Derivatives 21/22

Differentiation of univariate functions

Difference quotient: $\frac{\delta y}{\delta x} \quad \frac{f(x+\delta x) - f(x)}{\delta x}$ computes the slope of the secant line through two points on the graph of $f(x)$ → average slope of f between x and $x+\delta x$.

In the limit of $\delta x \to 0$, we obtain the tangent of f at x, if f is differentiable

Derivative: $\frac{df}{dx} = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon}$

The derivative of f points in the direction of steepest ascent

Example:

$y = f(x) = 2x$

$$\boxed{\begin{array}{l}\text{Difference quotient} \quad \delta x = 1, \; x_0 = 1 \\ \frac{\delta y}{\delta x} = \frac{f(1+1) - f(1)}{1} = \frac{4-2}{1} = 2\end{array}}$$

$$f'(x) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon} = \frac{\cancel{2x} + 2\epsilon - \cancel{2x}}{\epsilon} = 2$$

Exercise

$f(x) = 2x^2 - 16x + 35$

$$f'(x) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon} = \frac{2(x+\epsilon)^2 - 16(x+\epsilon) + 35 - (2x^2 - 16x + 35)}{\epsilon}$$

$$= \frac{\cancel{2x^2} + 2\epsilon^2 + 4x\epsilon \overset{-16\epsilon}{\cancel{-16(x+\epsilon)}} + \cancel{35} - \cancel{2x^2} + \cancel{16x} - \cancel{35}}{\epsilon}$$

$$= \underset{\to 0}{2\epsilon} + 4x - 16 = 4x - 16$$

Basic differentiation rules

Product rule:	$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$	(5.29)
Quotient rule:	$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$	(5.30)
Sum rule:	$(f(x) + g(x))' = f'(x) + g'(x)$	(5.31)
Chain rule:	$(g(f(x)))' = (g \circ f)'(x) = g'(f(x))f'(x)$	(5.32)

$y = f(x) \quad z = g(y)$

$$(g(f(x)))' = \frac{dz}{dy} \cdot \frac{dy}{dx}$$

Exercise (chain rule)

$h(x) = (2x+1)^4 = g(f(x))$

$f(x) = 2x + 1$ compute $f'(x) = 2$

$g(f) = f^4$ derivatives $g'(f) = 4f^3$

$h'(x) = g'(f(x))\, f'(x) = 4 f(x)^3 \cdot 2 = 4(2x+1)^3 \cdot 2 = \underline{8(2x+1)^3}$

TODO AT HOME

Exercises x: variable t, a, b, c: constants (numbers)

$\frac{d}{dx} 2 =$ $\frac{d}{dx} 2 \cdot 3 =$ $\frac{d}{dx} x$ $\frac{d}{dx} c \cdot x =$

$\frac{d}{dx} c =$ $\frac{d}{dx} c \cdot t$ $\frac{d}{dx} x^2$ $\frac{d}{dx} (c \cdot x)^2$

$\frac{d}{dx} (t - x) =$ $\frac{d}{dx} (t - cx) =$

$\frac{d}{dx} (t - x)^2$ $\frac{d}{dx} (t - cx)^2$

$\frac{d}{dx} (cx + ab)$ $\frac{d}{dx} \left(t - (cx + ab)\right) =$ $\frac{d}{dx} \left(t - (cx - ab)\right)^2$

Exercises x : variable t, a, b, c : constants (numbers)

$\frac{d}{dx} 2 = 0$ $\frac{d}{dx} 2 \cdot 3 = 0$ $\frac{d}{dx} x = 1$ $\frac{d}{dx} c \cdot x = c$

chain rule

$\frac{d}{dx} c = 0$ $\frac{d}{dx} c \cdot t = 0$ $\frac{d}{dx} x^2 = 2x$ $\frac{d}{dx} (c \cdot x)^2 = 2cx \cdot c$

$\frac{d}{dx} (t - x) = -1$ $\frac{d}{dx} (t - cx) = -c$

chain rule

$\frac{d}{dx} (t - x)^2 = 2(t-x) \cdot (-1)$ $\frac{d}{dx} (t - cx)^2 = 2(t - cx) \cdot (-c)$

$\frac{d}{dx} (cx + ab) = c$ $\frac{d}{dx} \left(t - (cx + ab)\right) = 0 - c - 0$ $\frac{d}{dx} \left(t - (cx - ab)\right)^2 = 2\left(t - cx + ab\right) \cdot (-c)$

Multivariate Calculus

General case of functions depending on ≥ 1 variables $\vec{x} \in \mathbb{R}^n$

e.g. $f(x_1, x_2)$

DEF: $$\frac{\partial f}{\partial x_1} = \lim_{\epsilon \to 0} \frac{f(x_1+\epsilon, x_2, \dots, x_n) - f(\vec{x})}{\epsilon}$$

...

$$\frac{\partial f}{\partial x_n} = \lim_{\epsilon \to 0} \frac{f(x_1 \dots x_n+\epsilon) - f(\vec{x})}{\epsilon}$$

GRADIENT: $\nabla_{\vec{x}} f = \frac{df}{d\vec{x}} = \left[\frac{\partial f(\vec{x})}{\partial x_1}, \dots, \frac{\partial f(\vec{x})}{\partial x_n} \right] \rightarrow$ may be considered row vector $\in \mathbb{R}^{1 \times n}$

Exercise : partial derivatives using chain rule

N.B. when computing the partial der. w.r.t. one variable, consider other variables as constants.

$f(x,y) = (x+2y^3)^2$

$g(f) = f^2 \qquad f(x,y) = x+2y^3$

Chain rule

$\boxed{g'(f(x))\, f'(x)}$

$$\frac{\partial f}{\partial x} = 2(x+2y^3) \underbrace{\frac{\partial}{\partial x}(x+2y^3)}_{=1} = 2(x+2y^3)$$

$$\frac{\partial f}{\partial y} = 2(x+2y^3) \frac{\partial}{\partial y}(x+2y^3) = 2(x+2y^3)\, 6y^2 = 12(x+2y^3)\, y^2$$

$$\nabla_{(x,y)} f = [\, 2(x+2y^3),\ 12(x+2y^3)\, y^2 \,]$$

Exercise (AC)

$z = f(x_1, x_2, x_3) = x_1^2 + x_2 x_3$

$$\frac{\partial z}{\partial x_1} = \overset{2x_1}{?} \qquad \frac{\partial z}{\partial x_2} = \overset{x_3}{?} \qquad \frac{\partial z}{\partial x_3} = \overset{x_2}{?} \qquad \nabla_{\vec{x}} z = [\qquad]$$

Exercise 2 (AC)

$f(x_1, x_2) = x_1^2 x_2 + x_1 x_2^3$

chain + power, sum

$$\frac{\partial f(x_1,x_2)}{\partial x_1} = 2x_1 x_2 + x_2^3$$

$$\frac{\partial f(x_1,x_2)}{\partial x_2} = x_1^2 + 3x_1 x_2^2$$

PRODUCT RULE

$$\frac{df}{d\vec{x}} = \nabla_{\vec{x}} f = [\qquad]$$

Rules of Partial Differentiation

We can deal with gradients in a symbolic way. In this case, the gradients now involve vectors and matrices

$$\text{Product rule:} \quad \frac{\partial}{\partial \boldsymbol{x}}\big(f(\boldsymbol{x})g(\boldsymbol{x})\big) = \frac{\partial f}{\partial \boldsymbol{x}} g(\boldsymbol{x}) + f(\boldsymbol{x})\frac{\partial g}{\partial \boldsymbol{x}} \qquad (5.46)$$

$$\text{Sum rule:} \quad \frac{\partial}{\partial \boldsymbol{x}}\big(f(\boldsymbol{x}) + g(\boldsymbol{x})\big) = \frac{\partial f}{\partial \boldsymbol{x}} + \frac{\partial g}{\partial \boldsymbol{x}} \qquad (5.47)$$

$$\text{Chain rule:} \quad \frac{\partial}{\partial \boldsymbol{x}}(g \circ f)(\boldsymbol{x}) = \frac{\partial}{\partial \boldsymbol{x}}\big(g(f(\boldsymbol{x}))\big) = \frac{\partial g}{\partial f}\frac{\partial f}{\partial \boldsymbol{x}} \qquad (5.48)$$

Chain Rule as Matrix Mulciplication

Consider a function $f:\mathbb{R}^2\to\mathbb{R}$ of 2 variables x_1 and x_2. Let $x_1(t)$ and $x_2(t)$ be themselves functions of t.

We can apply the chain rule for multivariate functions:

$$\frac{dF}{dt}=\frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t}+\frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t}=\begin{bmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2}\end{bmatrix}\underbrace{\begin{bmatrix}\frac{\partial x_1(t)}{\partial t}\\ \frac{\partial x_2(t)}{\partial t}\end{bmatrix}}$$

GRADIENT OF VECTOR-VALUED FUNCTION. WE SEE IT AS A VECTOR OF FUNCTIONS

Gradient is always a row vector!

$$\begin{bmatrix}\nabla_t x_1\\ \nabla_t x_2\end{bmatrix}$$

EXAMPLE $h:\mathbb{R}\to\mathbb{R}$ $\quad h(t)=(f\circ g)(t)$

$F(x_1,x_2)=x_1^2+2x_2$ $\quad x_1=\sin(t)$ $\quad x_2=\cos(t)$ $\quad x(t)=\begin{bmatrix}\sin(t)\\ \cos(t)\end{bmatrix}=\vec{x}=\begin{bmatrix}x_1\\ x_2\end{bmatrix}$

$\to \mathbb{R}^2\to\mathbb{R}$ $\qquad \mathbb{R}\to\mathbb{R}^2$

$$\frac{dh}{dt}=\frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t}+\frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t}=2x_1\frac{\overset{\sin(t)}{\partial x_1}}{\partial t}+2\frac{\overset{\cos(t)}{\partial x_2}}{\partial t}=2\sin(t)\cos(t)+2(-\sin(t))$$
$$=2\sin(t)(\cos(t)-1)$$

in matrix notation

$$\frac{dh}{dt}=\frac{\partial f}{\partial\vec{x}}\frac{\partial\vec{x}}{\partial t}=\begin{bmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2}\end{bmatrix}\begin{bmatrix}\frac{\partial x_1}{\partial t}\\ \frac{\partial x_2}{\partial t}\end{bmatrix}=\begin{bmatrix}2x_1 & 2\end{bmatrix}\begin{bmatrix}\cos(t)\\ -\sin(t)\end{bmatrix}$$

JACOBIAN

Same example, but now $g(s,t)=\begin{bmatrix}s\sin(t)\\ s\cos(t)\end{bmatrix}=\begin{bmatrix}g_1\\ g_2\end{bmatrix}$

The derivative is $\frac{\partial t}{\partial\vec{x}}=\nabla_{\vec{x}}t$

$\nabla_{(s,t)}g_1$

$$\frac{dh}{dt}=\frac{\partial f}{\partial\vec{x}}\frac{\partial\vec{x}}{\partial(s,t)}=\begin{bmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2}\end{bmatrix}\begin{bmatrix}\frac{\partial g_1}{\partial s} & \frac{\partial g_1}{\partial t}\\ \frac{\partial g_2}{\partial s} & \frac{\partial g_2}{\partial t}\end{bmatrix}$$

$\frac{\partial g}{\partial(s,t)}=\nabla_{(s,t)}g_1$

$\nabla_{(s,t)}g_2=\nabla_{(s,t)}g_2$

$$=\begin{bmatrix}2x_1 & 2\end{bmatrix}\begin{bmatrix}\sin(t) & s\cdot\cos(t)\\ \cos(t) & -s\cdot\sin(t)\end{bmatrix}$$

JACOBIAN