SHANGHAI JIAO TONG UNIVERSITY, UM-SJTU JOINT INSTITUTE

Introduction to Numerical Analysis
HW8

Yu Cang
018370210001

July 23, 2018

1 QUESTION 1

(a) It's convex as $f''(x) = e^x > 0$.

(b) Let

$$RHS \triangleq tf(x) + (1-t)f(y) \tag{1.1}$$

$$LHS \triangleq -f(tx + (1-t)y) \tag{1.2}$$

It's concave as

$$\begin{aligned} & RHS - LHS \\ =& [tx_1x_2 + (1-t)y_1y_2] - [tx_1 + (1-t)y_1][tx_2 + (1-t)y_2] \\ =& -t(1-t)(x_1y_2 + x_2y_1) < 0 \end{aligned} \tag{1.3}$$

(c) Convex.

(d) Convex.

(e) Convex.

(f) Concave.

2 QUESTION 2

(a) *Proof.* Suppose $x = (x_1, x_2, ..., x_m)$, then it's left to prove that

$$\frac{\partial f}{\partial x_i} = 0 \tag{2.1}$$

for $i = 1, 2, ..., m$ at $x = x^*$. □

The partial derivative of f at $x = x^*$ is defined as

$$\left.\frac{\partial f}{\partial x_i}\right|_{x=x^*} = \lim_{\Delta x_i \to 0} \frac{f(x_1, x_2, ..., x_i^* + \Delta x_i, ..., x_m) - f(x_1, x_2, ..., x_i^*, ..., x_m)}{\Delta x_i} \tag{2.2}$$

Since $f(x^*)$ is the local minimum, $f(x_1, x_2, ..., x_i^* + \Delta x_i, ..., x_m) \geq f(x_1, x_2, ..., x_i^*, ..., x_m)$ however Δx_i changes.
Thus

$$\frac{f(x_1, x_2, ..., x_i^* + \Delta x_i, ..., x_m) - f(x_1, x_2, ..., x_i^*, ..., x_m)}{\Delta x_i} \leq 0 \tag{2.3}$$

if $\Delta x_i < 0$.
And

$$\frac{f(x_1, x_2, ..., x_i^* + \Delta x_i, ..., x_m) - f(x_1, x_2, ..., x_i^*, ..., x_m)}{\Delta x_i} \geq 0 \tag{2.4}$$

if $\Delta x_i > 0$.
Hence, $\left.\frac{\partial f}{\partial x_i}\right|_{x=x^*} = 0$ as f is continuously differentiable. Therefore, $\triangledown f(x^*) = 0$.

(b) *Proof.* With taylor expansion of $f(x)$ at $x = x^*$

$$\begin{aligned} & f(x) - f(x^*) \geq 0 \qquad (f(x^*) \text{ is the local minimum}) \\ \Leftrightarrow & \triangledown f(x^*)\Delta x + \frac{1}{2}\Delta\xi H(x^*)\Delta\xi^T \geq 0 \\ \Leftrightarrow & \frac{1}{2}\Delta\xi H(x^*)\Delta\xi^T \geq 0 \qquad (\text{gradient be } 0) \end{aligned} \tag{2.5}$$

where $d(\Delta\xi, x^*) \leq d(x, x^*)$, and the gradient is 0 as been proved above. Thus, $\Delta\xi H(x^*)\Delta\xi^T \geq 0$ for any $\Delta\xi$ within the local neighbourhood of $x = x^*$.
Hence $H(x^*)$ is semi-positive finite.

□

(c) *Proof.* ($\Leftarrow$)The global minimum must be a local minimum, and the gradient is therefore 0 as f is differentiable.

($\Rightarrow$)As has been proved above, $\triangledown f(x^*) = 0$ indicates that $f(x^*)$ is the local minimum. It's left to prove that local minimum is also global minimum for a convex function. Suppose there exist x' s.t.

$$f(x') < f(x^*) \tag{2.6}$$

Then, by convexity

$$f(tx' + (1-t)x^*) \leq tf(x') + (1-t)f(x^*) < \leq tf(x^*) + (1-t)f(x^*) = f(x^*) \tag{2.7}$$

When $t \to 1$, the inequality above contradicts local optimality of x^*.
Thus, the assumption fails, which indicates that $f(x^*)$ is the global minimum for f.

□

(d) *Proof.* (⇒)By Taylor expansion

$$f(x+d) = f(x) + \nabla f(x)d + \frac{1}{2}d^T H(x)d + O(||d||^2) \tag{2.8}$$

If the Hessian matrix H is semi-definite positive, then

$$d^T H(x)d \geq 0 \tag{2.9}$$

Thus

$$f(x+d) \geq f(x) + \nabla f(x)d \tag{2.10}$$

which is the so called 1-st order condition, and it implies that the function f is convex.

(⇐)By Taylor expansion

$$f(x+td) = f(x) + t\nabla f(x)d + \frac{t^2}{2}d^T H(x)d + O(||d||^2) \tag{2.11}$$

With the 1-st order condition

$$f(x+td) \geq f(x) + t\nabla f(x)d \tag{2.12}$$

Thus

$$\frac{t^2}{2}d^T H(x)d + O(t||d||^2) \geq 0 \tag{2.13}$$

Dividing it by $\frac{t^2}{2}$ and set $t \to 0$, it gives out that for any $d \in \mathbb{R}^n$, $d^T H(x)d \geq 0$.

□

3 Question 4

The genetic algorithm is adopted as there exists intensive oscillation when x approaches 0, and optimization algorithms dependent on the gradient info are easy to fall into local minimum.
The global minimum found is $x_0 = 0.217, f(x_0) = -0.0249$.

4 Question 6

The plot of $g(y)$ is given as follows

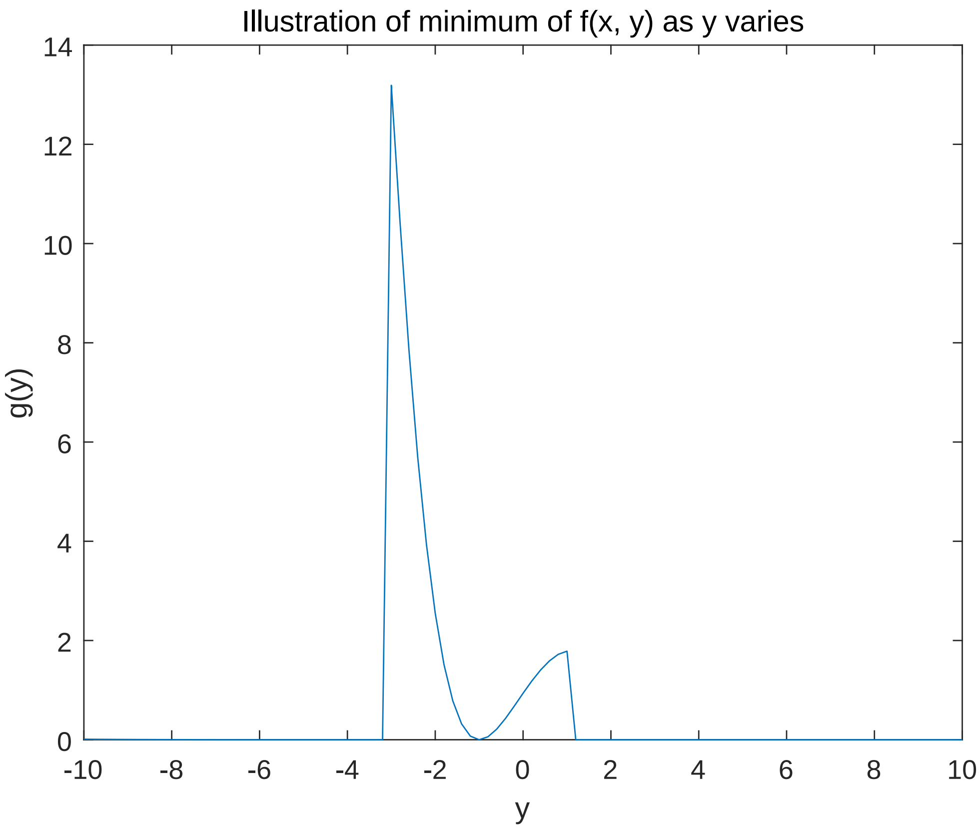
Illustration of minimum of f(x, y) as y varies
g(y)
y
14
12
10
8
6
4
2
0
-10
-8
-6
-4
-2
0
2
4
6
8
10

5 Question 7

According to the geometric relations, the allowable length of the ladder is given as

$$L(\alpha) = \min_{\beta}\left(\frac{1}{sin(\beta)} + \frac{1}{sin(\alpha+\beta)}\right) \tag{5.1}$$

By solving the minimisation problem, the plot of L versus α is given as follows