MAE 598 Design optimization

HW 1

1

Q2) let x and $b \in \mathbb{R}^n$ be vector and $A \in \mathbb{R}^{n \times n}$ be a square matrix

Define $f: \mathbb{R}^n \to \mathbb{R}$ as $f(x) = b^T x + x^T A x$

a) what is the gradient and Hession of $f(x)$ w.r.t x?

Sol) gradient $\boxed{\nabla f(x) = \frac{\partial f(x)}{\partial (x)}}$

$$\nabla f(x) = \frac{\partial}{\partial x}(b^T x + x^T A x)$$

$$\Rightarrow \frac{\partial b^T x}{\partial x} \Rightarrow \begin{bmatrix} \frac{\partial b^T x}{x_1} \\ \frac{\partial b^T x}{x_2} \\ \vdots \\ \frac{\partial b^T x}{x_n} \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{\partial}{\partial x_1}(b_1 x_1 + b_2 x_2 \cdots + b_n x_n) \\ \vdots \\ \frac{\partial}{\partial x_n}(b_1 x_1 + b_2 x_2 \cdots + b_n x_n) \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = b$$

$$\Rightarrow \quad x^T A x = \begin{bmatrix} x_1 \cdots x_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \begin{bmatrix} (a_{11}x_1 + \cdots a_{n1}x_n) & \cdots & (a_{1n}x_1 + \cdots a_{nn}x_n) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \left[\sum_{i=1}^{n} a_{i1} x_i \cdots \sum_{i=1}^{n} a_{in} x_i \right] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= x_1 \sum_{i=1}^{n} a_{i1} x_i \cdots x_n \sum_{i=1}^{n} a_{in} x_i$$

$$\boxed{\Rightarrow \sum_{j=1}^{n} \sum_{i=1}^{n} x_i a_{ij} x_j}$$

To computing the partial derivative we can simplified the eq as

$$\sum_{i=1}^{n} (a_{ii} x_i^2 + \sum_{j \neq i} x_i a_{ij} x_j)$$

$\Rightarrow$ for simplicity let consider the k^{th} row and do partial derivative

$$\frac{\partial x^T A x}{\partial x_k} = \frac{\partial}{\partial x_k} \sum_{i=1}^{n} (a_{ii} x_i^2 + \sum_{j \neq 1} x_i a_{ij} x_j)$$

$$\Rightarrow 2 a_{kk} x_k + \sum_{j \neq k} x_j a_{jk} + \sum_{j \neq k} a_{kj} x_j$$

we can re-write as

$$\boxed{\Rightarrow \sum_{j=1}^{n} x_j a_{jk} + \sum_{j=1}^{n} a_{kj} x_j}$$

convert Back into matrix from

$$\underbrace{\begin{bmatrix} \sum_{j=1}^{n} x_j a_{j1} \\ \vdots \\ \sum_{j=1}^{n} x_j a_{jn} \end{bmatrix}}_{A^T X} + \underbrace{\begin{bmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj} x_j \end{bmatrix}}_{AX} \Rightarrow (A^T + A) X$$

so $\boxed{\nabla f(x) = b + (A^T + A)x}$ = Gradient

⇒ If A is symmetric $(A^T = A)$ then $(A^T + A) = (A + A) = 2A$

So $\boxed{\nabla f(x) = b + 2Ax}$ = Gradient when A is symmetric

⇒ Hession matrix

$$\nabla^2 f(x) = \frac{\partial^2 f(x)}{\partial x \partial x^T}$$

Now we know the k^{th} element gradient

$$\boxed{\nabla f(x) = b_k + \sum_{j=1}^{n} x_j a_{jk} + \sum_{j=1}^{n} a_{kj} x_j}$$

So second partial deravitive for k^m row

$$\nabla^2 f(x) = 0 + a_{k'k} + a_{kk'}$$

which mean that

$$\nabla^2 f(x) = \begin{bmatrix} a_{11} + a_{11} & a_{21} + a_{12} & \cdots & a_{n1} + a_{1n} \\ \vdots & \ddots & & \\ a_{1n} + a_{n1} & & \cdots & a_{nn} + a_{nn} \end{bmatrix}$$

$\boxed{\nabla^2 f(x) = A + A^T}$ = Hessian

If A is symmetric

then $\boxed{\nabla^2 f(x) = 2A}$ = Hessian

b) Derive the first and second order Taylor's approximation of f(x) at x=0, Are these approximation exact

$\Rightarrow$ 1^{st} order

$$f(x) = f(x_0) + \frac{\partial f}{\partial x}\Big|_{x_0} (x - x_0)$$

$$\Rightarrow f(x) \approx f(x_0) + \nabla f\Big|_{x_0}^{T} (x - x_0)$$

$$\Rightarrow \cancel{0} \; b^T x_0 + x_0^T A x_0 + (b + (A^T + A)x_0)^T (x - x_0)$$

$$= 0 + b^T x$$

$$\boxed{= b^T x}$$

$\leftarrow$ these approximation is not exact because for quadratic linear 2nd order is exact

2^{nd} order

$$f(x) \approx f(x_0) + \nabla f\Big|_{x_0}^{T} (x - x_0) + \frac{1}{2}(x - x_0)^T H\Big|_{x_0} (x - x_0)$$

$$\Rightarrow 0 + b^T x + \frac{1}{2} x^T (A + A^T) x$$

$$\boxed{\Rightarrow b^T x + \frac{1}{2} x^T (A + A^T) x}$$

these approximation is ~~are~~ exact Because the function is quadratic

~~If~~ A is symmetric

$$\boxed{= b^T x + \frac{1}{2} x^T (2A) x}$$

c) what are the necessary and sufficient condition for A to be positive definite?

→ Since $A \in \mathbb{R}^{n\times n}$, so the necessary and sufficent condition for A to be positive definite is

→ All the eigen value of A must be ~~greater than~~ Positive $\boxed{\lambda > 0}$

→ $\boxed{x^T A x > 0}$ for all x [other than - zero vector]

→ All upper left determinant must be > 0.

d) what are the necessary and sufficent condition for A to have full rank

→ The $|A| \neq 0$. {Determinant of $A \neq 0$.}

↳ All rows are linearly indipendent

e) If there exist $y \in \mathbb{R}^n$ and $y \neq 0$ such that $A^T y = 0$ then what are the condition on b for $Ax = b$ to have solution for x

$$A^T y = 0 \qquad \therefore y \in N(A^T)$$

$$Ax = b$$

$$\cancel{\Rightarrow} \quad \cancel{(A^T y)^T = 0^T} \quad \cancel{\Rightarrow y^T A = 0^T}$$

mean $b \in C(A)$

↑

$\boxed{\therefore y \text{ is perpendicular to } b}$

$$x^T(A^T y) = 0 \Rightarrow (Ax)^T y = 0 \Rightarrow \boxed{b^T y = 0}$$

.e to the Recent . . . . . . Formulate an optimization problem to determine the minimum grocery cost to satisfy the nutrition need.

example matrix

$$\begin{array}{c|cccccc}
\text{unit Price} & \text{Nutrition} \rightarrow & y_1 & y_2 & \cdots & y_j \cdots & y_M \\
 & \text{Food} \downarrow & & & & & \\
c_1 & f_1 & a_{11} & a_{12} & \cdots & & a_{1n} \\
c_2 & f_2 & a_{21} & a_{22} & \cdots & & a_{2n} \\
\vdots & \vdots & \vdots & \ddots & & & \vdots \\
c_j & f_j & & & a_{ij} & \ddots & \\
\vdots & \vdots & & & & & \\
c_N & f_N & a_{n1} & a_{n2} & \cdots & & a_{nn}
\end{array}$$

$\Rightarrow$ objective function $\Rightarrow$ minimize the cost

$$= \boxed{\min \sum_{i=1}^{N} f_i c_i}$$

Constraints : $\boxed{\sum_{i=1}^{N} f_i a_{ij} \geq b_j} \Rightarrow \boxed{\sum_{i=1}^{N} f_i a_{ij} - b_j \geq 0}$

for all $j = 1, 2 \ldots M$

$\boxed{f_i \geq 0}$ for all $i = 1 \ldots N$