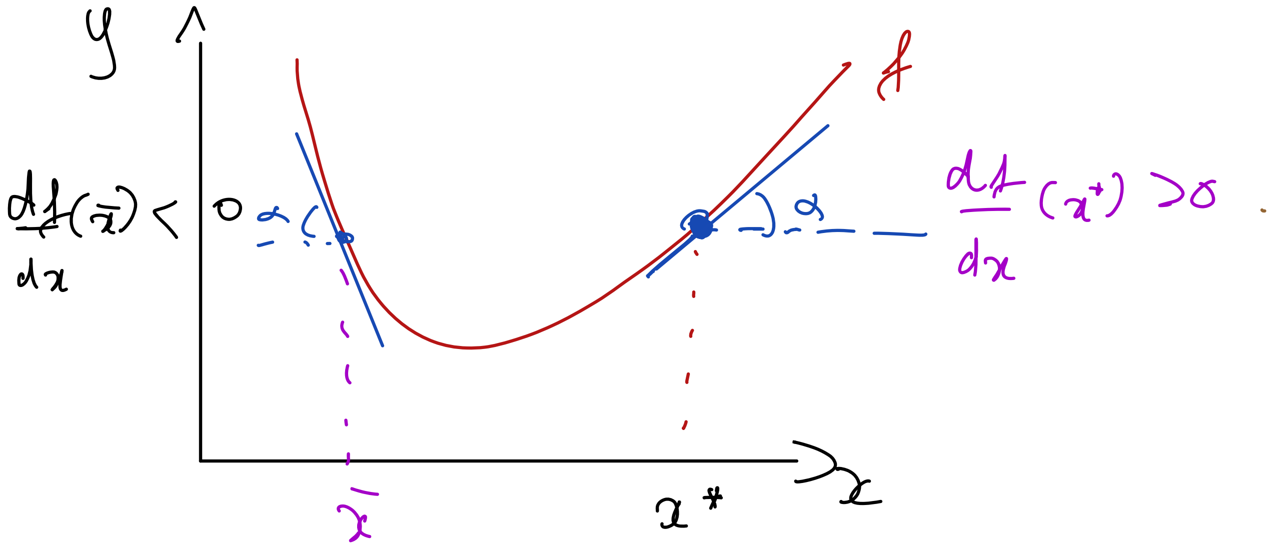


$$\tan \alpha = \frac{df}{dx}(x^*)$$

$$= \lim_{h \to 0} \frac{f(x^* + h) - f(x^*)}{h}$$

Observation: in one dimension ($x \in \mathbb{R}$) the value of the derivative tells us the "direction" where f increases. Specifically if $\frac{df}{dx}(x) > 0$ then moving in the direction h that aligns ($\frac{df}{dx}(x) \cdot h > 0$) with the gradient the function values increases.

Ex: at x^* we know that $\frac{df}{dx}(x^*) > 0$

thus if $h > 0$ (meaning we move in the positive direction) the function increases

- Now at $x = \bar{x}$ we have $\frac{df}{dx}(\bar{x}) < 0$

$\Rightarrow$ if we move in the negative direction, i.e. $h < 0$, then the function value increases

---

Introduction to Optimization

① one dimensional case : $x \in \mathbb{R}$

Let $f : \mathbb{R} \to \mathbb{R}$: is twice differentiable

Def: we say that x^* is a minimum/minimizer of f if there exists a $\delta > 0$:

$$f(x) \geq f(x^*) \; , \; \forall x \in (x^* - \delta, x^* + \delta)$$

If x^* is a (local) minimum, what can we say about its derivative?

Thm: (Necessary conditions)

Let: $f : \mathbb{R} \to \mathbb{R}$, twice diff. If x^* is a (local) minimizer, then

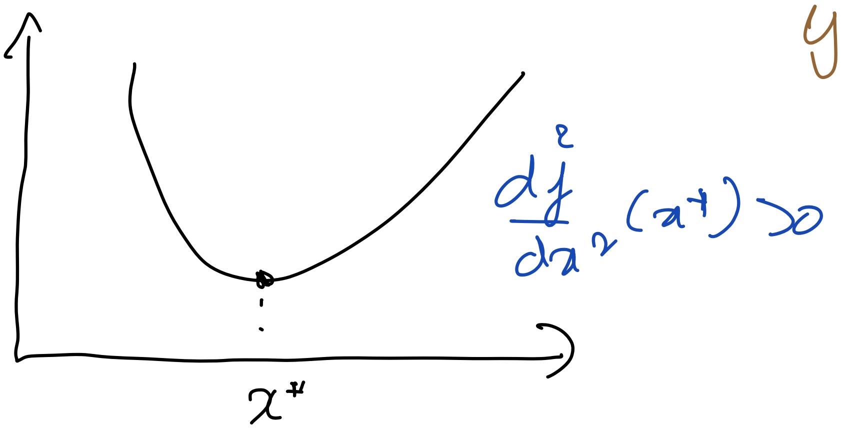
1) $\frac{df}{dx}(x^*) = 0$

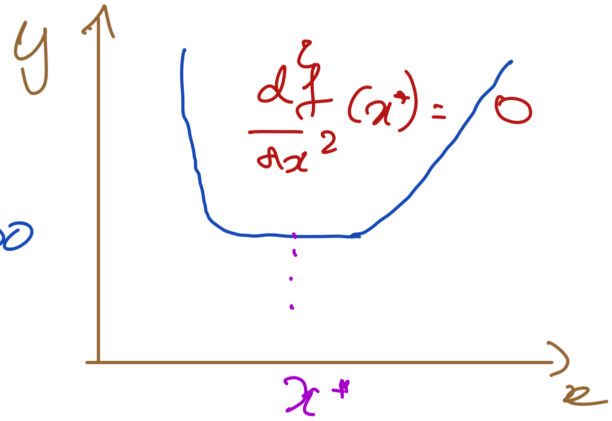
2) $\frac{d^2f}{dx^2}(x^*) \geq 0$

Proof: – use contradiction

- Continuity (sign-permanent property of continuous function)

- Note that we can only say that the curvature at x^*, i.e. $\frac{d^2f}{dx^2}(x^*)$, is non-negative if x^* is a minimizer





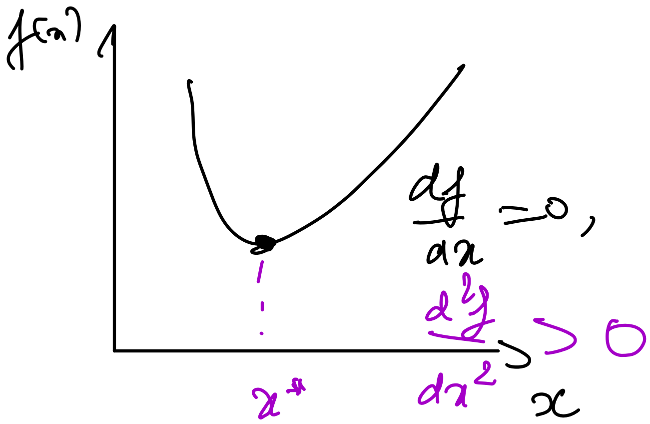
Thm. (sufficient conds)

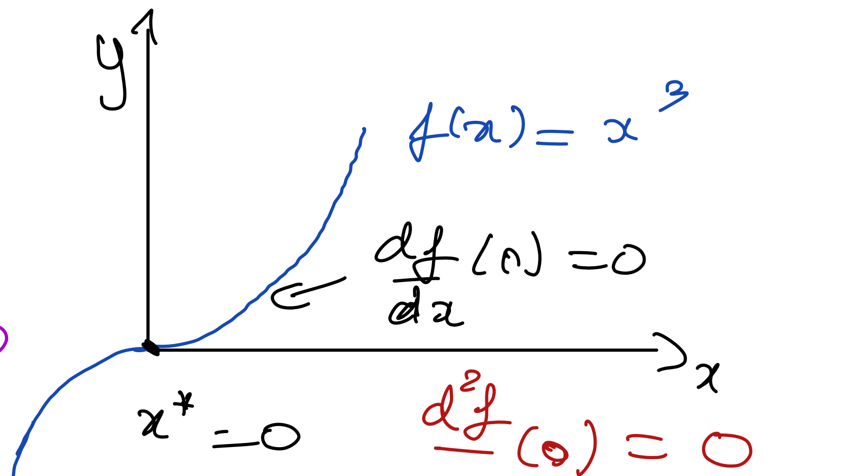
If at x^* we have

1) $\frac{df}{dx}(x^*) = 0$, and

2) $\frac{d^2f}{dx^2}(x^*) > 0$

Then x^* is a (local) minimizer





dx^2

② n-dimensional problem : $\vec{x} \in \mathbb{R}^n$:

Let $f : \mathbb{R}^n \to \mathbb{R}$, twice differentiable.

Def : (ball / neighborhood) in $\mathbb{R}^n$

$$B_\delta(\vec{x}^*) \overset{\text{def}}{:=} \text{ball of radius } \delta \text{ centered at } \vec{x}^*$$

$$= \{ \vec{x} : \|\vec{x} - \vec{x}^*\| < \delta \}$$

$$\|\vec{x} - \vec{x}^*\| = \left(\sum_{i=1}^{n} (x_i - x_i^*)^2 \right)^{1/2}$$

(of course when $n = 1 \Rightarrow \|x - x^*\| = |x - x^*|$

then: $|x - x^*| < \delta \iff x \in (x^* - \delta, x^* + \delta)$)

Def : $\vec{x}^*$ is a (local) minimizer of f if there exists $\delta > 0$: such that

$$f(\vec{x}) \geq f(\vec{x}^*) \quad \forall \vec{x} \in B_\delta(\vec{x}^*)$$

Thm : (Necessary conds)

- If f attains its (local) minimum at $\vec{x}^*$, then

1) $\vec{\nabla} f(\vec{x}^*) = \vec{0}$

2) $\nabla^2 f(\vec{x}^*) = H(\vec{x}^*) \geq 0$

where

1) $\vec{\nabla} f(\vec{x}) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right]^T$

2) $$\nabla^2 f(\vec{x}) = \begin{bmatrix} \vec{\nabla}\left(\frac{\partial f}{\partial x_1}\right)^T \\ \vdots \\ \vec{\nabla}\left(\frac{\partial f}{\partial x_n}\right)^T \end{bmatrix}$$

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

3) $H(\vec{x}) \succcurlyeq 0$ means the Hessian matrix is semi-positive definite

$\Updownarrow$

$$\vec{x}^T H \vec{x} \geq 0 \quad \forall \vec{x} \in \mathbb{R}^n$$

Note that H is strictly positive definite iff

$$\vec{x}^T H \vec{x} > 0 \quad \forall \vec{x} \neq \vec{0}$$

Proof: Proof of $\vec{\nabla} f(\vec{x}^*) = \vec{0}$

we have by def: $\exists\ \delta > 0$:

$$f(\vec{x}) \geq f(\vec{x}^*) \quad \forall \vec{x} \in B_\delta(\vec{x}^*)$$

$\Downarrow$

$\Downarrow$ taking $\vec{p}$: $\|\vec{p}\| \leq 1$

$0<|\varepsilon|<\delta$ $\quad \vec{x}^* + \varepsilon \vec{p} \in B_\delta(\vec{x}^*)$

$$f(\vec{x}^* + \varepsilon \vec{p}) \geq f(\vec{x}^*) \quad \forall \; 0<|\varepsilon|<\delta, \; \|\vec{p}\| \leq 1$$

$\Downarrow$ define $F(\varepsilon) := f(\vec{x}^* + \varepsilon \vec{p})$

$$F(\varepsilon) \geq F(0) \quad \text{for any } 0<|\varepsilon|<\delta, \; \|\vec{p}\| \leq 1$$

$\Downarrow$

$\exists \, \delta$ we have

$$F(\varepsilon) \geq F(0) \quad \forall \; -\delta < \varepsilon < \delta$$

$\Downarrow$ by def

0 is a minimum of $F(\varepsilon)$

$\Downarrow$ necessary

$$\frac{dF(0)}{d\varepsilon} = 0$$

$F(\varepsilon) = f(\vec{x}^* + \varepsilon \vec{p})$

$\Updownarrow$

$$\left. \frac{d f(\vec{x}^* + \varepsilon \vec{p})}{d\varepsilon} \right|_{\varepsilon = 0} = 0$$

but $\frac{df}{d\varepsilon}(\vec{x}^* + \varepsilon \vec{p}) =$

$$\frac{d}{d\varepsilon} f(\underline{x_1^* + \varepsilon p_1}, \dots, x_n^* + \varepsilon p_n)$$

$$= \frac{\partial f}{\partial x_1} \cdot p_1 + \dots + \frac{\partial f}{\partial x_n} \cdot p_n$$

$= \nabla f(\vec{x}^* + \varepsilon \vec{p}) \cdot \vec{p}$

Thus

$$\left.\frac{d f(\vec{x}^* + \varepsilon \vec{p})}{d\varepsilon}\right|_{\varepsilon=0} = 0$$

$$\Updownarrow$$

$$\left.\vec{\nabla} f(\vec{x}^* + \varepsilon \vec{p}) \cdot \vec{p}\right|_{\varepsilon=0} = 0$$

$$\Updownarrow$$

$$\vec{\nabla} f(\vec{x}^*) \cdot \vec{p} = 0 \quad \forall \|\vec{p}\| \leq 1$$

$$\Updownarrow$$

$$\vec{\nabla} f(\vec{x}^*) \cdot \vec{p} = 0 \quad \forall \vec{p}$$

$\Downarrow$ pick $\vec{p} = \vec{e}^{\,i} = [0, 0, .. \underset{i\text{th}}{1}, 0 ..]$

$$\vec{\nabla} f(\vec{x}^*) \cdot \vec{e}^{\,i} = 0 \quad \forall i = 1, .., n$$

$$\Downarrow$$

$$\frac{\partial f}{\partial x_i}(\vec{x}^*) = 0 \quad \forall i = 1, .., n$$

$$\Downarrow$$

$$\vec{\nabla} f(\vec{x}^*) = \vec{0}$$

different approach pick $\vec{p} = \vec{\nabla} f(\vec{x}^*)$

then $\vec{\nabla} f(\vec{x}^*) \cdot \vec{p} = 0$

$$\Downarrow$$

$$\vec{\nabla} f(\vec{x}^*) \cdot \vec{\nabla} f(\vec{x}^*) = 0$$

$$\Uparrow$$

$$\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2 = 0$$

$$\Downarrow$$

$$\frac{\partial f}{\partial x_i} = 0 \Rightarrow \vec{\nabla} f(\vec{x}^*) = \vec{0}$$

HW: show that $\nabla^2 f(\vec{x}^*) \geq 0$