Problem set 0 #.
CS 229.

①

(1)

(b)

Let $f(x) = g(h(x))$, where $g: R \to R$ is differentiable and $h: R^n \to R$ is differentiable. What is $\nabla f(x)$?

Soln:-

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

Here $f(x) = (g \circ h)(x)$. $g: R \to R$ and differentiable. and $h: R^n \to R$ is differentiable.

$\nabla f(x) = \nabla g \circ h(x) = ?$

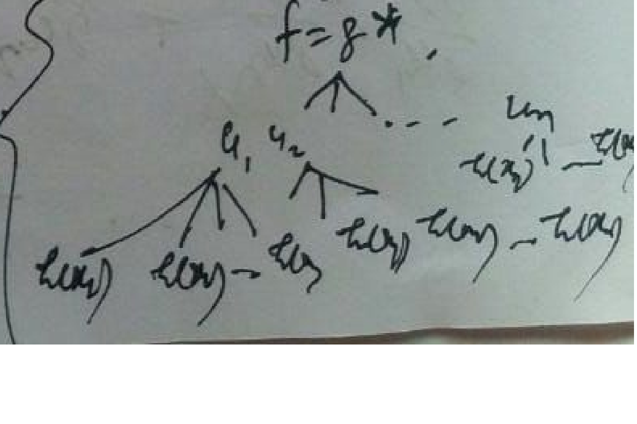
So, $f(x)$ is a composite f^n of $g \circ h(x)$.
g is the function of h
h is " " x.

Let. $f(x) = g(u)$, where $u = h(x)$.

$f = g(u)$ | u has n variable $\{u_1, u_2, u_3, \ldots u_n\}$
$u = h(x)$. where $h: R^n \to R$ $\{x_1, x_2, \ldots x_n\}$

$u = \{h(x_1), h(x_2) \ldots h(x_n)\}$

$f = g(u) = g(h(x)) = g \circ h(x)$

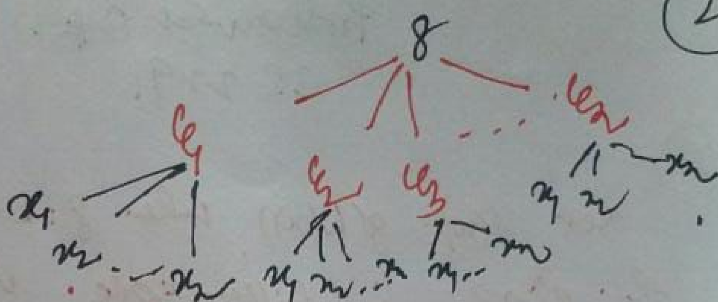


$\nabla f(x) = \nabla_x g \circ f = ?$

8r

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}^T$$

from our calc



Due to definition of gradient.

$$\nabla \begin{bmatrix} \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} & \cdots & \frac{\partial g}{\partial x_n} \end{bmatrix}^T$$

$$= \Big[\frac{\partial g}{\partial u_1} \cdot \frac{\partial u_1}{\partial x_1} + \frac{\partial g}{\partial u_2} \cdot \frac{\partial u_2}{\partial x_1} + \frac{\partial g}{\partial u_3} \cdot \frac{\partial u_3}{\partial x_1} + \cdots + \frac{\partial g}{\partial u_n} \cdot \frac{\partial u_n}{\partial x_1},$$

$$\frac{\partial g}{\partial u_1} \cdot \frac{\partial u_1}{\partial x_2} + \frac{\partial g}{\partial u_2} \cdot \frac{\partial u_2}{\partial x_2} + \frac{\partial g}{\partial u_3} \cdot \frac{\partial u_3}{\partial x_2} + \cdots + \frac{\partial g}{\partial u_n} \cdot \frac{\partial u_n}{\partial x_2},$$

$$\frac{\partial g}{\partial u_1} \cdot \frac{\partial u_1}{\partial x_3} + \frac{\partial g}{\partial u_2} \cdot \frac{\partial u_2}{\partial x_3} + \frac{\partial g}{\partial u_3} \cdot \frac{\partial u_3}{\partial x_3} + \cdots + \frac{\partial g}{\partial u_n} \cdot \frac{\partial u_n}{\partial x_3},$$

$$\cdots$$

$$\frac{\partial g}{\partial u_1} \cdot \frac{\partial u_1}{\partial x_n} + \frac{\partial g}{\partial u_2} \cdot \frac{\partial u_2}{\partial x_n} + \frac{\partial g}{\partial u_3} \cdot \frac{\partial u_3}{\partial x_n} + \cdots + \frac{\partial g}{\partial u_n} \cdot \frac{\partial u_n}{\partial x_n} \Big]^T$$

then $u_1 = h_1(x_1)$, $u_2 = h_2(x_2)$, $u_3 = h(x_3)$ ... $u_n = h(x_n)$

So, we can write it

$$= \Big[\frac{\partial g}{\partial u_1} \cdot \frac{\partial h(x_1)}{\partial x_1} + \frac{\partial g}{\partial u_2} \cdot \frac{\partial h(x_2)}{\partial x_1} + \dots + \frac{\partial g}{\partial u_n} \cdot \frac{\partial h(x_n)}{\partial x_1},$$

$$\frac{\partial g}{\partial u_1} \cdot \frac{\partial h(x_1)}{\partial x_2} + \frac{\partial g}{\partial u_2} \cdot \frac{\partial h_2(x_2)}{\partial x_2} + \dots + \frac{\partial g}{\partial u_n} \cdot \frac{\partial h(x_n)}{\partial x_2},$$

$$\text{— — — — —}$$

$$\frac{\partial g}{\partial u_1} \cdot \frac{\partial h_1(x_1)}{\partial x_n} + \frac{\partial g}{\partial u_2} \cdot \frac{\partial h(x_2)}{\partial x_n} + \dots + \frac{\partial g}{\partial u_n} \cdot \frac{\partial h(x_n)}{\partial u_n} \Big].$$

Nicely we see that

$$= \left[\frac{\partial g}{\partial u_1}, \frac{\partial g}{\partial u_2} \dots \frac{\partial g}{\partial u_n}\right] \begin{bmatrix} \frac{\partial h_1(x_1)}{\partial x_1} & \frac{\partial h_1(x_1)}{\partial x_n} & \dots & \frac{\partial h(x_1)}{\partial x_n} \\ \frac{\partial h_2(x_2)}{\partial x_2} & \frac{\partial h(x_2)}{\partial x_n} & \dots & \frac{\partial h(x_2)}{\partial x_n} \\ \vdots & & & \\ \frac{\partial h_n(x_n)}{\partial x_1} & \frac{\partial h(x_n)}{\partial x_2} & \dots & \frac{\partial h(x_n)}{\partial x_n} \end{bmatrix}$$

$$= \left(\nabla_u g\right)\left(\nabla_{x_{1,2,\dots n}} h(x)\right)$$

given a product of two matrices.