# Solutions to Mathematical Problems

Junjun Chen
Department of Mathematics, Lanzhou University

August 1, 2025

---

1.

**Problem.** Let $f$ be a continuously differentiable function on $\mathbb{R}^n$. Suppose there exists a positive constant $L$ such that $\nabla f$ is $L$-Lipschitz continuous, namely

$$\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2 \quad \text{for all} \quad x, y \in \mathbb{R}^n.$$

(a) Prove that

$$\inf_{y\in\mathbb{R}^n} f(y) \leq f(x) - \frac{1}{2L}\|\nabla f(x)\|_2^2 \quad \text{for all} \quad x \in \mathbb{R}^n.$$

(b) If in addition $f$ is convex, prove that

$$f(x) - f(y) - [\nabla f(x)]^T(x - y) \leq -\frac{1}{2L}\|\nabla f(x) - \nabla f(y)\|_2^2.$$

**Solution**.

(a) *Proof.* We first establish a fundamental inequality for $f$. Since $\nabla f$ is $L$-Lipschitz continuous, for any $x, y \in \mathbb{R}^n$:

$$\begin{aligned} f(y) &= f(x) + \int_0^1 \langle \nabla f(x + \tau(y - x)),\, y - x\rangle d\tau \\ &= f(x) + \langle \nabla f(x),\, y - x\rangle + \int_0^1 \langle \nabla f(x + \tau(y - x)) - \nabla f(x),\, y - x\rangle d\tau. \end{aligned}$$

This implies:

$$
\begin{aligned}
&|f(y) - f(x) - \langle \nabla f(x),\, y - x\rangle| \\
&= \left| \int_0^1 \langle \nabla f(x + \tau(y-x)) - \nabla f(x),\, y - x\rangle d\tau \right| \\
&\leq \int_0^1 |\langle \nabla f(x + \tau(y-x)) - \nabla f(x),\, y - x\rangle|\, d\tau \\
&\leq \int_0^1 \|\nabla f(x + \tau(y-x)) - \nabla f(x)\|_2 \cdot \|y - x\|_2 d\tau \quad \text{(by Cauchy-Schwarz)} \\
&\leq \int_0^1 L\tau \|y - x\|_2^2 d\tau \quad \text{(by } L\text{-Lipschitz continuity)} \\
&= \frac{L}{2}\|y - x\|_2^2.
\end{aligned}
$$

Thus, we have the quadratic bound:

$$
f(y) \leq f(x) + \langle \nabla f(x),\, y - x\rangle + \frac{L}{2}\|y - x\|_2^2 \quad \forall x, y \in \mathbb{R}^n. \tag{1}
$$

To prove the main result, fix $x \in \mathbb{R}^n$ and choose the specific point:

$$
y^* = x - \frac{1}{L}\nabla f(x).
$$

Substituting $y^*$ into (1):

$$
\begin{aligned}
f(y^*) &\leq f(x) + \left\langle \nabla f(x),\, \left(x - \frac{1}{L}\nabla f(x)\right) - x \right\rangle + \frac{L}{2}\left\| \left(x - \frac{1}{L}\nabla f(x)\right) - x \right\|_2^2 \\
&= f(x) + \left\langle \nabla f(x),\, -\frac{1}{L}\nabla f(x) \right\rangle + \frac{L}{2}\left\| -\frac{1}{L}\nabla f(x) \right\|_2^2 \\
&= f(x) - \frac{1}{L}\langle \nabla f(x),\, \nabla f(x)\rangle + \frac{L}{2} \cdot \frac{1}{L^2}\|\nabla f(x)\|_2^2 \\
&= f(x) - \frac{1}{L}\|\nabla f(x)\|_2^2 + \frac{1}{2L}\|\nabla f(x)\|_2^2 \\
&= f(x) - \frac{1}{2L}\|\nabla f(x)\|_2^2.
\end{aligned}
$$

Since $y^* \in \mathbb{R}^n$, the infimum satisfies:

$$
\inf_{z \in \mathbb{R}^n} f(z) \leq f(y^*) \leq f(x) - \frac{1}{2L}\|\nabla f(x)\|_2^2.
$$

This holds for all $x \in \mathbb{R}^n$, completing the proof. □

**Remark.** The proof of upper and lower bounds $f$ is referenced from [Lectures on Convex Optimization, Yurii Nesterov, 2010].

(b) *Proof.* Assume $f$ is convex and $\nabla f$ is $L$-Lipschitz continuous. Fix arbitrary $x, y \in \mathbb{R}^n$. Define the auxiliary function:

$$h(z) = f(z) - f(x) - \langle \nabla f(x), z - x \rangle.$$

This function has the following properties:

i. $h$ is convex (it can easily verified by first-order sufficient condition).

ii. $\nabla h(z) = \nabla f(z) - \nabla f(x)$ (by direct computation).

iii. $\nabla h$ is $L$-Lipschitz continuous since:

$$\|\nabla h(z_1) - \nabla h(z_2)\|_2 = \|\nabla f(z_1) - \nabla f(z_2)\|_2 \leq L\|z_1 - z_2\|_2.$$

iv. $h(x) = 0$ and $x$ is a minimum point because:

$$\nabla h(x) = 0 \quad \text{and} \quad h(z) \geq 0 \quad \forall z \in \mathbb{R}^n \quad \text{(by convexity of } f\text{)}.$$

Applying the result from part (a) to $h$ at point $y$, we have:

$$\inf_{z \in \mathbb{R}^n} h(z) \leq h\left(y - \frac{1}{L}\nabla h(y)\right) \leq h(y) - \frac{1}{2L}\|\nabla h(y)\|_2^2.$$

Since $x$ achieves the infimum ($h(x) = \inf_z h(z) = 0$):

$$\begin{aligned} 0 = h(x) &\leq h\left(y - \frac{1}{L}\nabla h(y)\right) \\ &\leq h(y) - \frac{1}{2L}\|\nabla h(y)\|_2^2. \end{aligned}$$

Thus:

$$0 \leq h(y) - \frac{1}{2L}\|\nabla h(y)\|_2^2,$$

which implies:

$$h(y) \geq \frac{1}{2L}\|\nabla h(y)\|_2^2.$$

Substituting back the definitions of $h$ and $\nabla h$:

$$f(y) - f(x) - \langle \nabla f(x), y - x \rangle \geq \frac{1}{2L}\|\nabla f(y) - \nabla f(x)\|_2^2.$$

This completes the proof for all $x, y \in \mathbb{R}^n$. □

2.

**Problem.** Given a symmetric matrix $A \in \mathbb{R}^{n \times n}$ and a vector $b \in \mathbb{R}^n$, define

$$q(x) = \frac{1}{2}x^T Ax - b^T x, \quad x \in \mathbb{R}^n.$$

Prove that the following statements are equivalent.

(a) $q$ is bounded from below.

(b) $A \succeq 0$ and $b \in \text{range}(A)$.

(c) $q$ has a local minimum.

(d) $q$ has a global minimum.

**Solution**. *Proof.* We establish the equivalence by proving the cycle of implications: (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (a).

(a) $\Rightarrow$ (b): Suppose $q$ is bounded from below, i.e., there exists a constant $C \in \mathbb{R}$ such that $q(x) \geq C$ for all $x \in \mathbb{R}^n$.

First, we prove that $A \succeq 0$. Assume, to the contrary, that $A$ has at least one negative eigenvalue. Let $\lambda < 0$ be such an eigenvalue with a corresponding eigenvector $v \in \mathbb{R}^n$, $v \neq 0$, normalized so that $\|v\| = 1$. Consider points $x = tv$ for $t \in \mathbb{R}$. Then,

$$q(tv) = \frac{1}{2}(tv)^\top A(tv) - b^\top(tv) = \frac{1}{2}t^2 v^\top Av - tb^\top v = \frac{1}{2}\lambda t^2 - t(b^\top v),$$

since $v^\top Av = \lambda v^\top v = \lambda\|v\|^2 = \lambda$. As $t \to \pm\infty$, the dominant term is $\frac{1}{2}\lambda t^2$ because $\lambda < 0$. Specifically,

$$\lim_{t\to\infty} q(tv) = -\infty \quad \text{and} \quad \lim_{t\to-\infty} q(tv) = -\infty,$$

regardless of the value of $b^\top v$. This contradicts the boundedness of $q$ from below. Therefore, $A$ has no negative eigenvalues, and since $A$ is symmetric, $A$ is positive semidefinite, i.e., $A \succeq 0$.

Next, we prove that $b \in \text{range}(A)$. Since $A$ is symmetric, the fundamental theorem of linear algebra gives $\text{range}(A) = (\ker(A))^\perp$, so $\mathbb{R}^n = \text{range}(A) \oplus \ker(A)$, and this decomposition is orthogonal. Suppose, for contradiction, that $b \notin \text{range}(A)$. Then $b$ has a nonzero orthogonal projection onto $\ker(A)$, i.e., there exists a vector $w \in \ker(A)$, $w \neq 0$, such that $b^\top w \neq 0$. Consider points $x = tw$ for $t \in \mathbb{R}$. Since $w \in \ker(A)$, $Aw = 0$, and thus

$$q(tw) = \frac{1}{2}(tw)^\top A(tw) - b^\top(tw) = \frac{1}{2}t^2 w^\top Aw - tb^\top w = -tb^\top w,$$

as $w^\top Aw = w^\top(Aw) = w^\top 0 = 0$. If $b^\top w > 0$, then $\lim_{t\to\infty} q(tw) = -\infty$. If $b^\top w < 0$, then $\lim_{t\to-\infty} q(tw) = -\infty$. In either case, this contradicts the boundedness of $q$ from below. Hence, $b \in \text{range}(A)$.

(b) $\Rightarrow$ (c): Assume $A \succeq 0$ and $b \in \text{range}(A)$. Since $b \in \text{range}(A)$, there exists a vector $x^* \in \mathbb{R}^n$ such that $Ax^* = b$. The gradient of $q$ is

$$\nabla q(x) = Ax - b,$$

and the Hessian is $\nabla^2 q(x) = A$, which is constant and positive semidefinite by assumption. At $x^*$, $\nabla q(x^*) = Ax^* - b = 0$. To verify that $x^*$ is a local minimum, note that for any

direction $d \in \mathbb{R}^n$, the second-order Taylor expansion around $x^*$ is

$$q(x^* + d) = q(x^*) + \nabla q(x^*)^\top d + \frac{1}{2} d^\top \nabla^2 q(x^*) d = q(x^*) + 0 \cdot d + \frac{1}{2} d^\top A d.$$

Since $A \succeq 0$, $d^\top A d \geq 0$ for all $d$, so $q(x^* + d) \geq q(x^*)$ for all $d \in \mathbb{R}^n$. This implies that $x^*$ is actually a global minimum and hence a local minimum.

(c) $\Rightarrow$ (d): Suppose $q$ has a local minimum at some point $x^* \in \mathbb{R}^n$. By the first-order necessary condition for a local minimum (since $q$ is continuously differentiable), $\nabla q(x^*) = 0$, which implies

$$Ax^* - b = 0, \quad \text{so} \quad Ax^* = b.$$

Now, we first show that $A \succeq 0$. Since $x^*$ is a local minimum, there exists $\delta > 0$ such that for all $d \in \mathbb{R}^n$ with $\|d\| < \delta$, $q(x^* + d) \geq q(x^*)$. Substituting $x = x^* + d$ and using $Ax^* = b$, we compute:

$$\begin{aligned} q(x^* + d) &= \frac{1}{2}(x^* + d)^\top A(x^* + d) - b^\top (x^* + d) \\ &= \frac{1}{2}(x^*)^\top A x^* + \frac{1}{2} d^\top A x^* + \frac{1}{2}(x^*)^\top A d + \frac{1}{2} d^\top A d - b^\top x^* - b^\top d. \end{aligned}$$

As $A$ is symmetric, $d^\top A x^* = (x^*)^\top A d$, and since $Ax^* = b$, this simplifies to:

$$\begin{aligned} q(x^* + d) &= \left[\frac{1}{2}(x^*)^\top A x^* - b^\top x^*\right] + d^\top (Ax^*) - b^\top d + \frac{1}{2} d^\top A d \\ &= q(x^*) + d^\top b - b^\top d + \frac{1}{2} d^\top A d. \end{aligned}$$

Since $d^\top b$ and $b^\top d$ are scalars and equal, $d^\top b = b^\top d$, so:

$$q(x^* + d) = q(x^*) + \frac{1}{2} d^\top A d.$$

The local minimum condition requires $q(x^* + d) \geq q(x^*)$ for all $\|d\| < \delta$, which implies $\frac{1}{2} d^\top A d \geq 0$ for all such $d$. For any nonzero $d_0 \in \mathbb{R}^n$, set $d = t d_0$ with $t > 0$ small enough so that $\|t d_0\| < \delta$. Then:

$$\frac{1}{2}(t d_0)^\top A (t d_0) = \frac{1}{2} t^2 d_0^\top A d_0 \geq 0,$$

so $d_0^\top A d_0 \geq 0$. As $d_0$ is arbitrary, $A \succeq 0$.

With $A \succeq 0$ and $Ax^* = b$, we now show that $x^*$ is a global minimum. For any $x \in \mathbb{R}^n$, set $d = x - x^*$. Then:

$$q(x) = q(x^* + d) = q(x^*) + \frac{1}{2} d^\top A d \geq q(x^*),$$

since $d^\top A d \geq 0$. Thus, $x^*$ is a global minimum.

(d) $\Rightarrow$ (a): Suppose $q$ has a global minimum at some point $x^*$. Then, by definition, $q(x) \geq q(x^*)$ for all $x \in \mathbb{R}^n$. Setting $C = q(x^*)$, we have $q(x) \geq C$ for all $x$, so $q$ is bounded from below.

Since all implications (a) $\Rightarrow$ (b), (b) $\Rightarrow$ (c), (c) $\Rightarrow$ (d), and (d) $\Rightarrow$ (a) hold, the statements (a), (b), (c), and (d) are equivalent. $\square$

3.

**Problem.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. For $t \in \mathbb{R}$, define

$$\mathcal{L}(t) = \{x \in \mathbb{R}^n : f(x) \leq t\}.$$

Suppose that there exists a certain $t_0 \in \mathbb{R}$ such that $\mathcal{L}(t_0)$ is nonempty and bounded. Show that $\mathcal{L}(t)$ is bounded for all $t \in \mathbb{R}$.

**Solution**. *Proof.* Assume that there exists $t_0 \in \mathbb{R}$ such that the level set $\mathcal{L}(t_0) = \{x \in \mathbb{R}^n : f(x) \leq t_0\}$ is nonempty and bounded. We must show that $\mathcal{L}(t)$ is bounded for all $t \in \mathbb{R}$.

First, if $t < t_0$, then $\mathcal{L}(t) \subseteq \mathcal{L}(t_0)$. Since $\mathcal{L}(t_0)$ is bounded, $\mathcal{L}(t)$ is also bounded. Thus, it suffices to consider $t > t_0$.

Suppose, for contradiction, that there exists some $t_1 > t_0$ such that $\mathcal{L}(t_1)$ is unbounded. Then there is a sequence $\{x_k\} \subset \mathbb{R}^n$ with:

$$f(x_k) \leq t_1 \quad \text{and} \quad \lim_{k\to\infty} \|x_k\| = \infty.$$

Fix $x_0 \in \mathcal{L}(t_0)$, so $f(x_0) \leq t_0$. Since $\mathcal{L}(t_0)$ is bounded and $\|x_k\| \to \infty$, the sequence $\{x_k\}$ cannot lie entirely in $\mathcal{L}(t_0)$. Otherwise, it would contradict the boundedness of $\mathcal{L}(t_0)$. Thus, by passing to a subsequence if necessary, assume $f(x_k) > t_0$ for all $k$.

For each $k$, define:

$$\lambda_k = \frac{t_0 - f(x_0)}{f(x_k) - f(x_0)}.$$

We now consider two cases based on whether $t_0 > f(x_0)$ or $t_0 = f(x_0)$.

*Case 1: There exists a point $x_0 \in \mathcal{L}(t_0)$ such that $f(x_0) < t_0$.* Since $f(x_k) > t_0 > f(x_0)$ and $f(x_k) \leq t_1 < \infty$, we have $\lambda_k \in (0, 1)$. Define:

$$y_k = (1 - \lambda_k)x_0 + \lambda_k x_k.$$

By convexity of $f$:

$$f(y_k) \leq (1 - \lambda_k)f(x_0) + \lambda_k f(x_k) = t_0,$$

so $y_k \in \mathcal{L}(t_0)$. Now compute:

$$\|y_k - x_0\| = \lambda_k \|x_k - x_0\| = \frac{t_0 - f(x_0)}{f(x_k) - f(x_0)} \|x_k - x_0\|.$$

Since $f(x_k) \leq t_1$ and $f(x_k) > t_0$:

$$f(x_k) - f(x_0) \leq t_1 - f(x_0), \quad \text{so} \quad \lambda_k \geq \frac{t_0 - f(x_0)}{t_1 - f(x_0)} > 0.$$

As $\|x_k - x_0\| \to \infty$, we have $\|y_k - x_0\| \to \infty$, so $\{y_k\}$ is an unbounded sequence in $\mathcal{L}(t_0)$, contradiction.

*Case 2: For all $x \in \mathcal{L}(t_0)$, we have $f(x) = t_0$.*

This condition implies that the set $\{x \in \mathbb{R}^n : f(x) < t_0\}$ is empty. By definition, this means that $t_0$ is the global minimum value of the function $f$. Therefore, the level set $\mathcal{L}(t_0)$ is precisely the set of all global minimizers of $f$.

$$\mathcal{L}(t_0) = \arg\min_{x \in \mathbb{R}^n} f(x).$$

By our initial assumption, this set of minimizers is nonempty and bounded. A fundamental result in convex analysis states that if a convex function has a nonempty and bounded set of global minimizers, then the function must be coercive. A function $f$ is coercive if

$$\lim_{\|x\| \to \infty} f(x) = \infty.$$

However, our contradiction assumption introduced a sequence $\{x_k\}$ such that $\|x_k\| \to \infty$ while $f(x_k) \leq t_1$ for all $k$. The existence of such a sequence contradicts the fact that $f$ is coercive.

Since both possible cases lead to a contradiction, our initial assumption that there exists an unbounded level set $\mathcal{L}(t_1)$ for some $t_1 > t_0$ must be false. Therefore, $\mathcal{L}(t)$ is bounded for all $t \in \mathbb{R}$. □

4.

**Problem.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $K \subset \mathbb{R}^n$ be a compact set. Prove that $f$ is Lipschitz continuous on $K$.

**Solution**. Since the tight set $K$ is a bounded closed set, it is always possible to find a ball covering $K$. We only need to show that $f$ is Lipschitz in the ball.

*Proof.* First we show it is true for closed cube. Let $Q := [-L, L]^n$ be a cube, with vertices $V = \{v_k\}_{k=1}^{2^n}$. We can write any point $x \in Q$ as a convex combination of the vertices:

$$x = \sum_{k=1}^{2^n} \lambda_k v_k, \quad \text{where } 0 \leq \lambda_k \leq 1 \text{ and } \sum \lambda_k = 1.$$

Hence

$$f(x) \leq \sum_{k=1}^{2^n} \lambda_k f(v_k) \leq \max_{v_k \in V} f(v_k) < \infty,$$

and thus $M := \sup_Q f < \infty$. To derive a lower bound, again select any point $x \in Q$ and write

$$0 = \frac{1}{2}x + \frac{1}{2}(-x).$$

Then

$$f(0) \le \frac{1}{2}f(x) + \frac{1}{2}f(-x) \le \frac{1}{2}f(x) + \frac{1}{2}M;$$

and so

$$f(x) \ge 2f(0) - M.$$

Therefore $\inf_Q f \ge 2f(0) - M$. These estimates are valid for each cube $Q$ as above, and hence $f$ is locally bounded.

Next we can prove for $B(r)$. If $x, y \in B(r)$ and $x \ne y$, select $\mu > 0$ so that

$$z := x + \mu(y - x) \in \partial B(2r).$$

Then $\mu = \frac{|z-x|}{|y-x|} > 1$ and $y = \frac{1}{\mu}z + \left(1 - \frac{1}{\mu}\right)x$. Hence

$$\begin{aligned} f(y) &\le \frac{1}{\mu}f(z) + \left(1 - \frac{1}{\mu}\right)f(x) \\ &= f(x) + \frac{1}{\mu}\left(f(z) - f(x)\right) \\ &\le f(x) + C|y - x| \end{aligned}$$

for $C := \frac{2}{r}\sup_{B(2r)}|f|$, since $|z - x| \ge r$. Interchanging $x, y$, we find that

$$|f(y) - f(x)| \le C|y - x| \quad (x, y \in B(r)).$$

. □

**Remark.** This proof is referenced to [Measure Theorem and Fine Properties of Function, Lawrence C. Evans and Ronald F. Gariepy, 2015].

5.

**Problem.** Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable convex function, $\nabla f$ is $L$-Lipschitz continuous, and $x^*$ is a minimizer of $f$. Prove that

$$\|x - t\nabla f(x) - x^*\|_2 \le \|x - x^*\|_2$$

for all $t \in [0, 2/L]$.

**Solution**. *Proof.* We need to show that $\|x - t\nabla f(x) - x^*\|_2 \le \|x - x^*\|_2$ for all $t \in [0, 2/L]$. Since the norm is non-negative, it suffices to prove the equivalent inequality for squares:

$$\|x - t\nabla f(x) - x^*\|_2^2 \le \|x - x^*\|_2^2.$$

Define the difference:

$$\Delta(t) = \|x - t\nabla f(x) - x^*\|_2^2 - \|x - x^*\|_2^2.$$

Expanding the first term:

$$\begin{aligned}\Delta(t) &= \|(x - x^*) - t\nabla f(x)\|_2^2 - \|x - x^*\|_2^2 \\ &= \left(\|x - x^*\|_2^2 - 2t\langle x - x^*, \nabla f(x)\rangle + t^2\|\nabla f(x)\|_2^2\right) - \|x - x^*\|_2^2 \\ &= -2t\langle x - x^*, \nabla f(x)\rangle + t^2\|\nabla f(x)\|_2^2.\end{aligned}$$

Thus, $\Delta(t) = t^2\|\nabla f(x)\|_2^2 - 2t\langle x - x^*, \nabla f(x)\rangle$.

If $\nabla f(x) = 0$, then the first-order optimality condition of convex function for the minimizer $x^*$ implies $x = x^*$, so $\Delta(t) = 0$ and the inequality holds trivially. Now assume $\nabla f(x) \neq 0$. Set $A = \|\nabla f(x)\|_2^2 > 0$ and $B = \langle x - x^*, \nabla f(x)\rangle$. By convexity of $f$ and the optimality of $x^*$, we have $B \geq 0$. Thus,

$$\Delta(t) = At^2 - 2Bt.$$

This is a quadratic in $t$ with positive leading coefficient $A > 0$, so $\Delta(t) \leq 0$ for $t$ between its roots. The roots are $t = 0$ and $t = 2B/A$, so $\Delta(t) \leq 0$ for all $t \in [0, 2B/A]$.

To ensure $\Delta(t) \leq 0$ for all $t \in [0, 2/L]$, it suffices to show $[0, 2/L] \subseteq [0, 2B/A]$, i.e., $2B/A \geq 2/L$, which simplifies to:

$$B \geq \frac{A}{L}.$$

We now prove this inequality. Since $\nabla f$ is $L$-Lipschitz continuous and $f$ is convex, the following cocoercivity property holds for all $x, y \in \mathbb{R}^n$:

$$\langle \nabla f(x) - \nabla f(y), x - y\rangle \geq \frac{1}{L}\|\nabla f(x) - \nabla f(y)\|_2^2. \tag{2}$$

Setting $y = x^*$ and noting $\nabla f(x^*) = 0$ (by the first-order optimality condition for the minimizer), we obtain:

$$\langle \nabla f(x) - 0, x - x^*\rangle \geq \frac{1}{L}\|\nabla f(x) - 0\|_2^2,$$

which simplifies to:

$$\langle \nabla f(x), x - x^*\rangle \geq \frac{1}{L}\|\nabla f(x)\|_2^2,$$

i.e., $B \geq A/L$.

Therefore, for all $t \in [0, 2/L]$, we have $t \leq 2/L \leq 2B/A$, so $t \in [0, 2B/A]$, implying $\Delta(t) \leq 0$. This completes the proof.

*Justification of* (2)*:* Recall the conclusion of problem 1(b), we have

$$f(y) - f(x) - \langle \nabla f(x), y - x\rangle \geq \frac{1}{2L}\|\nabla f(y) - \nabla f(x)\|^2.$$

Exchanging the order of x and y, we get

$$f(x) - f(y) - \langle \nabla f(y), x - y\rangle \geq \frac{1}{2L}\|\nabla f(x) - \nabla f(y)\|^2.$$

Adding the two equations gives

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \frac{1}{L} \|\nabla f(x) - \nabla f(y)\|^2.$$

□

6.

**Problem.** Find a convex function that is differentiable on an open convex set but not continuously differentiable on the same set — or prove that such a function does not exist.

**Solution**. *Proof.* We prove that no such function exists; that is, if $f : U \to \mathbb{R}$ is convex and differentiable on an open convex set $U \subseteq \mathbb{R}^n$, then $\nabla f$ must be continuous on $U$.

Fix $x_0 \in U$ and a sequence $\{x_k\} \subset U$ with $x_k \to x_0$. Since $f$ is differentiable on $U$, the gradient $\nabla f(x)$ exists at every $x \in U$. We will show $\nabla f(x_k) \to \nabla f(x_0)$.

For any $h \in \mathbb{R}^n$ and sufficiently small $t > 0$ such that $x_0 + th \in U$ and $x_k + th \in U$ for large $k$, convexity implies:

$$\frac{f(x_k + th) - f(x_k)}{t} \geq \langle \nabla f(x_k), h \rangle,$$

as the difference quotient decreases to the directional derivative. Similarly,

$$\frac{f(x_0 + th) - f(x_0)}{t} \geq \langle \nabla f(x_0), h \rangle.$$

By continuity of $f$ (convex functions on open sets are locally Lipschitz), for any $\epsilon > 0$, there exists $K \in \mathbb{N}$ such that for all $k \geq K$ and sufficiently small $t > 0$:

$$\left| \frac{f(x_k + th) - f(x_k)}{t} - \frac{f(x_0 + th) - f(x_0)}{t} \right| < \epsilon.$$

Thus,

$$\langle \nabla f(x_k), h \rangle \leq \frac{f(x_k + th) - f(x_k)}{t} < \frac{f(x_0 + th) - f(x_0)}{t} + \epsilon.$$

As $t \to 0^+$, the right side converges to $\langle \nabla f(x_0), h \rangle + \epsilon$, so:

$$\langle \nabla f(x_k), h \rangle \leq \langle \nabla f(x_0), h \rangle + \epsilon. \quad (1)$$

Applying the same argument to $-h$:

$$\langle \nabla f(x_k), -h \rangle \leq \langle \nabla f(x_0), -h \rangle + \epsilon,$$

which gives:

$$\langle \nabla f(x_k), h \rangle \geq \langle \nabla f(x_0), h \rangle - \epsilon. \quad (2)$$

Combining (1) and (2):

$$|\langle \nabla f(x_k) - \nabla f(x_0), h \rangle| \leq \epsilon.$$

Now let $h = e_j$ (the $j$-th standard basis vector) for $j = 1, \ldots, n$. Then:

$$|\nabla_j f(x_k) - \nabla_j f(x_0)| \leq \epsilon \quad \forall j.$$

Hence, $\nabla f(x_k) \to \nabla f(x_0)$ as $k \to \infty$, proving continuity at $x_0$. As $x_0$ is arbitrary, $\nabla f$ is continuous on $U$.

Therefore, there exists no convex function that is differentiable but not continuously differentiable on an open convex set. □

7.

**Problem.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice continuously differentiable function. Given any $d \in \mathbb{R}^n$ with $\|d\|_2 = 1$, the function $t \mapsto f(td)$ has a local minimum at $t^* = 0$. Is it guaranteed that $f$ has a local minimum at $x^* = 0$?

**Solution**. The conditions do not guarantee that $f$ has a local minimum at 0. We provide a counterexample for $n = 2$. Define the twice continuously differentiable function:

$$f(x, y) = (y - x^2)(y - 2x^2).$$

This function is a polynomial, hence smooth.

We first show that $f$ is locally minimal along any lines. For any unit vector $d = (d_1, d_2)$ with $\|d\|_2 = 1$, define $\phi_d(t) = f(td_1, td_2)$. Substituting:

$$\phi_d(t) = (td_2 - t^2 d_1^2)(td_2 - 2t^2 d_1^2) = d_2^2 t^2 - 3d_1^2 d_2 t^3 + 2d_1^4 t^4.$$

The derivatives are:

$$\phi_d'(t) = 2d_2^2 t - 0d_1^2 d_2 t^2 + 8d_1^4 t^3,$$
$$\phi_d''(t) = 2d_2^2 - 18d_1^2 d_2 t + 24d_1^4 t^2.$$

At $t = 0$:

$$\phi_d'(0) = 0, \quad \phi_d''(0) = 2d_2^2.$$

There are 2 cases.

*Case1:* If $d_2 \neq 0$, then $\phi_d''(0) = 2d_2^2 > 0$, so $\phi_d$ has a strict local minimum at $t = 0$.

*Case2:* If $d_2 = 0$, then $d_1 = \pm 1$ and $\phi_d(t) = 2t^4$. Since $\phi_d(t) \geq 0$ with equality only at $t = 0$, a strict local minimum occurs at $t = 0$.

Thus, $t \mapsto f(td)$ has a local minimum at $t = 0$ for every unit vector $d$.

Next, we show $f$ does not take local minimum at 0. Consider points along the curve $y = \frac{3}{2}x^2$:

$$f\left(x, \tfrac{3}{2}x^2\right) = \left(\tfrac{3}{2}x^2 - x^2\right)\left(\tfrac{3}{2}x^2 - 2x^2\right) = \left(\tfrac{1}{2}x^2\right)\left(-\tfrac{1}{2}x^2\right) = -\tfrac{1}{4}x^4 \leq 0.$$

Since $f(0, 0) = 0$ and $f\left(x, \frac{3}{2}x^2\right) < 0$ for all $x \neq 0$, every neighborhood of 0 contains points where $f$ is negative. Therefore, $f$ does not have a local minimum at 0.

The existence of directional minima along every line through 0 does not imply a local minimum for $f$ at 0.

8.

**Problem.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice continuously differentiable function. Suppose that there exists a unique point $x^* \in \mathbb{R}^n$ such that $\nabla f(x^*) = 0$. In addition, $x^*$ is a local minimizer of $f$. Is it guaranteed that $x^*$ is a global minimizer of $f$?

**Solution**. The assertion is disproven by the following counterexample for $n = 2$. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) = x^2 + y^2(1 - x)^3.$$

This function is twice continuously differentiable since it is a polynomial. We will show that:

- The only critical point is $(x^*, y^*) = (0, 0)$.
- $(0, 0)$ is a strict local minimizer.
- $(0, 0)$ is not a global minimizer, as $f$ attains lower values elsewhere.

The gradient of $f$ is computed as:

$$\begin{aligned}\nabla f(x, y) &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \\ &= \left(2x - 3y^2(1 - x)^2, 2y(1 - x)^3\right).\end{aligned}$$

Set $\nabla f(x, y) = (0, 0)$:

$$2y(1 - x)^3 = 0, \tag{3}$$

$$2x - 3y^2(1 - x)^2 = 0. \tag{4}$$

Equation (3) implies either $y = 0$ or $x = 1$.

*Case1:* If $y = 0$, then equation (4) simplifies to $2x = 0$, so $x = 0$. Thus, $(x, y) = (0, 0)$ is a solution.

*Case2:* If $x = 1$, then equation (4) becomes $2(1) - 3y^2(1 - 1)^2 = 2 \neq 0$ for all $y$, which contradicts $\nabla f = 0$. Hence, no solution exists when $x = 1$.

Therefore, $(0, 0)$ is the unique critical point.

To confirm local minimality, compute the Hessian matrix at $(x, y)$:

$$H_f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix},$$

where

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(2x - 3y^2(1 - x)^2\right) = 2 + 6y^2(1 - x), \\ \frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(2y(1 - x)^3\right) = 2(1 - x)^3, \\ \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(2y(1 - x)^3\right) = -6y(1 - x)^2.\end{aligned}$$

At $(0,0)$:

$$H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

This matrix is positive definite (eigenvalues are both 2, which are positive). Since $f$ is twice continuously differentiable, the second-order sufficient condition implies that $(0,0)$ is a strict local minimizer.

Finally we show that $(0,0)$ is not a global minimum. Consider the point $(2,3)$:

$$f(2,3) = (2)^2 + (3)^2(1-2)^3 = 4 + 9 \cdot (-1) = -5.$$

Since $f(0,0) = 0$ and $f(2,3) = -5 < 0$, the value at $(2,3)$ is less than at $(0,0)$. Thus, $(0,0)$ is not a global minimizer.

9.

**Problem.** Let $\{X_k\}$ be a sequence of independent random variables such that

(a) for each $k \geq 1$, $X_k$ is either 0 or 1;

(b) there exists a constant $p \in (0,1)$ such that $\mathbb{P}(X_k = 1) \geq p$ for each $k \geq 1$.

For all $t \in [0,p]$, prove that

$$\mathbb{P}\left(\sum_{k=1}^{n} X_k \leq tn\right) \leq \exp\left[-\frac{(p-t)^2}{2p}n\right].$$

Provide an interpretation for this bound.

**Solution**. *Proof.* Let $S_n = \sum_{k=1}^{n} X_k$. Since the $X_k$ are independent Bernoulli random variables with success probabilities $p_k \geq p$, we derive the Chernoff's bound. For any $\lambda > 0$,

$$\mathbb{P}(S_n \leq tn) \leq e^{\lambda tn}\mathbb{E}[e^{-\lambda S_n}].$$

By independence,

$$\mathbb{E}[e^{-\lambda S_n}] = \prod_{k=1}^{n} \mathbb{E}[e^{-\lambda X_k}].$$

For each $k$, $\mathbb{E}[e^{-\lambda X_k}] = p_k e^{-\lambda} + (1-p_k)$. Since $p_k \geq p$ and $x \mapsto 1 - x(1-e^{-\lambda})$ is decreasing for $\lambda > 0$,

$$\mathbb{E}[e^{-\lambda X_k}] \leq 1 - p(1-e^{-\lambda}) = 1 - p + pe^{-\lambda}.$$

Thus,

$$\mathbb{P}(S_n \leq tn) \leq e^{\lambda tn}(1-p+pe^{-\lambda})^n = \left[e^{\lambda t}(1-p+pe^{-\lambda})\right]^n.$$

Minimize the expression over $\lambda > 0$. Consider the derivative of $f(\lambda) = \left[e^{\lambda t}(1-p+pe^{-\lambda})\right]$:

$$f'(\lambda) = (1-p)t \cdot e^{\lambda t} + p(t-1) \cdot e^{\lambda(t-1)}.$$

Set $f'(\lambda) = 0$:

$$\lambda^* = \ln\left(\frac{p(1-t)}{(1-p)t}\right).$$

Since $t \in [0, p]$, we can derive $\lambda > \lambda^*, f'(\lambda) > 0$ and $\lambda < \lambda^*, f'(\lambda) < 0$. By substituting this critical point into $f$, the minimum value is:

$$\inf_{\lambda>0}\left[e^{\lambda t}(1-p+pe^{-\lambda})\right] = e^{-d(t\|p)}, \quad \text{where} \quad d(t\|p) = t\ln\frac{t}{p} + (1-t)\ln\frac{1-t}{1-p}.$$

Thus,

$$\mathbb{P}(S_n \le tn) \le \exp\left(-nd(t\|p)\right).$$

We now show $d(t\|p) \ge \frac{(p-t)^2}{2p}$ for $0 \le t \le p$. Define

$$h(t) = d(t\|p) - \frac{(p-t)^2}{2p}.$$

The second derivative of $d(t\|p)$ is $d''(t) = \frac{1}{t} + \frac{1}{1-t}$. By Taylor's theorem with Lagrange remainder (the zero-order term and first-order term are zero),

$$d(t\|p) = \frac{1}{2}d''(\xi)(t-p)^2 \quad \text{for some} \quad \xi \in (t, p).$$

Since $\xi \le p$ and $d''(\xi) = \frac{1}{\xi} + \frac{1}{1-\xi} \ge \frac{1}{p}$ (because $\xi \le p$ implies $\frac{1}{\xi} \ge \frac{1}{p}$),

$$d(t\|p) \ge \frac{1}{2}\cdot\frac{1}{p}(p-t)^2 = \frac{(p-t)^2}{2p}.$$

Therefore $h(t) \ge 0$, which implies

$$\mathbb{P}(S_n \le tn) \le \exp\left(-nd(t\|p)\right) \le \exp\left[-\frac{(p-t)^2}{2p}\,n\right],$$

completing the proof. □

Interpretation:

This bound quantifies the exponential decay rate for the probability that the sum $S_n$ falls below a fraction $t$ of trials when each trial has success probability at least $p > t$. The exponent $-\frac{(p-t)^2}{2p}n$ shows:

- The probability decays exponentially with $n$, reflecting concentration around the minimum mean $pn$.
- The deviation $(p-t)^2$ appears quadratically, describes the effect of the difference between $t$ and $p$ on the declining rate.

This bound is useful for guaranteeing minimum performance in applications like randomized algorithms and statistical testing. Consider a vaccine trial with $n = 100$ participants. Let $p = 0.75$: Minimum probability of immune response per participant.

Set $t = 0.7$: Threshold for acceptable success rate (70% positive responses).
The Chernoff bound gives:

$$\mathbb{P}(\text{Successes} \leq 70) \leq \exp\left[-\frac{(0.75-0.7)^2}{2 \times 0.75} \times 100\right] = e^{-1/6} \approx 0.846$$

This quantifies the risk of the vaccine underperforming the minimum efficacy standard.

10.

**Problem.** Recall that a consistent matrix norm on $\mathbb{R}^{n\times n}$ is a function $\psi : \mathbb{R}^{n\times n} \to \mathbb{R}$ that satisfies the following conditions.

(a) Absolute homogeneity: $\psi(\alpha A) = |\alpha|\psi(A)$ for all $A \in \mathbb{R}^{n\times n}$ and $\alpha \in \mathbb{R}$.

(b) Triangle inequality: $\psi(A + B) \leq \psi(A) + \psi(B)$ for all $A, B \in \mathbb{R}^{n\times n}$.

(c) Positive definiteness: $\psi(A) \geq 0$ for all $A \in \mathbb{R}^{n\times n}$, and $\psi(A) = 0$ if and only if $A = 0$.

(d) Consistency: $\psi(AB) \leq \psi(A)\psi(B)$ for all $A, B \in \mathbb{R}^{n\times n}$.

For any $A \in \mathbb{R}^{n\times n}$, let $\rho(A)$ denote the spectral radius of $A$. Is $\rho$ a consistent matrix norm on $\mathbb{R}^{n\times n}$? If yes, give a proof. Otherwise, which of the four conditions does $\rho$ violate (please name all of them)?

**Solution**. The spectral radius $\rho(A)$ of a matrix $A \in \mathbb{R}^{n\times n}$ is defined as the maximum modulus of its eigenvalues, i.e.,

$$\rho(A) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}.$$

It is **not** a consistent matrix norm because it violates the triangle inequality (b), positive definiteness (c), and consistency (d). However, it satisfies absolute homogeneity (a). We provide counterexamples for each violated condition and a proof for the satisfied condition.

*Proof.* (c) Positive definiteness: $\rho(A) \geq 0$ for all $A$, but $\rho(A) = 0$ does not imply $A = 0$.

Counterexample: Consider $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. The characteristic polynomial is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 \\ 0 & \lambda \end{bmatrix} = \lambda^2.$$

The eigenvalues are $\lambda = 0$ (with multiplicity 2), so $\rho(A) = 0$. However, $A \neq 0$.

(b) Triangle inequality: $\rho(A + B) \leq \rho(A) + \rho(B)$ does not hold for all $A, B$.

Counterexample: Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. Then $\rho(A) = 0$ (eigenvalues $0, 0$) and $\rho(B) = 0$ (eigenvalues $0, 0$). Now,

$$A + B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The characteristic polynomial is $\det(\lambda I - (A + B)) = \lambda^2 - 1$, with eigenvalues $\lambda = \pm 1$. Thus $\rho(A + B) = 1$. However,

$$\rho(A) + \rho(B) = 0 + 0 = 0 < 1 = \rho(A + B),$$

violating the triangle inequality.

(d) Consistency: $\rho(AB) \leq \rho(A)\rho(B)$ does not hold for all $A, B$.

Counterexample: Using the same $A$ and $B$ as above,

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

The eigenvalues are 1 and 0, so $\rho(AB) = 1$. However,

$$\rho(A)\rho(B) = 0 \cdot 0 = 0 < 1 = \rho(AB),$$

violating consistency.

However, condition (a) is satisfied.

(a) Absolute homogeneity: $\rho(\alpha A) = |\alpha|\rho(A)$ for all $\alpha \in \mathbb{R}$ and $A \in \mathbb{R}^{n \times n}$.

Proof: Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A$. Then the eigenvalues of $\alpha A$ are $\alpha\lambda_1, \ldots, \alpha\lambda_n$. The spectral radius is

$$\rho(\alpha A) = \max_{1 \leq i \leq n} |\alpha \lambda_i| = |\alpha| \max_{1 \leq i \leq n} |\lambda_i| = |\alpha|\rho(A).$$

Thus, absolute homogeneity holds.

Since $\rho$ violates conditions (b), (c), and (d), it is not a consistent matrix norm. □

11.

**Problem.** For any $x \in \mathbb{R}^n$, define

$$\|x\|_p = \left[ \sum_{i=1}^{n} |x_i|^p \right]^{1/p}, \quad p \in (0, \infty).$$

(a) Given $p \in (0, 1]$, prove that $\|x + y\|_p^p \leq \|x\|_p^p + \|y\|_p^p$ for all $x, y \in \mathbb{R}^n$.

(b) Given $p \in (0, 1]$, prove that $\|x + y\|_p \leq 2^{\frac{1}{p} - 1}(\|x\|_p + \|y\|_p)$ for all $x, y \in \mathbb{R}^n$.

(c) Given $p \in (0, 1]$, prove that $\|x + y\|_p \geq \|x\|_p + \|y\|_p$ for all $x, y \in \mathbb{R}^n$ whose entries are all nonnegative.

(d) Given $x \in \mathbb{R}^n$, prove that $\|x\|_p$ is a decreasing function of $p \in (0, \infty)$.

(e) Given $x \in \mathbb{R}^n \setminus \{0\}$, prove that $\log \|x\|_p$ is a convex function of $p \in (0, \infty)$.

(f) Recall that, for a matrix $A \in \mathbb{R}^{n\times n}$, $\|A\|_p$ is defined by

$$\|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p.$$

As a function of $p \in (0, +\infty)$, is $\|A\|_p$ increasing, decreasing, or neither?

**Solution**.

(a) *Proof.* First, prove the scalar inequality: for any $a, b \geq 0$ and $p \in (0, 1]$,

$$(a + b)^p \leq a^p + b^p.$$

The function $f(t) = t^p$ is concave on $[0, \infty)$ for $p \in (0, 1]$ because its second derivative $f''(t) = p(p-1)t^{p-2} \leq 0$. Set $\lambda = \frac{a}{a+b}$ and $1-\lambda = \frac{b}{a+b}$. By Jensen's inequality for concave functions,

$$f(\lambda \cdot (a + b) + (1 - \lambda) \cdot 0) \geq \lambda f(a + b) + (1 - \lambda)f(0).$$

Since $f(0) = 0$, this simplifies to

$$a^p \geq \lambda(a + b)^p = \frac{a}{a + b}(a + b)^p,$$

and similarly,

$$b^p \geq (1 - \lambda)(a + b)^p = \frac{b}{a + b}(a + b)^p.$$

Adding these inequalities gives

$$a^p + b^p \geq (a + b)^p.$$

Now, for vectors $x, y \in \mathbb{R}^n$, apply the scalar inequality component-wise. For each $i = 1, \ldots, n$,

$$|x_i + y_i|^p \leq (|x_i| + |y_i|)^p \leq |x_i|^p + |y_i|^p,$$

where the first inequality holds by the triangle inequality for absolute values, and the second is the scalar result. Summing over all components,

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq \sum_{i=1}^{n} (|x_i|^p + |y_i|^p) = \sum_{i=1}^{n} |x_i|^p + \sum_{i=1}^{n} |y_i|^p = \|x\|_p^p + \|y\|_p^p.$$

Thus, $\|x + y\|_p^p \leq \|x\|_p^p + \|y\|_p^p$. □

(b) *Proof.* From part (a), $\|x+y\|_p^p \le \|x\|_p^p + \|y\|_p^p$. Set $a = \|x\|_p$ and $b = \|y\|_p$. It suffices to show that

$$(a^p + b^p)^{1/p} \le 2^{\frac{1}{p}-1}(a+b)$$

for all $a, b \ge 0$. Since the expression is homogeneous, assume without loss of generality that $a + b = 1$ (if $a = b = 0$, the inequality holds trivially; otherwise, scale by $a + b$). Then, we need to show

$$(a^p + b^p)^{1/p} \le 2^{\frac{1}{p}-1}.$$

Given $a + b = 1$ and $b = 1 - a$, define $g(a) = a^p + (1-a)^p$ for $a \in [0, 1]$. The goal is to maximize $g(a)^{1/p}$, but since $p > 0$, maximizing $g(a)$ suffices. Compute the derivative:

$$g'(a) = pa^{p-1} - p(1-a)^{p-1}.$$

Set $g'(a) = 0$, yielding $a^{p-1} = (1-a)^{p-1}$. Since $p-1 \le 0$, this implies $a = 1-a$, so $a = \frac{1}{2}$. Check the second derivative or behavior: for $p \in (0,1)$, $g''(a) = p(p-1)[a^{p-2}+(1-a)^{p-2}] < 0$, so $g(a)$ is concave and has a maximum at $a = \frac{1}{2}$. At this point,

$$g\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^p + \left(\frac{1}{2}\right)^p = 2 \cdot 2^{-p} = 2^{1-p}.$$

Thus, $a^p + b^p \le 2^{1-p}$ when $a + b = 1$. Therefore,

$$(a^p + b^p)^{1/p} \le (2^{1-p})^{1/p} = 2^{\frac{1}{p}-1}.$$

Applying this to $a = \|x\|_p$ and $b = \|y\|_p$,

$$\|x\|_p^p + \|y\|_p^p \le 2^{1-p}(\|x\|_p + \|y\|_p)^p.$$

Combining with part (a),

$$\|x+y\|_p^p \le \|x\|_p^p + \|y\|_p^p \le 2^{1-p}(\|x\|_p + \|y\|_p)^p.$$

Taking the $p$-th root (which preserves inequalities since $p > 0$),

$$\|x+y\|_p \le 2^{\frac{1}{p}-1}(\|x\|_p + \|y\|_p).$$

□

(c) *Proof.* Assume $x, y \in \mathbb{R}^n$ have nonnegative entries, so $|x_i| = x_i$ and $|y_i| = y_i$. The function $f(t) = t^p$ is concave on $[0, \infty)$ for $p \in (0, 1]$. For each component $i$ and any $\lambda \in (0, 1)$, by concavity,

$$\begin{aligned}(x_i + y_i)^p &= f\left(\lambda \cdot \frac{x_i}{\lambda} + (1-\lambda) \cdot \frac{y_i}{1-\lambda}\right) \\ &\ge \lambda f\left(\frac{x_i}{\lambda}\right) + (1-\lambda) f\left(\frac{y_i}{1-\lambda}\right) \\ &= \lambda \left(\frac{x_i}{\lambda}\right)^p + (1-\lambda)\left(\frac{y_i}{1-\lambda}\right)^p.\end{aligned}$$

Simplifying,

$$(x_i + y_i)^p \geq \lambda^{1-p} x_i^p + (1-\lambda)^{1-p} y_i^p.$$

Sum over all components:

$$\begin{aligned}\|x+y\|_p^p &= \sum_{i=1}^n (x_i + y_i)^p \\ &\geq \sum_{i=1}^n \left[\lambda^{1-p} x_i^p + (1-\lambda)^{1-p} y_i^p\right] \\ &= \lambda^{1-p}\|x\|_p^p + (1-\lambda)^{1-p}\|y\|_p^p.\end{aligned}$$

Now, choose $\lambda = \frac{\|x\|_p}{\|x\|_p + \|y\|_p}$ and $1-\lambda = \frac{\|y\|_p}{\|x\|_p + \|y\|_p}$. If $\|x\|_p = \|y\|_p = 0$, the inequality holds trivially; otherwise, $\lambda \in (0,1)$. Substituting,

$$\|x+y\|_p^p \geq \left(\frac{\|x\|_p}{\|x\|_p + \|y\|_p}\right)^{1-p} \|x\|_p^p + \left(\frac{\|y\|_p}{\|x\|_p + \|y\|_p}\right)^{1-p} \|y\|_p^p.$$

Simplify the right-hand side:

$$\left(\frac{\|x\|_p}{\|x\|_p + \|y\|_p}\right)^{1-p} \|x\|_p^p = \|x\|_p^p \cdot \frac{(\|x\|_p + \|y\|_p)^{p-1}}{\|x\|_p^{p-1}} = \|x\|_p \cdot (\|x\|_p + \|y\|_p)^{p-1},$$

and similarly for the other term,

$$\left(\frac{\|y\|_p}{\|x\|_p + \|y\|_p}\right)^{1-p} \|y\|_p^p = \|y\|_p \cdot (\|x\|_p + \|y\|_p)^{p-1}.$$

Adding these,

$$\|x+y\|_p^p \geq [\|x\|_p + \|y\|_p]\,(\|x\|_p + \|y\|_p)^{p-1} = (\|x\|_p + \|y\|_p)^p.$$

Taking the $p$-th root (and noting it is increasing),

$$\|x+y\|_p \geq \|x\|_p + \|y\|_p.$$

□

(d) *Proof.* Let $x \in \mathbb{R}^n$. If $x = 0$, then $\|x\|_p = 0$ for all $p$, so the function is constant (hence decreasing). Assume $x \neq 0$. To show $\|x\|_p$ is decreasing in $p$, fix $p_1, p_2$ with $0 < p_1 < p_2 < \infty$. We need $\|x\|_{p_2} \leq \|x\|_{p_1}$.

By homogeneity, normalize so that $\|x\|_{p_1} = 1$. Specifically, define $y = \frac{x}{\|x\|_{p_1}}$. Then

$$\|y\|_{p_1} = \left(\sum_{i=1}^n \left|\frac{x_i}{\|x\|_{p_1}}\right|^{p_1}\right)^{1/p_1} = \left(\frac{\sum_{i=1}^n |x_i|^{p_1}}{\|x\|_{p_1}^{p_1}}\right)^{1/p_1} = \left(\frac{\|x\|_{p_1}^{p_1}}{\|x\|_{p_1}^{p_1}}\right)^{1/p_1} = 1.$$

Since $\|y\|_{p_1} = 1$, we have $\sum_{i=1}^n |y_i|^{p_1} = 1$. For each component, since $|y_i|^{p_1} \leq 1$ (because if $|y_i| > 1$ for some $i$, then $\sum |y_j|^{p_1} > 1$, contradiction), and since $p_2 > p_1 > 0$, the function $t \mapsto t^{p_2/p_1}$ is increasing for $t \geq 0$. Thus,

$$|y_i|^{p_2} = (|y_i|^{p_1})^{p_2/p_1} \leq |y_i|^{p_1},$$

because $0 \leq |y_i|^{p_1} \leq 1$ and $p_2/p_1 > 1$. Summing over components,

$$\sum_{i=1}^n |y_i|^{p_2} \leq \sum_{i=1}^n |y_i|^{p_1} = 1.$$

Therefore,

$$\|y\|_{p_2}^{p_2} = \sum_{i=1}^n |y_i|^{p_2} \leq 1 = \|y\|_{p_1}^{p_1}.$$

Taking the $p_2$-th root (and noting it preserves inequality),

$$\|y\|_{p_2} \leq 1 = \|y\|_{p_1}.$$

Since $y = x/\|x\|_{p_1}$, by homogeneity of norms,

$$\left\| \frac{x}{\|x\|_{p_1}} \right\|_{p_2} \leq \left\| \frac{x}{\|x\|_{p_1}} \right\|_{p_1} \implies \frac{\|x\|_{p_2}}{\|x\|_{p_1}} \leq 1 \implies \|x\|_{p_2} \leq \|x\|_{p_1}.$$

Thus, $\|x\|_p$ is decreasing in $p$. □

(e) *Proof.* Fix a non-zero vector $x \in \mathbb{R}^n \setminus \{0\}$. Define the function $\phi(p) = \log \|x\|_p = \frac{1}{p} \log \left(\sum_{i=1}^n |x_i|^p\right)$ for $p \in (0, \infty)$. To prove convexity, we must show that for any $p_1, p_2 > 0$ and $\lambda \in (0, 1)$,

$$\phi(\lambda p_1 + (1-\lambda)p_2) \leq \lambda \phi(p_1) + (1-\lambda)\phi(p_2).$$

Without loss of generality, assume $p_1 \leq p_2$ (the case $p_1 > p_2$ follows by symmetry). Let $r = \lambda p_1 + (1-\lambda)p_2$. Since $p_1 \leq p_2$ and $\lambda \in (0, 1)$, we have $p_1 \leq r \leq p_2$. Then

$$\phi(r) = \frac{1}{r} \log \left( \sum_{i=1}^n |x_i|^r \right) = \frac{1}{r} \log \left( \sum_{i=1}^n |x_i|^{\lambda p_1 + (1-\lambda)p_2} \right).$$

Note that $|x_i|^{\lambda p_1 + (1-\lambda)p_2} = |x_i|^{\lambda p_1} \cdot |x_i|^{(1-\lambda)p_2}$. By Hölder's inequality with conjugate exponents $m = 1/\lambda$ and $k = 1/(1-\lambda)$ (satisfying $1/m + 1/k = \lambda + (1-\lambda) = 1$),

$$\sum_{i=1}^n a_i b_i \leq \left( \sum_{i=1}^n a_i^m \right)^{1/m} \left( \sum_{i=1}^n b_i^k \right)^{1/k},$$

where $a_i = |x_i|^{\lambda p_1}$ and $b_i = |x_i|^{(1-\lambda)p_2}$. Substituting,

$$\sum_{i=1}^n |x_i|^{\lambda p_1} |x_i|^{(1-\lambda)p_2} \leq \left( \sum_{i=1}^n \left( |x_i|^{\lambda p_1} \right)^m \right)^{1/m} \left( \sum_{i=1}^n \left( |x_i|^{(1-\lambda)p_2} \right)^k \right)^{1/k}.$$

Simplifying the exponents:

$$\left(|x_i|^{\lambda p_1}\right)^m = |x_i|^{\lambda p_1 \cdot 1/\lambda} = |x_i|^{p_1}, \quad \left(|x_i|^{(1-\lambda)p_2}\right)^k = |x_i|^{(1-\lambda)p_2 \cdot 1/(1-\lambda)} = |x_i|^{p_2},$$

so

$$\sum_{i=1}^n |x_i|^r \leq \left(\sum_{i=1}^n |x_i|^{p_1}\right)^\lambda \left(\sum_{i=1}^n |x_i|^{p_2}\right)^{1-\lambda}.$$

Since the logarithm is monotonic increasing,

$$\begin{aligned}\log\left(\sum_{i=1}^n |x_i|^r\right) &\leq \log\left(\left(\sum_{i=1}^n |x_i|^{p_1}\right)^\lambda \left(\sum_{i=1}^n |x_i|^{p_2}\right)^{1-\lambda}\right) \\ &= \lambda \log\left(\sum_{i=1}^n |x_i|^{p_1}\right) + (1-\lambda)\log\left(\sum_{i=1}^n |x_i|^{p_2}\right).\end{aligned}$$

By the definition of $\phi$,

$$\log\left(\sum_{i=1}^n |x_i|^{p_1}\right) = p_1\phi(p_1), \quad \log\left(\sum_{i=1}^n |x_i|^{p_2}\right) = p_2\phi(p_2),$$

so

$$\phi(r) \leq \frac{1}{r}\left(\lambda p_1\phi(p_1) + (1-\lambda)p_2\phi(p_2)\right) = \frac{\lambda p_1}{r}\phi(p_1) + \frac{(1-\lambda)p_2}{r}\phi(p_2).$$

Let $w_1 = \frac{\lambda p_1}{r}$ and $w_2 = \frac{(1-\lambda)p_2}{r}$. Since $r = \lambda p_1 + (1-\lambda)p_2$, we have $w_1 + w_2 = 1$ with $w_1, w_2 \geq 0$. Thus,

$$\phi(r) \leq w_1\phi(p_1) + w_2\phi(p_2).$$

We now prove $w_1\phi(p_1) + w_2\phi(p_2) \leq \lambda\phi(p_1) + (1-\lambda)\phi(p_2)$. Consider the difference:

$$\lambda\phi(p_1) + (1-\lambda)\phi(p_2) - (w_1\phi(p_1) + w_2\phi(p_2)) = (\lambda - w_1)\phi(p_1) + (1-\lambda-w_2)\phi(p_2).$$

Substituting $w_1$ and $w_2$:

$$\lambda - w_1 = \lambda\left(1 - \frac{p_1}{r}\right) = \lambda \cdot \frac{r-p_1}{r}, \quad 1-\lambda-w_2 = (1-\lambda)\left(1-\frac{p_2}{r}\right) = (1-\lambda)\cdot\frac{r-p_2}{r}.$$

Since $p_1 \leq p_2$,

$$r - p_1 = (1-\lambda)(p_2 - p_1) \geq 0, \quad r - p_2 = \lambda(p_1 - p_2) \leq 0.$$

The function $p \mapsto \|x\|_p$ is decreasing for fixed $x \neq 0$, so $\phi(p) = \log\|x\|_p$ is also decreasing. Thus, $\phi(p_1) \geq \phi(p_2)$. Now,

$$(\lambda - w_1)\phi(p_1) + (1-\lambda-w_2)\phi(p_2) = \underbrace{\lambda\frac{r-p_1}{r}}_{\geq 0}\phi(p_1) + \underbrace{(1-\lambda)\frac{r-p_2}{r}}_{\leq 0}\phi(p_2).$$

Using $\phi(p_1) \geq \phi(p_2)$,

$$\begin{aligned}\lambda\frac{r-p_1}{r}\phi(p_1) + (1-\lambda)\frac{r-p_2}{r}\phi(p_2) &\geq \lambda\frac{r-p_1}{r}\phi(p_2) + (1-\lambda)\frac{r-p_2}{r}\phi(p_2) \\ &= \phi(p_2)\cdot\frac{1}{r}\Big[\lambda(r-p_1) + (1-\lambda)(r-p_2)\Big].\end{aligned}$$

Simplifying the expression in brackets:

$$\lambda(r-p_1)+(1-\lambda)(r-p_2) = \lambda r-\lambda p_1+r-\lambda r-(1-\lambda)p_2 = r-(\lambda p_1+(1-\lambda)p_2) = r-r = 0.$$

Thus,

$$\phi(p_2)\cdot\frac{0}{r} = 0,$$

so

$$(\lambda - w_1)\phi(p_1) + (1-\lambda-w_2)\phi(p_2) \geq 0,$$

which implies

$$w_1\phi(p_1) + w_2\phi(p_2) \leq \lambda\phi(p_1) + (1-\lambda)\phi(p_2).$$

Combining inequalities,

$$\phi(r) \leq w_1\phi(p_1) + w_2\phi(p_2) \leq \lambda\phi(p_1) + (1-\lambda)\phi(p_2).$$

This holds for all $p_1, p_2 > 0$ and $\lambda \in (0,1)$, with equality cases handled trivially. Therefore, $\phi(p) = \log \|x\|_p$ is convex on $(0,\infty)$. □

(f) The operator norm $\|A\|_p$ is *neither* monotonic increasing nor decreasing in $p$. Counterexamples:

- *Identity matrix:* $A = I_2$. Then $\|A\|_p = 1$ for all $p$, constant.
- *Non-monotonic case 1*: $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

$$\|A\|_1 = \max_j \sum_i |a_{ij}| = \max(1,2) = 2,$$

$$\|A\|_2 = \sqrt{\rho(A^TA)} = \sqrt{\frac{3+\sqrt{5}}{2}} \approx 1.618,$$

$$\|A\|_\infty = \max_i \sum_j |a_{ij}| = \max(2,1) = 2.$$

  Thus $\|A\|_1 = 2 > \|A\|_2 \approx 1.618 < \|A\|_\infty = 2$.

- *Non-monotonic case 2*: $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$

$$\|A\|_1 = \max(3,7) = 7,$$

$$\|A\|_2 = \sqrt{\frac{23+\sqrt{221}}{2}} \approx 5.549,$$

$$\|A\|_\infty = \max(5,5) = 5.$$

  Thus $\|A\|_1 = 7 > \|A\|_2 \approx 5.549 > \|A\|_\infty = 5$.

The norms exhibit non-monotonic behavior in both directions.

12.

**Problem.** For any matrix $A \in \mathbb{R}^{n\times n}$ and any vector $x \in \mathbb{R}^n$, prove that $\max_{\|d\|\le 1} \|A(x+d)\| \ge \|A\|$. Here, $\|\cdot\|$ denotes a vector norm on $\mathbb{R}^n$ and the operator norm on $\mathbb{R}^{n\times n}$ induced by this vector norm.

**Solution**. *Proof.* Let $\|A\|$ be the operator norm induced by the vector norm, defined as

$$\|A\| = \max_{\|z\|=1} \|Az\|.$$

By the definition of the maximum, there exists a vector $y \in \mathbb{R}^n$ with $\|y\| = 1$ such that

$$\|Ay\| = \|A\|.$$

Consider the vectors $x+y$ and $x-y$. We claim that at least one of the following inequalities holds:

$$\|A(x+y)\| \ge \|A\| \quad \text{or} \quad \|A(x-y)\| \ge \|A\|.$$

Suppose, for contradiction, that both inequalities are false, i.e.,

$$\|A(x+y)\| < \|A\| \quad \text{and} \quad \|A(x-y)\| < \|A\|.$$

By the triangle inequality for the vector norm,

$$\begin{aligned} \|A(x+y) + A(x-y)\| &\le \|A(x+y)\| + \|A(x-y)\|, \\ \|2Ax\| &< \|A\| + \|A\| = 2\|A\|, \end{aligned}$$

which implies $\|Ax\| < \|A\|$. Similarly,

$$\begin{aligned} \|A(x+y) - A(x-y)\| &\le \|A(x+y)\| + \|A(x-y)\|, \\ \|2Ay\| &< \|A\| + \|A\| = 2\|A\|, \end{aligned}$$

which implies $\|Ay\| < \|A\|$. However, this contradicts the choice of $y$ since $\|Ay\| = \|A\|$.

Thus, for at least one choice of $c \in \{-1, 1\}$, we have

$$\|A(x+cy)\| \ge \|A\|.$$

Now, define $x = cy$. Then $\|x\| = \|cy\| = |c|\|y\| = 1 \le 1$, and

$$\|A(x+x)\| = \|A(x+cy)\| \ge \|A\|.$$

Since $x$ satisfies $\|x\| \le 1$, and the maximum over a set is at least any value in the set, we conclude

$$\max_{\|x\|\le 1} \|A(x+x)\| \ge \|A(x+x)\| \ge \|A\|.$$

□

13.

**Problem.** Consider matrices $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$.

(a) Show that $AB$ and $BA$ share the same set of nonzero eigenvalues.

Optional Requirements:

- Give a proof without using determinants or matrix decomposition.
- Give a proof from a geometric point of view.
- Give a proof from an algebraic point of view.

(b) If $\lambda$ is a nonzero eigenvalue of $AB$ and $BA$, show that the geometric multiplicity of $\lambda$ is the same with respect to $AB$ and $BA$.

(c) Prove the same conclusion as above for the algebraic multiplicity.

**Solution**.

(a) *Proof. Algebraic proof (without determinants):*
Let $\lambda \neq 0$ be an eigenvalue of $AB$ and let $x \in \mathbb{C}^m$ be a corresponding eigenvector. By definition, $x \neq 0$ and $ABx = \lambda x$.

Left-multiplying by $B$, we get $B(ABx) = B(\lambda x)$, which can be rewritten as:

$$(BA)(Bx) = \lambda(Bx).$$

We must show that the vector $Bx$ is nonzero. Assume for contradiction that $Bx = 0$. Then the eigenvalue equation for $AB$ becomes $A(Bx) = A(0) = 0$. This implies $\lambda x = 0$. Since $x$ is an eigenvector and thus nonzero, we must have $\lambda = 0$. This contradicts our assumption that $\lambda \neq 0$. Therefore, $Bx \neq 0$.

This shows that $Bx$ is an eigenvector of $BA$ corresponding to the same eigenvalue $\lambda$. Thus, any nonzero eigenvalue of $AB$ is also an eigenvalue of $BA$.

The converse holds by symmetry. If we start with an eigenvalue $\mu \neq 0$ of $BA$ with eigenvector $y \in \mathbb{C}^n$, the same argument shows that $Ay$ is a nonzero eigenvector of $AB$ for the same eigenvalue $\mu$. Consequently, $AB$ and $BA$ have the same set of nonzero eigenvalues. □

(b) *Proof.* Let $\lambda \neq 0$ be an eigenvalue of both $AB$ and $BA$. Let $E_\lambda(AB)$ and $E_\lambda(BA)$ denote the corresponding eigenspaces. The geometric multiplicity, denoted by $\mathrm{gm}(\lambda)$, is the dimension of the eigenspace. We aim to show that $\mathrm{gm}_{AB}(\lambda) = \mathrm{gm}_{BA}(\lambda)$.

First, we will show that $\mathrm{gm}_{AB}(\lambda) \leq \mathrm{gm}_{BA}(\lambda)$. Let $k = \mathrm{gm}_{AB}(\lambda)$, and let $\{x_1, x_2, \ldots, x_k\}$ be a basis for the eigenspace $E_\lambda(AB)$. By definition, for each $x_i$, we have $ABx_i = \lambda x_i$.

Consider the set of $k$ vectors $\{Bx_1, Bx_2, \ldots, Bx_k\}$. These vectors lie in the eigenspace $E_\lambda(BA)$, since for each $i$:

$$(BA)(Bx_i) = B(ABx_i) = B(\lambda x_i) = \lambda(Bx_i).$$

Now, we must show that this set of eigenvectors is linearly independent. Suppose there exist scalars $c_1, \ldots, c_k$ such that:

$$\sum_{i=1}^{k} c_i B x_i = 0.$$

Left-multiplying by $A$ gives:

$$A\left(\sum_{i=1}^{k} c_i B x_i\right) = A(0) \implies \sum_{i=1}^{k} c_i (AB x_i) = 0.$$

Substituting $ABx_i = \lambda x_i$, we have:

$$\sum_{i=1}^{k} c_i (\lambda x_i) = \lambda \left(\sum_{i=1}^{k} c_i x_i\right) = 0.$$

Since $\lambda \neq 0$, it follows that $\sum_{i=1}^{k} c_i x_i = 0$. As $\{x_1, \ldots, x_k\}$ is a basis, it is a linearly independent set, which implies that all scalars must be zero, i.e., $c_i = 0$ for all $i$.

Therefore, the set $\{Bx_1, \ldots, Bx_k\}$ is a linearly independent set of $k$ vectors in $E_\lambda(BA)$. This implies that the dimension of $E_\lambda(BA)$ must be at least $k$. Thus,

$$\mathrm{gm}_{AB}(\lambda) = k \leq \dim(E_\lambda(BA)) = \mathrm{gm}_{BA}(\lambda).$$

The reverse inequality, $\mathrm{gm}_{BA}(\lambda) \leq \mathrm{gm}_{AB}(\lambda)$, follows from a symmetrical argument by swapping the roles of $A$ and $B$.

Combining the two inequalities, we conclude that $\mathrm{gm}_{AB}(\lambda) = \mathrm{gm}_{BA}(\lambda)$. □

(c) *Proof.* We only need to show that $AB$ and $BA$ have the same characteristic polynomials.

Define

$$P = \begin{pmatrix} I & O \\ -A & I \end{pmatrix}, \quad Q = \begin{pmatrix} I & \lambda B \\ A & I \end{pmatrix}.$$

Then

$$PQ = \begin{pmatrix} I & \lambda B \\ O & I - \lambda AB \end{pmatrix}, \quad QP = \begin{pmatrix} I - \lambda BA & \lambda B \\ O & I \end{pmatrix}.$$

Since $|PQ| = |QP|$, it follows that

$$|I - \lambda AB| = |I - \lambda BA|.$$

Since the characteristic polynomials are the same, $AB$ and $BA$ have the same algebraic multiplicity.

□

14.

**Problem.** Consider a polynomial $p \in \mathbb{C}[x]$ and a matrix $A \in \mathbb{C}^{n \times n}$.

(a) For any $\lambda \in \mathbb{C}$, show that $\lambda$ is an eigenvalue of $A$ if and only if $p(\lambda)$ is an eigenvalue of $p(A)$.

Optional Requirements:

- Give a proof without using determinants or matrix decomposition.
- Give a proof from a geometric point of view.
- Give a proof from an algebraic point of view.

(b) Suppose that the eigenvalues of $A$ are $\lambda_1, \lambda_2, \ldots, \lambda_n$, multiple eigenvalues counted with multiplicity. Show that the eigenvalues of $p(A)$ are $p(\lambda_1), p(\lambda_2), \ldots, p(\lambda_n)$, multiple eigenvalues counted with multiplicity.

**Solution**.

(a) *Proof.* ($\Rightarrow$) Let $\lambda$ be an eigenvalue of $A$. By definition, there exists a non-zero eigenvector $v \in \mathbb{C}^n$, such that $Av = \lambda v$. It is obvious that $A^k v = \lambda^k v$ for all $k \geq 0$.

Let the polynomial be $p(x) = \sum_{k=0}^{m} c_k x^k$ for $c_k \in \mathbb{C}$. The corresponding matrix polynomial is $p(A) = \sum_{k=0}^{m} c_k A^k$.

Applying $p(A)$ to the eigenvector $v$:

$$\begin{aligned} p(A)v &= \left(\sum_{k=0}^{m} c_k A^k\right) v = \sum_{k=0}^{m} c_k (A^k v) \\ &= \sum_{k=0}^{m} c_k (\lambda^k v) = \left(\sum_{k=0}^{m} c_k \lambda^k\right) v \\ &= p(\lambda)v. \end{aligned}$$

Since $v$ is a non-zero vector, the relation $p(A)v = p(\lambda)v$ shows that $p(\lambda)$ is an eigenvalue of $p(A)$, with $v$ as a corresponding eigenvector.

($\Leftarrow$) Let $\mu$ be an eigenvalue of $p(A)$. By definition, the matrix $p(A) - \mu I$ is singular.

Consider a new polynomial $q(x) = p(x) - \mu$. Then $q(A) = p(A) - \mu I$ is a singular matrix.

By the Fundamental Theorem of Algebra, $q(x)$ can be factored over $\mathbb{C}$ into linear terms. If $\deg(p) = m$, then

$$q(x) = c \prod_{j=1}^{m} (x - r_j),$$

where $c$ is the leading coefficient and $r_1, \ldots, r_m$ are the roots of $q(x)$. By definition of a root, $q(r_j) = 0$, which implies $p(r_j) = \mu$ for all $j \in \{1, \ldots, m\}$.

Substituting the matrix $A$ into the polynomial $q(x)$ yields the matrix expression:

$$q(A) = p(A) - \mu I = c \prod_{j=1}^{m} (A - r_j I),$$

Since $q(A)$ is singular, its determinant is zero. Using the property $\det(XY) = \det(X)\det(Y)$:

$$\det(q(A)) = \det\left(c\prod_{j=1}^{m}(A - r_j I)\right) = c^n \prod_{j=1}^{m}\det(A - r_j I) = 0.$$

For this product of scalars to be zero, at least one of its factors must be zero. Thus, there must exist an index $j_0 \in \{1, \ldots, m\}$ such that

$$\det(A - r_{j_0} I) = 0.$$

This equation implies that $r_{j_0}$ is an eigenvalue of $A$. Let us set $\lambda = r_{j_0}$.

We know that for this root $r_{j_0}$, we have $p(r_{j_0}) = \mu$. Therefore, we have found an eigenvalue $\lambda$ of $A$ such that $p(\lambda) = \mu$. □

(b) *Proof.* Let the eigenvalues of the matrix $A \in \mathbb{C}^{n\times n}$ be $\lambda_1, \lambda_2, \ldots, \lambda_n$, counted with their algebraic multiplicities. By the Jordan Normal Form theorem, any square matrix with complex entries is similar to a Jordan matrix. Thus, there exists an invertible matrix $P$ and a Jordan matrix $J$ such that:

$$A = PJP^{-1}$$

The Jordan matrix $J$ is a block diagonal matrix, where the diagonal entries are the eigenvalues of $A$. Specifically,

$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix}.$$

Each block $J_i$ is a Jordan block corresponding to an eigenvalue $\lambda_i$ of the form:

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}.$$

Let $p(x) = \sum_{k=0}^{m} c_k x^k$ be a polynomial in $\mathbb{C}[x]$. We apply this polynomial to the matrix $A$.

$$p(A) = \sum_{k=0}^{m} c_k A^k.$$

Substituting $A = PJP^{-1}$, we observe that for any non-negative integer $k$, $A^k = (PJP^{-1})^k = PJ^kP^{-1}$. Therefore,

$$p(A) = \sum_{k=0}^{m} c_k (PJ^kP^{-1}) = P\left(\sum_{k=0}^{m} c_k J^k\right)P^{-1} = Pp(J)P^{-1}.$$

This equation shows that the matrix $p(A)$ is similar to the matrix $p(J)$. Similar matrices have the same characteristic polynomial, and thus the same eigenvalues with the same algebraic multiplicities. Our task is now reduced to finding the eigenvalues of $p(J)$.

Since $J$ is a block diagonal matrix, any polynomial in $J$ is also a block diagonal matrix, with the polynomial applied to each block:

$$p(J) = \begin{pmatrix} p(J_1) & 0 & \cdots & 0 \\ 0 & p(J_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p(J_k) \end{pmatrix}.$$

The eigenvalues of a block diagonal matrix are the union of the eigenvalues of its diagonal blocks. We therefore need to determine the eigenvalues of each block $p(J_i)$.

Each Jordan block $J_i$ is an upper triangular matrix. Any polynomial of an upper triangular matrix is also an upper triangular matrix. The eigenvalues of an upper triangular matrix are its diagonal entries. Let's find the diagonal entries of $p(J_i)$. The diagonal entries of $J_i$ are all $\lambda_i$. For any power $k$, the diagonal entries of the upper triangular matrix $J_i^k$ are all $\lambda_i^k$. Consequently, the diagonal entries of $p(J_i) = \sum_{k=0}^{m} c_k J_i^k$ are all equal to $\sum_{k=0}^{m} c_k \lambda_i^k = p(\lambda_i)$.

Since $p(J_i)$ is an upper triangular matrix, its eigenvalues are its diagonal entries, which are all $p(\lambda_i)$. The size of the block $p(J_i)$ is the same as the size of $J_i$, so the algebraic multiplicity is preserved.

By combining the eigenvalues from all blocks $p(J_i)$, we find that the eigenvalues of $p(J)$ are precisely $\{p(\lambda_1), p(\lambda_2), \ldots, p(\lambda_n)\}$, counted with multiplicity. Since $p(A)$ has the same eigenvalues as $p(J)$, this completes the proof. □

15.

**Problem.** Let $n > 1$. Define $A \in \mathbb{R}^{n \times n}$ to be the matrix with entries

$$A_{i,j} = \begin{cases} 1 & \text{if } i = j,\ i, j = 1, 2, \ldots, n, \\ x & \text{if } i \neq j. \end{cases}$$

(a) Find the eigenvalues of $A$. Specify their multiplicities.

(b) Prove that $A$ is positive definite if and only if $-1/(n-1) < x < 1$.

**Solution**.

(a) *Proof. Proof based on matrix decomposition:*

The matrix $A$ can be expressed as:

$$A = (1 - x)I_n + x\mathbf{1}\mathbf{1}^\top$$

where $I_n$ is the $n \times n$ identity matrix and $\mathbf{1} = (1, 1, \ldots, 1)^\top$.

Consider the eigenvalues:

- For eigenvector $v_1 = \mathbf{1}$:

  $A\mathbf{1} = [(1-x)I_n + x\mathbf{1}\mathbf{1}^\top]\mathbf{1} = (1-x)\mathbf{1} + x\mathbf{1}(\mathbf{1}^\top\mathbf{1}) = (1-x)\mathbf{1} + nx\mathbf{1} = [1+(n-1)x]\mathbf{1}.$

  Thus $\lambda_1 = 1 + (n-1)x$ is an eigenvalue.

- For any vector $v \perp \mathbf{1}$ (i.e., $\mathbf{1}^\top v = 0$):

  $$Av = [(1-x)I_n + x\mathbf{1}\mathbf{1}^\top]v = (1-x)v + x\mathbf{1}(\mathbf{1}^\top v) = (1-x)v.$$

  Thus $\lambda_2 = 1 - x$ is an eigenvalue. The eigenspace has dimension $n-1$ since $\{v : \mathbf{1}^\top v = 0\}$ is $(n-1)$-dimensional.

The eigenvalues are:

$$\lambda_1 = 1 + (n-1)x \quad \text{(multiplicity 1)}, \qquad \lambda_2 = 1 - x \quad \text{(multiplicity } n-1\text{)}.$$

*Proof based on determinant's computation:*

The characteristic polynomial is given by $\det(\lambda I - A)$. Consider:

$$\lambda I - A = \begin{bmatrix} \lambda - 1 & -x & \cdots & -x \\ -x & \lambda - 1 & \cdots & -x \\ \vdots & \vdots & \ddots & \vdots \\ -x & -x & \cdots & \lambda - 1 \end{bmatrix}.$$

Add all rows to the first row:

$$\begin{bmatrix} \lambda - 1 - (n-1)x & \lambda - 1 - (n-1)x & \cdots & \lambda - 1 - (n-1)x \\ -x & \lambda - 1 & \cdots & -x \\ \vdots & \vdots & \ddots & \vdots \\ -x & -x & \cdots & \lambda - 1 \end{bmatrix}.$$

Factor out $\lambda - 1 - (n-1)x$ from the first row:

$$= (\lambda - 1 - (n-1)x) \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -x & \lambda - 1 & \cdots & -x \\ \vdots & \vdots & \ddots & \vdots \\ -x & -x & \cdots & \lambda - 1 \end{bmatrix}.$$

Subtract $x$ times the first row from each subsequent row:

$$= (\lambda - 1 - (n-1)x) \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & \lambda - 1 + x & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda - 1 + x \end{bmatrix}.$$

The determinant is now:

$$\det(\lambda I - A) = (\lambda - 1 - (n-1)x) \cdot \det \begin{bmatrix} \lambda - 1 + x & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda - 1 + x \end{bmatrix}_{(n-1)\times(n-1)}.$$

Thus:

$$\det(\lambda I - A) = (\lambda - 1 - (n-1)x)\,(\lambda - 1 + x)^{n-1}.$$

Setting equal to zero gives eigenvalues:

$$\lambda_1 = 1 + (n-1)x \quad \text{(multiplicity 1)}, \qquad \lambda_2 = 1 - x \quad \text{(multiplicity } n-1\text{)}.$$

□

(b) *Proof.* For positive definiteness, all eigenvalues must be positive:

$$\begin{aligned} & 1 + (n-1)x > 0 \quad \text{and} \quad 1 - x > 0 \\ \iff & x > -\frac{1}{n-1} \quad \text{and} \quad x < 1 \\ \iff & -\frac{1}{n-1} < x < 1. \end{aligned}$$

Thus the strict inequalities are necessary and sufficient. □

16.

**Problem.** Suppose that $m \geq n$. Define $\mathcal{S} = \{X \in \mathbb{C}^{m\times n} : X^H X = I_n\}$. Given $X \in \mathbb{C}^{m\times n}$, let $\operatorname{dist}(X, \mathcal{S})$ be the distance from $X$ to $\mathcal{S}$ in Frobenius norm.

(a) Prove that $\operatorname{dist}(X, \mathcal{S}) \leq \|I_n - X^H X\|_F$.

(b) Prove that there does not exist a constant $C$ such that $\|I_n - X^H X\|_F \leq C \operatorname{dist}(X, \mathcal{S})$ for all $X \in \mathbb{C}^{m\times n}$.

**Solution**.

(a) *Proof.* Let $X \in \mathbb{C}^{m\times n}$. We use the thin singular value decomposition (SVD) of $X$,

$$X = U\Sigma V^{\mathrm{H}},$$

where $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$ is a diagonal matrix with non-negative singular values $\sigma_i \geq 0$, and $U, V \in \mathbb{C}^{n\times n}$ are unitary matrices (i.e., $UU^{\mathrm{H}} = VV^{\mathrm{H}} = I_n$).

Consider the matrix $Y = UV^{\mathrm{H}}$. We first show that $Y \in \mathcal{S}$.

$$Y^{\mathrm{H}}Y = (UV^{\mathrm{H}})^{\mathrm{H}}(UV^{\mathrm{H}}) = VU^{\mathrm{H}}UV^{\mathrm{H}} = VI_nV^{\mathrm{H}} = VV^{\mathrm{H}} = I_n.$$

Thus, $Y$ is an element of $\mathcal{S}$.

By the definition of the distance from a point to a set, we have

$$\operatorname{dist}(X, \mathcal{S}) = \inf_{Z \in \mathcal{S}} \|X - Z\|_F \leq \|X - Y\|_F.$$

Let us compute $\|X - Y\|_F$. Using the unitary invariance of the Frobenius norm, we get

$$\begin{aligned} \|X - Y\|_F &= \|U\Sigma V^{\mathrm{H}} - UI_n V^{\mathrm{H}}\|_F \\ &= \|U(\Sigma - I_n)V^{\mathrm{H}}\|_F \\ &= \|\Sigma - I_n\|_F = \left(\sum_{i=1}^{n} (\sigma_i - 1)^2\right)^{1/2}. \end{aligned}$$

Next, we evaluate $\|I_n - X^{\mathrm{H}}X\|_F$. First, we find $X^{\mathrm{H}}X$:

$$X^{\mathrm{H}}X = (U\Sigma V^{\mathrm{H}})^{\mathrm{H}}(U\Sigma V^{\mathrm{H}}) = V\Sigma^{\mathrm{H}}U^{\mathrm{H}}U\Sigma V^{\mathrm{H}} = V\Sigma^2 V^{\mathrm{H}}.$$

Again, by unitary invariance,

$$\begin{aligned} \|I_n - X^{\mathrm{H}}X\|_F &= \|I_n - V\Sigma^2 V^{\mathrm{H}}\|_F \\ &= \|V(I_n - \Sigma^2)V^{\mathrm{H}}\|_F \\ &= \|I_n - \Sigma^2\|_F = \left(\sum_{i=1}^{n} (1 - \sigma_i^2)^2\right)^{1/2}. \end{aligned}$$

Now we establish the inequality between the norms. For each singular value $\sigma_i \geq 0$, we have

$$(1 - \sigma_i^2)^2 = ((1 - \sigma_i)(1 + \sigma_i))^2 = (1 - \sigma_i)^2(1 + \sigma_i)^2.$$

Since $\sigma_i \geq 0$, it follows that $1 + \sigma_i \geq 1$, and thus $(1 + \sigma_i)^2 \geq 1$. This implies

$$(1 - \sigma_i^2)^2 \geq (1 - \sigma_i)^2 = (\sigma_i - 1)^2.$$

Summing over $i = 1, \ldots, n$ and taking the square root, we obtain

$$\|I_n - \Sigma^2\|_F \geq \|\Sigma - I_n\|_F.$$

Combining our results, we have the following chain of inequalities:

$$\operatorname{dist}(X, \mathcal{S}) \leq \|X - Y\|_F = \|\Sigma - I_n\|_F \leq \|I_n - \Sigma^2\|_F = \|I_n - X^{\mathrm{H}}X\|_F.$$

This completes the proof. □

(b) *Proof.* To prove that no such constant $C$ exists, we show that the ratio $\frac{\|I_n - X^{\mathrm{H}}X\|_F}{\operatorname{dist}(X, \mathcal{S})}$ is unbounded.

Take $X$ and $Y$ such that $X = U\Sigma V^H$, $Y = UV^H$, $\Sigma = \operatorname{diag}(\sigma, \sigma, \ldots, \sigma)$.

By (a), we have

$$\|I_n - X^H X\|_F = \|I_n - \Sigma^2\|_F = \left(\sum_{i=1}^{n}(1-\sigma^2)^2\right)^{\frac{1}{2}} = \sqrt{n}|1-\sigma^2|$$

and

$$\|X - Y\|_F = \|I_n - \Sigma\|_F = \left(\sum_{i=1}^{n}(1-\sigma^2)\right)^{\frac{1}{2}} = \sqrt{n}|1-\sigma|.$$

Thus,

$$\frac{\|I_n - X^H X\|_F}{\operatorname{dist}(X, \mathcal{S})} \geq \frac{\|I_n - X^H X\|_F}{\|I_n - \Sigma\|_F} = \frac{\sqrt{n}|1-\sigma^2|}{\sqrt{n}|1-\sigma|} = |1+\sigma|.$$

Take $\sigma \to +\infty$, this ratio is unbounded. □

17.

**Problem.** Let $A \in \mathbb{C}^{n\times n}$ be a nonsingular matrix, and

$$J = \begin{pmatrix} 0 & A \\ A^H & 0 \end{pmatrix}.$$

(a) If the eigenvalues of $A^H A$ are $\sigma_1, \ldots, \sigma_n$, multiplicity included, prove that the eigenvalues of $J$ are $\sqrt{\sigma_1}, -\sqrt{\sigma_1}, \ldots, \sqrt{\sigma_n}, -\sqrt{\sigma_n}$, multiplicity included.

(b) Consider $n \times n$ complex matrices $U_1, U_2, V_1, V_2$, and $\Sigma$. Suppose that $\Sigma$ is a diagonal matrix whose diagonal entries are all positive. If

$$J = \begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix} \begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix}^H$$

is an eigenvalue decomposition of $J$, prove that

$$A = 2U_1 \Sigma V_1^H = -2U_2 \Sigma V_2^H.$$

**Solution**.

(a) *Proof.* Since $A$ is nonsingular, $A^H A$ is positive definite with eigenvalues $\sigma_i > 0$. Consider the block structure of $J$. For any eigenvalue $\lambda$ of $J$ with eigenvector $\begin{pmatrix} u \\ v \end{pmatrix}$:

$$J\begin{pmatrix} u \\ v \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \end{pmatrix} \implies \begin{cases} Av = \lambda u \\ A^H u = \lambda v \end{cases}.$$

Substituting gives:

$$A^H A v = \lambda^2 v, \quad A A^H u = \lambda^2 u.$$

Thus $\lambda^2$ is an eigenvalue of $A^H A$, so $\lambda = \pm\sqrt{\sigma_k}$ for some $k$.

Now let $v_k$ be an eigenvector of $A^H A$ for eigenvalue $\sigma_k$. Define:

$$w_k^+ = \begin{pmatrix} \sigma_k^{-1/2} A v_k \\ v_k \end{pmatrix}, \quad w_k^- = \begin{pmatrix} -\sigma_k^{-1/2} A v_k \\ v_k \end{pmatrix}.$$

Direct computation shows:

$$Jw_k^+ = \sqrt{\sigma_k} w_k^+, \quad Jw_k^- = -\sqrt{\sigma_k} w_k^-.$$

The $2n$ vectors $\{w_k^+, w_k^-\}_{k=1}^n$ are linearly independent because: the $v_k$ are linearly independent and the map $v_k \mapsto Av_k$ is bijective ($A$ nonsingular).

Thus $J$ has eigenvalues $\pm\{\sqrt{\sigma_k}\}_{i=1}^n$ . □

(b) *Proof.* The eigenvalue decomposition of $J$ is:

$$J = \begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix} \begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix}^{\mathrm{H}},$$

where the block matrix $Q = \begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix}$ is unitary ($QQ^{\mathrm{H}} = Q^{\mathrm{H}}Q = I_{2n}$), and $\Sigma$ is diagonal with positive entries.

Let the columns of $\begin{pmatrix} U_1 \\ V_1 \end{pmatrix}$ be the eigenvectors corresponding to the positive eigenvalues in $\Sigma$, whose diagonal entries are $\{\lambda_k\}_{k=1}^n$. Let the $k$-th column be $q_k = \begin{pmatrix} u_k \\ v_k \end{pmatrix}$. From part (a), we know that if $q_k$ is an eigenvector for $\lambda_k$, then the vector $q_k' = \begin{pmatrix} -u_k \\ v_k \end{pmatrix}$ is an eigenvector for $-\lambda_k$. Thus, we can choose $\begin{pmatrix} U_2 \\ V_2 \end{pmatrix}$ formed by $q_k'$.

Therefore, we can choose eigenvectors such that:

$$V_1 = V_2, \tag{5}$$
$$U_2 = -U_1. \tag{6}$$

The $(1,2)$-block of $J$ is $A$. Expanding the decomposition:

$$J = \begin{pmatrix} U_1\Sigma U_1^{\mathrm{H}} - U_2\Sigma U_2^{\mathrm{H}} & U_1\Sigma V_1^{\mathrm{H}} - U_2\Sigma V_2^{\mathrm{H}} \\ V_1\Sigma U_1^{\mathrm{H}} - V_2\Sigma U_2^{\mathrm{H}} & V_1\Sigma V_1^{\mathrm{H}} - V_2\Sigma V_2^{\mathrm{H}} \end{pmatrix}.$$

The $(1,2)$-block gives:

$$A = U_1\Sigma V_1^{\mathrm{H}} - U_2\Sigma V_2^{\mathrm{H}}.$$

Substituting (5) and (6):

$$A = U_1\Sigma V_1^{\mathrm{H}} - (-U_1)\Sigma V_1^{\mathrm{H}} = U_1\Sigma V_1^{\mathrm{H}} + U_1\Sigma V_1^{\mathrm{H}} = 2U_1\Sigma V_1^{\mathrm{H}}.$$

Similarly,

$$-2U_2\Sigma V_2^{\mathrm{H}} = -2(-U_1)\Sigma V_1^{\mathrm{H}} = 2U_1\Sigma V_1^{\mathrm{H}} = A.$$

Thus, $A = 2U_1\Sigma V_1^{\mathrm{H}} = -2U_2\Sigma V_2^{\mathrm{H}}$.

□

18.

**Problem.** (a) If $2 \leq m \leq n+1$, show that there exists $\{v_1, v_2, \ldots, v_m\} \subset \mathbb{R}^n$ such that $v_i^T v_j < 0$ for all distinct indices $i, j \in \{1, 2, \ldots, m\}$.

(b) If $m > n+1$, show that there does not exist $\{v_1, v_2, \ldots, v_m\} \subset \mathbb{R}^n$ such that $v_i^T v_j < 0$ for all distinct indices $i, j \in \{1, 2, \ldots, m\}$.

**Solution**.

(a) *Proof.* We proceed by induction on the dimension $n$. For any given $n$, it is sufficient to prove the existence for the maximal case, $m = n + 1$, as any subset of the constructed set will also satisfy the condition.

For $n = 1$, we construct a set of $m = 1+1 = 2$ vectors in $\mathbb{R}^1$. Let $v_1 = (1)$ and $v_2 = (-1)$. Their inner product is $v_1^T v_2 = -1 < 0$. Thus, the statement holds for $n = 1$.

We establish the inductive hypothesis first. Assume that for some integer $k \geq 1$, the proposition holds for the dimension $n = k$. That is, there exists a set of $k + 1$ vectors, $\{\alpha_1, \ldots, \alpha_{k+1}\} \subset \mathbb{R}^k$, such that $\alpha_i^T \alpha_j < 0$ for all $i \neq j$.

Then, we do the inductive step. We need to show that the proposition holds for the dimension $n = k+1$. Our goal is to construct a set of $(k+1)+1 = k+2$ vectors in $\mathbb{R}^{k+1}$ with the desired property, using the set from our hypothesis.

We embed the vectors from $\mathbb{R}^k$ into $\mathbb{R}^{k+1}$ by appending a zero as the final coordinate. Let

$$\tilde{\alpha}_i = (\alpha_i, 0) \in \mathbb{R}^{k+1} \quad \text{for } i = 1, \ldots, k+1.$$

The inner products are preserved: $\tilde{\alpha}_i^T \tilde{\alpha}_j = \alpha_i^T \alpha_j < 0$ for $i \neq j$.

Next, we introduce a new vector in $\mathbb{R}^{k+1}$ that is orthogonal to the subspace containing the embedded vectors. Let $u = (0, \ldots, 0, 1)$ be the standard basis vector for the $(k+1)$-th dimension.

We now construct the new set of $k+2$ vectors, $\{v_1, \ldots, v_{k+2}\}$, by perturbing the embedded vectors. For a small real number $\varepsilon > 0$, we define:

$$\begin{aligned} v_i &:= \tilde{\alpha}_i + \varepsilon u \quad \text{for } i = 1, \ldots, k+1 \\ v_{k+2} &:= -u. \end{aligned}$$

We verify that all inner products of distinct vectors in this new set are negative. There are two cases to consider.

*Case 1:* Inner product of $v_i$ and $v_j$ for $i, j \in \{1, \ldots, k+1\}$ and $i \neq j$.

$$\begin{aligned} v_i^T v_j &= (\tilde{\alpha}_i + \varepsilon u)^T (\tilde{\alpha}_j + \varepsilon u) \\ &= \tilde{\alpha}_i^T \tilde{\alpha}_j + \varepsilon(\tilde{\alpha}_i^T u) + \varepsilon(u^T \tilde{\alpha}_j) + \varepsilon^2(u^T u) \\ &= \alpha_i^T \alpha_j + 0 + 0 + \varepsilon^2 \|u\|^2 \\ &= \alpha_i^T \alpha_j + \varepsilon^2. \end{aligned}$$

By the inductive hypothesis, $\alpha_i^T \alpha_j < 0$. Since there are a finite number of such pairs, we can choose $\varepsilon > 0$ to be small enough such that $\alpha_i^T \alpha_j + \varepsilon^2 < 0$ for all pairs $(i, j)$ with $i \neq j$.

*Case 2:* Inner product of $v_i$ and $v_{k+2}$ for $i \in \{1, \ldots, k+1\}$.

$$\begin{aligned} v_i^T v_{k+2} &= (\tilde{\alpha}_i + \varepsilon u)^T (-u) \\ &= -(\tilde{\alpha}_i^T u) - \varepsilon(u^T u) \\ &= 0 - \varepsilon \|u\|^2 = -\varepsilon. \end{aligned}$$

Since $\varepsilon > 0$, this inner product is strictly negative.

We have successfully constructed a set of $k+2$ vectors in $\mathbb{R}^{k+1}$ that satisfies the condition. Thus, the statement holds for $n = k + 1$.

By the principle of mathematical induction, the proposition is true for all integers $n \geq 1$. □

(b) *Proof.* We proceed by contradiction. Assume there exists such a set of vectors with $m > n+1$. This implies the existence of a subset of $n+2$ vectors, $\{v_1, \ldots, v_{n+2}\}$, satisfying the property.

The set of the first $n+1$ vectors, $\{v_1, \ldots, v_{n+1}\}$, must be linearly dependent in $\mathbb{R}^n$. Thus, there exist scalars $k_1, \ldots, k_{n+1}$, not all zero, such that:

$$\sum_{i=1}^{n+1} k_i v_i = 0. \tag{7}$$

Taking the inner product of this equation with the vector $v_{n+2}$ yields:

$$\left( \sum_{i=1}^{n+1} k_i v_i \right)^T v_{n+2} = \sum_{i=1}^{n+1} k_i (v_i^T v_{n+2}) = 0.$$

By the initial hypothesis, each inner product $v_i^T v_{n+2}$ is strictly negative. For this weighted sum to be zero, the set of coefficients $\{k_i\}_{i=1}^{n+1}$ must contain both positive and negative values.

Let us partition the indices into two non-empty sets: $I^+ = \{i \mid k_i > 0\}$ and $I^- = \{i \mid k_i < 0\}$. We can rearrange the linear dependence relation (7) to separate the terms based on the sign of their coefficients:

$$\sum_{i \in I^+} k_i v_i = - \sum_{j \in I^-} k_j v_j.$$

Define the vector $u = \sum_{i \in I^+} k_i v_i$. We now compute the squared norm of $u$:

$$\|u\|^2 = u^T u = \left( \sum_{i \in I^+} k_i v_i \right)^T \left( - \sum_{j \in I^-} k_j v_j \right) = - \sum_{i \in I^+} \sum_{j \in I^-} k_i k_j (v_i^T v_j).$$

Consider any term in the final summation. For $i \in I^+$ and $j \in I^-$, we have $k_i > 0$, $k_j < 0$, and by hypothesis, $v_i^T v_j < 0$. The product $k_i k_j (v_i^T v_j)$ is therefore strictly positive.

Since the double summation consists entirely of strictly positive terms, its result is strictly positive. This implies:

$$\|u\|^2 = -(\text{a strictly positive number}) < 0.$$

This contradicts the fundamental property that the squared norm of any real vector must be non-negative. The initial assumption must therefore be false. □

19.

**Problem.** Given $A \in \mathbb{R}^{m \times m}$ and $B \in \mathbb{R}^{n \times n}$, prove that the equation

$$AX - XB = C, \quad X \in \mathbb{R}^{m \times n}$$

has a unique solution for all $C \in \mathbb{R}^{m \times n}$ if and only if $A$ and $B$ do not share any eigenvalue.

[When $n = 1$, $B$ is a scalar while $X$ and $C$ are $m$-dimensional vectors; in this case, the conclusion says nothing but $(A - BI)X = C$ has a unique solution for all $C \in \mathbb{R}^m$ if and only if $B$ is not an eigenvalue of $A$.]

**Solution**. *Proof.* Let $V = \mathbb{R}^{m \times n}$ be the vector space of $m \times n$ real matrices. We define a linear operator $T : V \to V$ by

$$T(X) = AX - XB.$$

The statement that the equation $AX - XB = C$ has a unique solution for every $C$ is equivalent to the statement that the linear operator $T$ is invertible. For a linear operator on a finite-dimensional vector space, this is equivalent to its kernel (or null space) being trivial, i.e., $\ker(T) = \{0\}$. The proof proceeds in two parts. For full generality, we consider the eigenvalues in $\mathbb{C}$.

($\Leftarrow$) If $\sigma(A) \cap \sigma(B) = \emptyset$, then the solution is unique.

We will show that if the spectra are disjoint, then $\ker(T) = \{0\}$. Let $X \in \mathbb{R}^{m \times n}$ be a matrix such that $T(X) = 0$, which means

$$AX = XB.$$

Let $\lambda \in \sigma(B)$ be an eigenvalue of $B$. Let $J = P^{-1}BP$ be the Jordan Canonical Form of $B$ over $\mathbb{C}$, where $P$ is an invertible matrix in $\mathbb{C}^{n \times n}$. The equation becomes $AX = X(PJP^{-1})$,

which can be rearranged to $A(XP) = (XP)J$. Let $Y = XP$. The matrix $Y \in \mathbb{C}^{m\times n}$ is zero if and only if $X$ is zero. The transformed equation is

$$AY = YJ.$$

Let $J$ be composed of Jordan blocks $J_k(\lambda_k)$ on its diagonal, where $\lambda_k$ are the eigenvalues of $B$. Let us analyze the equation for one such block. Let $y_1, \ldots, y_r$ be the columns of $Y$ corresponding to a Jordan block $J_r(\lambda)$. The equation $AY = YJ_r(\lambda)$ expands to:

$$\begin{aligned} Ay_1 &= \lambda y_1 \\ Ay_2 &= y_1 + \lambda y_2 \\ &\vdots \\ Ay_r &= y_{r-1} + \lambda y_r \end{aligned}$$

These can be rewritten as:

$$\begin{aligned} (A - \lambda I)y_1 &= 0 \\ (A - \lambda I)y_2 &= y_1 \\ &\vdots \\ (A - \lambda I)y_r &= y_{r-1} \end{aligned}$$

By our initial assumption, $\sigma(A) \cap \sigma(B) = \emptyset$. Since $\lambda \in \sigma(B)$, it follows that $\lambda \notin \sigma(A)$. Therefore, the matrix $(A - \lambda I)$ is invertible.

From the first equation, $(A - \lambda I)y_1 = 0$, the invertibility of $(A - \lambda I)$ implies that $y_1 = 0$. Substituting $y_1 = 0$ into the second equation gives $(A - \lambda I)y_2 = 0$, which in turn implies $y_2 = 0$. Continuing this process inductively, we find that $y_j = 0$ for all $j = 1, \ldots, r$.

Since this holds for any Jordan block of $J$, all columns of $Y$ must be zero. Thus, $Y = 0$. As $Y = XP$ and $P$ is invertible, we conclude that $X = 0$. Therefore, $\ker(T) = \{0\}$, and the operator $T$ is invertible, guaranteeing a unique solution for every $C$.

($\Rightarrow$) If the solution is unique for every $C$, then $\sigma(A) \cap \sigma(B) = \emptyset$.

We prove the contrapositive: if $\sigma(A) \cap \sigma(B) \neq \emptyset$, then the operator $T$ is not invertible because its kernel is non-trivial.

Assume there is a common eigenvalue $\lambda \in \sigma(A) \cap \sigma(B)$. Since $\lambda \in \sigma(A)$, there exists a non-zero right eigenvector $v \in \mathbb{C}^m$ such that $Av = \lambda v$.

Since $\lambda \in \sigma(B)$, we know that $\det(B - \lambda I) = 0$. As $\det(M) = \det\left(M^T\right)$, it follows that $\det\left(B^T - \lambda I\right) = 0$, which implies that $\lambda$ is also an eigenvalue of $B^T$. Therefore, there exists a non-zero eigenvector $w \in \mathbb{C}^n$ for $B^T$ corresponding to $\lambda$:

$$B^T w = \lambda w.$$

Taking the transpose of this equation yields $(B^T w)^T = (\lambda w)^T$, we obtain the left eigenvector relationship for $B$:

$$w^T B = \lambda w^T.$$

Now, we construct the matrix $X = vw^T$. Since $v \neq 0$ and $w \neq 0$, $X$ is a non-zero matrix. We verify that this $X$ is in the kernel of $T$:

$$\begin{aligned} T(X) &= AX - XB \\ &= A(vw^T) - (vw^T)B \\ &= (Av)w^T - v(w^TB) \\ &= (\lambda v)w^T - v(\lambda w^T) \\ &= \lambda(vw^T) - \lambda(vw^T) = 0. \end{aligned}$$

We have found a non-zero matrix $X$ such that $T(X) = 0$. This means that $\ker(T)$ is non-trivial, and thus the operator $T$ is not invertible. This completes the proof of the contrapositive. □

20.

**Problem.** Let $X$ be a random variable and $f$ be a convex function on $\mathbb{R}$. Suppose that both $X$ and $f(X)$ have finite expectations. Prove Jensen's inequality:

$$f(\mathbb{E}(X)) \leq \mathbb{E}[f(X)].$$

**Solution**. *Proof.* Let $\mu = \mathbb{E}[X]$, which is well-defined since $X$ has finite expectation. Since convex functions have subgradients in the interior of the domain of definition, for any $x \in \mathbb{R}$ there exists a subgradient $c \in \mathbb{R}$ at $\mu$ satisfying:

$$f(x) \geq f(\mu) + c(x - \mu).$$

This inequality holds pointwise for all realizations of $X$. Applying the expectation operator to both sides and using linearity of expectation:

$$\begin{aligned} \mathbb{E}[f(X)] &\geq \mathbb{E}\left[f(\mu) + c(X - \mu)\right] \\ &= \mathbb{E}[f(\mu)] + c \cdot \mathbb{E}[X - \mu] \\ &= f(\mu) + c \cdot (\mathbb{E}[X] - \mu) \\ &= f(\mu) + c \cdot 0 \\ &= f(\mathbb{E}[X]), \end{aligned}$$

where $\mathbb{E}[f(\mu)] = f(\mu)$ since $f(\mu)$ is deterministic, and $\mathbb{E}[X - \mu] = \mathbb{E}[X] - \mu = 0$ by the definition of $\mu$. □

21.

**Problem.** For any convex function $f$ on $[0, 1]$, prove that

$$f\left(\frac{1}{2}\right) \leq \int_0^1 f(x)\,dx \leq \frac{1}{2}\left[f(0) + f(1)\right].$$

**Solution**. We first recall the general Hadamard inequality for convex functions, which is established in the proof provided, and then we state that the conclusion to be proved is its special case.

Let $f$ be a convex function on $[a, b]$. Then for any subinterval $[x_1, x_2] \subseteq [a, b]$,

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} f(x)dx \leq \frac{1}{2}\left[f(x_1) + f(x_2)\right].$$

*Proof.* The left inequality follows from convexity and integration. For all $x \in [x_1, x_2]$, convexity implies:

$$f\left(\frac{x + (x_1 + x_2 - x)}{2}\right) = f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x) + f(x_1 + x_2 - x)}{2}.$$

Integrating both sides over $[x_1, x_2]$:

$$\int_{x_1}^{x_2} f\left(\frac{x_1 + x_2}{2}\right) dx \leq \frac{1}{2} \int_{x_1}^{x_2} \left[f(x) + f(x_1 + x_2 - x)\right] dx$$
$$f\left(\frac{x_1 + x_2}{2}\right)(x_2 - x_1) \leq \frac{1}{2}\left[\int_{x_1}^{x_2} f(x)dx + \int_{x_1}^{x_2} f(x_1 + x_2 - x)dx\right].$$

By the substitution $u = x_1 + x_2 - x$, the second integral equals $\int_{x_1}^{x_2} f(u)du$, so:

$$f\left(\frac{x_1 + x_2}{2}\right)(x_2 - x_1) \leq \frac{1}{2}\left[\int_{x_1}^{x_2} f(x)dx + \int_{x_1}^{x_2} f(x)dx\right] = \int_{x_1}^{x_2} f(x)dx.$$

Dividing by $x_2 - x_1$ yields the left inequality.

The right inequality follows from the convexity definition and a change of variables. Let $\lambda \in [0, 1]$, and set $x = (1 - \lambda)x_1 + \lambda x_2$. Then:

$$f(x) = f((1 - \lambda)x_1 + \lambda x_2) \leq (1 - \lambda)f(x_1) + \lambda f(x_2).$$

Substituting $dx = (x_2 - x_1)d\lambda$ when integrating over $[x_1, x_2]$:

$$\begin{aligned}
\int_{x_1}^{x_2} f(x)dx &= (x_2 - x_1) \int_0^1 f((1 - \lambda)x_1 + \lambda x_2)d\lambda \\
&\leq (x_2 - x_1) \int_0^1 \left[(1 - \lambda)f(x_1) + \lambda f(x_2)\right] d\lambda \\
&= (x_2 - x_1)\left[f(x_1) \int_0^1 (1 - \lambda)d\lambda + f(x_2) \int_0^1 \lambda d\lambda\right] \\
&= (x_2 - x_1)\left[f(x_1) \cdot \frac{1}{2} + f(x_2) \cdot \frac{1}{2}\right] \\
&= \frac{x_2 - x_1}{2}\left[f(x_1) + f(x_2)\right].
\end{aligned}$$

Dividing by $x_2 - x_1$ gives the right inequality.

Now, specialize to the interval $[0, 1]$ by setting $x_1 = 0$ and $x_2 = 1$. The general inequality becomes:

$$f\left(\frac{0 + 1}{2}\right) \leq \frac{1}{1 - 0} \int_0^1 f(x)dx \leq \frac{1}{2}\left[f(0) + f(1)\right].$$

Simplifying:

$$f\left(\frac{1}{2}\right) \leq \int_0^1 f(x)dx \leq \frac{1}{2}\left[f(0)+f(1)\right].$$

This is the desired inequality for $[0,1]$. □

22.

**Problem.** Let $X$ be a random variable. Suppose that $f$ and $g$ are two increasing functions such that $f(X)$ and $g(X)$ are both bounded. Prove

$$\mathbb{E}[f(X)g(X)] \geq \mathbb{E}[f(X)]\mathbb{E}[g(X)].$$

**Solution**. *Proof.* The boundedness of $f(X)$ and $g(X)$ ensures all expectations exist. Let $X_1$ and $X_2$ be independent copies of $X$.

We begin by stating a conclusion, i.e., let $f$ and $g$ be increasing functions, then for any realizations $x_1, x_2$ of $X_1, X_2$:

$$(f(x_1)-f(x_2))(g(x_1)-g(x_2)) \geq 0.$$

Since $f$ and $g$ are increasing:

- If $x_1 \geq x_2$, then $f(x_1) \geq f(x_2)$ and $g(x_1) \geq g(x_2)$, so the product is non-negative.
- If $x_1 < x_2$, then $f(x_1) \leq f(x_2)$ and $g(x_1) \leq g(x_2)$, so $(f(x_1)-f(x_2)) \leq 0$ and $(g(x_1)-g(x_2)) \leq 0$, and their product is non-negative.

We show the proof for discrete random variable. Assume $X$ takes values in $\{x_1,\ldots,x_n\}$ with $P(X=x_i)=p_i$. Then:

$$\begin{aligned}
&\mathbb{E}[f(X)g(X)] - \mathbb{E}[f(X)]\mathbb{E}[g(X)] \\
&= \sum_{i=1}^n f(x_i)g(x_i)p_i - \left(\sum_{i=1}^n f(x_i)p_i\right)\left(\sum_{j=1}^n g(x_j)p_j\right) \\
&= \sum_{i,j=1}^n p_i p_j \left[f(x_i)g(x_i) - f(x_i)g(x_j)\right] \\
&= \frac{1}{2}\sum_{i,j=1}^n \Big[f(x_i)g(x_i)p_ip_j + f(x_j)g(x_j)p_jp_i \\
&\quad - f(x_i)g(x_j)p_ip_j - f(x_j)g(x_i)p_jp_i\Big] \\
&= \frac{1}{2}\sum_{i,j=1}^n p_ip_j(f(x_i)-f(x_j))(g(x_i)-g(x_j)).
\end{aligned}$$

By the conclusion above, each term $(f(x_i)-f(x_j))(g(x_i)-g(x_j)) \geq 0$, and since $p_ip_j \geq 0$, the entire sum is non-negative. Note that the continuous case is similar, doing nothing but replacing the summation with an integral. □

23.

**Problem.** Suppose that $\{a_k\}$ and $\{b_k\}$ are monotone real sequences with the same monotonicity. Let $n$ be a nonnegative integer. Prove that

$$\sum_{k=0}^{n} a_k b_{n-k} \leq \frac{1}{n+1} \left( \sum_{k=0}^{n} a_k \right) \left( \sum_{k=0}^{n} b_k \right) \leq \sum_{k=0}^{n} a_k b_k.$$

Give as many proofs as possible.

**Solution**. *First Method*

*Proof.* Assume without loss of generality that both sequences are non-decreasing. We first establish the rearrangement inequality: for any permutation $\gamma$ of $\{0, 1, \ldots, n\}$,

$$\sum_{k=0}^{n} a_k b_{n-k} \leq \sum_{k=0}^{n} a_k b_{\gamma(k)} \leq \sum_{k=0}^{n} a_k b_k.$$

To prove the right inequality (the left is similar), define $s_k = b_k - b_{\gamma(k)}$ and $t_m = \sum_{k=0}^{m} s_k$ for $0 \leq m \leq n$. Since $\gamma$ is a permutation, $\sum_{k=0}^{n} s_k = 0$, so $t_n = 0$. By monotonicity and rearrangement, $t_m \leq 0$ for all $0 \leq m \leq n-1$. Applying Abel summation:

$$\begin{aligned} \sum_{k=0}^{n} a_k s_k &= a_n t_n - \sum_{k=0}^{n-1} (a_{k+1} - a_k) t_k \\ &= 0 - \sum_{k=0}^{n-1} \underbrace{(a_{k+1} - a_k)}_{\geq 0} \underbrace{t_k}_{\leq 0} \geq 0, \end{aligned}$$

hence $\sum_{k=0}^{n} a_k b_k - \sum_{k=0}^{n} a_k b_{\gamma(k)} = \sum_{k=0}^{n} a_k s_k \geq 0$, proving $\sum_{k=0}^{n} a_k b_{\gamma(k)} \leq \sum_{k=0}^{n} a_k b_k$.

Let $S = \left(\sum_{i=0}^{n} a_i\right) \left(\sum_{j=0}^{n} b_j\right)$ and let $\mathfrak{S}$ be the set of all permutations of $\{0, 1, \ldots, n\}$ ($|\mathfrak{S}| = (n+1)!$). For each $\sigma \in \mathfrak{S}$, define $T_\sigma = \sum_{k=0}^{n} a_k b_{\sigma(k)}$. By the rearrangement inequality,

$$\sum_{k=0}^{n} a_k b_{n-k} \leq T_\sigma \leq \sum_{k=0}^{n} a_k b_k \quad \forall \sigma \in \mathfrak{S}. \tag{8}$$

Summing (8) over all $\sigma \in \mathfrak{S}$:

$$\left( \sum_{k=0}^{n} a_k b_{n-k} \right) (n+1)! \leq \sum_{\sigma \in \mathfrak{S}} T_\sigma \leq \left( \sum_{k=0}^{n} a_k b_k \right) (n+1)!.$$

The middle term simplifies as:

$$\begin{aligned} \sum_{\sigma \in \mathfrak{S}} T_\sigma &= \sum_{\sigma \in \mathfrak{S}} \sum_{k=0}^{n} a_k b_{\sigma(k)} \\ &= \sum_{k=0}^{n} a_k \left( \sum_{\sigma \in \mathfrak{S}} b_{\sigma(k)} \right) \\ &= \sum_{k=0}^{n} a_k \left( n! \sum_{j=0}^{n} b_j \right) \\ &= n! \left( \sum_{i=0}^{n} a_i \right) \left( \sum_{j=0}^{n} b_j \right) = n! S, \end{aligned}$$

since for fixed $k$, $\sum_{\sigma\in\mathfrak{S}} b_{\sigma(k)} = n!\sum_{j=0}^{n} b_j$. Substituting gives:

$$\left(\sum_{k=0}^{n} a_k b_{n-k}\right)(n+1)! \le n!S \le \left(\sum_{k=0}^{n} a_k b_k\right)(n+1)!.$$

Dividing by $(n+1)!$ yields the result:

$$\sum_{k=0}^{n} a_k b_{n-k} \le \frac{S}{n+1} \le \sum_{k=0}^{n} a_k b_k.$$

□

*Second Method*

*Proof.* Assume both sequences are non-decreasing.

Consider the difference:

$$D_{\text{right}} = \sum_{k=0}^{n} a_k b_k - \frac{1}{n+1}\left(\sum_{i=0}^{n} a_i\right)\left(\sum_{j=0}^{n} b_j\right).$$

This can be rewritten as:

$$\begin{aligned} D_{\text{right}} &= \sum_{k} a_k b_k - \frac{1}{n+1}\left(\sum_{i} a_i\right)\left(\sum_{j} b_j\right) \\ &= \frac{1}{2(n+1)}\left[2(n+1)\sum_{k} a_k b_k - 2\sum_{i} a_i \sum_{j} b_j\right] \\ &= \frac{1}{2(n+1)}\left[\sum_{i,j} a_i b_i + \sum_{i,j} a_j b_j - \sum_{i,j} a_i b_j - \sum_{i,j} a_j b_i\right] \\ &= \frac{1}{2(n+1)} \sum_{0\le i,j\le n} (a_i - a_j)(b_i - b_j). \end{aligned}$$

For each pair $(i,j)$, $(a_i - a_j)(b_i - b_j) \ge 0$ since both sequences are non-decreasing. Thus $D_{\text{right}} \ge 0$. We have proved the left inequality, and we prove the right inequality below.

Define the reversed sequence $c_k = b_{n-k}$. Since $\{b_k\}$ is non-decreasing, $\{c_k\}$ is non-increasing. The difference is:

$$D_{\text{left}} = \frac{1}{n+1}\left(\sum_{i=0}^{n} a_i\right)\left(\sum_{j=0}^{n} c_j\right) - \sum_{k=0}^{n} a_k c_k.$$

Similarly:

$$D_{\text{left}} = \frac{1}{2(n+1)} \sum_{0\le i,j\le n} (a_i - a_j)(c_i - c_j).$$

For $i < j$, we have $a_i \le a_j$ but $c_i \ge c_j$ (since $\{c_k\}$ is non-increasing), so $(a_i - a_j) \le 0$ and $(c_i - c_j) \ge 0$, making $(a_i - a_j)(c_i - c_j) \le 0$. For $i > j$, we have $a_i \ge a_j$ but $c_i \le c_j$,

so $(a_i - a_j) \geq 0$ and $(c_i - c_j) \leq 0$, making $(a_i - a_j)(c_i - c_j) \leq 0$. Thus $D_{\text{left}} \geq 0$. Since $\sum_{j=0}^{n} c_j = \sum_{j=0}^{n} b_{n-j} = \sum_{j=0}^{n} b_j$, we have:

$$\sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{n} a_k c_k \leq \frac{1}{n+1} \left( \sum_{j=0}^{n} a_j \right) \left( \sum_{j=0}^{n} b_j \right).$$

Since the left inequality and the right inequality hold, we complete the proof. □

24.

**Problem.** Prove that a sequence $\{x_k\} \subset \mathbb{R}$ converges if $\sum_{|k|>\epsilon} |x_k - x_{k+1}| < \infty$ for all $\epsilon > 0$. Is the converse proposition true?

**Solution**. This proposition is correct, but the converse is not.

(a) *proof of Proposition.* Assume, to the contrary, that the sequence $\{x_k\}$ does not converge. Since $\mathbb{R}$ is complete, $\{x_k\}$ is not a Cauchy sequence. Consequently, there exists $\delta > 0$ such that for every $N \in \mathbb{N}$, there are integers $m, n > N$ with $|x_m - x_n| \geq \delta$. For any such pair $(m, n)$, at least one of $|x_m|$ or $|x_n|$ must satisfy $|x_m| \geq \delta/2$ or $|x_n| \geq \delta/2$; otherwise, if both were less than $\delta/2$, we would have $|x_m - x_n| < \delta$, contradicting the assumption. Thus the set $\{k : |x_k| \geq \delta/2\}$ is infinite. Let $\{k_i\}_{i=1}^{\infty}$ be a strictly increasing sequence of indices such that $|x_{k_i}| \geq \delta/2$ for all $i$.

Fix $\epsilon = \delta/8 > 0$. By hypothesis, the series $\sum_{k:|x_k|>\epsilon} |x_k - x_{k+1}|$ converges. Define $S = \{k : |x_k| > \delta/8\}$ and select $N_0 \in \mathbb{N}$ such that

$$\sum_{\substack{k \geq N_0 \\ k \in S}} |x_k - x_{k+1}| < \frac{\delta}{8}.$$

Without loss of generality, we may assume $k_i > N_0$ for all $i$ by passing to a subsequence.

For each $i$, define $n_i$ as the smallest integer greater than $k_i$ satisfying $|x_{n_i}| \leq \delta/8$, provided such an integer exists. We consider two exhaustive cases:

*Case 1:* For some $i$, no such $n_i$ exists. Then $|x_k| > \delta/8$ for all $k \geq k_i$, so $k \in S$ for all $k \geq k_i$. The series $\sum_{k=k_i}^{\infty} |x_k - x_{k+1}|$ converges as a tail of a convergent series. For any integers $m > n \geq k_i$,

$$|x_n - x_m| \leq \sum_{j=n}^{m-1} |x_j - x_{j+1}|.$$

The right-hand side tends to zero as $n, m \to \infty$ by the Cauchy criterion for series convergence. Hence $\{x_k\}_{k=k_i}^{\infty}$ is a Cauchy sequence and converges in $\mathbb{R}$, contradicting the non-convergence hypothesis.

*Case 2:* The integer $n_i$ exists for every $i$. By minimality of $n_i$, we have $|x_k| > \delta/8$ for all $k_i \leq k < n_i$, so these indices belong to $S$. Moreover,

$$|x_{k_i} - x_{n_i}| \geq |x_{k_i}| - |x_{n_i}| \geq \frac{\delta}{2} - \frac{\delta}{8} = \frac{3\delta}{8}.$$

Applying the triangle inequality yields

$$|x_{k_i} - x_{n_i}| \leq \sum_{j=k_i}^{n_i-1} |x_j - x_{j+1}|,$$

so $\sum_{j=k_i}^{n_i-1} |x_j - x_{j+1}| \geq 3\delta/8$. We now construct a strictly increasing sequence $\{i_l\}$ inductively: Set $i_1 = 1$, and for each $l \geq 1$, select $i_{l+1}$ such that $k_{i_{l+1}} > n_{i_l}$ (possible since $\{k_i\}$ increases to infinity). The intervals $[k_{i_l}, n_{i_l} - 1]$ are pairwise disjoint, and

$$\sum_{k \in S} |x_k - x_{k+1}| \geq \sum_{l=1}^{\infty} \sum_{j=k_{i_l}}^{n_{i_l}-1} |x_j - x_{j+1}| \geq \sum_{l=1}^{\infty} \frac{3\delta}{8} = \infty,$$

contradicting the convergence of $\sum_{k \in S} |x_k - x_{k+1}|$.

Both cases lead to contradictions, so $\{x_k\}$ must converge. □

(b) Consider the sequence $x_k = \sum_{j=1}^{k} \frac{(-1)^{j+1}}{j}$ for $k \geq 1$.

*Claim 1:* $\{x_k\}$ converges.
*Proof:* The series $\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j}$ is the alternating harmonic series. Since the absolute values $\frac{1}{j}$ decrease monotonically to 0, by the Leibniz's Test, the series converges.

*Claim 2:* For $\epsilon = 0.1$, the sum $\sum_{|x_k| > \epsilon} |x_k - x_{k+1}|$ diverges.
*Proof:* First, observe that:

$$|x_k - x_{k+1}| = \left| \frac{(-1)^{k+2}}{k+1} \right| = \frac{1}{k+1}.$$

Since $\lim_{k \to \infty} x_k = \ln 2 > 0.1$ (by Maclaurin's series of $\ln(1+x)$), there exists $N \in \mathbb{N}$ such that for all $k \geq N$,

$$|x_k| > 0.1.$$

Thus, the set $\{k : |x_k| > 0.1\}$ contains all $k \geq N$. We then have:

$$\sum_{|x_k| > 0.1} |x_k - x_{k+1}| \geq \sum_{k=N}^{\infty} |x_k - x_{k+1}| = \sum_{k=N}^{\infty} \frac{1}{k+1}.$$

The series $\sum_{k=N}^{\infty} \frac{1}{k+1}$ is the tail of the harmonic series, which diverges. Hence, the sum diverges.

The sequence $\{x_k\}$ converges but violates the condition $\sum_{|x_k| > \epsilon} |x_k - x_{k+1}| < \infty$ for $\epsilon = 0.1$. Therefore, the converse of the original proposition is false.

25.

**Problem.** Let $\{a_k\}$ and $\{b_k\}$ be nonnegative real sequences. For each index $k \geq 0$, one of the following two conditions holds:

(a) $a_k \leq b_k$ and $a_{k+1} = 2a_k$;

(b) $a_{k+1} = a_k/2$.

Prove that

$$\sum_{k=0}^{\infty} a_k \leq 2a_0 + 4\sum_{k=0}^{\infty} b_k.$$

**Solution**. *Proof.* Let the given conditions be denoted by (a) and (b). We will first establish a key inequality that holds for each index $k \geq 0$:

$$2a_{k+1} \leq a_k + 4b_k.$$

We verify this by considering the two possible conditions for the index $k$.

*Case1:* If condition (a) holds for $k$, then $a_{k+1} = 2a_k$ and $a_k \leq b_k$. Substituting $a_{k+1}$ into the inequality, we need to show that $2(2a_k) \leq a_k + 4b_k$, which simplifies to $3a_k \leq 4b_k$. Since we are given $a_k \leq b_k$ and $a_k \geq 0$, it follows that $3a_k \leq 3b_k \leq 4b_k$, so the inequality holds.

*Case2:* If condition (b) holds for $k$, then $a_{k+1} = a_k/2$. Substituting $a_{k+1}$ into the inequality, we need to show that $2(a_k/2) \leq a_k + 4b_k$, which simplifies to $a_k \leq a_k + 4b_k$. This is true because $b_k \geq 0$.

Thus, the inequality $2a_{k+1} \leq a_k + 4b_k$ is valid for all $k \geq 0$.

Now, we prove by induction on $n$ that for all integers $n \geq 0$, the following statement $P(n)$ holds:

$$P(n): \quad \sum_{k=0}^{n} a_k + 2a_{n+1} \leq 2a_0 + 4\sum_{k=0}^{n} b_k.$$

For the base case $n = 0$, we must show that $a_0 + 2a_1 \leq 2a_0 + 4b_0$, which is equivalent to the inequality $2a_1 \leq a_0 + 4b_0$. This is precisely the key inequality we established above for the case $k = 0$. Therefore, $P(0)$ is true.

For the inductive step, assume that $P(n)$ is true for some integer $n \geq 0$. We want to prove $P(n+1)$. We have

$$\begin{aligned}
\sum_{k=0}^{n+1} a_k + 2a_{n+2} &= \left(\sum_{k=0}^{n} a_k\right) + a_{n+1} + 2a_{n+2} \\
&\leq \left(2a_0 + 4\sum_{k=0}^{n} b_k - 2a_{n+1}\right) + a_{n+1} + 2a_{n+2} \quad \text{(by the hypothesis } P(n)\text{)} \\
&= 2a_0 + 4\sum_{k=0}^{n} b_k - a_{n+1} + 2a_{n+2}.
\end{aligned}$$

To complete the proof of $P(n+1)$, we must show that this expression is less than or equal to $2a_0 + 4\sum_{k=0}^{n+1} b_k$. This requires showing

$$-a_{n+1} + 2a_{n+2} \leq 4b_{n+1},$$

which is equivalent to $2a_{n+2} \leq a_{n+1} + 4b_{n+1}$. This is again our key inequality for the index $k = n + 1$, which we have already proven to be true. Thus, if $P(n)$ holds, then $P(n + 1)$ holds. By the principle of mathematical induction, $P(n)$ is true for all $n \geq 0$.

From the proven statement $P(n)$, we have

$$\sum_{k=0}^{n} a_k \leq \sum_{k=0}^{n} a_k + 2a_{n+1} \leq 2a_0 + 4\sum_{k=0}^{n} b_k,$$

where the first inequality holds since $a_{n+1} \geq 0$. This gives us the partial sum inequality

$$\sum_{k=0}^{n} a_k \leq 2a_0 + 4\sum_{k=0}^{n} b_k$$

for all $n \geq 0$. Since all terms $a_k$ and $b_k$ are non-negative, the series $\sum a_k$ and $\sum b_k$ consist of non-negative terms. The sequence of partial sums for each series is therefore non-decreasing. Taking the limit as $n \to \infty$, we conclude that

$$\sum_{k=0}^{\infty} a_k \leq 2a_0 + 4\sum_{k=0}^{\infty} b_k.$$

If $\sum b_k$ diverges, the inequality holds trivially. If $\sum b_k$ converges, the right-hand side is finite, which implies that the partial sums of $\sum a_k$ are bounded above, and thus $\sum a_k$ also converges, validating the limit operation. □

26.

**Problem.** Suppose that $X \subset \mathbb{R}^n$ is a compact set, and $T : X \to X$ is a continuous operator satisfying

$$\|T(x) - T(y)\| < \|x - y\| \quad \text{for all distinct } x, y \in X.$$

(a) Show that $T$ has a unique fixed point.

(b) For any $x_0 \in X$, show that the fixed point iteration

$$x_{k+1} = T(x_k)$$

converges to the fixed point.

**Solution**.

(a) *Proof.* Let us first establish the existence of a fixed point. Define a function $f : X \to \mathbb{R}$ by $f(x) = \|x - T(x)\|$. Since $T$ is a continuous map on $X$ and the norm $\|\cdot\|$ is a continuous function, $f$ is continuous on $X$. As $X$ is a compact set, the Extreme Value Theorem guarantees that $f$ attains its minimum value on $X$. Let $x^* \in X$ be a point such that $f(x^*) = \min_{x \in X} f(x)$. Let this minimum value be $d$.

We claim that $d = 0$. We argue by contradiction. Assume $d > 0$. This implies $x^* \neq T(x^*)$. Let $y = T(x^*)$. Since $T : X \to X$, we have $y \in X$. By the given condition, as $x^* \neq y$,

$$f(y) = \|y - T(y)\| = \|T(x^*) - T(T(x^*))\| = \|T(x^*) - T(y)\| < \|x^* - y\|.$$

Substituting $y = T(x^*)$ into the right-hand side, we get

$$f(y) < \|x^* - T(x^*)\| = f(x^*) = d.$$

This contradicts the fact that $d$ is the minimum value of $f$ on $X$. Therefore, our assumption that $d > 0$ must be false. Hence, $d = 0$, which means $\|x^* - T(x^*)\| = 0$, so $T(x^*) = x^*$. Thus, $x^*$ is a fixed point of $T$.

To prove uniqueness, suppose $x^*$ and $y^*$ are two distinct fixed points of $T$, so $T(x^*) = x^*$, $T(y^*) = y^*$, and $x^* \neq y^*$. Then we have

$$\|x^* - y^*\| = \|T(x^*) - T(y^*)\|.$$

However, the given condition states that $\|T(x) - T(y)\| < \|x - y\|$ for all distinct $x, y \in X$. Since $x^* \neq y^*$, this leads to a contradiction. Thus, the fixed point is unique. □

(b) *Proof.* Let $x^*$ be the unique fixed point of $T$. Consider the sequence $\{x_k\}_{k=0}^{\infty}$ defined by the iteration $x_{k+1} = T(x_k)$ for any initial point $x_0 \in X$. Let $d_k = \|x_k - x^*\|$ for $k \geq 0$.

If $x_k = x^*$ for some $k$, then $x_{k+1} = T(x_k) = T(x^*) = x^*$, and all subsequent terms will be $x^*$. The sequence converges to $x^*$.

Now, assume $x_k \neq x^*$ for all $k$. Then for each $k$,

$$d_{k+1} = \|x_{k+1} - x^*\| = \|T(x_k) - T(x^*)\| < \|x_k - x^*\| = d_k.$$

The sequence $\{d_k\}$ is a strictly decreasing sequence of non-negative real numbers. Therefore, it must converge to some limit $M \geq 0$. We want to show that $M = 0$.

We argue by contradiction. Assume $M > 0$. The sequence $\{x_k\}$ lies in the compact set $X$, so there must exist a convergent subsequence $\{x_{k_j}\}_{j=1}^{\infty}$. Let $\lim_{j\to\infty} x_{k_j} = z$ for some $z \in X$. By the continuity of the norm,

$$\|z - x^*\| = \lim_{j\to\infty} \|x_{k_j} - x^*\| = \lim_{j\to\infty} d_{k_j} = M.$$

Since $M > 0$, it follows that $z \neq x^*$.

Now consider the sequence $\{x_{k_j+1}\}$. Since $x_{k_j+1} = T(x_{k_j})$ and $T$ is continuous, we have

$$\lim_{j\to\infty} x_{k_j+1} = T(\lim_{j\to\infty} x_{k_j}) = T(z).$$

The limit of the corresponding distance sequence $\{d_{k_j+1}\}$ must also be $M$. Thus,

$$\|T(z) - x^*\| = \lim_{j\to\infty} \|x_{k_j+1} - x^*\| = \lim_{j\to\infty} d_{k_j+1} = M.$$

So we have established both $\|z - x^*\| = M$ and $\|T(z) - T(x^*)\| = \|T(z) - x^*\| = M$. However, since $z \neq x^*$, the defining property of the operator $T$ requires that

$$\|T(z) - T(x^*)\| < \|z - x^*\|.$$

Substituting our findings, this implies $M < M$, which is a clear contradiction. Therefore, our assumption that $M > 0$ must be false. We conclude that $M = 0$.

Since the limit of the sequence $\{d_k\} = \{\|x_k - x^*\|\}$ is 0, this implies that the sequence $\{x_k\}$ converges to $x^*$. □

27.

**Problem.** Let $f : [0, 1] \to [0, 1]$ be a continuous function. Consider the fixed point iteration $x_{k+1} = f(x_k)$ with a certain $x_0 \in [0, 1]$. If $x_k - x_{k+1} \to 0$, is it guaranteed that $\{x_k\}$ converges?

**Solution**. Here we prove that under the assumption that $f$ has a unique fixed point.

*Proof.* Since $f : [0, 1] \to [0, 1]$ is continuous, define the auxiliary function $g(x) = f(x) - x$. The function $g$ is continuous on $[0, 1]$ as the difference of continuous functions. Note that $g(0) = f(0) - 0 \geq 0$ since $f(0) \in [0, 1]$, and $g(1) = f(1) - 1 \leq 0$ since $f(1) \in [0, 1]$. By the intermediate value theorem, there exists $\xi \in [0, 1]$ such that $g(\xi) = 0$, so $f(\xi) = \xi$, confirming that $f$ has at least one fixed point.

Assume that $f$ has a unique fixed point, denoted $\xi$. Since $[0, 1]$ is compact, the sequence $\{x_k\}$ has a convergent subsequence $\{x_{k_j}\}$ with limit $z \in [0, 1]$. By continuity of $f$,

$$\lim_{j\to\infty} (x_{k_j} - x_{k_j+1}) = \lim_{j\to\infty} (x_{k_j} - f(x_{k_j})) = z - f(z).$$

Given that $\lim_{k\to\infty}(x_k - x_{k+1}) = 0$, it follows that $z - f(z) = 0$, so $z$ is a fixed point of $f$. Since $\xi$ is the unique fixed point, $z = \xi$. Thus, every convergent subsequence of $\{x_k\}$ converges to $\xi$. As $[0, 1]$ is compact, this implies that $\{x_k\}$ converges to $\xi$. □

28.

**Problem.** Suppose that $f$ is a twice differentiable function on $[0, 1]$ satisfying

$$f'(0) = 0 = f'(1).$$

Show that there exists a number $\xi \in (0, 1)$ such that

$$|f''(\xi)| = 4|f(0) - f(1)|.$$

**Solution**. *Proof.* To prove the conclusion, We need to assume that $f''$ is continuous on $[0, 1]$. The proof relies on applying Taylor's theorem and the Intermediate Value Theorem.

By Taylor's theorem with the Lagrange form of the remainder, we can expand $f(1/2)$ around the points $x = 0$ and $x = 1$. Expanding around $x = 0$, there exists $\xi_1 \in (0, 1/2)$ such that:

$$f\left(\frac{1}{2}\right) = f(0) + f'(0)\left(\frac{1}{2}\right) + \frac{f''(\xi_1)}{2!}\left(\frac{1}{2}\right)^2.$$

Given that $f'(0) = 0$, this simplifies to:

$$f\left(\frac{1}{2}\right) = f(0) + \frac{1}{8}f''(\xi_1). \quad (1)$$

Similarly, expanding around $x = 1$, there exists $\xi_2 \in (1/2, 1)$ such that:

$$f\left(\frac{1}{2}\right) = f(1) + f'(1)\left(\frac{1}{2} - 1\right) + \frac{f''(\xi_2)}{2!}\left(\frac{1}{2} - 1\right)^2.$$

Given that $f'(1) = 0$, this simplifies to:

$$f\left(\frac{1}{2}\right) = f(1) + \frac{1}{8}f''(\xi_2). \quad (2)$$

Equating (1) and (2), we have:

$$f(0) + \frac{1}{8}f''(\xi_1) = f(1) + \frac{1}{8}f''(\xi_2),$$

which rearranges to:

$$f''(\xi_2) - f''(\xi_1) = 8(f(0) - f(1)).$$

Let $C = 4|f(0) - f(1)|$. Taking the absolute value of the equation above gives:

$$|f''(\xi_2) - f''(\xi_1)| = 8|f(0) - f(1)| = 2C.$$

By the triangle inequality, $|f''(\xi_2) - f''(\xi_1)| \leq |f''(\xi_2)| + |f''(\xi_1)|$. Thus,

$$2C \leq |f''(\xi_1)| + |f''(\xi_2)|.$$

Let $|f''(\xi_{\max})| = \max\{|f''(\xi_1)|, |f''(\xi_2)|\}$. Then $|f''(\xi_1)| + |f''(\xi_2)| \leq 2|f''(\xi_{\max})|$. Combining these inequalities, we get:

$$2C \leq 2|f''(\xi_{\max})| \implies C \leq |f''(\xi_{\max})|.$$

This shows there is a point, $\xi_{\max} \in \{\xi_1, \xi_2\} \subset (0, 1)$, where the magnitude of the second derivative is at least $C$.

Next, we use the given condition $f'(0) = f'(1)$. Since $f'$ is differentiable (and thus continuous) on $[0, 1]$, by Rolle's Theorem, there exists a number $c \in (0, 1)$ such that $f''(c) = 0$.

We have established two facts:

1. There exists $c \in (0, 1)$ such that $|f''(c)| = 0$.

2. There exists $\xi_{\max} \in (0, 1)$ such that $|f''(\xi_{\max})| \geq C$.

If $C = 0$, then $f(0) = f(1)$, and we can choose $\xi = c$ to satisfy $|f''(c)| = 0$. If $C > 0$, we have $0 \leq C \leq |f''(\xi_{\max})|$. Since we assumed $f''$ is continuous, the function $|f''|$ is also continuous on the closed interval between $c$ and $\xi_{\max}$. By the Intermediate Value Theorem, $|f''|$ must take on every value between $|f''(c)| = 0$ and $|f''(\xi_{\max})|$. Since $C$ is such a value, there must exist a number $\xi$ in the interval between $c$ and $\xi_{\max}$ (and thus in $(0, 1)$) such that:

$$|f''(\xi)| = C = 4|f(0) - f(1)|.$$

This completes the proof. □

29.

**Problem.** Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function.

(a) Suppose that $\lim_{k\to\infty} f(k + x) \to 0$ for all $x \in \mathbb{R}$. Is it guaranteed that $f(x) \to 0$ when $x \to +\infty$?

(b) Suppose that $\lim_{k\to\infty} f(kx) \to 0$ for all $x > 0$. Is it guaranteed that $f(x) \to 0$ when $x \to +\infty$?

**Solution**.

(a) The statement is false. We provide a counterexample by constructing a continuous function $f : \mathbb{R} \to \mathbb{R}$ such that $\lim_{k\to\infty} f(k + x) = 0$ for all $x \in \mathbb{R}$ (where $k \in \mathbb{Z}$), but the limit $\lim_{x\to+\infty} f(x)$ is not zero.

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a continuous "hat" function defined by $\phi(t) = \max(0, 1 - |t|)$. The support of $\phi$ is $[-1, 1]$, and its maximum value is $\phi(0) = 1$.

For each integer $n \geq 2$, we define a function $f_n : \mathbb{R} \to \mathbb{R}$ by scaling and translating $\phi$. Let $x_n = n + \frac{1}{n}$ and let the width of the support be $w_n = \frac{1}{n^2}$. We define

$$f_n(x) := \phi\left(\frac{x - x_n}{w_n/2}\right) = \phi\left(2n^2\left(x - \left(n + \frac{1}{n}\right)\right)\right).$$

Each $f_n$ is continuous, attains a maximum value of 1 at $x = x_n$, and has a compact support, namely the interval $I_n = [x_n - w_n/2, x_n + w_n/2] = [n + \frac{1}{n} - \frac{1}{2n^2}, n + \frac{1}{n} + \frac{1}{2n^2}]$. For $n \geq 2$, the intervals $I_n$ are disjoint.

We define the function $f : \mathbb{R} \to \mathbb{R}$ as the sum

$$f(x) = \sum_{n=2}^{\infty} f_n(x).$$

Since the supports $I_n$ are disjoint, for any $x \in \mathbb{R}$, at most one term in the sum is non-zero. The function $f$ is therefore well-defined. To see that $f$ is continuous, consider any compact set $K \subset \mathbb{R}$. $K$ can only intersect a finite number of the supports $I_n$. On $K$, $f$ is a finite sum of continuous functions, and is therefore continuous. Since continuity is a local property, $f$ is continuous on $\mathbb{R}$.

First, we show that $\lim_{x\to+\infty} f(x) \neq 0$. Consider the sequence of points $(x_n)_{n\geq 2}$ where $x_n = n + \frac{1}{n}$. Clearly, $x_n \to +\infty$ as $n \to \infty$. By construction, we have

$$f(x_n) = f_n(x_n) = \phi(0) = 1 \quad \forall n \geq 2.$$

Since we have found a sequence of points tending to infinity for which the function value is constantly 1, it is not possible that $\lim_{x\to+\infty} f(x) = 0$.

Next, we prove that $\lim_{k\to\infty} f(k + x) = 0$ for all $x \in \mathbb{R}$, where $k$ is an integer. Fix an arbitrary $x \in \mathbb{R}$. We want to show that there exists an integer $K$ such that for all integers $k > K$, $f(k + x) = 0$. The value $f(k + x)$ is non-zero if and only if the point $k + x$ lies in the support $I_n$ of some $f_n$ for an integer $n \geq 2$. This condition is equivalent to

$$\left|(k + x) - \left(n + \frac{1}{n}\right)\right| \leq \frac{1}{2n^2}.$$

Let $m = n - k$ be an integer. The inequality can be rewritten as

$$\left|x - m - \frac{1}{k + m}\right| \leq \frac{1}{2(k + m)^2}.$$

We need to show that for a fixed $x$, this inequality holds for only a finite number of pairs of integers $(k, m)$. Let us consider large $k$. For the LHS to be small, $n$ must be close to $k$, so $m = n - k$ must be bounded. Let's analyze the inequality for any fixed integer $m$. As $k \to \infty$, the RHS tends to 0, while the LHS tends to $|x - m|$.

*Case1:* If $x$ is not an integer, then $|x - m| > 0$ for all integers $m$. For any fixed $m$, there exists an integer $K_m$ such that for all $k > K_m$, the inequality fails, as the LHS will be bounded away from zero while the RHS approaches zero.

*Case2:* If $x$ is an integer, say $x = m_0$. Since the case of $m \neq m_0$ is similar as previous case, we only check the case $m = m_0$. The inequality becomes:

$$\left|x - x - \frac{1}{k + x}\right| \leq \frac{1}{2(k + x)^2} \implies \frac{1}{|k + x|} \leq \frac{1}{2(k + x)^2}.$$

This simplifies to $2|k + x| \leq 1$. This inequality is false for any sufficiently large integer $k$.

This shows that for any fixed $x \in \mathbb{R}$, the point $k + x$ can only lie in the support $I_n$ for a finite number of integers $k$. To be more precise, for a given $x$, there exists an integer $K$ such that for all integers $k > K$, the point $k + x$ does not belong to any support interval $I_n$. Thus, for all $k > K$, $f(k + x) = 0$. This implies, by definition of the limit of a sequence, that

$$\lim_{k\to\infty} f(k + x) = 0.$$

We have successfully constructed a function that satisfies the premises but not the conclusion, thus disproving the original statement.

(b) *Proof.* We aim to show that for any given $\epsilon > 0$, there exists a real number $T > 0$ such that for all $y > T$, we have $|f(y)| \leq \epsilon$.

Let $\epsilon > 0$ be fixed. For each integer $N \geq 1$, we define the set $E_N$ as follows:

$$E_N = \left\{ x \in \mathbb{R}^+ : |f(kx)| \leq \epsilon \text{ for all integers } k \geq N \right\}.$$

Each set $E_N$ is closed in the metric space $\mathbb{R}^+ = (0, \infty)$. To see this, note that for any fixed integer $k$, the function $g_k(x) = f(kx)$ is continuous on $\mathbb{R}^+$. The set $S_k = \{x \in \mathbb{R}^+ : |f(kx)| \leq \epsilon\}$ is the preimage of the closed interval $[-\epsilon, \epsilon]$ under the continuous map $|g_k|$, and is therefore closed. Since $E_N = \bigcap_{k=N}^{\infty} S_k$, it is an intersection of closed sets and is thus closed.

The hypothesis $\lim_{k\to\infty} f(kx) = 0$ for every $x > 0$ implies that for each $x \in \mathbb{R}^+$, there exists an integer $N_x$ such that for all $k \geq N_x$, we have $|f(kx)| \leq \epsilon$. By definition, this means $x \in E_{N_x}$. Consequently, every point in $\mathbb{R}^+$ belongs to at least one $E_N$, which gives the decomposition:

$$\mathbb{R}^+ = \bigcup_{N=1}^{\infty} E_N.$$

The space $\mathbb{R}^+$, being an open subset of the complete metric space $\mathbb{R}$, is itself a complete metric space. By the Baire Category Theorem, $\mathbb{R}^+$ cannot be written as a countable union of nowhere-dense sets. Since each $E_N$ is a closed set, at least one of them must not be nowhere-dense. Let this set be $E_{N_0}$. A closed set is not nowhere-dense if and only if its interior is non-empty. Therefore, there exists an integer $N_0 \geq 1$ and an open interval $(a, b)$ with $0 < a < b$ such that $(a, b) \subseteq E_{N_0}$.

This implies that for every $x \in (a, b)$ and for every integer $k \geq N_0$, we have $|f(kx)| \leq \epsilon$.

Next, we show that a sufficiently large ray $(T, \infty)$ can be covered by scaled versions of the interval $(a, b)$. Choose an integer $M \geq N_0$ large enough to satisfy: $M > \frac{a}{b-a}$.

This implies $M(b-a) > a$, which rearranges to $Mb > (M+1)a$. This inequality guarantees that the interval $(Ma, Mb)$ overlaps with the next one, $((M+1)a, (M+1)b)$.

Such an integer exists; we can take $M = \max(N_0, \lfloor \frac{a}{b-a} \rfloor + 1)$. The sequence of overlapping intervals $\bigcup_{k=M}^{\infty}(ka, kb)$ forms a continuous ray starting at $Ma$. Thus, we have $(Ma, \infty) = \bigcup_{k=M}^{\infty}(ka, kb)$.

Let $T = Ma$. We now show that for any $y > T$, $|f(y)| \leq \epsilon$. If $y > T$, then $y \in (Ma, \infty)$. By the covering property, there must exist an integer $k \geq M$ such that $y \in (ka, kb)$. Let $x_0 = y/k$. The condition $y \in (ka, kb)$ is equivalent to $a < x_0 < b$, so $x_0 \in (a, b)$. Since $(a, b) \subseteq E_{N_0}$, we know that for any integer $j \geq N_0$, it holds that $|f(jx_0)| \leq \epsilon$.

We can express $y$ as $y = kx_0$. The multiplier for $x_0$ is $k$. By our choice of $M$, we have $k \geq M \geq N_0$. This satisfies the condition $k \geq N_0$ required by the definition of $E_{N_0}$. Therefore, we can conclude:

$$|f(y)| = |f(kx_0)| \leq \epsilon.$$

Since our choice of $\epsilon > 0$ was arbitrary, we have shown that for any $\epsilon > 0$, there exists $T > 0$ such that $y > T \implies |f(y)| \leq \epsilon$. This is precisely the definition of $\lim_{y\to\infty} f(y) = 0$. $\square$

30.

**Problem.** Suppose that $f$ is a continuous function over $[0,1]$ and

$$\int_0^x [f(t)]^2 dt \leq f(x) \quad \text{for all } x \in [0,1].$$

(a) Show that

$$\min_{x\in[0,1]} f(x) \leq 2.$$

(b) Is the bound in (a) tight or not?

**Solution**.

(a) *Proof.* Define $g(x) = \int_0^x [f(t)]^2 dt$. By hypothesis, $g(x) \leq f(x)$ for all $x \in [0,1]$. Note that:

- $g(0) = 0$ and $g$ is non-decreasing since $g'(x) = [f(x)]^2 \geq 0$.
- $f(x) \geq g(x) \geq 0$ for all $x$, so $f$ is non-negative.

Let $m = \min_{x\in[0,1]} f(x) \geq 0$. If $m = 0$, the result holds trivially. Assume $m > 0$. Then:

$$g(x) \geq \int_0^x m^2 dt = m^2 x \quad \forall x \in [0,1].$$

Since $g(x) \leq f(x)$ and $g'(x) = [f(x)]^2$, we have:

$$g'(x) = [f(x)]^2 \geq [g(x)]^2 \quad \forall x \in [0,1].$$

For any $a \in (0,1)$, consider $h(x) = 1/g(x)$ on $[a,1]$. Then:

$$h'(x) = -\frac{g'(x)}{[g(x)]^2} \leq -1.$$

Integrating from $a$ to 1:

$$h(1) - h(a) \leq -(1-a) \implies \frac{1}{g(1)} - \frac{1}{g(a)} \leq -(1-a).$$

Rearranging gives:

$$g(a) \leq \frac{1}{1-a} \quad \forall a \in (0,1).$$

Combining with the earlier inequality $g(a) \geq m^2 a$:

$$m^2 a \leq g(a) \leq \frac{1}{1-a} \quad \forall a \in (0,1).$$

Choosing $a = \frac{1}{2}$:

$$m^2 \cdot \frac{1}{2} \leq \frac{1}{1-\frac{1}{2}} = 2 \implies m^2 \leq 4 \implies m \leq 2.$$

□

(b) *Proof.* The proof consists of two parts. First, we demonstrate that no function satisfying the conditions can attain a minimum of 2. Second, we construct a family of functions $\{f_\epsilon\}_{\epsilon>0}$ that satisfy the conditions and whose minima can be made arbitrarily close to 2.

Let $f$ be a continuous function on $[0,1]$ satisfying the given integral inequality, and let $m = \min_{x\in[0,1]} f(x)$. We know from part (a) that $m \leq 2$. We proceed by contradiction to show that $m \neq 2$.

Assume $m = 2$. Let $g(x) = \int_0^x [f(t)]^2\,dt$. Since $f(t) \geq m = 2$ for all $t \in [0,1]$, we can establish a lower bound for $g(x)$:

$$g(x) = \int_0^x [f(t)]^2\,dt \geq \int_0^x 2^2\,dt = 4x \quad \text{for all } x \in [0,1].$$

From the hypothesis, $f(x) \geq g(x)$. Since $f$ is continuous, $g$ is continuously differentiable with $g'(x) = [f(x)]^2$. Therefore,

$$g'(x) = [f(x)]^2 \geq [g(x)]^2.$$

Since $m = 2$, $f$ is not identically zero, which implies $g(x) > 0$ for all $x \in (0,1]$. We can thus divide by $[g(x)]^2$:

$$\frac{g'(x)}{[g(x)]^2} \geq 1 \quad \text{for all } x \in (0,1].$$

Let $h(x) = -1/g(x)$. The inequality becomes $h'(x) \geq 1$. For any $a \in (0,1)$, we can integrate $h'(t) \geq 1$ over $[a,1]$:

$$h(1) - h(a) = \int_a^1 h'(t)\,dt \geq \int_a^1 1\,dt = 1 - a.$$

Substituting back $h(x) = -1/g(x)$, we have:

$$-\frac{1}{g(1)} + \frac{1}{g(a)} \geq 1 - a \implies \frac{1}{g(a)} \geq \frac{1}{g(1)} + 1 - a.$$

Since $g(1) = \int_0^1 [f(t)]^2\,dt \geq \int_0^1 4\,dt = 4$, the term $1/g(1)$ is positive. Thus, we have a strict inequality:

$$\frac{1}{g(a)} > 1 - a \implies g(a) < \frac{1}{1-a}.$$

We have now established for any $a \in (0,1)$ that $4a \leq g(a)$ and $g(a) < 1/(1-a)$. Combining these gives:

$$4a < \frac{1}{1-a} \implies 4a(1-a) < 1.$$

This inequality must hold for all $a \in (0,1)$. However, for $a = 1/2$, the left side is $4(1/2)(1-1/2) = 1$. The statement $1 < 1$ is false. This is a contradiction. Therefore, our initial assumption that $m = 2$ is false. We conclude that $\min_{x\in[0,1]} f(x) < 2$.

To show that 2 is the supremum of the possible minima, we construct a family of functions that satisfy the hypothesis and whose minima can be made arbitrarily close to 2.

For any $\delta \in (1/2, 1]$, define the function $f_\delta$ on $[0, 1]$ as:

$$f_\delta(x) = \begin{cases} \frac{1}{\delta} & \text{if } 0 \le x \le \delta, \\ \frac{1}{2\delta - x} & \text{if } \delta < x \le 1. \end{cases}$$

The function $f_\delta$ is well-defined and continuous on $[0, 1]$. Continuity at $x = \delta$ is confirmed by checking the limit from both sides: $f_\delta(\delta) = 1/\delta$ and $\lim_{x\to\delta^+} f_\delta(x) = 1/(2\delta - \delta) = 1/\delta$. The minimum value of $f_\delta$ on $[0, 1]$ is $1/\delta$, which occurs on the interval $[0, \delta]$.

We must verify that $f_\delta$ satisfies the integral inequality $\int_0^x [f_\delta(t)]^2\, dt \le f_\delta(x)$.

- For $x \in [0, \delta]$:

$$\int_0^x [f_\delta(t)]^2\, dt = \int_0^x \left(\frac{1}{\delta}\right)^2 dt = \frac{x}{\delta^2}.$$

  The condition is $\frac{x}{\delta^2} \le \frac{1}{\delta}$, which simplifies to $x \le \delta$. This holds true in this interval.

- For $x \in (\delta, 1]$:

$$\begin{aligned} \int_0^x [f_\delta(t)]^2\, dt &= \int_0^\delta \left(\frac{1}{\delta}\right)^2 dt + \int_\delta^x \frac{1}{(2\delta - t)^2}\, dt \\ &= \frac{\delta}{\delta^2} + \left[\frac{1}{2\delta - t}\right]_\delta^x \\ &= \frac{1}{\delta} + \left(\frac{1}{2\delta - x} - \frac{1}{2\delta - \delta}\right) \\ &= \frac{1}{2\delta - x}. \end{aligned}$$

  In this interval, $f_\delta(x) = \frac{1}{2\delta - x}$, so the inequality holds with equality.

Thus, $f_\delta$ satisfies the hypothesis for any $\delta \in (1/2, 1]$. The minimum value is $m_\delta = 1/\delta$. As $\delta \to (1/2)^+$, the minimum $m_\delta \to 2$. For any given $\epsilon > 0$, we can choose $\delta \in (1/2, 1/(2-\epsilon))$ to find a function whose minimum is greater than $2 - \epsilon$.

Since no function can have a minimum of 2, but we can construct functions with minima arbitrarily close to 2, we conclude that 2 is the supremum of the set of possible minimum values. □

31.

**Problem.** Show that

$$\min_{\|x\|_2=1} \|Ax\|_\infty \le \frac{1}{n}\|A\|_F$$

for all matrix $A \in \mathbb{R}^{n\times n}$, or find a counterexample.

**Solution**. I can only prove a weaker upper bound:

$$\min_{\|x\|_2=1} \|Ax\|_\infty \le \frac{1}{\sqrt{n}}\|A\|_F.$$

*Proof.* The proof relies on an averaging argument over the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$. Let $x$ be a vector chosen uniformly at random from $S^{n-1}$. We compute the expected value of $\|Ax\|_2^2$. By definition, we have $\|Ax\|_2^2 = x^T A^T Ax$. By the linearity of expectation,

$$\mathbb{E}[\|Ax\|_2^2] = \mathbb{E}[x^T A^T Ax] = \sum_{i,j=1}^{n} (A^T A)_{ij} \mathbb{E}[x_i x_j].$$

For a random vector on the unit sphere, we have $\mathbb{E}[x_i x_j] = \frac{1}{n}\delta_{ij}$, where $\delta_{ij}$ is the Kronecker delta. This is because $\mathbb{E}[x_i^2] = 1/n$ due to symmetry and the constraint $\sum \mathbb{E}[x_i^2] = 1$, while $\mathbb{E}[x_i x_j] = 0$ for $i \neq j$. Substituting this in, we find

$$\mathbb{E}[\|Ax\|_2^2] = \frac{1}{n}\sum_{i=1}^{n}(A^T A)_{ii} = \frac{1}{n}\operatorname{Tr}\big(A^T A\big) = \frac{1}{n}\|A\|_F^2.$$

Since the average value of the non-negative function $f(x) = \|Ax\|_2^2$ is $\frac{1}{n}\|A\|_F^2$, there must exist at least one vector $x_0 \in S^{n-1}$ for which its value is no greater than the average:

$$\|Ax_0\|_2^2 \leq \frac{1}{n}\|A\|_F^2.$$

For any vector $y \in \mathbb{R}^n$, the infinity norm and Euclidean norm are related by $\|y\|_\infty \leq \|y\|_2$. Applying this to the vector $y = Ax_0$, we get $\|Ax_0\|_\infty^2 \leq \|Ax_0\|_2^2$. Combining these inequalities yields

$$\|Ax_0\|_\infty^2 \leq \|Ax_0\|_2^2 \leq \frac{1}{n}\|A\|_F^2.$$

Taking the square root gives $\|Ax_0\|_\infty \leq \frac{1}{\sqrt{n}}\|A\|_F$. By the definition of the minimum, we have

$$\min_{\|x\|_2=1} \|Ax\|_\infty \leq \|Ax_0\|_\infty \leq \frac{1}{\sqrt{n}}\|A\|_F.$$

This completes the proof of the bound. □

32.

**Problem.** Show that there exists a set $\mathcal{S} \subset \mathbb{R}^n$ satisfying the following conditions.

(a) $\|x\|_2 = 1$ for all $x \in \mathcal{S}$.

(b) $|x^T y| \leq \epsilon$ for all distinct $x, y \in \mathcal{S}$.

(c) The cardinality of $\mathcal{S}$ is at least $\exp(cn\epsilon^2)$ with a certain absolute constant $c > 0$ that you must specify. [An absolute constant is a number that maintains the same value wherever it appears, e.g., 1, $\pi$, and $\log 2$.]

In theory, if a set of unit vectors in $\mathbb{R}^n$ are pairwise orthogonal, then the cardinality of the set cannot exceed $n$. Use the existence of $\mathcal{S}$ to explain why we cannot rely on such a theory in numerical computations.

**Solution**. *Proof.* We begin by sampling $M$ vectors, $\mathcal{X} = \{X_1, X_2, \ldots, X_M\}$, independently and uniformly from the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$. The size $M$ will be determined later. This construction immediately satisfies condition (a).

We define a "bad pair" as a pair of distinct vectors $(X_i, X_j)$ for which their inner product's absolute value exceeds $\epsilon$, i.e., $|X_i^T X_j| > \epsilon$.

A standard result from the concentration of measure on the sphere states that the probability $p$ of any single pair being a bad pair is exponentially small in the dimension $n$:

$$p = P\left(|X_i^T X_j| > \epsilon\right) \leq 2e^{-\frac{n\epsilon^2}{2}}.$$

Let $N_{bad}$ be the random variable that counts the total number of bad pairs in our set $\mathcal{X}$. By the linearity of expectation, the expected number of bad pairs is the number of pairs multiplied by the probability $p$:

$$E[N_{bad}] = \binom{M}{2} p < \frac{M^2}{2} p \leq \frac{M^2}{2} \cdot 2e^{-\frac{n\epsilon^2}{2}} = M^2 e^{-\frac{n\epsilon^2}{2}}.$$

We now apply the deletion algorithm: for each bad pair $(X_i, X_j)$, we remove one of the vectors (say, $X_j$) from the set $\mathcal{X}$. The final set, which we call $\mathcal{S}$, consists of the remaining vectors.

By construction, $\mathcal{S}$ satisfies condition (b). The number of vectors removed is at most $N_{bad}$. Therefore, the size of the resulting set is $|\mathcal{S}| \geq M - N_{bad}$.

A random variable must take on a value no smaller than its expectation for at least one outcome. Therefore, there must exist a specific choice of vectors for which the size of the resulting set $\mathcal{S}$ is at least its expected value:

$$|\mathcal{S}| \geq E[|\mathcal{S}|] \geq E[M - N_{bad}] = M - E[N_{bad}].$$

To ensure the final set is large, we choose $M$ such that the expected number of removed vectors, $E[N_{bad}]$, is small relative to $M$. A good choice is to ensure $E[N_{bad}] \leq M/2$. Let's choose:

$$M = \left\lfloor \frac{1}{2} e^{\frac{n\epsilon^2}{2}} \right\rfloor.$$

For this choice of $M$ (assuming $n\epsilon^2$ is large enough so $M \geq 2$, specifically $n\epsilon^2 \geq 4\ln(2)$), we can bound the expected number of bad pairs:

$$E[N_{bad}] < M^2 e^{-\frac{n\epsilon^2}{2}} \leq M \cdot \left(\frac{1}{2} e^{\frac{n\epsilon^2}{2}}\right) \cdot e^{-\frac{n\epsilon^2}{2}} = \frac{M}{2}.$$

The second inequality holds because $M \leq \frac{1}{2} e^{\frac{n\epsilon^2}{2}}$. This implies that the expected size of our final set is:

$$E[|\mathcal{S}|] > M - \frac{M}{2} = \frac{M}{2}.$$

Therefore, there exists a set $\mathcal{S}$ with cardinality:

$$|\mathcal{S}| \geq \frac{M}{2} = \frac{1}{2} \left\lfloor \frac{1}{2} e^{\frac{n\epsilon^2}{2}} \right\rfloor.$$

To satisfy the condition $|\mathcal{S}| \geq \exp(cn\epsilon^2)$, we need to show that for some $c > 0$, this bound is larger. Let's specify the constant $c = 1/4$. For $n\epsilon^2 \geq 4\ln(2)$, it is true that $\frac{1}{4}e^{n\epsilon^2/2} \geq e^{n\epsilon^2/4}$. For $n\epsilon^2 < 4\ln(2)$, we can always find two mutually orthogonal vectors since $n \geq 2$ (specifically standard basis vector $e_i, e_j$). In this case,$|\mathcal{S}| \geq 2 > e^{n\epsilon^2/4}$.

Thus, the existence is proven with $c = 1/4$. □

Explanation for numerical computations:

The existence of $\mathcal{S}$ shows that there are sets of unit vectors of size exponentially large in $n$ (for fixed $\epsilon > 0$) that are pairwise $\epsilon$-orthogonal. In contrast, exactly orthogonal sets have size at most $n$. Numerical computations cannot achieve exact orthogonality due to rounding errors. Thus, algorithms relying on orthogonality may produce vectors that are only approximately orthogonal. The exponential growth of $\epsilon$-orthogonal sets implies that such numerical errors can allow "orthogonal" sets much larger than $n$, violating theoretical linear independence assumptions and potentially causing instability.