Problem Set #0: Linear Algebra, Multivariable Calculus, and Probability

Sumanth Palla

March 31, 2024

1 Problem 1

Gradients and Hessians

a. Let $f(x) = \frac{1}{2}x^T Ax + b^T x$, where A is symmetric matrix and $b \in \mathbb{R}^n$ is a vector. Now, $\nabla f(x)$ denotes the Gradient of the function f which maps a n-dimensional point to a Real line.

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

Let, A be a symmetric matrix defined by

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \text{ and } A = A^T$$

Therefore,

$$= \nabla f(x) = \nabla(\frac{1}{2}x^T Ax + b^T x) = \nabla(\frac{1}{2}x^T Ax) + \nabla(b^T x)$$

$$\nabla(\frac{1}{2}x^T Ax) = \frac{1}{2}\nabla x^T Ax = \frac{1}{2}\nabla(x^T) \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + a_{12}x_2 + \ldots + a_n n x_n \end{bmatrix}$$

$$= \frac{1}{2}\nabla \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + a_{12}x_2 + \ldots + a_n n x_n \end{bmatrix}$$

$$= \frac{1}{2}\nabla x_1(a_{11}x_1 + \ldots + a_{1n}x_n) + x_2(a_{21}x_1 + \ldots + a_{2n}) + \ldots + x_n(a_{n1}x_1 + \ldots a_{nn}x_n)$$

$$= \frac{1}{2} \begin{bmatrix} \frac{\partial}{\partial x_1}(x_1(a_{11}x_1 + \ldots + a_{1n}x_n) + x_2(a_{21}x_1 + \ldots + a_{2n}) + \ldots + x_n(a_{n1}x_1 + \ldots a_{nn}x_n)) \\ \vdots \\ \frac{\partial}{\partial x_n}(x_1(a_{11}x_1 + \ldots + a_{1n}x_n) + x_2(a_{21}x_1 + \ldots + a_{2n}) + \ldots + x_n(a_{n1}x_1 + \ldots a_{nn}x_n)) \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 2(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n) \\ \vdots \\ 2(a_{n1}x_1 + a_{12}x_2 + \ldots + a_n n x_n) \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + a_{12}x_2 + \ldots + a_n n x_n \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{j=1}^{j=n} a_{1j}x_1 \\ \sum_{j=1}^{j=n} a_{2j}x_2 \\ \vdots \\ \sum_{j=1}^{j=n} a_{nj}x_n \end{bmatrix}$$

$$= \mathbf{Ax}$$

$$\therefore \frac{1}{2}\nabla x^T A x = Ax$$

Now,

$$\nabla b^T x = \begin{bmatrix} b_1 & b_2 & b_3 & \ldots & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_nx_n$$

$$= \begin{bmatrix} \frac{\partial}{\partial x_1}(b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_nx_n) \\ \vdots \\ \frac{\partial}{\partial x_n}(b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_nx_n) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix} = b$$

$$\therefore \nabla b^T x = b$$

$$\therefore \nabla_x f(x) = Ax + b$$

b. Given, $f(x) = g(h(x))$ where $g : \mathbb{R} \to \mathbb{R}$ is differentiable and $h : \mathbb{R}^n \to \mathbb{R}$ is differentiable

$$\text{by chain rule, } \nabla_x f(x) = g'(h(x))\nabla_x h(x)$$

c. Given, $f(x) = \frac{1}{2}x^T Ax + b^T x$, A is symmetric and $b \in \mathbb{R}^n$.

$$f(x) = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix} \begin{bmatrix} x_1a_{11} + x_2a_{12} + \ldots + x_na_{1n} \\ \vdots \\ x_na_{n1} + x_na_{n2} + \ldots + x_na_{nn} \end{bmatrix} + (b_1x_1 + b_2x_2 + \ldots + b_nx_n)$$

$$f(x) = x_1(x_1a_{11} + x_2a_{12} + \ldots + x_na_{1n}) + x_2(x_1a_{21} + x_2a_{22} + \ldots + x_na_{2n}) + \ldots + x_n(x_1a_{n1} + x_2a_{n2} + \ldots + x_na_{nn})$$
$$+(b_1x_1 + b_2x_2 + \ldots + b_nx_n)$$

$$f(x) = a_{11}x_1^2 + a_{22}x_2^2 + \ldots + a_{nn}x_n^2 + a_{12}x_1x_2 + a_{21}x_2x_1 + \ldots + b_1x_1 + b_2x_2 + \ldots + b_nx_n$$

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

$$\therefore \nabla_x^2 f(x) = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} = A$$

d. Given $f(x) = g(a^T x)$ where $g : \mathbb{R} \to \mathbb{R}$ is continuously differentiable and $a \in \mathbb{R}^n$.

$$\nabla_x f(x) = g'(a^T x) \nabla_x (a^T x) = g'(a^T x) a \; [\because \nabla a^T x = a]$$
$$\nabla_x^2 f(x) = \nabla_x (\nabla_x f(x)) = \nabla (g'(a^T x) a) = g''(a^T x) a a^T$$

2 Problem 2

Positive Definite matrices

A matrix $A \in \mathbb{R}^{n \times n}$ is *positive semi-definite* PSD, denoted $A \succ 0$, if $A = A^T$ and $x^T A x \geq 0 \; \forall x \in \mathbb{R}^n$
A matrix A is *positive definite* denoted $A \succeq 0$ if $A = A^T$ (and) $x^T A x > 0 \; \forall x \neq 0$
a. Given $z \in \mathbb{R}^n$ is a

$$\begin{aligned} A &= z z^T \\ x^T A x &= x^T z z^T x \\ &= (z^T x)^2 \geq 0 \end{aligned}$$

Hence, we can say that given matrix A is *positive semi-definite* PSD, denoted $A \succeq 0$ for given $z \in \mathbb{R}^n$.

b. Given, $A = z z^T$, Null space is defined as vector space of x for a given vector $x \in \mathbb{R}^n$, where $Ax = 0$, i.e, $z z^T x = 0$ From previous problem, for a given $z \in \mathbb{R}^n$, but A is PSD, therefore, $A \succeq 0$ and $Ax = 0$ hence the only solution is $x = 0$
c. $A \in \mathbb{R}^{n \times n}$ is PSD. Let, $B \in \mathbb{R}^{m \times n}$ is arbitary matrix. For an arbitary vector, p $\in \mathbb{R}^m$ and q $\in \mathbb{R}^n$, $B = p q^T$

$$\begin{aligned} B &= p q^T \\ B A B^T &= p q^T A q p^T \\ B A B^T &= \end{aligned}$$

Prolem 3

Eigen Vectors, Eigenvalues, and the spectral theorem

a. A matrix is $A \in \mathbb{R}$