Assignment 2
COMP3670

September 20, 2020

name: Xuecheng Zhang
UID:u6284513

31/8/2020

Exercise 1:

1.As **0** is a zero vector in V, also because of X is subspace of V, which means **0** $\in$ X. As $\forall$ x $\in$ X <x,**0**> = <x,-u + u> = -<x,u> + <x,u> = 0, then **0** $\in X^T$. It implies **0** $\in X \cap X^T$

Assume that $\alpha \in X \cap X^T \backslash \{0\}$. As $\alpha \in X \cap X^T$, it is true that $<\alpha,\alpha> = 0$, $\alpha \notin \{0\}$. However, as in the inner product operation, $<\alpha,\alpha>$ not equal to zero unless $\alpha = \mathbf{0}$. Therefore, it can only contain one element in the intersection of X and X^T, which is **0**

2. Assume $v \in Y^T$. Then for all x $\in Y$, $< x, v > = 0$. Because of X $\subseteq$ Y, for all x $\in X$, $< x, v > = 0$, which implies that $v \in X^T$. Therefore, Y^T is the subset of $X^T \implies Y^T \subseteq X^T$.

Exercise 2:

1.

$$< v - \frac{< v, u >}{< u, u >} u, u > = < v, u > - < \frac{< v, u >}{< u, u >} u, u > \tag{1}$$

$$= < v, u > - \frac{< v, u >}{< u, u >} < u, u > = < v, u > - < v, u > = 0 \tag{2}$$

Therefore, $v - proj_u(v)$ and u are orthogonal.

2.

For absolutely homogeneous,

$$||\lambda x|| = \sqrt{< \lambda x, \lambda x >} = \sqrt{\lambda^2 < x, x >} = ||\lambda||\sqrt{< x, x >} = |\lambda|||x|| \tag{3}$$

For positive definite, as in inner product, if and only x = **0**, the inner product $< x, x > = 0$ and if x $\neq$ **0**, the inner product $< x, x > > 0$. Thus, it satisfies positive definite property.

For triangle inequality,

$$||x + y|| = \sqrt{< x + y, x + y >} = \sqrt{< x, x > + < x, y > + < y, x > + < y, y >} \tag{4}$$

$$\leq \sqrt{< x, x > + < y, y > + 2||x||||y||} \tag{5}$$

$$= \sqrt{< x, x > + < y, y > + 2\sqrt{< x, x >< y, y >}} \tag{6}$$

$$= \sqrt{(\sqrt{< x, x >} + \sqrt{< y, y >})^2} = \sqrt{< x, x >} + \sqrt{< y, y >} = ||x|| + ||y|| \tag{7}$$

Exercise 3:

a)To compute the gradient df/dx,we first determine the dimension of df/dx: Since f: $\mathbb{R}^n \implies \mathbb{R}^1$, it follows that $df/dx \in \mathbb{R}^{1*n}$

f(x) = $\sum_{i=1}^{i=N} c_i x_i \implies$

$$\frac{\partial f(x)}{\partial x} = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} & \frac{\partial f(x)}{\partial x_2} & \cdots & \frac{\partial f(x)}{\partial x_N} \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & \cdots & c_N \end{pmatrix} = c^T \tag{8}$$

b)To compute the gradient dg/dx,we first determine the dimension of dg/dx: Since g: $\mathbb{R}^n \implies \mathbb{R}^1$, it follows that $dg/dx \in \mathbb{R}^{1*n}$
let f(x) $= c^T x + \mu^2$

$$\frac{\partial g(x)}{\partial x} = \frac{\partial g}{\partial f}\frac{\partial f}{\partial x} = \frac{1}{2\sqrt{c^T x + \mu^2}} \cdot \frac{\partial f}{\partial x} \tag{9}$$

$$= \frac{1}{2\sqrt{c^T x + \mu^2}} \cdot \begin{pmatrix} \frac{\partial \sum_{i=1}^{N} x_i c_i + \mu^2}{\partial x_1} & \frac{\partial \sum_{i=1}^{N} x_i c_i + \mu^2}{\partial x_2} & \cdots & \frac{\partial \sum_{i=1}^{N} x_i c_i + \mu^2}{\partial x_N} \end{pmatrix} \tag{10}$$

$$= \frac{1}{2\sqrt{c^T x + \mu^2}} \cdot \begin{pmatrix} c_1 & c_2 & \cdots & c_N \end{pmatrix} = \frac{c^T}{2\sqrt{c^T x + \mu^2}} \tag{11}$$

2. To compute the gradient dl/dx,we first determine the dimension of dl/dx: Since l: $\mathbb{R}^n \implies \mathbb{R}^1$, it follows that $dl/dx \in \mathbb{R}^{1*n}$

Let $l_1(x) = ||Ax - b||_2^2$ and $l_2(x) = \lambda ||x||_2^2$

$$l_1(x)_i = ||Ax - b||_2^2 = \sqrt{\sum_{j=1}^{N} (A_{ij} x_j - b_i)^2}^{\,2} \tag{12}$$

$$= \sum_{j=1}^{N} (A_{ij} x_j - b_i)^2 \tag{13}$$

Let f(x) $= A_{ij} x_j - b_i$

We can derive that:

$$\frac{l_1(x)_i}{dx_j} = \frac{\partial l}{\partial f} \cdot \frac{\partial f}{\partial x} = 2\sum_{j=1}^{N} (A_{ij} x_j - b_i) \cdot \frac{\partial f}{\partial x} \tag{14}$$

$$= 2\sum_{j=1}^{N} (A_{ij} x_j - b_i) \cdot \begin{pmatrix} \frac{\partial (A_{i1} x_1 - b_i)}{\partial x_1} & \frac{\partial (A_{i2} x_2 - b_i)}{\partial x_2} & \cdots & \frac{\partial (A_{iN} x_N - b_i)}{\partial x_N} \end{pmatrix} \tag{15}$$

$$= 2\sum_{j=1}^{N} A_{ij} x_j \cdot A_{ij} - 2\sum_{j=1}^{N} b_i \cdot A_{ij} = 2\sum_{j=1}^{N} (A_{ij} x_j) \cdot A_{ij} - 2\sum_{j=1}^{N} b_i \cdot A_{ij} \tag{16}$$

Therefore, $l_1(x) = 2(Ax)^T A - 2b^T A = 2(x^T A^T A - b^T A)$

$$l_2(x) = \lambda ||x||_2^2 = \lambda \sqrt{\sum_{i=1}^{N} (x_i - 0)^2}^{\,2} = \lambda \sum_{i=1}^{N} x_i^2 \tag{17}$$

We can derive that:

$$\frac{l_2(x)}{dx_j} = \begin{pmatrix} \frac{\partial \lambda \sum_{i=1}^{N} x_i^2}{\partial x_1} & \frac{\partial \lambda \sum_{i=1}^{N} x_i^2}{\partial x_2} & \cdots & \frac{\partial \lambda \sum_{i=1}^{N} x_i^2}{\partial x_N} \end{pmatrix} \tag{18}$$

$$= \begin{pmatrix} 2\lambda x_1 & 2\lambda x_2 & \cdots & 2\lambda x_N \end{pmatrix} = 2\lambda x^T \tag{19}$$

Therefore, the formula is proved