ML Written Homework 1

Student ID: b10201054

September 22, 2025

1 Linear Algebra Recap

(a) (i) $\Rightarrow$ (ii) $\mathbf{A}$ is positive semi-definite implies $\mathbf{A}$ is symmetric. Since $\mathbf{A}$ is symmetric, $\mathbf{A}$ is diagonalizable. Suppose $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of $\mathbf{A}$, there are eigenvectors $v_1, v_2, ..., v_n$ such that $\mathbf{A}v_k = \lambda_k v_k$.

By the definition of positive semi-definite, we have

$$\langle \mathbf{A}v_k, v_k \rangle = \langle \lambda_k v_k, v_k \rangle = \lambda_k \|v_k\|^2 \geq 0$$

Hence, all eigenvalues $\lambda_k \geq 0$.

(ii) $\Rightarrow$ (iii) $\mathbf{A}$ is symmetric and all of eigenvalues λ_k are non negative. Since $\mathbf{A}$ is symmetric, $\mathbf{A}$ is diagonalizable. That is, there are P, D such that $\mathbf{A} = P^T DP$, where D $= diag(\lambda_1, \lambda_2, ..., \lambda_n)$.

Since $\lambda_k \geq 0$, we can find $e_k = \sqrt{\lambda_k}$ and $E = diag(e_1, e_2, ..., e_n)$ so that $D = E^2$. Now we have

$$\mathbf{A} = P^T DP = P^T EEP = (P^T E)(EP) = (EP)^T (EP)$$

Hence, $\mathbf{A} = \mathbf{B}^T\mathbf{B}$ where $\mathbf{B} = EP$

(iii) $\Rightarrow$ (i) $\mathbf{A} = \mathbf{B}^T\mathbf{B}$, then we have

$$u^T\mathbf{A}u = u^T\mathbf{B}^T\mathbf{B}u = (\mathbf{B}u)^T(\mathbf{B}u) = \|\mathbf{B}u\|^2 \geq 0$$

for all $u \in V$. Hence, $\mathbf{A}$ is positive semi-definite.

(b) Note that $\left\langle \mathbf{A}^T\mathbf{A}x, x \right\rangle = \langle \mathbf{A}x, \mathbf{A}x \rangle = \|\mathbf{A}x\|^2 \geq 0$ for all $x \in V$. By (a), $\mathbf{A}^T\mathbf{A}$ is positive semi-definite. Hence the eigenvalues of $\mathbf{A}^T\mathbf{A}$ are nonnegative.

(c) For $x \in Null(\mathbf{A})$, we have $\mathbf{A}^T\mathbf{A}x = 0$, so that $x \in Null(\mathbf{A}^T\mathbf{A})$.

For $x \in Null(\mathbf{A}^T\mathbf{A})$,

$$\left\langle \mathbf{A}^T\mathbf{A}x, x \right\rangle = \langle \mathbf{A}x, \mathbf{A}x \rangle = \|\mathbf{A}x\|^2 = 0$$

which implies $\mathbf{A}x = 0$. Hence $x \in Null(\mathbf{A})$. By the above, we can find that $Null(\mathbf{A}) = Null(\mathbf{A}^T\mathbf{A})$.

(d) $\langle \mathbf{A}v_i, \mathbf{A}v_j \rangle = \langle \mathbf{A}^T\mathbf{A}v_i, v_j \rangle = \langle \lambda_i v_i, v_j \rangle = 0$ for any $i \neq j$. Hence $\{\mathbf{A}v_1, ..., \mathbf{A}v_r\}$ is an orthogonal set.

(e) Suppose V is a $n \times n$ matrix whose i-th column is v_i, then V is an orthogonal matrix where $\mathbf{A}^T\mathbf{A} = VDV^T$. Let U be a $m \times m$ matrix with its i-th column u_i, where $u_i = \frac{\mathbf{A}v_i}{\sigma_i}$, and Σ be a $m \times n$ matrix with its diagonal entries σ_i.

Then we have the relation:

$$U\Sigma = \mathbf{A}V \Rightarrow \mathbf{A} = U\Sigma V^T$$

By (d), $\{\mathbf{A}v_1, ..., \mathbf{A}v_r\}$ is an orthogonal set, we have $\{\mathbf{u}_1, ..., \mathbf{u}_r\}$ is also an orthogonal set. Hence U is also an orthogonal matrix.

2 Definition of Derivative as Linear Operator

Assume $f(n), g(n), h(n)$ are all positive function.

(a) $f(n) \in O(g(n))$ and $g(n) \in O(h(n))$ gives us $f(n) \leq c_1 g(n), g(n) \leq c_2 h(n)$. Hence we have

$$f(n) \leq c_1 g(n) \leq c_1 c_2 h(n) \leq ch(n)$$

This gives us $f(n) \in O(h(n))$

(b) $f(n) \in O(g(n))$ means $f(n) \leq c_1 g(n)$. Since $f(n), g(n)$ are positive, we have c_1 also be positive. Hence we have

$$g(n) \geq \frac{1}{c_1} f(n)$$

This gives us $g(n) \in \Omega(f(n))$

(c) $f(n) \in \omega(g(n))$ means for all constants $c > 0$, there exists an n_0 such that $f(n) > cg(n)$ when $n \geq n_0$.

Then we have for all constants $c > 0$, there exists an n_0 such that $g(n) < \frac{1}{c} f(n)$ when $n \geq n_0$. This gives us $g(n) \in o(f(n))$.

3 Matrix Calculus

(a) Here we use limit supremum to prove.

$$\limsup_{n\to\infty} \frac{2^{2n}}{2^{2n+1024}} = \limsup_{n\to\infty} \frac{1}{2^{1024}} = \frac{1}{2^{1024}} < \infty$$

Hence $2^{2n} \in O(2^{2n+1024})$

(b) Note that $\log_{1024} n^2 = 2\log_{1024} n = 2\frac{\log n}{\log 1024}, \log_2 n^{1024} = 1024 \log_2 n = 1024 \frac{\log n}{\log 2}$ Then we have

$$\limsup_{n\to\infty} \frac{\log_{1024} n^2}{\log_2 n^{1024}} = \limsup_{n\to\infty} \frac{\frac{2\log n}{\log 1024}}{\frac{1024 \log n}{\log 2}} = \limsup_{n\to\infty} \frac{2\log 2}{1024 \log 1024} = \frac{1}{5120} \neq 0$$

Hence $\log_{1024} n^2 \notin o(\log_2 n^{1024})$

(c) From the Stirling's approximation, we have $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ when n is large enough. Then we get

$$\log(n!) \sim \log\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right) = n(\log n - \log e) + 0.5(\log n + \log 2\pi)$$

$$\frac{\log(n!)}{n\log n} = 1 - \frac{\log e}{\log n} + \frac{1}{2n} + \frac{\log 2\pi}{2n\log n}$$

Hence, we have $\lim_{n\to\infty}\frac{\log(n!)}{n\log n} = 1$. By the result of 1.(c), we have

$$\log(n!) \in \Theta(n\log n)$$

(d) $x^{2.5} = \lfloor x^{2.5}\rfloor + \{x^{2.5}\}$, where $\{x^{2.5}\}$ is the fraction part of $x^{2.5}$. Then we have

$$\lim_{x\to\infty}\frac{\lfloor x^{2.5}\rfloor}{x^{2.5}} = \lim_{x\to\infty}\frac{x^{2.5} - \{x^{2.5}\}}{x^{2.5}} = 1$$

By the result of 1.(c), we have $\lfloor x^{2.5}\rfloor \in \Theta(x^{2.5}) \Rightarrow \lfloor x^{2.5}\rfloor \in \Omega(x^{2.5})$

(e) $\left\lceil\frac{x}{2}\right\rceil = \frac{x}{2} + 1 - \left\{\frac{x}{2}\right\}$, where $\left\{\frac{x}{2}\right\}$ is the fraction part of $\frac{x}{2}$.

$$\frac{x\left\lceil\frac{x}{2}\right\rceil}{x^2} = \frac{x(\frac{x}{2} + 1 - \left\{\frac{x}{2}\right\})}{x^2} = \frac{1}{2} + \frac{1}{x} - \frac{\left\{\frac{x}{2}\right\}}{x}$$

Then we have

$$x\left\lceil\frac{x}{2}\right\rceil = 2x^2 + \left(1 - \left\{\frac{x}{2}\right\}\right)x$$

When $x > 1$, we know that $x < x^2$, hence the $\left(1 - \left\{\frac{x}{2}\right\}\right)x$ part has the upper-bound x^2. So that we finally get the result

$$x\left\lceil\frac{x}{2}\right\rceil \le 3x^2$$

for $x > 1$. This gives us $x\left\lceil\frac{x}{2}\right\rceil \in O(x^2)$

4 Closed-Form Linear Regression Solution

$$f_{10}(n) \succ f_9(n) \succ f_4(n) \succ f_{12}(n) \succ f_2(n) \succ f_3(n)$$
$$f_3(n) \sim f_{13}(n) \succ f_1(n) \succ f_5(n) \succ f_{11}(n) \succ f_7(n) \succ f_6(n)$$

$f_{10} \succ f_9$:

$$\lim_{n\to\infty}\frac{\log\log n}{10^{10^{10}}} = \infty$$

$f_9 \succ f_4$:

$$\lim_{n\to\infty}\frac{\sqrt{\log n}}{\log\log n} = \lim_{n\to\infty}\frac{\frac{1}{2x\sqrt{\log x}}}{\frac{1}{x\log x}} = \lim_{n\to\infty}\frac{\sqrt{\log x}}{2} = \infty$$

$f_4 \succ f_{12}$:

$$\lim_{n\to\infty} \frac{(\log n)^{1.5}}{\sqrt{\log n}} = \lim_{n\to\infty} \log n = \infty$$

$f_{12} \succ f_2$:

$$10^{\log n^2} = (e^{\log 10})^{2\log n} = (e^{\log n})^{2\log 10} = n^{2\log 10}$$

$$\lim_{n\to\infty} \frac{n^{2\log 10}}{(\log n)^{1.5}} = \lim_{n\to\infty} \frac{(2\log 10) n^{(2\log 10)-1}}{\frac{1.5(\log n)^{0.5}}{n}} = \lim_{n\to\infty} \frac{(2\log 10) n^{2\log 10}}{1.5(\log n)^{0.5}} = \infty$$

$f_2 \succ f_3$:

$$\lim_{n\to\infty} \frac{n^6}{n^{2\log 10}} = \infty \text{ since } 6 > 2\log 10$$

$f_3 \sim f_{13}$:

$$\sum_{k=1}^{n} k^5 = \frac{1}{12} n^2 (n+1)^2 (2n^2 + 2n - 1) = \frac{1}{6} n^6 + \frac{1}{2} n^5 + \frac{5}{12} n^4 - \frac{1}{12} n^2$$

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} k^5}{n^6} = \frac{1}{6}$$

$f_{13} \succ f_8$: Suppose d is a constant, we have

$$d \log n < (\log\log n)(\log n) \Rightarrow n^d < n^{\log\log n}$$

$$\lim_{n\to\infty} \frac{n^{\log\log n}}{\sum_{k=1}^{n} k^5} = \lim_{n\to\infty} \frac{n^{\log\log n}}{\frac{1}{6}n^5 + \frac{1}{2}n^4 + \frac{5}{12}n^3 - \frac{1}{12}n} = \infty$$

$f_8 \succ f_1$:

$$(\log\log n)(\log n) < \log n \log n < n < n \log \frac{3}{2} \Rightarrow n^{\log\log n} < (\frac{3}{2})^n$$

$$\lim_{n\to\infty} \frac{n(\frac{3}{2})^n}{n^{\log\log n}} = \infty$$

$f_1 \succ f_5$:

$$\lim_{n\to\infty} \frac{3^n}{n(\frac{3}{2})^n} = \lim_{n\to\infty} \frac{2^n}{n} = \infty$$

$f_5 \succ f_{11}$

$$\lim_{n\to\infty} \frac{n!}{3^n} = \lim_{n\to\infty} \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{3^n} = \lim_{n\to\infty} \sqrt{2\pi n}\left(\frac{n}{3e}\right)^n = \infty$$

$f_{11} \succ f_7$

$$\lim_{n\to\infty} \frac{n^n}{n!} = \lim_{n\to\infty} \frac{n^n}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n} = \lim_{n\to\infty} \frac{e^n}{\sqrt{2\pi n}} = \infty$$

$f_7 \succ f_6$

$$6^n \log 6 > n \log n \Rightarrow 6^{6^n} > n^n$$

5 Noise and Regularization

(a) Suppose c_1, c_2 are two constants such that $f(n) \leq c_1 h(n), g(n) \leq c_2 h(n)$, then we have

$$f(n) + g(n) \leq (c_1 + c_2)h(n) = ch(n)$$

Hence, $f(n) + g(n) \in O(h(n))$

(b) Since f, g are positive, we have $|f(n) - g(n)| \leq \max(f(n), g(n))$. Then either $|f(n) - g(n)| \in O(f(n))$ or $|f(n) - g(n)| \in O(g(n))$ would be satisfied. This results that $|f(n) - g(n)| \in O(h(n))$ because $f(n), g(n) \in O(h(n))$

(c) Since f, g are positive, suppose c_1 is a constant that $f(n) \leq c_1 g(n)$. Square both sides of the inequation, we have $(f(n))^2 \leq (c_1)^2 (g(n))^2$. Hence, $(f(n))^2 \in O((g(n))^2)$

(d) Counterexample: $f(n) = n, g(n) = \frac{n}{2}$, then we have $2^{f(n)} = 2^n, 2^{g(n)} = (\sqrt{2})^n$.

$$\lim_{n \to \infty} \frac{2^n}{(\sqrt{2})^n} = \lim_{n \to \infty} (\sqrt{2})^n = \infty$$

Hence, $2^{f(n)} \notin O(2^{g(n)})$

(e) Counterexample: Suppose $f(n) = \log n, g(n) = n, h(n) = n^2$, we have $f(g(n)) = \log n, f(h(n)) = 2 \log n$. Then $f(g(n)) \notin o(f(h(n)))$.

Problem 6

(a) The loop from line 4 to line 7 runs n times, and for each loop, the line 6 runs k times. Hence the time complexity is

$$\sum_{k=1}^{n} k\Theta(1) = \Theta(n^2)$$

(b) The loop from line 4 to line 7 runs n times, and for each loop, the line 6 runs k times. When the line 6 runs the first time of the loop, the time complexity is $\Theta(\min(x, k)) = \Theta(1)$. After the first time of the loop, the time complexity is $\Theta(\min(x, k)) = \Theta(k)$ Sum up the complexity and we have the total time complexity of the k_{th} loop be $(k-1)\Theta(k) + \Theta(1) = \Theta(k^2)$. Now compute n times, we get the time complexity

$$\sum_{k=1}^{n} \Theta(k^2) = \Theta(k^3)$$