Problem Set 0 #
CS 229.

① ⓐ Let $f(x) = \frac{1}{2}x^T A x + b^T x$ where A is a symmetric matrix and $b \in R^n$ is a vector. What is $\nabla f(x)$?

Sol^n:- matrix $A \in R^{n \times n}$ is symmetric. $A^T = A$.

$A - A^T = 0$ $A + A^T = 2A$

Let $A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}$ given $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$

$\nabla f(x) = ?$ where $f : R^n \to R$

given, $\nabla f(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) \\ \frac{\partial}{\partial x_2} f(x) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x) \end{bmatrix}$

given $f(x) = \frac{1}{2}x^T A x + b^T x$, where $b \in R^n$ is a vector. So,

$= \frac{1}{2} \cdot \underbrace{[x_1, x_2, x_3 \cdots x_n]}_{1 \times n} \underbrace{\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}}_{n \times n} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_{n \times 1}$

$+ \underbrace{[b_{11}, b_{12}, b_{13} \cdots b_{1n}]}_{1 \times n} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_{n \times 1}$

→ first multiply (Ax) after that $\frac{1}{2}x^T(Ax)$

$$\Rightarrow f(x) = \frac{1}{2}\Big[A_{11}x_1^2 + A_{12}x_1x_2 + A_{13}x_1x_3 + \dots A_{1n}x_n^2 +$$
$$A_{21}x_2x_1 + A_{22}x_2^2 + A_{23}x_2x_3 + \dots A_{2n}x_n^2 +$$
$$A_{31}x_3x_1 + A_{32}x_3x_2 + A_{33}x_3^2 + \dots A_{3n}x_n^2$$
$$\dots$$
$$A_{n1}x_nx_1 + A_{n2}x_nx_2 + \dots + A_{nn}x_n^2\Big]$$
$$+ \Big[b_{11}x_1 + b_{12}x_2 + b_{13}x_3 + \dots + b_{1n}x_n \Big]$$

c) $$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

then $f(x) = f_1(x) + f_2(x)$.

$f_2(x) = b_{11}x_1 + b_{12}x_2 + b_{13}x_3 + \dots b_{1n}x_3$

$f_1(x) = \frac{1}{2}\Big[A_{11}x_1^2 + A_{12}x_1x_2 + \dots$

$\dots + A_{nn}x_n^2 \Big)$

$A^T = A$. $A_{ij}x_ix_j =$

$A_{ij} = A_{ji}$ for all i, j

$$\frac{\partial f_1(x)}{\partial x_1} = 2A_{11}x_1 + x_2(A_{12} + A_{21}) + x_3(A_{13} + A_{31}) + \dots$$

$$\frac{\partial f_1(x)}{\partial x_2} = x_1[A_{12} + A_{21}] + 2x_2A_{22} + x_3(A_{23} + A_{32}) \dots$$

$$\vdots$$

$$\frac{\partial f_n(x)}{\partial x_n} = x_1[A_{1n} + A_{n1}) + \dots + 2x_nA_{nn}$$

$$\frac{\partial f_2(x)}{\partial x} = \nabla f_2(x) = \begin{bmatrix} \frac{\partial f_2(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f_2(x)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{12} \\ \vdots \\ \vdots \\ b_{1n} \end{bmatrix}$$

other term = 0

When we use $\nabla f_2(x)$ only those term constant as $\frac{\partial f_2(x)}{\partial x_1}$ only b_{11} term found and other term = 0. Similarly b_{12} ... & other we get.

So,

$a_{ij} = a_{ji}$	$a_{13} = a_{31}$. so on.
$a_{12} = a_{21}$	$a_{1n} = a_{n1}$.

$$\nabla f(x) = \nabla f_1(x) + \nabla f_2(x)$$

$$\Rightarrow \frac{1}{2} \begin{bmatrix} 2a_{11} & a_{12}+a_{21} & \cdots & a_{1n}+a_{n1} \\ a_{12}+a_{21} & 2a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ & & & 2a_{nn} \\ a_{1n}+a_{n1} & \cdots & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{12} \\ \vdots \\ b_{1n} \end{bmatrix}$$

~~2 common ... of $\nabla f_1(x)$ factor we get~~

$$\Rightarrow \frac{1}{2} \begin{bmatrix} 2a_{11} & 2a_{12} & 2a_{13} & \cdots & 2a_{1n} \\ 2a_{21} & 2a_{22} & & & \vdots \\ 2a_{31} & & \ddots & & \\ \vdots & & & & 2a_{nn} \\ 2a_{n1} & \cdots & & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{12} \\ \vdots \\ b_{1n} \end{bmatrix}$$

$$\begin{bmatrix} \vdots & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{pmatrix} \vdots \\ x_n \end{pmatrix} \begin{pmatrix} \vdots \\ b_n \end{pmatrix}$$

$$= Ax + b$$

(c) An $f(x) = \frac{1}{2}x^T Ax + b^T x$, where A is a symmetric and $b \in R^n$ is a vector. What is $\nabla^2 f(x)$?

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} f(x) & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2}{\partial x_1 \partial x_n} f(x) \\ \frac{\partial^2}{\partial x_2 \partial x_1} f(x) & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2}{\partial x_2 \partial x_n} f(x) \\ \vdots & & & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_1} f(x) & \frac{\partial^2}{\partial x_n \partial x_2} f(x) & \cdots & \frac{\partial^2}{\partial x_n^2} f(x) \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$\nabla^2 f(x) = A. \text{ ans}$$