IE 411 HW # 1

Efe Eren Ceyani
21903359

Notation

$x \to$ scalar, $\underline{x} \to$ vector, $X \to$ matrix

Q1 Recall norm properties:

1-) $\|\underline{x}\| \geq 0, \ \forall \underline{x} \in \mathbb{R}^n$

2-) $\|\underline{x}\| = 0 \iff \underline{x} = \underline{0}$

3-) $\|\lambda \underline{x}\| = |\lambda| \|\underline{x}\|, \forall \underline{x} \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$

4-) $\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|, \ \forall \underline{x}, \underline{y} \in \mathbb{R}^n$

For $p = \frac{1}{2}$, p-norm is given by

$\|\underline{x}\|_{\frac{1}{2}} = \left(\sum_{i=1}^{n} \sqrt{|x_i|} \right)^2$, where x_i is the ith element of $\underline{x}$.

Consider $\underline{x}, \underline{y} \in \mathbb{R}^n$, where $\underline{x} = \underline{e_1}$, and $\underline{y} = \underline{e_2}$, i.e.,

$$\underline{x} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \underline{y} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\|\underline{x}\|_{\frac{1}{2}} = \left(\sum_{i=1}^{n} \sqrt{|x_i|} \right)^2 = 1, \ \|\underline{y}\|_{\frac{1}{2}} = \left(\sum_{i=1}^{n} \sqrt{|y_i|} \right)^2 = 1$

$\|\underline{x} + \underline{y}\|_{\frac{1}{2}} = \left(\sum_{i=1}^{n} \sqrt{|x_i + y_i|} \right)^2 = 4$

$\implies \|\underline{x} + \underline{y}\|_{\frac{1}{2}} > \|\underline{x}\| + \|\underline{y}\|$, which contradicts property 4.

Hence, for $p = \frac{1}{2}$, p-norm is not a norm.

Q2 CS inequality: $|\underline{x}^T\underline{y}| \le \|\underline{x}\|_2\|\underline{y}\|_2, \forall \underline{x}, \underline{y} \in \mathbb{R}$

a) $\underline{x} = \underline{0}$

$$|\underline{x}^T\underline{y}| = \left|\sum_{i=1}^{n} 0 \cdot y_i\right| = |0| = 0$$

$$\|\underline{x}\|_2\|\underline{y}\|_2 = \|\underline{0}\|_2\|\underline{y}\|_2 = 0\|\underline{y}\|_2 = \|0\,\underline{y}\|_2 = \|\underline{0}\|_2 = 0$$

Hence, CS inequality is satisfied. Also, since $\underline{x} = 0\underline{y}$, where $k=0$, the inequality becomes an equality.

b) $$\frac{1}{\|\underline{x}\|_2^2}\left\| \|\underline{x}\|_2^2\underline{y} - (\underline{x}^T\underline{y})\underline{x}\right\|_2^2 = A$$

Recall $\|\underline{x}\|_2 = \sqrt{\langle \underline{x}, \underline{x}\rangle_2}$ since ℓ_2 norm is an inner product induced norm.

$$A = \frac{1}{\|\underline{x}\|_2^2}\left\langle \|\underline{x}\|_2^2\underline{y} - (\underline{x}^T\underline{y})\underline{x}, \|\underline{x}\|_2^2\underline{y} - (\underline{x}^T\underline{y})\underline{x}\right\rangle_2$$

Recall inner product properties:
- $\langle \underline{x}, \underline{y}\rangle = \langle \underline{y}, \underline{x}\rangle^*$
- $\langle \alpha\underline{x}_1 + \beta\underline{x}_2, \underline{y}\rangle = \alpha\langle \underline{x}_1, \underline{y}\rangle + \beta\langle \underline{x}_2, \underline{y}\rangle$
- $\langle \underline{x}, \underline{x}\rangle > 0$, $\langle \underline{x}, \underline{x}\rangle = 0 \Leftrightarrow \underline{x} = \underline{0}$

for $\forall \underline{x}, \underline{x}_1, \underline{x}_2, \underline{y} \in \mathbb{C}^n$, $\alpha, \beta \in \mathbb{C}$

This doesn't matter for $\mathbb{R}$

Linear in first element

Example: $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$ but $\langle x, \lambda y\rangle = \langle \lambda y, x\rangle^* = \lambda^*\langle y, x\rangle^* = \lambda^*\langle x, y\rangle$

Second property

$$= \frac{1}{\|\underline{x}\|_2^2}\left(\langle \|\underline{x}\|_2^2\underline{y}, \|\underline{x}\|_2^2\underline{y} - (\underline{x}^T\underline{y})\underline{x}\rangle_2 - \langle (\underline{x}^T\underline{y})\underline{x}, \|\underline{x}\|_2^2\underline{y} - (\underline{x}^T\underline{y})\underline{x}\rangle_2\right)$$

First & second

$$= \frac{1}{\|\underline{x}\|_2^2}\left(\langle \|\underline{x}\|_2^2\underline{y}, \|\underline{x}\|_2^2\underline{y}\rangle_2 - \langle \|\underline{x}\|_2^2\underline{y}, (\underline{x}^T\underline{y})\underline{x}\rangle_2 - \langle (\underline{x}^T\underline{y})\underline{x}, \|\underline{x}\|_2^2\underline{y}\rangle_2 + \langle (\underline{x}^T\underline{y})\underline{x}, (\underline{x}^T\underline{y})\underline{x}\rangle_2\right)$$

First & second & $\underline{x}, \underline{y} \in \mathbb{R}^n$

$$= \frac{1}{\|\underline{x}\|_2^2}\left(\|\underline{x}\|_2^4\|\underline{y}\|_2^2 - \|\underline{x}\|_2^2(\underline{x}^T\underline{y})(\underline{y}^T\underline{x}) - (\underline{x}^T\underline{y})\|\underline{x}\|_2^2(\underline{x}^T\underline{y}) + (\underline{x}^T\underline{y})^2\|\underline{x}\|_2^2\right)$$

$$= \|\underline{x}\|_2^2\|\underline{y}\|_2^2 - (\underline{x}^T\underline{y})^2 - \cancel{(\underline{x}^T\underline{y})^2} + \cancel{(\underline{x}^T\underline{y})^2}$$

$$= \boxed{\|\underline{x}\|_2^2\|\underline{y}\|_2^2 - (\underline{x}^T\underline{y})^2}$$

Solution does not have an magnitude because $\underline{x}, \underline{y} \in \mathbb{R}^n$

$$\frac{1}{\|\underline{x}\|_2^2}\|(\|\underline{x}\|_2^2\underline{y}-(\underline{x}^T\underline{y})\underline{x}\|_2^2 = \|\underline{x}\|_2^2\|\underline{y}\|_2^2 - (\underline{x}^T\underline{y})^2$$

Notice that RHS has individual elements of CS inequality squared. LHS is positive (or non-negative in general) for non-zero vectors, so

$$\|\underline{x}\|_2^2\|\underline{y}\|_2^2 \geq (\underline{x}^T\underline{y})^2$$

$$\Rightarrow \|\underline{x}\|_2\|\underline{y}\|_2 \geq |\underline{x}^T\underline{y}| \quad \text{(C.S. ineq.)}$$

d) $\|\underline{x}\|_2\|\underline{y}\|_2 = |\underline{x}^T\underline{y}|$

$$\Rightarrow \frac{1}{\|\underline{x}\|_2^2}\|(\|\underline{x}\|_2^2\underline{y}-(\underline{x}^T\underline{y})\underline{x})\|_2^2 = 0$$

$$\Rightarrow \|(\|\underline{x}\|_2^2\underline{y}-(\underline{x}^T\underline{y})\underline{x})\|_2^2 = 0$$

$$\Rightarrow \|\underline{x}\|_2^2\underline{y} = (\underline{x}^T\underline{y})\underline{x}$$

$$\Rightarrow \underline{y} = \frac{(\underline{x}^T\underline{y})\underline{x}}{\|\underline{x}\|_2^2} \Rightarrow h = \frac{\underline{x}^T\underline{y}}{\|\underline{x}\|_2^2} = \frac{\langle \underline{x},\underline{y}\rangle}{\langle \underline{x},\underline{x}\rangle}$$

↳ This is basically the projection of $\underline{y}$ onto $\underline{x}$. If a vector is equal to its projection of onto another, then two vectors are linearly dependent, hence $\underline{y} = h\underline{x}$, $h \in \mathbb{R}$.

Q3 $T \in \mathbb{R}^{2\times 2}$, $\forall \underline{x} \in \mathbb{R}^2$

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \implies T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

i.e., T is a permutation matrix.

Let $\underline{x}$ be an eigenvector of T.

$$T\underline{x} = \lambda \underline{x}$$

$$(T - \lambda I)\underline{x} = 0$$

For non-zero $\underline{x} \implies \det(T - \lambda I) = 0$

$$\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = 0$$

$$\implies (\lambda - 1)(\lambda + 1) = 0$$

$$\lambda_1 = 1, \ \lambda_2 = -1$$

$$T\underline{x}_1 = \underline{x}_1 \implies \underline{x}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$T\underline{x}_2 = -\underline{x}_2 \implies \underline{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Q4 $\operatorname{rank}(A) = 1 \iff \dim(N(A)) = n-1 \implies n-1$ zero eigenvalues

T is symmetric (diagonalizable)

Consider $A^T A = A^2 = nA$

Let $\underline{x}$ be an eigenvector of A

$$A\underline{x} = \lambda \underline{x}$$

$$AA\underline{x} = \lambda A\underline{x} \implies \lambda = n$$

$$AA\underline{x} = nA\underline{x} = \lambda A\underline{x}$$

Hence, $\lambda_1 = n$, $\lambda_j = 0$, $j = 2, \dots, n$

5 $f: \mathbb{R}^2 \to \mathbb{R}$, $f(\underline{r}) = x^2 + y^2 + 2x - 3y$ where $\underline{r} = \begin{bmatrix} x \\ y \end{bmatrix}$.

$S = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$

$S \subseteq \mathbb{R}^2$, and it is nonempty, closed, and bounded, hence we can apply Weierstrass theorem (also f is continuous over S). We know that f has global min/max over S. We also know that if an optimal point resides in the interior of S, i.e., $\underline{r}^* \in \text{int}(S)$, then $\nabla f(\underline{r}^*) = 0$.

$\nabla f(\underline{r}) = \begin{bmatrix} 2x + 2 \\ 2y - 3 \end{bmatrix} \Rightarrow \underline{r}^* = \begin{bmatrix} -1 \\ 1.5 \end{bmatrix}$

However $\underline{r}^* \notin S$ since $\|\underline{r}\|_2^2 = \frac{13}{4} > 1$, which is not in the ball.

So, we can infer that optimal point resides on bound(S), i.e., $x^2 + y^2 = 1$.

Our minimization problem becomes $\min_{\underline{r}} (1 + 2x - 3y) = \min_{\underline{r}} g(x,y)$

$x^2 + y^2 = 1 \Rightarrow x^2 = 1 - y^2 \Rightarrow x = \mp\sqrt{1-y^2}$.

$g_1(y) = 1 + 2\sqrt{1-y^2} - 3y$, $g_2(y) = 1 - 2\sqrt{1-y^2} - 3y$

$g_1'(y) = \frac{-2y}{\sqrt{1-y^2}} - 3 = 0$, $g_2'(y) = \frac{2y}{\sqrt{1-y^2}} - 3 = 0$

$2y = -3\sqrt{1-y^2}$ $\quad$ $2y = 3\sqrt{1-y^2}$

$4y^2 = 9(1-y^2)$ $\quad$ $4y^2 = 9 - 9y^2$

$y^2 = \frac{9}{13} \Rightarrow \boxed{y = \mp\frac{3}{\sqrt{13}}}$

From y, we also get $x = \mp\sqrt{1 - \frac{9}{13}} = \boxed{\mp\frac{2}{\sqrt{13}}}$

Since in original f formulation, x has a plus sign, and y has a minus sign, min f will occur at $\boxed{\underline{r} = \begin{bmatrix} -2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}}$.

Check if $\underline{r} \in S$:

$x^2 + y^2 = \frac{4}{13} + \frac{9}{13} = 1$ ✓ $\quad (x,y) = \left(-\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}\right)$

Also, $f\left(-\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}\right) \approx -2.61$

Q6 $f(\underline{r}) = 2x - 3y$, $S = \{(x,y) \in \mathbb{R}^2 \mid 2x^2 + 5y^2 \leq 1\}$

$S \subseteq \mathbb{R}^2$ and it is nonempty, closed, and bounded, also f is continuous on S. Hence Weierstrass theorem is applicable; we know that f has a optimum point over S.

If f has a global max/min $\underline{r}^* \in int(S)$, then $\nabla f(\underline{r}^*) = 0$.

$\nabla f(\underline{r}^*) = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, $\underline{r}^* \notin S$. $(2(2)^2 + 5(-3)^2 > 1)$

Hence, optimum points reside on bound(S); i.e.,

$$2x^2 + 5y^2 = 1$$

$$\Rightarrow x = \mp \sqrt{\frac{1-5y^2}{2}}.$$

New optimization target $g_1(y) = \sqrt{2 - 10y^2} - 3y$

$g_2(y) = -\sqrt{2-10y^2} - 3y$

$$\frac{dg_1(y)}{dy} = \frac{-10y}{\sqrt{2-10y^2}} - 3 = 0 \Rightarrow 10y = -3\sqrt{2-10y^2}$$

$$100y^2 = 18 - 90y^2$$

$$\Rightarrow \boxed{y = \mp \frac{3}{\sqrt{95}}}$$

Then, $x = \mp\sqrt{\frac{1-5y^2}{2}} \Rightarrow \boxed{x = \mp\sqrt{\frac{5}{19}} = \mp \frac{5}{\sqrt{95}}}$

Due to signs in f, global max at $(x,y) = \left(\frac{5}{\sqrt{95}}, -\frac{3}{\sqrt{95}}\right)$

global min at $(x,y) = \left(-\frac{5}{\sqrt{95}}, \frac{3}{\sqrt{95}}\right)$

$f\left(\frac{5}{\sqrt{95}}, -\frac{3}{\sqrt{95}}\right) \approx 1.95$, $f\left(-\frac{5}{\sqrt{95}}, \frac{3}{\sqrt{95}}\right) \approx -1.95$

Also, these points are in S since

$2\left(\frac{5}{\sqrt{95}}\right)^2 + 5\left(-\frac{3}{\sqrt{95}}\right)^2 = 1$, $2\left(-\frac{5}{\sqrt{95}}\right)^2 + 5\left(-\frac{3}{\sqrt{95}}\right)^2 = 1$

1 Coerciveness: $\lim_{\|\underline{x}\|\to\infty} f(\underline{x}) = \infty$

a) $f(x_1, x_2) = 2x_1^2 - 8x_1x_2 + x_2^2$

$= 2x_1^2 - 4x_1x_2 - 4x_1x_2 + x_2^2$

$= x_1(2x_1 - 4x_2) + x_2(-4x_1 + x_2)$

$= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} 2 & -4 \\ -4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ $\left(\text{Also } \underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right)$

$= \underline{x}^T A \underline{x}$, where $A = \begin{bmatrix} 2 & -4 \\ -4 & 1 \end{bmatrix}$

If f is coercive, then it must have a global minimum point on $\mathbb{R}^2$.

If there is a global minimum point $\underline{x}^*$, it should satisfy $\nabla f(\underline{x}^*) = 0$ & $\nabla^2 f(\underline{x}^*) \succeq 0$

Check eigenvalues of A: $tr(A) = 3$, $det(A) = -14$ $\Rightarrow$ they have different signs

$\Rightarrow$ A is indefinite $\Rightarrow$ f cannot be coercive due to the previous reasoning.

b) $f(x_1, x_2) = 4x_1^2 + 2x_1x_2 + 2x_2^2$

$= 4x_1^2 + x_1x_2 + x_1x_2 + 2x_2^2$

$= x_1(4x_1 + x_2) + x_2(x_1 + 2x_2)$

$= \underline{x}^T A \underline{x}$, where $A = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$ $\Rightarrow$ A is positive definite

$\Rightarrow$ f is coercive

c) $f(x_1, x_2) = x_1^4 + x_2^4$

$= \|\underline{x}\|_4^4$, where $\|\underline{x}\|_4 = (x_1^4 + x_2^4)^{1/4}$

$\geq \|\underline{x}\|_4$

$\lim_{\|\underline{x}\|_4\to\infty} f(\underline{x}) \geq \lim_{\|\underline{x}\|_4\to\infty} \|\underline{x}\|_4 = \infty \Rightarrow$ f is coercive

d) $f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$

Consider $\underline{x} = \begin{bmatrix} x_1 \\ -x_1 \\ 0 \end{bmatrix}$. For $x_1 \to \infty$, $\|\underline{x}\| \to \infty$,

however $\lim_{\substack{\|\underline{x}\|\to\infty \\ x_1\to\infty}} f(\underline{x}) = 0$

$\Rightarrow$ f is not coercive