1 Question 1

(4 points)

1. Let $\{x^k\}$ be a sequence in $\mathbb{R}^n$ that converges to x^* with Q-sublinear convergence rate, briefly argue that $\log(||x^k - x^*||)$ always lies <u>above</u> any given line with negative slope, i.e., let $y = -ax, a > 0$, then for $k > N$ where N is sufficiently large, we always have $\log(||x^k - x^*||) > -ak$.
2. (Bonus) Briefly argue what happens the other way around? i.e, if $\log(||x^k - x^*||)$ lies above any given line of negative slope for sufficiently large k, what can we say about the sequence $\{x^k\}$?

Proof:

Since $\{x^k\}$ is Q-sublinear,

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 1$$

$$\lim_{k\to\infty} \log\left(\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}\right) = \log 1 = 0$$

$$\lim_{k\to\infty} [\log(\|x^{k+1} - x^*\|) - \log(\|x^k - x^*\|)] = 0$$

$$\forall \varepsilon > 0, \exists N > 0, s.t. -\varepsilon < \log(\|x^{k+1} - x^*\|) - \log(\|x^k - x^*\|) < \varepsilon, \forall k > N$$

Let $\varepsilon = a$, then

$$-a < \log(\|x^{k+1} - x^*\|) - \log(\|x^k - x^*\|)$$

$$-ak < a < \log(\|x^{k+1} - x^*\|) - \log(\|x^k - x^*\|) \text{ since } k > 0$$

Assume that there exists a > 0, such that for any N, $\log(\|x^k - x^*\|) \leq -ak$ for some k > N for a contradiction. We have 2 cases:

a. Let N be arbitrary, and let k goes to infinity, which is larger than N, and satisfies $\log(\|x^k - x^*\|) > -ak$, where k + 1 satisfies $\log(\|x^{k+1} - x^*\|) \leq -a(k+1)$, assuming there is such k. Then we have:

$$\lim_{k\to\infty} \log(\|x^k - x^*\|) \leq \lim_{k\to\infty} -ak = \lim_{k\to\infty} \log e^{-ak}$$

$$\lim_{k\to\infty} \|x^k - x^*\| \leq \lim_{k\to\infty} e^{-ak}$$

And

$$\lim_{k\to\infty} \log(\|x^{k+1} - x^*\|) > \lim_{k\to\infty} -a(k+1) = \lim_{k\to\infty} \log e^{-a(k+1)}$$

$$\lim_{k\to\infty} \|x^{k+1} - x^*\| > \lim_{k\to\infty} e^{-a(k+1)} = \lim_{k\to\infty} e^{-ak-a} = e^{-a} \lim_{k\to\infty} e^{-ak}$$

Thus,

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = \frac{\lim_{k\to\infty} \|x^{k+1} - x^*\|}{\lim_{k\to\infty} \|x^k - x^*\|} = \frac{e^{-a} \lim_{k\to\infty} e^{-ak}}{\lim_{k\to\infty} e^{-ak}} = e^{-a} < 1, \text{since a} > 0$$

However, since $\{x^k\}$ is Q-sublinear, $\lim_{k\to\infty} \frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|} = 1$ when k goes to infinity, so here is a contradiction.

b. If there is no k > N, where k goes to infinity, such that $\log(\|x^k - x^*\|) \leq -ak$ and $\log(\|x^{k+1} - x^*\|) > -a(k+1)$, for any N, which means that for any k that goes to infinity, $\log(\|x^k - x^*\|) \leq -ak$ and $\log(\|x^{k+1} - x^*\|) \leq -a(k+1) = -ak - a$. That said, let $y^k = \log\|x^k - x^*\| + ak$, $\lim_{k\to\infty} y^k \leq 0$.

$$y^{k+1} - y^k = \log(\|x^{k+1} - x^*\|) + ak + a - \log(\|x^k - x^*\|) - ak$$
$$= \log\left(\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}\right) + a$$

$$\lim_{k\to\infty}(y^{k+1} - y^k) = \lim_{k\to\infty}\left(\log\left(\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}\right) + a\right)$$
$$= \lim_{k\to\infty}\left(\log\left(\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}\right)\right) + a = \log\left(\lim_{k\to\infty}\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}\right) + a$$
$$= \log 1 + a = a > 0$$

2. Then $\{x^k\}$ may be sublinear or divergent.

2 Question 2

(4 points) Prove that the induced norm is also a norm, i.e, it satisfies the three properties.

Proof:

By definition, for any matrix $A \in R^{n\times n}$, the induced norm has

$$\|A\| = \max_{\|x\|=1} \|Ax\| \geq 0$$

That is because Ax is a vector, and the norms of vectors are always non-negative.

Then, we have

$$\|A\| = \max_{\|x\|\neq 0} \frac{\|Ax\|}{\|x\|} = 0$$

Iff

$$\forall y \in R^n \text{ such that } \|y\| \neq 0, \frac{\|Ay\|}{\|y\|} \leq 0$$

Iff

$$\forall y \in R^n \text{ such that } \|y\| \neq 0, \|Ay\| \leq 0$$

Since Ay is a vector, and the norms of vectors are non-negative, $\|Ay\| \geq 0$, so $\|Ay\| = 0$, if and only if $Ay = 0$ by the property of vector norms. That said, for any $y \in R^n$ such that $\|y\| \neq 0$, $Ay = 0$. Also, it is trivial to show that $Ay = 0$ when $y = 0$, so $null(A) = R^n$, which means A = 0.

Thus, $\|A\| = 0$ if and only if $A = 0$.

Then, for any $c\epsilon R$, we have

$$\|cA\| = \max_{\|x\|=1} \|cAx\| = \max_{\|x\|=1} \|c(Ax)\|$$

Since Ax is a vector, based on the property of vector norms and the fact that $|c| > 0$, we have

$$\|cA\| = \max_{\|x\|=1} \|c(Ax)\| = \max_{\|x\|=1} |c|\|Ax\| = |c| \max_{\|x\|=1} \|Ax\| = |c|\|A\|$$

Then, for any matrices $A \in R^{n\times n}, B \in R^{n\times n}$,

$$\|A + B\| = \max_{\|x\|=1} \|(A + B)x\| = \max_{\|x\|=1} \|(Ax) + (Bx)\|$$

Let x, y be optimal solutions to $\max_{\|x\|=1} \|(Ax) + (Bx)\|$, we have

$$\|A + B\| = \|(Ax) + (Bx)\|$$

Since Ax and Bx are vectors, based on the property of vector norms, we have

$$\|A + B\| = \|(Ax) + (Bx)\| \leq \|Ax\| + \|Bx\|$$

Thus,

$$\|A + B\| \leq \|Ax\| + \|Bx\| \leq \max_{\|a\|=1} \|Aa\| + \max_{\|b\|=1} \|Bb\| = \|A\| + \|B\|$$

Finally, we have proved for any matrix $A \in R^{n\times n}$, $\|A\| \geq 0$, $\|A\| = 0$ if and only if $A = 0$; For any $c\epsilon R$, $\|cA\| = |c| \max_{\|x\|=1} \|Ax\| = |c|\|A\|$; For any matrices $A \in R^{n\times n}, B \in R^{n\times n}$, $\|A + B\| \leq \|A\| + \|B\|$. That said, the induced norm is also a norm.

3 Question 3

(4 points) Let $A \in \mathbb{R}^{n\times n}$, prove

$$||A||_2^2 = ||AA^T||_2 = ||A^TA||_2.$$

Hint: use Proposition 2.5, 2.6 and Lemma 2.7.

Proof:

First, we prove $\|A^TA\|_2 \le \|A\|_2^2$:

By the definition of induced norm,

$$\|A^TA\|_2 = \max_{\|x\|_2=1} \|A^TAx\|_2 = \max_{\|x\|_2=1} \|A^T(Ax)\|_2$$

That said, let x be the optimal solution to $\max_{\|x\|_2=1} \|A^T(Ax)\|_2$, we have

$$\|A^TA\|_2 = \|A^T(Ax)\|_2$$

By Lemma 2.7,

$$\|A^TA\|_2 = \|A^T(Ax)\|_2 \le \|A^T\|_2\|Ax\|_2$$

Since $\|Ax\|_2 \le \max_{\|z\|_2=1} \|Az\|_2 = \|A\|_2$, we have

$$\|A^TA\|_2 \le \|A^T\|_2\|Ax\|_2 \le \|A^T\|_2\|A\|_2$$

By Proposition 2.6,

$$\|A^TA\|_2 \le \|A^T\|_2\|A\|_2 = \|A\|_2\|A\|_2 = \|A\|_2^2$$

Then, we prove $\|A^TA\|_2 \ge \|A\|_2^2$:

Let x be optimal to $\max_{\|x\|_2=1} \|Ax\|_2$, then by the definition of induced norm,

$$\|A\|_2^2 = \|Ax\|_2^2 = |x^TA^TAx|$$

Let y be optimal to $\max_{\|x\|_2=\|y\|_2=1} |y^TA^TAx|$. From Proposition 2.5, we know the x that is optimal to $\max_{\|x\|_2=1} \|Ax\|_2$ is exactly the x that is optimal to $\max_{\|x\|_2=\|y\|_2=1} |y^TA^TAx|$.

Therefore,

$$\|A\|_2^2 = |x^TA^TAx| \le |y^TA^TAx| = \|A^TA\|_2$$

Since $\|A^TA\|_2 \le \|A\|_2^2$ and $\|A\|_2^2 \le \|A^TA\|_2$, we have $\|A\|_2^2 = \|A^TA\|_2$.

Swap A^T and A, we can trivially get $\|A^T\|_2^2 = \|AA^T\|_2$. By Proposition 2.6, we have $\|A\|_2^2 = \|A^T\|_2^2 = \|AA^T\|_2$. Finally, we have

$$\|A\|_2^2 = \|AA^T\|_2 = \|A^TA\|_2$$

4 Question 4

(4 points) Let $A \in \mathbb{R}^{n\times n}$ be symmetric, prove $||A^k||_2 = ||A||_2^k, \forall k = 1, ..., n$.

Proof:

Let λ_{max} be the eigenvalue of A with the largest magnitude, and we know $\|A\|_2 = |\lambda_{max}|$ by Proposition 2.14 since A is symmetric. Hence,

$$\|A\|_2^k = |\lambda_{max}|^k = |\lambda_{max}^k|$$

Let m be the number of eigenvectors of A, then $\forall i$ that $i = 1, \dots, m,\ |\lambda_i| < |\lambda_{max}|$, where λ_i are eigenvalues of A. Note that since any power functions with positive powers are non-increasing, we know $\forall i$ that $i = 1, \dots, m,\ |\lambda_i^k| = |\lambda_i|^k < |\lambda_{max}|^k = |\lambda_{max}^k|$, since k $\geqslant$ 1.

Then, by Proposition 2.10, λ_i^k are all eigenvalues of A^K, i = 1, …, m. Since $\forall i$ that $i = 1, \dots, m,\ |\lambda_i^k| < |\lambda_{max}^k|,\ \lambda_{max}^k$ is the eigenvalue of A^k with the largest magnitude, namely $|\lambda_{max}^k| = \rho(A^k)$.

Then, since A is symmetric,

$$A = A^T$$
$$A^k = (A^T)^k = (A^k)^T$$

Therefore, A^k is also symmetric, so $\|A^k\|_2 = \rho(A^k) = |\lambda_{max}^k| = |\lambda_{max}|^k = \|A\|_2^k$ by Proposition 2.14.

5 Question 5

(4 points) Consider the quadratic inequality

$$x^T Ax + b^T x + c \leq 0$$

Let A be a $n \times n$ symmetric matrix, and let C be the set of all $x \in \mathbb{R}^n$ which satisfies the above ineuality. Prove if A is positive semidefinite, then C is a convex set.

Proof:

Assuming A is positive semidefinite. Let $x_1, x_2 \in C$, and $y = \theta x_1 + (1 - \theta)x_2, \forall\theta \in [0,1]$, then we have

By Proposition 2.20, $A = B^T B$ for some matrix $B \in R^{n\times r}$ of full column rank, where $r = rank(A)$, since A is symmetric and positive semidefinite. Then, we rearrange the inequality:

$$x^T Ax + b^T x + c = x^T B^T Bx + b^T x + c = (Bx)^T(Bx) + b^T x + c$$
$$= \|Bx\|_2^2 + b^T x + c \leq 0$$

Thus,

$$y^T Ay + b^T y + c = \|By\|_2^2 + b^T y + c$$
$$= \|B[\theta x_1 + (1-\theta)x_2]\|_2^2 + b^T[\theta x_1 + (1-\theta)x_2] + c$$
$$= \|\theta Bx_1 + (1-\theta)Bx_2\|_2^2 + b^T[\theta x_1 + (1-\theta)x_2] + c$$

By Pythagorean Theorem,

$$y^T Ay + b^T y + c = \|\theta Bx_1 + (1-\theta)Bx_2\|_2^2 + b^T[\theta x_1 + (1-\theta)x_2] + c$$
$$= \|\theta Bx_1\|_2^2 + \|(1-\theta)Bx_2\|_2^2 + b^T[\theta x_1 + (1-\theta)x_2] + c$$
$$= \theta^2\|Bx_1\|_2^2 + (1-\theta)^2\|Bx_2\|_2^2 + \theta b^T x_1 + (1-\theta)b^T x_2 + c$$

Since $x_1, x_2 \in C$, $\|Bx_1\|_2^2 + b^T x_1 + c \leq 0$ and $\|Bx_2\|_2^2 + b^T x_2 + c \leq 0$, we have

$$\theta^2\|Bx_1\|_2^2 \leq -\theta^2 b^T x_1 - \theta^2 c$$

And

$$(1-\theta)^2\|Bx_2\|_2^2 \leq -(1-\theta)^2 b^T x_2 - (1-\theta)^2 c$$

Therefore,

$$y^T Ay + b^T y + c = \theta^2\|Bx_1\|_2^2 + (1-\theta)^2\|Bx_2\|_2^2 + \theta b^T x_1 + (1-\theta)b^T x_2 + c$$
$$\leq -\theta^2 b^T x_1 - \theta^2 c - (1-\theta)^2 b^T x_2 - (1-\theta)^2 c + \theta b^T x_1$$
$$+ (1-\theta)b^T x_2 + c$$
$$= -\theta^2 b^T x_1 - \theta^2 c - (1 - 2\theta + \theta^2)b^T x_2 - (1 - 2\theta + \theta^2)c$$
$$+ \theta b^T x_1 + (1-\theta)b^T x_2 + c$$
$$= -\theta^2 b^T x_1 - \theta^2 c - b^T x_2 + 2\theta b^T x_2 - \theta^2 b^T x_2 - c + 2\theta c - \theta^2 c$$
$$+ \theta b^T x_1 + b^T x_2 - \theta b^T x_2 + c$$
$$= -\theta^2 b^T x_1 - \theta^2 c + \theta b^T x_2 - \theta^2 b^T x_2 + 2\theta c - \theta^2 c + \theta b^T x_1$$
$$= b^T x_1(-\theta^2 + \theta) + b^T x_2(-\theta^2 + \theta) + c(-\theta^2 + \theta) + c(-\theta^2 + \theta)$$
$$= (-\theta^2 + \theta)[(b^T x_1 + c) + (b^T x_2 + c)]$$

Note that since $0 \leq \theta \leq 1$, we have $\theta^2 = \theta\theta \leq \theta$, which means $-\theta^2 + \theta \geq 0$. Then, since $x_1, x_2 \in C$, we have

$$x_1^T Ax_1 + b^T x_1 + c \leq 0$$

$$b^T x_1 + c \leq -x_1^T A x_1$$

Since A is positive semidefinite, $x_1^T A x_1 \geq 0$ by definition, we have

$$b^T x_1 + c \leq -x_1^T A x_1 \leq 0$$

Similarly, we know $b^T x_2 + c \leq -x_2^T A x_2 \leq 0.$

Thus,

$$(b^T x_1 + c) + (b^T x_2 + c) \leq 0$$

$y^T A y + b^T y + c = (-\theta^2 + \theta)[(b^T x_1 + c) + (b^T x_2 + c)] \leq 0 \; since \; -\theta^2 + \theta \geq 0$

Therefore, $y \in C$, so C is convex.