$$f(x_0+h) \approx f(x_0) + [D_{x_0} f](h)$$

$$f(x_0+h) - f(x_0) \approx [D_{x_0} f](h)$$

$$f: R^m \to R^n$$

increment

$$x_0 \in R^m, \; h \in R^m, \; h = X - X_0$$

$$f(x) \in R^n$$

$$[D_{x_0} f](h) \in R^n$$

Examples, when $f(x) \in R^1$

1. $f : R^1 \to R^1$

$$f(x_0+h) = f(x_0) + [D_{x_0} f(x)](h)$$

$$f(x_0+h) - f(x_0) = f'(x) \cdot h$$

2. $f : R^m \to R$

$$f(x_0+h) - f(x_0) = [D_{x_0} f(x)](h)$$

$$= \sum \frac{\partial f(x_0)}{\;} \cdot h = \langle \partial f(x \;), h \rangle$$

$$= \sum_i x_i h_i = \left\langle \frac{\partial f(x)}{\partial x}, h \right\rangle$$

$$\nabla_{x_0} f$$

3. $f: R^{m\times n} \to R$

$$f(X_0+H_0) - f(X_0) = [D_{X_0} f(X_0)](H)$$

$$= \sum_{i,j} \frac{\partial f(X_0)}{\partial X_{ij}} \cdot H_{ij} = tr\left(\frac{\partial f(X_0)}{\partial X}\right)^T H$$

$$= \left\langle \frac{\partial f(X_0)}{\partial X}, H \right\rangle$$

$$\nabla_{x_0} f$$

$$f(x) = u(V(x)) = u \circ V$$

$$f(x_0+h) - f(x_0) = u(V(x_0+h)) - u(V(x_0))$$

Let $y = V(x)$, then

$$y_0 = V(x_0), \quad \Delta y = V(h)$$

$$u(y_0+\Delta y) - u(y_0) = [D u](\Delta y)$$

$u(y_0+\Delta y) - u(y_0) \approx [D_{y_0} \ldots$

$$\Delta y = V(h) = V(x_0+h) - V(x_0)$$

$$\Delta y \approx [D_{x_0} V(x_0)](h)$$

This is a <u>chain</u> method:

$$[D_{x_0} f(x_0)](h) =$$
$$[D_{V(x_0)} u(V(x_0))]\left([D_{x_0} V(x_0)](h)\right)$$

When u is linear:

$$u \sim L : L(y_0+\Delta y) = L(y_0) + L(\Delta y) \Rightarrow$$
$$\Rightarrow [D_{y_0} L(y_0)](\Delta y) = L(\Delta y)$$

Thus:

$$D_{x_0}[f(x_0)](h) = L\left([D_{x_0} V(x_0)](h)\right)$$

Ex. 7

$$f(X) = \langle a, x \rangle, \ a\text{-const}$$

$$[D f(x_0)](h) = \langle [D a](h), x \rangle +$$

$$\left[D_{x_0} f(x_0)\right](h) = \langle [D_{x_0} g](h), x_0 \rangle$$

$$+ \langle a, [D_{x_0} X_0](h) \rangle = \langle 0 \cdot h, X_0 \rangle$$

$$+ \langle a, h \rangle = 0 + \langle a, h \rangle = \langle \underset{\nabla_{x_0} f}{a}, h \rangle$$

Ex. 2.

$$f(x) = \langle Ax, x \rangle, \quad A - \text{const}$$

$$\left[D_{x_0} f(x_0)\right](h) = \langle Ah, x_0 \rangle + \langle Ax_0, h \rangle =$$

$$= \langle \nabla_{x_0} f, h \rangle \qquad \Big| \qquad \langle Ah, x_0 \rangle = (Ah)^T x_0 =$$

$$= A^T h^T x_0 = h^T (A^T x_0)$$

$$= \langle h, A^T x_0 \rangle = \langle A^T x_0, h \rangle$$

$$\downarrow ?$$

$$\nabla_{x_0} f = A^T x_0 + A x_0 = (A^T + A) x_0$$

Ex. 3 $f(x) = \|Ax - b\|^2 =$

$$= \langle Ax - b, Ax - b \rangle$$

$$\left[D_{x_0} f(x_0)\right](h) = \langle Ah, Ax - b \rangle + \langle Ax - b, Ah \rangle$$

$$= \langle \nabla_{x_0} f , h \rangle$$

$$\langle A h , Ax-b \rangle = h^T (A^T(Ax-b)) =$$

$$= \langle A^T(Ax-b), h \rangle$$

$$\nabla_{x_0} f = 2A^T(Ax-b)$$

$\boxed{\text{Ex. 4}}$ $$f(X) = X^{-1}, \quad [D_{x_0}(X_0^{-1})](M) - ?$$

$$X X^{-1} = I$$

$$D_{x_0} X X^{-1} = 0 = [D_{x_0}(X_0)](M) \cdot X^{-1}$$

$$+ X \cdot [D_{x_0}(X_0^{-1})](M) = 0$$

$$X \cdot [D_{x_0}(X_0^{-1})](M) = -[D_{x_0}(X_0)](M) \cdot X^{-1} \quad \Big| \cdot X^{-1}$$

$$\cancel{I} \cdot [D_{x_0}(X_0^{-1})](M) = -X^{-1} \cdot M \cdot X^{-1}$$

$\boxed{\text{Ex. 5}}$ $f(x) = \det(X)$

$$\nabla_{x_0} f = \left(\frac{\partial f}{\partial x_{ij}} \right)_{ij}$$

Expand det along ith row i

$$\det(X) = \sum_k X_{ik} \cdot (-1)^{i+k} M_{ik}$$

$$\frac{\partial f}{\partial x_{ij}} = (-1)^{i+j} \cdot M_{ij}$$

transpon them

$$X^{-1} = \frac{1}{\det(X} \cdot \left[(-1)^{i+j} M_{ji} \right)_{(i,j)}$$

$\cdot \det(X)$

$$X^{-T} = \frac{1}{\det(X)} \cdot \left((-1)^{i+j} M_{ij} \right)_{(i,j)}$$

$\frac{\partial f}{\partial x_{ij}}$

$$\nabla_{x_0} f = X^{-T} \cdot \det(X)$$

$[D_{X_0} f](M) = \langle X^{-T} \det(X), M \rangle$

Ex. 6

$$f(X) = \ln(\det(X))$$

$$[D_{X_0} f(X_0)](M) =$$

$$= [D_{\det(X_0)} \ln]([D_{X_0} \det(X_0)](M))$$

$$= \frac{1}{\det(X_0)} \cdot \langle X^{-T} \det(X_0), M \rangle$$

$$= \underbrace{\langle X^{-T}, M \rangle}_{D_{X_0} f}$$

Ex. 7. $f(X) = \operatorname{tr}(A X^T X)$

tr is linear

$[D_{X_0} f](M)$

$$[D_{x_0} f](M) =$$

$$= tr\left([D_{x_0} A X_0^T](M) X_0 + A X_0^T [D_{x_0} X](M)\right)$$

$$= tr\left(A [D_{x_0} X_0^T](M) X_0 + A X_0^T M\right) =$$

$$\boxed{[D_{x_0} X^T](M) = V(X_0 + M) - V(X_0) = (X_0 + M)^T - X_0^T = \cancel{X_0^T} + M^T - \cancel{X_0^T} = M^T}$$

$$= tr(A M^T X_0 + A X_0^T M)$$

$$= tr(A M^T X_0) + tr(A X_0^T M) =$$

$$= tr(X_0 A M^T) + tr(X_0 A^T M^T)$$

$$= tr((X_0 A + X_0 A^T) M^T)$$

$$\ldots A^T)^T M)$$

$$= \mathrm{tr}\left(\left(X_0(A+A^T)\right)^T M\right)$$

$$= \langle X_0(A+A^T), M \rangle$$

$$\nabla_{X_0} f$$