Assignment 1

1° Characterization of Functions

3. Compute the gradient and Hessian

a) $f(x) = (a^T x - d)^2$ where $a = (-1\ 3)^T,\ d = 2.5$

$$f(x) = (-x_1 + 3x_2 - 2.5)^2$$

$$\frac{\partial f(x)}{\partial x_1} = 2(x_1 - 3x_2 + 2.5)$$

$$\frac{\partial f}{\partial x_2} = 6(-x_1 + 3x_2 - 2.5)$$

$$\nabla f(x) = \begin{bmatrix} 2(x_1 - 3x_2 + 2.5) \\ \\ 6(-x_1 + 3x_2 - 2.5) \end{bmatrix}$$

$$\nabla^2 f(x) = \begin{bmatrix} 2 & -6 \\ -6 & 8 \end{bmatrix}$$

b) $f(x) = (x_1 - 2)^2 + x_1 x_2^2 - 2$

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 + x_2^2 - 4$$

$$\frac{\partial f(x)}{\partial x_2} = 2x_1 x_2$$

$$\nabla f(x) = \begin{bmatrix} 2x_1 + x_2^2 - 4 \\ \\ 2x_1 x_2 \end{bmatrix} \qquad \nabla^2 f(x) = \begin{bmatrix} 2 & 2x_2 \\ \\ 2x_2 & 2x_1 \end{bmatrix}$$

c) $f(x) = x_1^2 + x_1 \|x_1\|^2 + \|x\|^2$

$$f(x) = x_1^2 + x_1(x_1^2 + x_2^2) + (x_1^2 + x_2^2)$$

$$\frac{\partial f(x)}{\partial x_1} = 3x_1^2 + 4x_1 + x_2^2$$

$$\frac{\partial f(x)}{\partial x_2} = 2x_2 + 2x_1 x_2$$

$$\nabla f(x) = \begin{bmatrix} 3x_1^2 + 4x_1 + x_2^2 \\ 2x_2 + 2x_1 x_2 \end{bmatrix}$$

$$\nabla^2 f(x) = \begin{bmatrix} 6x_1 + 4 & 2x_2 \\ 2x_2 & 2x_1 + 2 \end{bmatrix}$$

d) $f(x) = \alpha x_1^2 - 2x_1 + \beta x_2^2$

$$\frac{\partial f(x)}{\partial x_1} = 2\alpha x_1 - 2 \qquad \frac{\partial f(x)}{\partial x_2} = 2\beta x_2$$

$$\nabla f(x) = \begin{bmatrix} 2\alpha x_1 - 2 \\ 2\beta x_2 \end{bmatrix} \qquad \nabla^2 f(x) = \begin{bmatrix} 2\alpha & 0 \\ 0 & 2\beta \end{bmatrix}$$

4. Determin the set of stationary points

a) $\nabla f(x) = 0$

$$\Leftrightarrow \begin{cases} 2(x_1 - 3x_2 + 2.5) = 0 \\ 6(-x_1 + 3x_2 - 2.5) = 0 \end{cases}$$

Stationary points: $\begin{pmatrix} 3c - 2.5 \\ c \end{pmatrix}$ $\quad c \in \mathbb{R}$

b) $\nabla f(x) = 0$

$$\Leftrightarrow \begin{cases} 2x_1 + x_2^2 - 4 = 0 & ① \\ 2x_1 x_2 = 0 & ② \end{cases}$$

From ② we know $x_1 = 0$ or $x_2 = 0$

if $x_1 = 0$, $x_2 = \pm 2$

if $x_2 = 0$, $x_1 = 2$

Stationary points: $\begin{pmatrix} 0 \\ 2 \end{pmatrix} \begin{pmatrix} 0 \\ -2 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix}$

c) $\nabla f(x) = 0$

$$\Leftrightarrow \begin{cases} 3x_1^2 + 4x_1 + x_2^2 = 0 & ① \\ 2x_2 + 2x_1x_2 = 0 & ② \end{cases}$$

② $\Rightarrow \quad 2x_2(1+x_1) = 0$

$\Rightarrow \quad x_2 = 0 \quad or \quad x_1 = -1$

1. $x_2 = 0$

① $\Rightarrow \quad 3x_1^2 + 4x_1 = 0 \quad \Rightarrow \quad x_1(3x_1+4) = 0$

$\Rightarrow \quad x_1 = 0 \quad or \quad x_1 = -\frac{4}{3}$

2. $x_1 = -1$

① $\Rightarrow \quad 3 - 4 + x_2^2 = 0 \quad \Rightarrow \quad x_2 = \pm 1$

Stationary points: $\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} -\frac{4}{3} \\ 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix}$

d) $\nabla f(x) = 0$

$$\Rightarrow \begin{cases} 2\alpha x_1 - 2 = 0 \\ 2\beta x_2 = 0 \end{cases}$$ $\alpha \neq 0$, otherwise there is no stationary point

If $\beta \neq 0$, $x_2 = 0$, $x_1 = \frac{1}{\alpha}$ $(\alpha \neq 0)$

If $\beta = 0$, $x_2 \in R$, $x_1 = \frac{1}{\alpha}$ $(\alpha \neq 0)$

If $\beta \neq 0$, stationary points: $\begin{pmatrix} \frac{1}{\alpha} \\ 0 \end{pmatrix}$

If $\beta = 0$, stationary points: $\begin{pmatrix} \frac{1}{\alpha} \\ c \end{pmatrix}$ $c \in \mathbb{R}$

5. For a) to c) characterize every stationary point whether it is a saddle point, (strict) local/global minimum or maximum

a) $\nabla^2 f(x) = \begin{bmatrix} 2 & -6 \\ -6 & 8 \end{bmatrix}$

By the property of eigen values, we know that:

$\lambda_1 + \lambda_2 = 10$

$\lambda_1 \lambda_2 = -20$

Therefore $\lambda_1 > 0$, $\lambda_2 < 0$, $\nabla^2 f(x)$ is indefinite for any x, then the stationary points are saddle points.

b) $\nabla^2 f(x) = \begin{bmatrix} 2 & 2x_2 \\ 2x_2 & 2x_1 \end{bmatrix}$ stationary points: $\begin{pmatrix} 0 \\ 2 \end{pmatrix} \begin{pmatrix} 0 \\ -2 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix}$

① $x^* = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ $\nabla^2 f(x^*) = \begin{bmatrix} 2 & 4 \\ 4 & 0 \end{bmatrix}$

$\lambda_1 + \lambda_2 = 2$; $\lambda_1 \lambda_2 = -16$

Therefore $\lambda_1 > 0$, $\lambda_2 < 0$, $\nabla^2 f(x^*)$ is indefinite, the stationary point $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ is a saddle point.

② $x^* = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$ $\nabla^2 f(x^*) = \begin{bmatrix} 2 & -4 \\ -4 & 0 \end{bmatrix}$

$\nabla^2 f(x^*) \succ 0$, $\begin{pmatrix} 0 \\ -2 \end{pmatrix}$ is a strict local minimum point.

③ $x^* = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ $\nabla^2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$

$\nabla^2 f(x^*) \succ 0$, $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ is a strict local minimal point.

c) $\nabla^2 f(x) = \begin{bmatrix} 6x_1 + 4 & 2x_2 \\ 2x_2 & 2x_1 + 2 \end{bmatrix}$

① $x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $\nabla^2 f(x^*) = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$

$\nabla^2 f(x^*) \succ 0$, $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a strict local minimal point.

② $x^* = \begin{pmatrix} -\frac{4}{3} \\ 0 \end{pmatrix}$ $\nabla^2 f(x^*) = \begin{bmatrix} -8 & 0 \\ 0 & -\frac{2}{3} \end{bmatrix}$

$\nabla^2 f(x^*) \prec 0$, $\begin{pmatrix} -\frac{4}{3} \\ 0 \end{pmatrix}$ is a strict local maximal point.

③ $x^* = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ $\nabla^2 f(x^*) = \begin{bmatrix} -2 & 2 \\ 2 & 0 \end{bmatrix} \Rightarrow \begin{array}{l} \lambda_1 + \lambda_2 = -2 \\ \lambda_1 \lambda_2 = -4 \end{array}$

$\nabla^2 f(x^*)$ is indefinite, $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ is a saddle point.

④ $x^* = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ $\nabla^2 f(x^*) = \begin{bmatrix} -2 & -2 \\ -2 & 0 \end{bmatrix} \Rightarrow \begin{array}{l} \lambda_1 + \lambda_2 = -2 \\ \lambda_1 \lambda_2 = -4 \end{array}$

$\nabla^2 f(x^*)$ is indefinite, $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$ is a saddle point.

6. For d) denote the intervals for α and β for which maxima/minima and saddle points are attained

$$\nabla^2 f(x) = 2\begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \qquad x^* = \begin{pmatrix} \frac{1}{\alpha} \\ 0 \end{pmatrix} \quad (\alpha \neq 0)$$

α and β are eigen values

① $\alpha \in (0, +\infty)$, $\beta \in [0, +\infty) \Rightarrow \nabla^2 f(x) \succeq 0$

x^* is a <u>global minimum point</u>.

② $\alpha \in (-\infty, 0)$, $\beta \in (-\infty, 0] \Rightarrow \nabla^2 f(x) \preceq 0$

x^* is a <u>global maximum point</u>.

③ $\alpha \in (0, +\infty)$, $\beta \in (-\infty, 0)$

or $\alpha \in (-\infty, 0)$, $\beta \in (0, +\infty)$

x^* is a <u>saddle point</u>.

2° Matrix Calculus

1. Compute the gradient of $f(x)$.

a) $f(x) = \frac{1}{4}\|x-b\|^4$

$$\Rightarrow f(x) = \frac{1}{4}\left[\sum_{i=1}^{n}(x_i-b_i)^2\right]^2$$

$$\frac{\partial f(x)}{\partial x_k} = \frac{\partial}{\partial x_k}\frac{1}{4}\left[\sum_{i=1}^{n}(x_i-b_i)^2\right]^2$$

$$= \frac{1}{4}\cdot 2\left[\sum_{i=1}^{n}(x_i-b_i)^2\right]\cdot 2(x_k-b_k)$$

$$= \left[\sum_{i=1}^{n}(x_i-b_i)^2\right](x_k-b_k)$$

$$\frac{\partial f(x)}{\partial x} = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \left[\sum_{i=1}^{n}(x_i-b_i)^2\right](x_1-b_1) \\ \left[\sum_{i=1}^{n}(x_i-b_i)^2\right](x_2-b_2) \\ \vdots \\ \left[\sum_{i=1}^{n}(x_i-b_i)^2\right](x_n-b_n) \end{bmatrix}$$

$$= \|x-b\|^2(x-b)$$

b) $f(x) = \sum_{i=1}^{n} g((Ax)_i)$

$$= \sum_{i=1}^{n}\left[\frac{1}{2}\left(\sum_{j=1}^{n} a_{ij}x_j\right)^2 + \sum_{j=1}^{n} a_{ij}x_j\right]$$

$$= \frac{1}{2}\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}x_j\right)^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_j$$

$$\frac{\partial f(x)}{\partial x_k} = \frac{1}{2}\frac{\partial}{\partial x_k}\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij} x_j\right)^2 + \frac{\partial}{\partial x_k}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_j$$

$$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} 2(a_{ij}x_j)a_{ik} + \sum_{i=1}^{n} a_{ik}$$

$$= \sum_{i=1}^{n} a_{ik}\left(\sum_{j=1}^{n} a_{ij}x_j\right) + \sum_{i=1}^{n} a_{ik}$$

Therefore, $\frac{\partial f(x)}{\partial x} = A^T A x + A^T u$

where $u = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1}$ $u \in \mathbb{R}^n$

c) $f(x) = (x \oslash b)^T D (x \oslash b)$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{1}{b_i} x_i d_{ij} x_j \frac{1}{b_j}$$

$$\frac{\partial f(x)}{\partial x_k} = \frac{\partial}{\partial x_k}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{1}{b_i} x_i d_{ij} x_j \frac{1}{b_j}$$

$$= \sum_{j=1}^{n} \frac{1}{b_k} d_{kj} x_j \frac{1}{b_j} + \sum_{i=1}^{n} \frac{1}{b_i} x_i d_{ik} \frac{1}{b_k}$$

$$= \left(\sum_{j=1}^{n} d_{k,j}\frac{x_j}{b_j} + \sum_{i=1}^{n} d_{ik}\frac{x_i}{b_i}\right)\frac{1}{b_k}$$

Therefore, $\frac{\partial f(x)}{\partial x} = \left[D(x \oslash b) + D^T(x \oslash b)\right] \oslash b$

2. Compute the Hessian of $f(x)$

a) $$\frac{\partial f(x)}{\partial x_k} = \left[\sum_{i=1}^{n} (x_i - b_i)^2\right](x_k - b_k)$$

$$\frac{\partial^2 f(x)}{\partial x_k^2} = \sum_{i=1}^{n} (x_i - b_i)^2 + 2(x_k - b_k)^2$$

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_l} = 2(x_k - b_k)(x_l - b_l)$$

$$\nabla^2 f(x) = \begin{bmatrix} \sum_{i=1}^{n}(x_i - b_i)^2 + 2(x_1 - b_1)^2 & \cdots & 2(x_1 - b_1)(x_n - b_n) \\ & \ddots & \\ & & \sum_{i=1}^{n}(x_i - b_i)^2 + 2(x_n - b_n)^2 \end{bmatrix}$$

b) $$\frac{\partial f(x)}{\partial x_k} = \sum_{i=1}^{n} a_{ik}\left(\sum_{j=1}^{n} a_{ij} x_j\right) + \sum_{i=1}^{n} a_{ik}$$

$$\frac{\partial^2 f(x)}{\partial x_k^2} = \sum_{i=1}^{n} a_{ik} \frac{\partial}{\partial x_k}\left(\sum_{j=1}^{n} a_{ij} x_j\right) + 0$$

$$= \sum_{i=1}^{n} a_{ik}^2$$

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_l} = \sum_{i=1}^{n} a_{ik} \frac{\partial}{\partial x_l}\left(\sum_{j=1}^{n} a_{ij} x_j\right) + 0$$

$$= \sum_{i=1}^{n} a_{ik} a_{il}$$

$$\nabla^2 f(x) = \begin{bmatrix} \sum_{i=1}^{n} a_{i1}^2 & \sum_{i=1}^{n} a_{i1} a_{i2} & \cdots & \sum_{i=1}^{n} a_{i1} a_{in} \\ & \sum_{i=1}^{n} a_{i2}^2 & & \\ \vdots & & \ddots & \vdots \\ & & & \sum_{i=1}^{n} a_{in}^2 \end{bmatrix}$$

c) $$\frac{\partial f(x)}{\partial x_k} = \sum_{j=1}^{n} \frac{1}{b_k} d_{kj} x_j \frac{1}{b_j} + \sum_{i=1}^{n} \frac{1}{b_i} x_i d_{ik} \frac{1}{b_k}$$

$$\frac{\partial^2 f(x)}{\partial x_k^2} = \frac{1}{b_k} d_{kk} \frac{1}{b_k} + \frac{1}{b_k} d_{kk} \frac{1}{b_k}$$

$$= \frac{2}{b_k^2} d_{kk}$$

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_l} = \frac{1}{b_k} d_{kl} \frac{1}{b_l} + \frac{1}{b_l} d_{lk} \frac{1}{b_k}$$

$$= \frac{1}{b_k b_l} (d_{kl} + d_{lk})$$

$$\nabla^2 f(x) = \begin{bmatrix} \frac{2}{b_1^2} d_{11} & \frac{2}{b_1 b_2} (d_{12} + d_{21}) & \cdots & \frac{2}{b_1 b_n} (d_{1n} + d_{n1}) \\ & \frac{2}{b_2^2} d_{22} & & \vdots \\ & & \ddots & \\ & & & \frac{2}{b_n^2} d_{nn} \end{bmatrix}$$

4 Scheduling Otimization Problem

a) The objective function is formulated as

$$\min_{x} C^T x$$

where $x = [x_0, x_1, \cdots x_{15}]^T$

$x_0, x_1, \cdots x_7$ are the number of instructions for processor $0, 1, \cdots, 7$ on CPUs.

$x_8, x_9, \cdots x_{15}$ are the number of instructions for processor $0, 1, \cdots, 7$ on GPUs.

$$C = [0.11\ 0.13\ 0.09\ 0.12\ 0.15\ 0.14\ 0.11\ 0.12\ 0.10\ 0.13\ 0.08\ 0.13\ 0.14\ 0.14\ 0.09\ 0.13]^T$$

b) Constraints are formulated as follows

① Equality constraints

$$A_{eq} = \left.\begin{bmatrix} \overbrace{1\ 0\ 0\ \cdots\ 1}^{8} & \overbrace{1\ 0\ 0\ \cdots\ 0}^{8} \\ 0\ 1\ \ \cdots\ 0 & 0\ 1\ 0\ \cdots\ 0 \\ 0\ 0\ 1\ \cdots\ 0 & 0\ 0\ 1\ \cdots\ 0 \\ \vdots\ \ \ddots\ \vdots & \vdots\ \ \ \ \ddots\ \vdots \\ 0\ 0\ 0\ \cdots\ 1 & 0\ 0\ 0\ \cdots\ 1 \end{bmatrix}\right\} 8 \iff A_{eq} = \begin{bmatrix} I_8 & I_8 \end{bmatrix}$$

$$b_{eq} = [1200\ 1500\ 1400\ 400\ 1000\ 800\ 760\ 1300]^T$$

$$A_{eq} x = b_{eq}$$

② Inequality constraints

$$A_{ub} = \begin{bmatrix} \overbrace{1\ 1\ \cdots\ 1}^{8} & \overbrace{0\ 0\ \cdots\ 0}^{8} \\ 0\ 0\ \cdots\ 0 & 1\ 1\ \cdots\ 1 \end{bmatrix} \quad b_{ub} = \begin{bmatrix} 4500 \\ 4500 \end{bmatrix}$$

$$A_{ub} x \leq b_{ub}$$

③ constraints for x

$$x \geq l$$

$$l = [480\ 600\ 560\ \overbrace{0\ 0 \cdots 0\ 0}^{13}]^T$$

c) $\Rightarrow$ see the code

d) /

$$M = \begin{bmatrix} 480 & 720 \\ 960 & 540 \\ 400 & 0 \\ 0 & 1000 \\ 800 & 0 \\ 0 & 760 \\ 1300 & 0 \end{bmatrix}$$

e) The solution fullfils the constraint

f) The total energy consumption

$$c^T x = 961.8\ [\mu Wh]$$

g)

$$x_0 + x_1 + x_2 + x_7 = 1200 + 1500 + 1400 + 1300$$
$$= 5400 > 4500$$

The constraint is not satisfied,

The linear programming problem is not solvable.

This is varified by the code as well.