CS560 Statistical Machine Learning: Homework 0

Brendan von Hofe

January 2019

Problem 1: Let $\boldsymbol{x} \in \mathbb{R}^d$. Show that if for all d-dimensional vectors $\boldsymbol{v}, \langle \boldsymbol{x}, \boldsymbol{v} \rangle = 0$, then $\boldsymbol{x}$ is a zero vector.

Proof. We assume $\boldsymbol{x}$ to be any vector that is not the zero vector. Therefore there exists at least one element in $\boldsymbol{x}$ that is non-zero. Let $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_d \end{bmatrix}$ where $\exists x_i \in x(x_i \neq 0)$ and $\boldsymbol{v} = \begin{bmatrix} 1 \\ 1 \\ ... \\ 1 \end{bmatrix}$. Therefore $\langle \boldsymbol{x}, \boldsymbol{v} \rangle = \sum_{i=1}^{d} x_i \neq 0$. By contradiction, $\boldsymbol{x}$ must be the zero-vector. ■

Problem 2: Let $\boldsymbol{A} \in \mathbb{R}^{dxd}$ be a square matrix. Show that if for all square matrices $\boldsymbol{B} \in \mathbb{R}^{dxd}$, $\boldsymbol{AB} = \boldsymbol{BA}$, then $\boldsymbol{A}$ is diagonal.

Proof. Let $n \in \mathbb{Z}$ and $A = nI_d$. Thus $A_{ij} = n$ when $i = j$ and 0 otherwise. For the commutative property to hold, $c_{ij} = \sum_{k=1}^{d} a_{ik}b_{kj} = \sum_{k=1}^{d} b_{ik}a_{kj}$. Since $a_{ik} = 0$ when $i \neq k$ and $a_{kj} = 0$ when $j \neq k$, $c_{ij} = a_{ii}b_{ij} = b_{ij}a_{ii} = nB_{ij}$. ■

Problem 3: Show that $\frac{1}{2}(e^x + e^{-x}) \leq e^{\frac{x^2}{2}}$ for all $x \in \mathbb{R}$, where e is the natural logarithm.

$\frac{1}{2}(e^x + e^{-x}) \leq e^{\frac{x^2}{2}}$

$1 + \frac{1}{2}x^2 + \frac{1}{24}x^4 + ... \leq 1 + \frac{1}{2}x^2 + \frac{3}{24}x^4 + ...$ (Taylor expansion of both sides around 0)

Both approximated polynomials have identical form besides the magnitude of the coefficients. The inequality is confirmed by the greater coefficients on the right-hand side.

Problem 4: Show that for all $x > 0$, $\log(1 + x) \leq x - \frac{x^2}{2} + \frac{x^3}{3}$

Proof. Since both sides evaluate to 0 at $x = 0$, we compare the first derivatives and show that the right-hand side is increasing faster over all $x > 0$. Replacing the expressions with their derivatives we get $\frac{1}{1+x} \leq 1 - x + x^2$. After algebraic manipulation, we have $1 \leq 1 + x^3$ which is vacuously true for $x > 0$. Therefore, the original inequality holds. ■

Problem 5: Let $\boldsymbol{x} \in \mathbb{R}^d$. Show that $||\boldsymbol{x}||_2 \leq ||\boldsymbol{x}||_1 \leq \sqrt{d}||\boldsymbol{x}||_2$.

Proof. We begin by proving $||\boldsymbol{x}||_2 \leq ||\boldsymbol{x}||_1$. By definition $\sqrt{\sum_{i=1}^d x_i^2} \leq \sum_{i=1}^d |x_i|$. Squaring both sides we get $\sum_{i=1}^d x_i^2 \leq (\sum_{i=1}^d |x_i|)^2$, which equals $\sum_{i=1}^d x_i x_i \leq \sum_{i=1}^d \sum_{j=1}^d |x_i||x_j|$. Therefore, after subtraction we get $0 \leq \sum_{i=1}^d \sum_{j=1}^d |x_i||x_j|$ where $i \neq j$. Second part is to come ■