CMPUT 340 Written Assignment 1

Problem 1 Yiyang Zuo(Student No: 1861800)

Solution:

a)

$$f(\hat{x}) = f(x+h) = \frac{x+h+2}{x+h+3} = 1 - \frac{1}{x+h+3}$$

$$\Rightarrow \text{Abs error} = f(\hat{x}) - f(x) = 1 - \frac{1}{x+h+3} - \left(1 - \frac{1}{x+3}\right)$$

$$= \frac{1}{x+3} - \frac{1}{x+h+3}$$

$$= \frac{h}{(x+3)(x+h+3)}$$

Choose $x = 3, h = 0.2$

$$\Rightarrow \text{Abs error} = \frac{0.2}{(3+3)\cdot(3+0.2+3)} = \frac{0.2}{6 \cdot 6.2} \approx 0.0054$$

b)

$$\text{relative error} = \frac{\text{Abs error}}{f(x)} = \frac{\frac{h}{(x+3)(x+h+3)}}{\frac{x+2}{x+3}} = \frac{h}{(x+2)(x+h+3)}$$

Choose $x = 3, h = 0.2$

$$\Rightarrow \text{relative error} = \frac{0.2}{(3+2)\cdot(3+0.2+3)} = \frac{0.2}{5 \cdot 6.2} \approx 0.0065$$

c) Let $y = f(x)$

$$f'(x) = \frac{d}{dx}\left(\frac{x+2}{x+3}\right) = \frac{d}{dx}\left(1 - \frac{1}{x+3}\right) = \frac{1}{(x+3)^2}$$

$$\frac{\Delta y}{y} = \frac{f(x+h) - f(x)}{f(x)} \approx \frac{hf'(x)}{f(x)}, \quad \frac{\Delta x}{x} = \frac{h}{x}$$

$$\Rightarrow \text{cond} = \frac{|\Delta y / y|}{|\Delta x / x|} = \frac{\left|\frac{hf'(x)}{f(x)}\right|}{\left|\frac{h}{x}\right|} = \left|\frac{xf'(x)}{f(x)}\right| = \left|\frac{\frac{x}{(x+3)^2}}{\frac{x+2}{x+3}}\right| = \left|\frac{x}{(x+3)(x+2)}\right|$$

d)

Let $g(x) = \dfrac{x}{(x+3)(x+2)}$. In order to make the function well-conditioned, indicating the condition number is near 1. To find the value such that $(\text{cond} - 1)$ is small, we need to find the interval for x whose value of $g(x)$ is greater than -1 and less than 1 and find a value making the condition number is near 1. Of course, (we can find the interval for x whose value of $g(x)$ is greater than 1 or less than -1 and find a value making the condition number is near 1.) In this problem, we choose the first case to illustrate.

Let $\frac{x}{(x+3)(x+2)} \in (-1,1) \Rightarrow x < -\sqrt{3}-3$ or $x > \sqrt{3}-3$

Choose $x = -4.74$, $\text{cond} = \left|\frac{-4.74}{-1.74 \cdot (-2.74)}\right| \approx 0.9942$, the reason for choosing this value is that $-\sqrt{3}-3 \approx -4.7321$.

e)

To find the interval of x whose condition number is highly sensitive, we have to figure out the interval for the function $g(x)$ increasing and decreasing.

$$g'(x) = \frac{d}{dx}\left(\frac{x}{(x+3)(x+2)}\right) = \frac{(x+3)(x+2) - x(2x+5)}{(x+3)^2(x+2)^2} = \frac{6-x^2}{(x+3)^2(x+2)^2}$$

Let $g'(x) = 0 \Rightarrow x = \pm\sqrt{6} \Rightarrow x \in (-\infty, -\sqrt{6}) \cup (\sqrt{6}, +\infty)$, $g(x)$ is decreasing.

Note that $x \neq -3$ and $x \neq -2$. Notice that we cannot determine if $g(x)$ is positive and negative by increasing and decreasing.

Let $g(x) < 0 \Rightarrow x < -3$ or $-2 < x < 0$; let $g(x) > 0 \Rightarrow -3 < x < -2$ or $x > 0$

Thus, when $x < -3$, condition number is increasing(due to absolute value), and when $x \to -3$, condition number is highly sensitive.

When $-\sqrt{6} < x < -2$, condition number is increasing, and when $x \to -2$, conditional number is highly sensitive.

When $-2 < x < 0$, condition number is increasing from negative value, indicating that condition is not highly sensitive.

When $0 < x < \sqrt{6}$, $g(x) > 0$ and condition number is increasing but when $x > \sqrt{6}$, $g(x) > 0$ and condition number is decreasing. There is a possibility that $g(\sqrt{6})$ is large. Verify $g(\sqrt{6}) = \frac{\sqrt{6}}{(\sqrt{6}+3)(\sqrt{6}+2)} = \frac{\sqrt{6}}{12+5\sqrt{6}} < 1$, indicating that the condition number is not highly sensitive.

In conclusion, when $x \to -3$ and $x \to -2$, this problem is highly sensitive.

Problem 2

Solution:

In terms of the formula $\sigma_i = \dfrac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}$, to avoid the overflow of exponential function to an entry z_i, we can find the maximum entry in vector $\mathbf{z}$, denoted by m. That is $m = \max\{z_1, z_2, ..., z_n\}$. Then subtract m to the entry z_i and z_j in the formula, which can prevent the overflow by reducing its power. This procedure outputs the same value as σ_i.

Proof:

Denote the new softmax function with subtraction by σ', where $\sigma'_i = \dfrac{e^{z_i - m}}{\sum_{j=1}^{n} e^{z_j - m}}$.

We can derive that

$$\sigma'_i = \frac{e^{z_i - m}}{\sum_{j=1}^{n} e^{z_j - m}} = \frac{\frac{1}{e^m} e^{z_i}}{\sum_{j=1}^{n} \frac{1}{e^m} e^{z_j}} = \frac{\frac{1}{e^m} e^{z_i}}{\frac{1}{e^m} \sum_{j=1}^{n} e^{z_j}} = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}$$, which equals to the original σ_i.

Problem 3

Solution:

Denote the given matrix by A.

$$A = \begin{pmatrix} 1 & -4 & 5 \\ -5 & 10 & -3 \\ 0 & -8 & 3 \end{pmatrix} \xrightarrow{R_1 \cdot (-5) + R_2} \begin{pmatrix} 1 & -4 & 5 \\ 0 & -10 & 22 \\ 0 & -8 & 3 \end{pmatrix} \xrightarrow{R_2 \cdot \left(-\frac{4}{5}\right) + R_3} \begin{pmatrix} 1 & -4 & 5 \\ 0 & -10 & 22 \\ 0 & 0 & -\frac{73}{5} \end{pmatrix} = U$$

Let $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -10 \\ -8 \end{pmatrix}$, $\mathbf{v}_3 = \left(-\frac{73}{5}\right)$, divide $\mathbf{v}_2$ by $-\frac{1}{10}$ and divide $\mathbf{v}_3$ by 1, we can obtain

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & \frac{4}{5} & 1 \end{pmatrix}$$

Problem 4

Proof:

Firstly, we need to know that $\mathrm{cond}(\mathrm{C}) = \|\mathrm{C}\|_\infty \|\mathrm{C}^{-1}\|_\infty$, where $\|\mathrm{C}\|_\infty = \max_{1 \leqslant i \leqslant m} \sum_{j=1}^{n} |c_{ij}|$

Since $\mathrm{C}\mathbf{g} = \mathbf{h}$ and $\mathrm{C}\hat{\mathbf{g}} = \mathbf{h} + \mathbf{\Delta h}$, we can derive that

$$\mathrm{C}(\hat{\mathbf{g}} - \mathbf{g}) = \mathbf{\Delta h} \Rightarrow \mathrm{C}\mathbf{\Delta g} = \mathbf{\Delta h} \Rightarrow \mathbf{\Delta g} = \mathrm{C}^{-1}\mathbf{\Delta h}$$

Thus, based on what we have derived above, we can derive that

$$\frac{\frac{\|\mathbf{\Delta g}\|}{\|\mathbf{g}\|}}{\frac{\|\mathbf{\Delta h}\|}{\mathbf{h}}} = \frac{\|\mathbf{\Delta g}\|\,\|\mathbf{h}\|}{\|\mathbf{g}\|\,\|\mathbf{\Delta h}\|} = \frac{\|\mathrm{C}^{-1}\mathbf{\Delta h}\|\,\|\mathrm{C}\mathbf{g}\|}{\|\mathbf{g}\|\,\|\mathbf{\Delta h}\|} \leqslant \frac{\|\mathrm{C}^{-1}\|\,\|\mathbf{\Delta h}\|\,\|\mathrm{C}\|\,\|\mathbf{g}\|}{\|\mathbf{g}\|\,\|\mathbf{\Delta h}\|} = \|\mathrm{C}^{-1}\|\,\|\mathrm{C}\| = \|\mathrm{C}^{-1}\|_\infty \|\mathrm{C}\|_\infty = \mathrm{cond}(\mathrm{C})$$

Thus, we can derive that $\frac{\|\mathbf{\Delta g}\|}{\|\mathbf{g}\|} \leqslant \mathrm{cond}(\mathrm{C}) \frac{\|\mathbf{\Delta h}\|}{\|\mathbf{h}\|}$