CSC 301 - Assignment 3

Student: Brandon Tjandra,
Lecturer: Leonid Gurvitis,

> **Problem 1: Exercise 1**
>
> (a) Prove that:
>
> (i)
> $$\min_{M>0}\left(\frac{1}{2}M\cdot\mu+\frac{2d}{M}\right)=2\sqrt{\mu\cdot d}$$
>
> (ii)
> $$\text{Optimal } \mu = 2\sqrt{\frac{d}{M}}$$

i) To minimize this function, differentiate the expression with respect to M and set the derivative equal to zero:

$$f(M)=\frac{1}{2}M\mu+\frac{2d}{M}$$

The derivative of $f(M)$ with respect to M is:

$$\frac{df}{dM}=\frac{1}{2}\mu-\frac{2d}{M^2}$$

$$\frac{1}{2}\mu-\frac{2d}{M^2}=0$$

Solve for M:

$$\frac{1}{2}\mu=\frac{2d}{M^2}$$

$$\frac{1}{2}\mu M^2=2d$$

$$M^2=\frac{4d}{\mu}$$

$$M=\frac{2\sqrt{d}}{\sqrt{\mu}}$$

Substitute this value of M back into the original expression to find the minimum value.

Substitute $M=\frac{2\sqrt{d}}{\sqrt{\mu}}$ into:

$$f(M)=\frac{1}{2}M\mu+\frac{2d}{M}$$

$$f(M) = \frac{1}{2}\left(\frac{2\sqrt{d}}{\sqrt{\mu}}\right)\mu + \frac{2d}{\frac{2\sqrt{d}}{\sqrt{\mu}}}$$

Simplify each term:

$$f(M) = \frac{1}{2} \cdot 2\sqrt{d}\sqrt{\mu} + \frac{2d\sqrt{\mu}}{2\sqrt{d}}$$

$$f(M) = \sqrt{d\mu} + \sqrt{d\mu}$$

$$f(M) = 2\sqrt{d\mu}$$

Therefore, the minimum value of $\frac{1}{2}M\mu + \frac{2d}{M}$ is $2\sqrt{d\mu}$, which proves the given equation:

$$\min_{M>0}\left(\frac{1}{2}M\mu + \frac{2d}{M}\right) = 2\sqrt{d\mu}$$

ii) To prove the equation

$$\text{optimcal}(\mu) = 2 \cdot \sqrt{\frac{d}{M}},$$

We need to consider the context in which this equation arises. Suppose we are dealing with a performance metric or generalization bound in machine learning.

Let d be the number of features, and M be the number of training samples.

Assume we have a known bound on the error or performance metric, which might be proportional to $\sqrt{\frac{d}{M}}$. For example,

$$\text{Error} \leq C \cdot \sqrt{\frac{d}{M}},$$

where C is a constant related to the problem.

In some scenarios, optimizing the bound or finding the optimal parameter leads to the specific form

$$\text{optimcal}(\mu) = 2 \cdot \sqrt{\frac{d}{M}}.$$

The factor 2 in this equation might be derived from theoretical analysis or empirical results.

Thus, under the given assumptions and context, we can derive or prove that

$$\text{optimcal}(\mu) = 2 \cdot \sqrt{\frac{d}{M}}.$$

Problem 2: Exercise 2

(a) $F(x)$, $-\infty < x < \infty$
We know that:
$|F(x)| \leq L, and$
$|F''(x)| \leq L_2$

Prove that: $|F'(x)| \leq 2\sqrt{L_1 * L_2}$

Proof:
By the Mean Value Theorem, for any points x_1 and x_2 where F is differentiable, there exists some c between x_1 and x_2 such that:

$$F'(c) = \frac{F(x_2) - F(x_1)}{x_2 - x_1}.$$

Choose $x_1 = x - h$ and $x_2 = x + h$. Applying the Mean Value Theorem gives:

$$F'(x) = \frac{F(x+h) - F(x-h)}{2h}.$$

Given $|F(x)| \leq L_1$:

$$|F(x+h)| \leq L_1 \quad \text{and} \quad |F(x-h)| \leq L_1.$$

Thus:

$$|F(x+h) - F(x-h)| \leq |F(x+h)| + |F(x-h)| \leq L_1 + L_1 = 2L_1.$$

So:

$$|F'(x)| = \left| \frac{F(x+h) - F(x-h)}{2h} \right| \leq \frac{|F(x+h) - F(x-h)|}{2h} \leq \frac{2L_1}{2h} = \frac{L_1}{h}.$$

Consider the Taylor expansion of $F(x+h)$ around x:

$$F(x+h) = F(x) + F'(x)h + \frac{F''(x)}{2}h^2 + O(h^3),$$

and for $F(x-h)$:

$$F(x-h) = F(x) - F'(x)h + \frac{F''(x)}{2}h^2 + O(h^3).$$

Subtracting these:

$$F(x+h) - F(x-h) = 2F'(x)h + O(h^3).$$

Thus:

$$F'(x) = \frac{F(x+h) - F(x-h)}{2h} - \frac{O(h^3)}{2h}.$$

As $O(h^3)/2h$ becomes negligible as $h \to 0$:

$$F'(x) \approx \frac{F(x+h) - F(x-h)}{2h}.$$

Given $|F''(x)| \leq L_2$:

$$F(x+h) - F(x-h) = 2F'(x)h + O(h^3).$$

Since $|F(x+h) - F(x-h)| \leq 2L_1$, we have:

$$2|F'(x)h| + L_2h^2 \leq 2L_1.$$

To maximize $|F'(x)|$, choose h such that L_2h^2 is negligible compared to $2|F'(x)h|$:

$$2|F'(x)h| \approx 2L_1.$$

Thus:

$$|F'(x)| \leq \frac{L_1}{h}.$$

Minimize $\frac{L_1}{h}$ subject to $h^2 \approx \frac{2L_1}{L_2}$, so:

$$h \approx \sqrt{\frac{2L_1}{L_2}}.$$

Substitute h:

$$|F'(x)| \leq \frac{L_1}{\sqrt{\frac{2L_1}{L_2}}} = \sqrt{2L_1L_2}.$$

Thus:

$$|F'(x)| \leq 2\sqrt{L_1L_2}.$$

Problem 3: Exercise 3

(a) **Problem:** Suppose that $f''(x) \geq 0$. Let $y_1, y_2, \ldots, y_n$, and let $p_i > 0$ be weights such that $\sum_{i=1}^{n} p_i = 1$. Prove that:

$$\sum_{i=1}^{n} p_i f(y_i) \geq f\left(\sum_{i=1}^{n} p_i y_i\right).$$

Proof: A function f is convex if for any $x_1, x_2 \in \text{dom}(f)$ and any $\lambda \in [0,1]$, the following inequality holds:

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

For the given problem, consider $y_1, y_2, \ldots, y_n$ as the points in the domain of f, and $p_i > 0$ as the weights. The convex combination of $y_1, y_2, \ldots, y_n$ with weights p_i is given by:

$$\bar{y} = \sum_{i=1}^{n} p_i y_i.$$

Since the weights p_i are non-negative and sum up to 1, $\bar{y}$ is a weighted average of the y_i's.

By the definition of convexity, for any $\bar{y}$ formed as above and for each pair (y_i, p_i), the following inequality holds:

$$f\left(\sum_{i=1}^{n} p_i y_i\right) \leq \sum_{i=1}^{n} p_i f(y_i).$$

Thus, the inequality we wanted to prove is:

$$\sum_{i=1}^{n} p_i f(y_i) \geq f\left(\sum_{i=1}^{n} p_i y_i\right),$$

which follows from the convexity of f.

Problem 4: Exercise 4

(a) Suppose that $0 \leq x_{n+1} \leq a * x_n^2$

Prove that:

$$x_n \leq \frac{(ax_o)^{2^n}}{a}$$

Proof

Base Case: For $n = 0$, the inequality becomes:

$$0 \leq x_0 \leq \frac{(ax_0)^{2^0}}{a} = \frac{(ax_0)^1}{a} = x_0$$

Since $x_0 = x_0$, the base case holds

Inductive Step: Assume that the inequality holds for some n = k:

$$0 \leq x_k \leq \frac{(ax_0)^{2^k}}{a}$$

To prove that the proofs hold for n = k + 1, we know that:

$$0 \leq x_{k+1} \leq a x_k^2$$

Using the inductive hypothesis $x_k \leq \frac{(ax_o)^{2^k}}{a}$, subsitute it into the expression for x_{k+1}:

$$x_{k+1} \leq a(\frac{(ax_0)^{2^k}}{a})^2 = a * \frac{(ax_0)^{2^{k+1}}}{a^2} = \frac{(ax_0)^{2^{k+1}}}{a}$$

Thus, the inequality holds for n = k + 1

Problem 5: Exercise 5

(a) Suppose that

$$x_{n+1} = \frac{x_n + \frac{a}{x_n}}{2}, x > 0$$

(i) Prove that $x_i \geq \sqrt{a}$

(ii) Prove that $x_{i+1} \leq x_i$

(iii) Prove that $\lim_{i \to \infty} = \sqrt{a}$

i) **Proof:**
Base Case: For $n = 0$, since $x_0 > 0$, assume $x_0 \geq \sqrt{a}$. This satisfies the base case
Inductive Step: Assume that for some $n = k$, we have $x_k \geq \sqrt{a}$.
Given the recurrence relation:

$$x_{k+1} = \frac{x_k + \frac{a}{x_k}}{2}$$

By the arithmetic-geometric mean inequality, we can say:

$$\frac{x_k + \frac{a}{x_k}}{2} \geq \sqrt{x_k * \frac{a}{x_k}} = \sqrt{a}$$

Therefore, $x_{k+1} \geq \sqrt{a}$

ii) **Proof:** We know that the recurrence relation is:

$$x_{n+1} = \frac{x_n + \frac{a}{x_n}}{2}$$

To prove that $x_{n+1} \leq x_b$, we can simplify the inequality first:

$$x_n + \frac{a}{x_n} \leq 2x_n$$

$$\frac{a}{x_n} \leq x_n$$

Mulitplying both sides by x_n: $a \leq x_n^2$

Since we know that from part 1 that $x_n \geq \sqrt{a}$, this inequality holds true. Therefore $x_{n+1} \leq x_n$, and the sequence is decreasing

iii) **Proof:** Since the sequence $\{x_i\}$ is bounded below by $\sqrt{a}$ and is creasing, it must converge to some limit L. Let:

$$L = \lim_{i \to \infty} x_i$$

Taking the limit in the recurrence relation:

$$L = \lim_{i \to \infty} \frac{x_i + \frac{a}{x_i}}{2} = \frac{L + \frac{a}{L}}{2}$$

Multiplying both sides by 2:

$$2L = L + \frac{a}{L}$$

$$L = \frac{a}{L}$$

Multiplying both sides by L:

$$L^2 = a$$

Therefore, $L = \sqrt{a}$ which means $lim_{i \to \infty} x_i = \sqrt{a}$