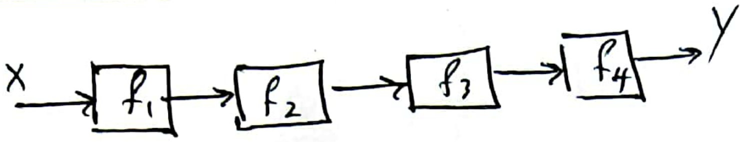
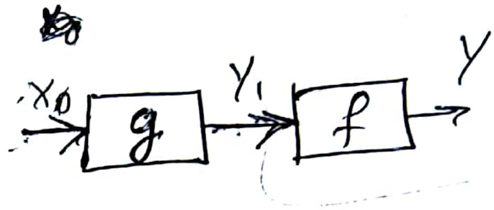
MA20 I

$f: \mathbb{R}^m \longrightarrow \mathbb{R}^n$ $\qquad f(x) \simeq f(x_0) + \underbrace{J(x_0)}_{\substack{\text{Jacobian} \\ n\times m}} (x - x_0)$

$$J = \begin{array}{cc} & j \\ i & \begin{bmatrix} & \dfrac{\partial f_i}{\partial x_j} & \end{bmatrix} \end{array}$$

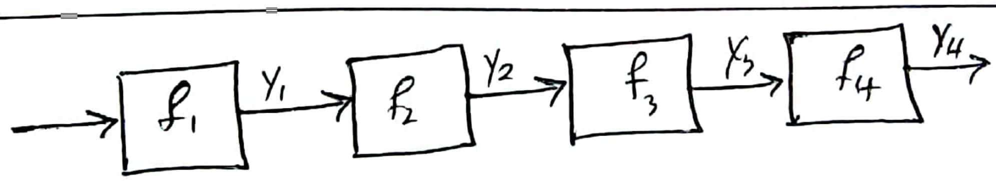


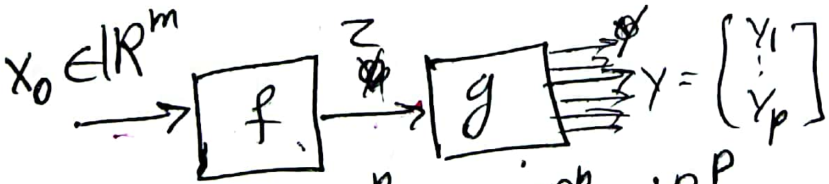


$$\frac{d}{dx} f(g(x))\Big|_{x_0} = (f\circ g)'(x_0)$$

$$= g'(x_0)\, f'(g(x_0))$$

$$= \frac{dg}{dx}\Big|_{x_0} \frac{df}{dy_1}\Big|_{g(x_0)}$$





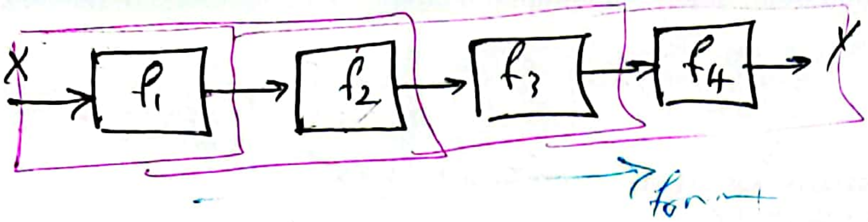
$f: \mathbb{R}^m \to \mathbb{R}^n \qquad g: \mathbb{R}^n \to \mathbb{R}^p$

$$\underbrace{\frac{\partial y}{\partial x}}_{p\times m}\Big|_{x_0} = \frac{\partial y}{\partial z}\Big|_{f(x_0)} \frac{\partial z}{\partial x}\Big|_{x_0}$$

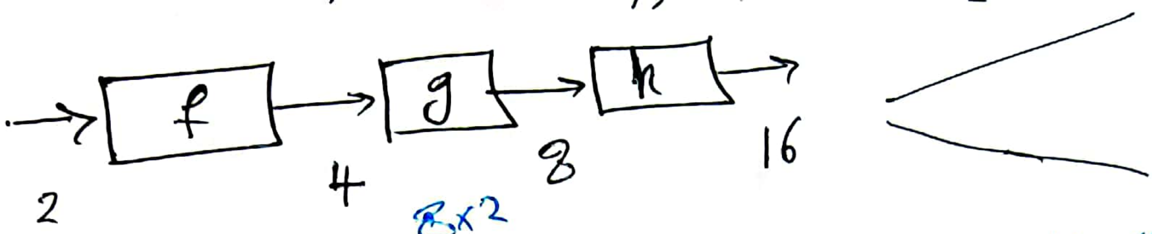
$$J_{g\circ f}\Big|_{x_0} = J_g\Big|_{f(x_0)} J_f\Big|_{x_0}$$

Scanned with CamScanner

MA II
20



$$J_{f_4 \circ f_3 \circ f_2 \circ f_1}\Big|_{x_0} = J_{f_4}(f_3(f_2(f_1(x_0))))\, J_{f_3}(f_2(f_1(x_0)))\, J_{f_2}(f_1(x_0))\, J_{f_1}(x_0) \qquad \cancel{J_{f_1}(x_0)}$$

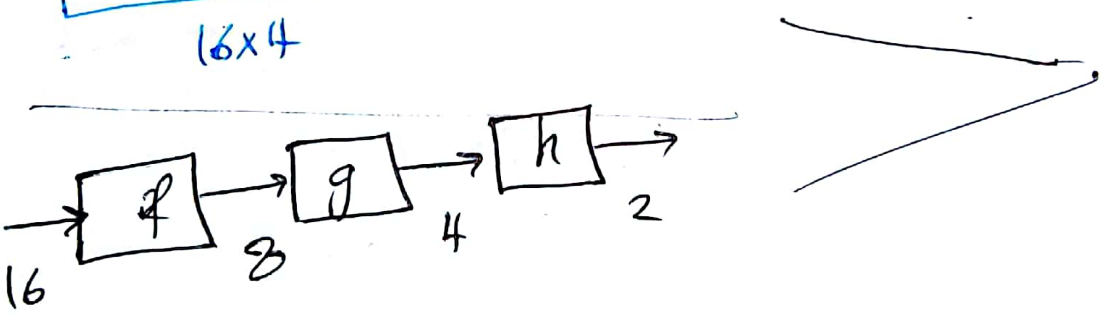


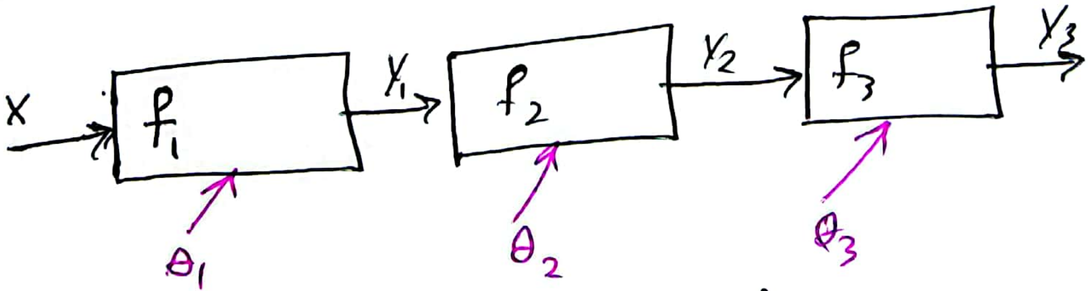
$$\underbrace{J_h \overbrace{J_g \; J_f}^{8\times 2}}_{16\times 4}$$

$16\times 8 \quad 8\times 4 \quad 4\times 2$

$16\times 8\times 4 + 16\times 4\times 2$

$8\times 4\times 2 + 16\times 8\times 2$

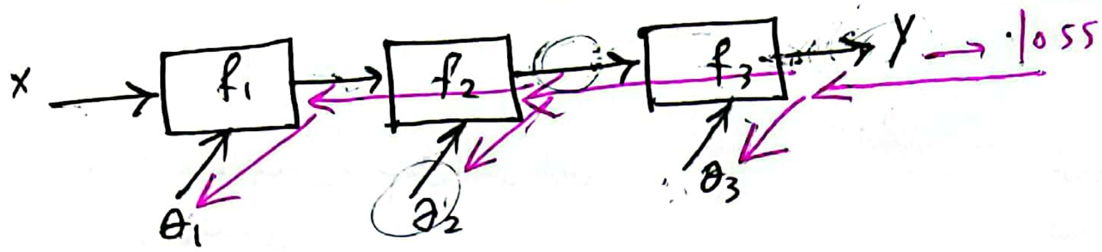




f_i: $\quad f_i(\theta_i, Y_{i-1})$

$f(\theta, x) \qquad f: \mathbb{R}^p \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$

$\frac{\partial f}{\partial \theta} \in \mathbb{R}^{m\times p} \qquad \frac{\partial f}{\partial x} \in \mathbb{R}^{m\times n}$



traing data: $x^1, x^2, \dots, x^N$

$y^1, y, \dots, y^N$

$$C(\theta_1, \theta_2, \theta_3) = \sum_{i=1}^{N} \left[L\left(y^i, f_3(\theta_3, f_2(\theta_2, f_1(\theta_1, x^i)))\right) \right]$$

$$\frac{\partial C}{\partial \theta_3} = \sum_{i=1}^{N} \frac{\partial L}{\partial y}\bigg|_{f_3(f_2(f_1(x^i)))} \frac{\partial f_3}{\theta_3}\bigg|_{f_2(f_1(x^i))}$$

$\vdots$

$$\frac{\partial C}{\partial \theta_1} = \sum_{i=1}^{N} \frac{\partial L}{\partial y}\bigg|_{f_3 \circ f_2 \circ f_1(x^i)} \frac{\partial f_3}{\partial y_2}\bigg|_{f_2(f_1(x^i))} \frac{\partial f_2}{y_1}\bigg|_{f_1(x^i)}$$

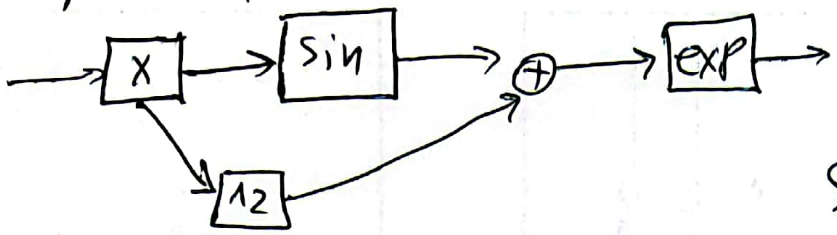
$$\frac{\partial f_1}{\theta_1} f_1(x^i)$$

$$\boxed{\frac{\partial L}{\partial y}} \frac{\partial y}{\partial \theta_3}$$

$$\boxed{\frac{\partial L}{\partial y} \frac{\partial y}{\partial y_2}} \frac{\partial y_2}{\partial \theta_2}$$

$$\boxed{\frac{\partial L}{\partial y} \frac{\partial y}{\partial y_2}} \frac{\partial y_2}{\partial y_1} \frac{\partial y_1}{\partial \theta_1}$$

$$\frac{d}{dx}\exp(\sin(x)+x^2) = (\cos(x)+2x)\exp(\sin x+x^2)$$ MA IV



Symbolic

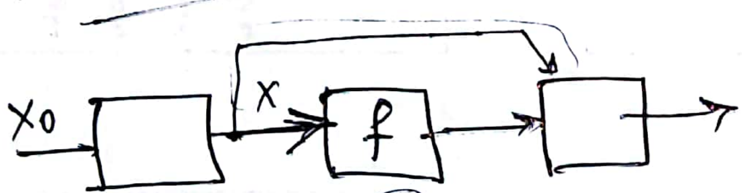
$f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ $\qquad f = f_3 \circ f_2 \circ f_1$

$$\frac{\partial f}{\partial x_i} \approx \frac{f(x+\delta e_i)-f(x)}{\delta}$$ $\qquad \delta = 10^{-8}$

Numeric

Algorithmic Differentiation

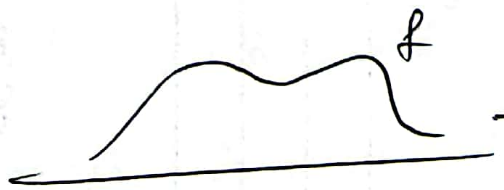
Automatic Differentiation



$f = ✓$

$\frac{\partial f}{\partial x} = ✓$

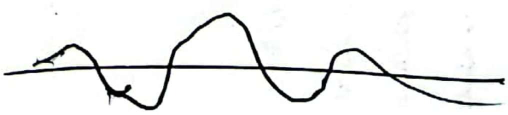
Second Derivative



$\xrightarrow{\frac{\partial f}{\partial x}}$

$\downarrow \; \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}$

$f''>0 \qquad f''<0$



MA 20 (V)

$f: \mathbb{R}^n \longrightarrow \mathbb{R}$

$D[u]f : \mathbb{R}^n \longrightarrow \mathbb{R}$

$D[v]D[u]f : \mathbb{R}^n \longrightarrow \mathbb{R}$

$D^2[\underline{v}, \underline{u}]$

$f: \mathbb{R}^m \longrightarrow \mathbb{R}$ خطی linear

$$f(x) = f\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}\right) = f\left(x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 1 \end{bmatrix} + \cdots + x_m \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}\right)$$

$$= f(x_1 \vec{e}_1 + x_2 \vec{e}_2 + \cdots + x_m \vec{e}_m)$$

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \qquad = x_1 \underbrace{f(\vec{e}_1)}_{a_1} + x_2 \underbrace{f(e_2)}_{a_2} + \cdots + x_m \underbrace{f(\vec{e}_m)}_{a_m}$$

$$= a^T x$$

$f \equiv a \longrightarrow$ تنسور مرتبه ۱ $\quad a = [a_1 \text{ — } a_2]$

آرایه تک بعدی

$f: \mathbb{R}^m \times \mathbb{R}^n \longrightarrow \mathbb{R} \qquad f(x, y) \qquad$ bilinear

$f: \mathbb{R}^3 \times \mathbb{R}^2 \longrightarrow \mathbb{R} \qquad f(x, y) \quad x \in \mathbb{R}^3 \quad y \in \mathbb{R}^2$

$$f(x, y) = f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = f(x_1 \vec{e}_1 + x_2 \vec{e}_2 + x_3 \vec{e}_3, y)$$

$$\cancel{x_1 f(e_1)} \; x_1 f(e_1, y) + x_2 f(e_2, y) + x_3 f(e_3, y) \qquad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} =$$

$$x_1 y_1 f(e_1, e'_1) + x_1 y_2 f(e_1, e'_2) \qquad y_1 e'_1 + y_2 e'_2$$

$$x_2 y_1 f(e_2, e'_1) + x_2 y_2 f(e_2, e'_2) \qquad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$x_3 y_1 f(e_3, e'_1) + x_3 y_2 f(e_3, e'_2)$$

M420 VI

$$A = \begin{bmatrix} f(e_1, e'_1) & f(e_1, e'_2) \\ f(e_2, e'_1) & f(e_2, e'_2) \\ f(e_3, e'_1) & f(e_3, e'_2) \end{bmatrix} \in \mathbb{R}^{3\times 2}$$

$$f(x, y) = f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right) = \sum_{i=1}^{3} \sum_{j=1}^{2} x_i \, y_j \, a_{ij}$$

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x^T A y$$

$$\boxed{\begin{array}{c} f : \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R} \\ \text{bilinear} \end{array}} \equiv \boxed{A \in \mathbb{R}^{3\times 2}}$$

$$\boxed{\begin{array}{c} f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R} \\ \text{bilinear} \end{array}} \equiv \boxed{A \in \mathbb{R}^{m\times n}}$$ آرایه دو بعدی

f or A is called second-order tensor

تنسور مرتبه ۲

$$f(x, y) = x^T A y = \sum_i \sum_j a_{ij} x^i y^j$$

$$f(x, y) = a_{ij} x^i y^j$$

$f(x, y, z)$ trilinear

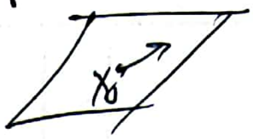
$$f : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$$

تنسور مرتبه ۳

$$f(x, y, z) = \sum \sum \sum a_{ijk} x^i y^j z^k$$

$f(x_1^{n_1}, x_2^{n_2}, \ldots, x_n^{n_n})$ multilinear

$$D[v]D[u]f : \mathbb{R}^n \to \mathbb{R}$$



$$= \boxed{D^2[v,u]f}$$ bilinear in (v,u) at any point x_0

$$D^2[v,u]f\Big|_{x_0} = v^T H u \qquad H = H(x_0)$$

Hessian Matrix

هشین

هسین

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & & & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$