$f: \mathbb{R}^2 \to \mathbb{R}$

$(x,y) \mapsto f(x,y)$

$f: \mathbb{R} \to \mathbb{R}^2$

$t \mapsto \langle x(t), y(t) \rangle$

$$Df = \nabla f = \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \rangle$$

$$Df = \frac{df}{dt} = \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix}$$

General chain rule:

$$\begin{tikzcd} \mathbb{R}^a \arrow[r, "g"] \arrow[rr, bend left, "f(g(\underline{x}))"] & \mathbb{R}^b \arrow[r, "f"] & \mathbb{R}^c \end{tikzcd}$$

$$\underline{x} \longrightarrow g(\underline{x}) \longrightarrow f(g(\underline{x}))$$

$$\underset{\substack{b \times a \\ \text{matrix}}}{Dg(\underline{x})} \quad \underset{\substack{c \times b \\ \text{matrix}}}{Df(\underline{g}(\underline{x}))}.$$

$$\underset{c \times a}{D(f(g(\underline{x})))} = \underset{c \times b}{Df(\underline{g}(x))}\ \underset{b \times a}{Dg(x)}$$

matrix multiplication

$$\begin{bmatrix} \frac{\partial f_i}{\partial x_j} \end{bmatrix} \begin{bmatrix} \frac{\partial g_j}{\partial x_k} \end{bmatrix}$$

<u>Example</u>

① $\mathbb{R} \xrightarrow{g} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}$

$t \qquad (x,y)$

$$Dg = g'(t) = \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} \qquad Df = \nabla f = \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \rangle$$

$$\begin{aligned} D(f(g(t))) &= Df(g(t)) \cdot Dg(t) \\ &= \nabla f(g(t)) \cdot g'(t) \\ &= \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \rangle \cdot \langle \frac{dx}{dt}, \frac{dy}{dt} \rangle \end{aligned}$$

$$\left(\text{or} \quad \begin{bmatrix} \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} \right).$$

$$= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

② $\mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}^2$

$x_1, x_2 \qquad y_1, y_2$

$$g(\underline{x}) = (g_1(x_1,x_2), g_2(x_1,x_2)) \qquad f(y_1,y_2) = (f_1(y_1,y_2), f_2(y_1,y_2))$$

$$Dg = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{bmatrix} \qquad Df = \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} \end{bmatrix}$$

$$D(f(g(\underline{x}))) = Df(g(\underline{x})) \cdot Dg(\underline{x})$$

$$Df\, Dg = \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}\frac{\partial f_1}{\partial y_1} + \frac{\partial g_1}{\partial x_2}\frac{\partial f_2}{\partial y_2} & \cdots \\ \cdots & \cdots \end{bmatrix}$$

<u>mnemonic</u> $f(x_1, \dots, x_n)$ $x_i(y_1, y_2, \dots, y_m)$

to find $\frac{\partial f_j}{\partial y_i} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y_i} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial y_i} + \dots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial y_i}$

"differentiate f wrt all variables, multiply by $\frac{\partial x_k}{\partial y_i}$ and add".

<u>Example</u> $f(x,y,z) = xy + z^3$ $\quad x = s+t \quad g: \mathbb{R}^2 \to \mathbb{R}$

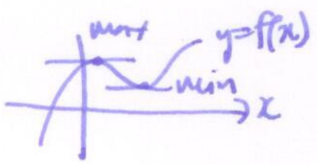
$f: \mathbb{R}^3 \to \mathbb{R}$ $\quad y = s-t$

$z = st \quad (s,t) \mapsto (x,y,z)$

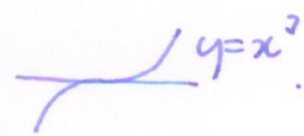
so $f(g(s,t))$ makes sense.

<u>Q</u> find $\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial s}$

§14.7 Optimization

<u>recall</u>: 1d 

max, min $\Rightarrow \frac{dy}{dx} = 0$

$\frac{dy}{dx} \not\Rightarrow$ max, min 

<u>2d</u> local max  local min  $\Rightarrow$ flat tangent plane $z = \text{const}$.

<u>recall</u>: tangent plane to $z = f(x,y)$ at (a,b) is

$$z = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$$

flat tangent plane $\Rightarrow$ $f_x(a,b) = 0$ <u>and</u> $f_y(a,b) = 0$.

<u>warning</u>: $f_x(a,b) = 0$ and $f_y(a,b) = 0 \not\Rightarrow$ local max or min.

Def^n^ A <u>critical point</u> (a,b) is a point s.t. $\frac{\partial f}{\partial x}(a,b)=0$ and $\frac{\partial f}{\partial y}(a,b)=0$ or at least one of $f_x(a,b)$, $f_y(a,b)$ does not exist.

Th^m^ If $f(x,y)$ has a local max or min at (a,b), then (a,b) is a critical point.

Example finding critical points

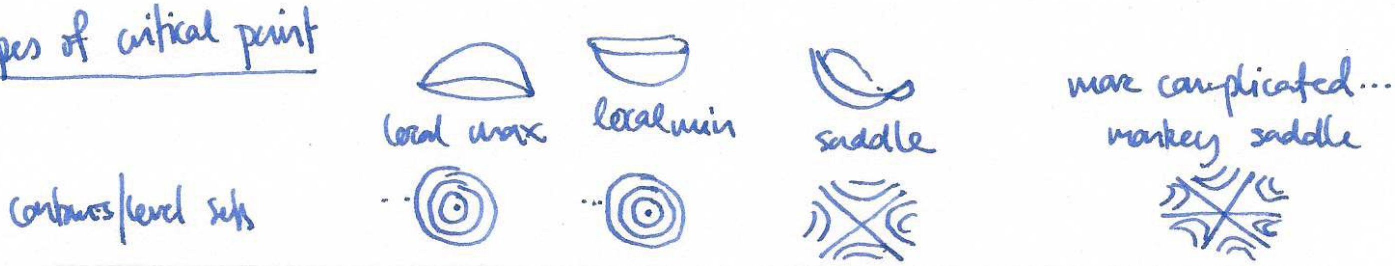
$$f(x,y) = x^2 - 2xy + 2y^2 + 3y + 1$$

$$\left.\begin{aligned} \frac{\partial f}{\partial x} &= 2x-2y = 0 \quad (1) \\ \frac{\partial f}{\partial y} &= -2x+4y+3 = 0 \quad (2) \end{aligned}\right\}$$

$(1)+(2)$: $2y+3=0$ $\quad y=-3/2 \quad x=-3/2$

exactly one critical point at $(-3/2,-3/2)$

Types of critical point

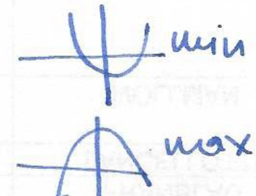


Q: how to tell which one? **A**: look at 2nd order / quadratic approximation.

1 var: $f(x) \simeq f(a) + \underbrace{f'(a)}_{=0 \text{ if critical point}}(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots$

$f''(a) > 0$

$f''(a) < 0$



$f''(a) = 0$ no information!

2 vars: **Th^m^** (2nd derivative test). Let (a,b) be a critical point for $f(x,y)$

let $D(a,b) = f_{xx}(a,b)\,f_{yy}(a,b) - f_{xy}(a,b)^2$, then:

1) if $D > 0$ then $f(a,b)$ is a <u>minimum</u> if $f_{xx}(a,b) > 0$
   <u>maximum</u> if $f_{xx}(a,b) < 0$

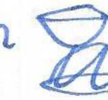
2) if $D < 0$ then $f(a,b)$ is a <u>saddle</u>

3) if $D = 0$ <u>no information</u>

Why does this work?

Quadratic approximation is $f(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) + \begin{bmatrix} x-a & y-b \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \begin{bmatrix} x-a \\ y-b \end{bmatrix}$

quadratic terms are: $f_{xx}(a,b)(x-a)^2 + 2f_{xy}(a,b)(x-a)(y-b) + f_{yy}(a,b)(y-b)^2$

this is a quadratic form, like: x^2+y^2  $-x^2-y^2$  xy or x^2-y^2 

up to change of coordinates, a symmetric matrix looks like

$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \longleftrightarrow \lambda_1 x^2 + \lambda_2 y^2$

$\lambda_1, \lambda_2 > 0$	local min
$\lambda_1, \lambda_2 < 0$	local max
different signs	saddle
at least one 0	don't know.

$D =$ det of matrix $=$ product of eigenvalues

so $D > 0$ same sign
$D < 0$ different sign.

<u>Example</u> find the critical points of $f(x,y) = (x^2+y^2)e^{-2x} = x^2e^{-2x} + y^2e^{-2x}$.

$f_x = 2xe^{-2x} + x^2(-2e^{-2x}) - 2y^2e^{-2x}$

$f_y = 2ye^{-2x}$

<u>solve</u> $\left.\begin{array}{l} f_x = 0 \\ f_y = 0 \end{array}\right\}$ $\begin{array}{l} e^{-2x}(2x - 2x^2 - 2y^2) \\ e^{-2x}(2y) = 0 \Rightarrow y = 0 \end{array}$ $\Rightarrow 2(x - x^2) = 0$

$x(x-1) = 0 \quad x = 0, 1.$

$(0,0) \quad (1,0).$

$\left.\begin{array}{l} f_{xx} = 2x.(-2e^{-2x}) + 2e^{-2x} + 2x \\ f_{yx} = -4ye^{-2x} \\ f_{yy} = 2e^{-2x} \end{array}\right\}$ $\begin{array}{l} D = f_{xx}f_{yy} - (f_{xy})^2 \\ D(0,0), D(0,1) \text{ etc. } \ldots \end{array}$