(s)-\lambda_{k+1}X(s)$. By part (1) $sX(s)$ is monic and of degree $k+1$. The coefficient of s^k is $\sum_{i=1}^{k}(-\lambda_i)-\lambda_k$ and the constant coefficient is $\Pi_{i=1}^{k+1}-\lambda_i$, note that the constant coefficient of $sX(s)$ is 0. Therefore, the relationship holds.

5. Schur's Theorem and Cayley-Hamilton Theorem

1. By Schur's Theorem, we can construct T such that $T=U^*AU$ where U is described in the theorem. $p(t)=\Pi_{i=1}^{n}(t-\lambda_i)$ is the characteristic polynomial of T. Since λ_i are at the diagonals of T, note that λ_i's don't need to be unique, $P(T)=\Pi_{i=1}^{n}(T-\lambda_iI)=0$. Evaluate

$$\begin{aligned} U^*P(A)U &= U^*(A-\lambda_1I)\dots(A-\lambda_n)U \\ &= U^*(A-\lambda_1I)UU^*\dots UU^*(A-\lambda_n)U \\ &= (U^*AI-\lambda_1U^*U)\dots(U^*AI-U^*\lambda_nU^*U) \text{ note } U^*U=I \\ &= (T-\lambda_1I)\dots(T-\lambda_nI) \\ &= P(T) \end{aligned} \tag{4}$$

   Reorganizing the equality in (4) yields

$$P(A)=UP(T)U^*(A)=0$$

   .

2. Evaluating $p_A(A)=A^n+a_{n-1}A^{n-1}+\dots a_1A+a_0=0$. Then

$$A^n = -a_{n-1}A^{n-1}+\dots-a_1A-a_0 \tag{5}$$

$$\begin{aligned} A^{n+1} &= A(-a_{n-1}A^{n-1}+\dots-a_1A-a_0) \\ &= -a_{n-1}A^n+\dots+-a_1A^2-a_0A \qquad (6) \\ &= (a_{n-1}^2+a_{n-1}a_{n-2})A^{n-1}+\dots+(a_2a_{n-1}-a_1)A^2+(a_1a_{n-1}-a_0)A+a_{n-1}a_0 \qquad (7) \\ A^{n+2} &= A(a_{n-1}^2A^{n-1}+\dots+(a_2a_{n-1}-a_1)A^2+(a_1a_{n-1}-a_0)A+a_{n-1}a_0) \\ &= a_{n-1}^2A^n+\dots+(a_2a_{n-1}-a_1)A^3+(a_1a_{n-1}-a_0)A^2+(a_{n-1}a_0A) \qquad (8) \\ &= -a_{n-1}^3A^{n-1}+\dots+(-a_2a_{n-1}^2+a_1a_{n-1}-a_0)A^2+(-a_{n-1}^2a_1+a_{n-1}a_0)A-a_{n-1}^2a_0 \end{aligned}$$

   (7) follows from plugging (5) into (6). Formally in the expansion of A^k where $k\geq n$ the coefficient of A^{n-1} is $(-a_{n-1})^{k-n+1}$ and the coefficient of A^i is $\sum_{j=1}^{i+1}(-1)^{j-1}a_{n-1}^{j-1}a_{j-1}$

6. Theory in Action for Big Data!: Eigenvectors as community detectors

1. $E(C)\in\mathcal{R}^{2N\times2N}$ where E is of the form:

$$E(C)_{ij}=\begin{cases} p \text{ if both i,j } \in\mathcal{I}_A \text{ or } \mathcal{I}_B \\ q \text{ if i } \in\mathcal{I}_A \text{ and j } \in\mathcal{I}_B \text{ or vice versa} \\ 0, \text{ if } C_{ij}=0 \end{cases} \tag{9}$$

2. $2\times$ rank C