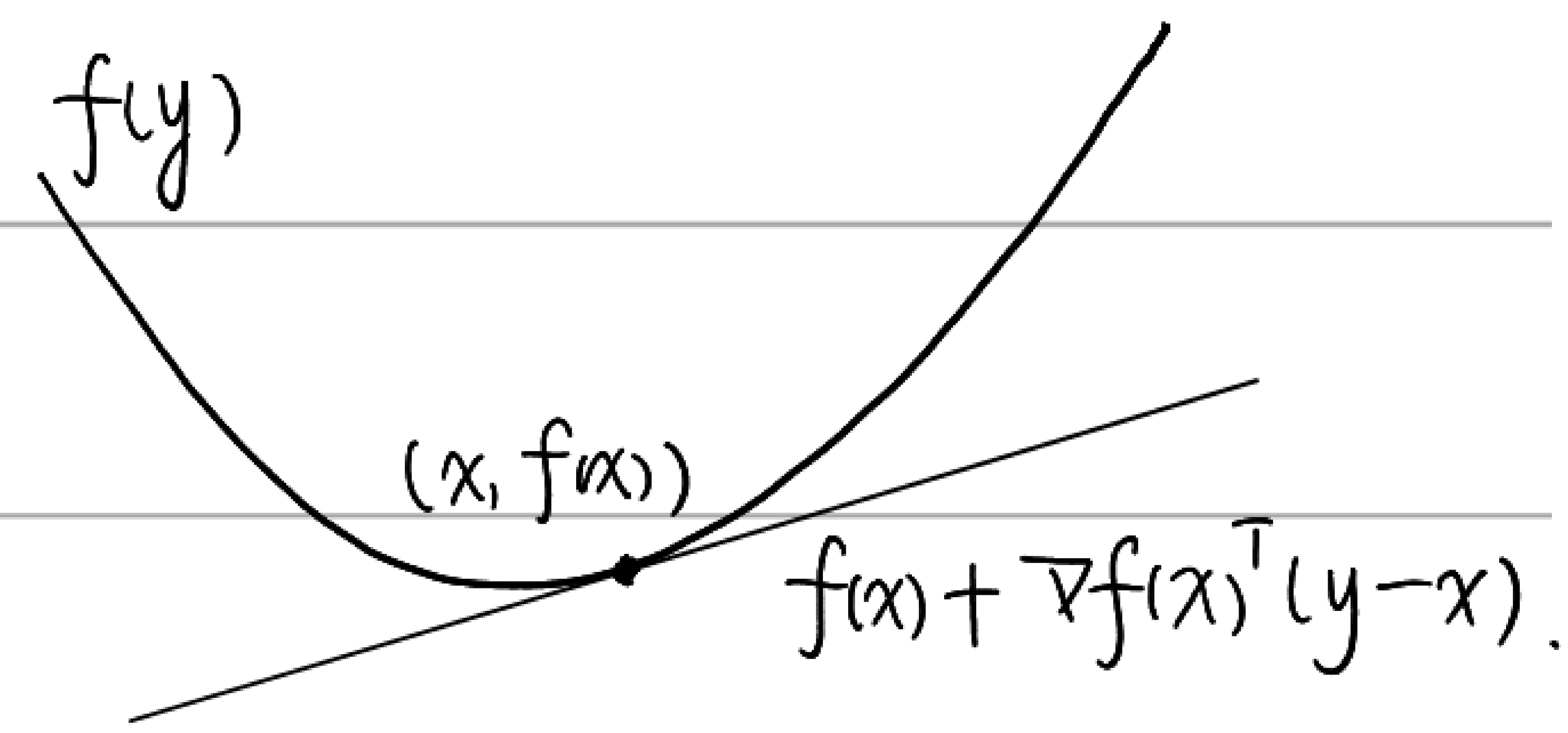
Lecture 7. Properties and characterization of convex functions

Zeroth-order condition:

f is convex iff f restricted to any direction is convex.

i.e. $\forall d \in \mathbb{R}^n$. $g(t) = f(x_0 + td)$ is convex.

First-order condition:



Suppose f is differentiable in an open convex set $\mathrm{dom} f$. Then.

f is convex iff $f(y) \geq f(x) + \nabla f(x)^T (y-x)$ $\forall x. y \in \mathrm{dom} f$.

Example (Bernoulli's inequality): $(1+x)^r \geq 1 + rx$ if $r \geq 1$. $x \geq -1$.

in particular, $e^{rx} > (1+1/k)^{krx} \ (\forall k. \forall r > 0) \overset{\text{let } k=1/x}{>} (1+x)^r$.

Remark: The first-order Taylor approximation is a global under-estimator of a convex function. and vice versa.

local information (value, gradient) $\Rightarrow$ global inequality.

In particular. $\nabla f(x) = 0 \Rightarrow f(y) \geq f(x)$ $\forall y \in \mathrm{dom} f$.

strictly convex: iff $f(y) > f(x) + \nabla f(x)^T (y-x)$ $\forall x \neq y$.

Note that $\mathrm{epi}(f)$ convex $\Rightarrow$ $f(x) + \nabla f(x)^T (y-x)$ supporting hyperplane.

Proof: "$\Rightarrow$". Fix $x, y \in \mathrm{dom}\, f$, and let $d = y - x$.

By Jensen's inequality, $f(x+td) \leq (1-t) f(x) + t f(y)$. $\forall t \in (0,1)$

$$f(x+td) - f(x) \leq t(f(y) - f(x)).$$

Taking the limit $t \to 0$, $\nabla f(x)^T \cdot d \leq f(y) - f(x)$

"$\Leftarrow$". Given x, y, θ, let $z = \theta x + \bar{\theta} y$.

the first-order condition $\Rightarrow \begin{cases} f(x) \geq f(z) + \nabla f(z)^T (x-z) \\ f(y) \geq f(z) + \nabla f(z)^T (y-z). \end{cases}$

$$\Rightarrow \theta f(x) + \bar{\theta} f(y) \geq f(z) = f(\theta x + \bar{\theta} y). \qquad \square$$

Exercise. strictly convex. "$\Leftarrow$" trivial but "$\Rightarrow$"?

Second-order condition:

Suppose f is twice differentiable in an open convex set $\mathrm{dom}\, f$.

Then f is convex iff $\nabla^2 f(x) \succeq 0$ at $\forall x \in \mathrm{dom}\, f$.

Proof: "$\Rightarrow$" $\forall x_0$. $g(x) = f(x) - (f(x_0) + \nabla f(x_0)^T (x - x_0)) \geq 0$.

x_0 is a minima $\Rightarrow \nabla^2 g(x_0) = \nabla^2 f(x_0) \succeq 0$.

$$\frac{1}{2} t^2 d^T \nabla^2 f(x_0) d + o(t^2 \|d\|^2) \geq 0. \quad t \to 0 \Rightarrow \nabla^2 f(x_0) \succeq 0.$$

"$\Leftarrow$". two problems: $\nabla^2 f(x_0) \succeq 0$ not sufficient; local $\to$ global.

Taylor expansion: $f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x-x_0)^2 + \cdots + R_n$.

Lagrange remainder: $R_n = \frac{1}{n!} f^{(n)}(x_0 + \theta(x-x_0))(x-x_0)^n$ for some $\theta \in (0,1)$.

Given $x, y \in \mathrm{dom}\, f$. let $d = y - x$.

$$f(y) = f(x+d) = f(x) + \nabla f(x)^T d + \frac{1}{2} d^T \nabla^2 f(x+td)\, d \quad \text{for some } t \in (0,1)$$

$\mathrm{dom}\, f$ convex $\Rightarrow$ f defined over segment $[x, y]$.

$\nabla^2 f(x+td) \succeq 0 \Rightarrow f(y) \geq f(x) + \nabla f(x)^T d$. $\square$

strictly convex: iff? "$\Leftarrow$" trivial. but "$\Rightarrow$"? $f(x) = x^4$

Exercise: give a proof or a counterexample. $f: \mathbb{R}^n \to \mathbb{R}$. $(n \geq 3)$

Example: negative entropy $f(x) = x \log x$. $f' = \log x + 1$. $f'' = \frac{1}{x}$.

Quadratic functions: $f(x) = \frac{1}{2} x^T Q x + w^T x + b$. with symmetric Q.

$f(x)$ is convex iff $Q \succeq 0$. is strictly convex iff $Q \succ 0$.

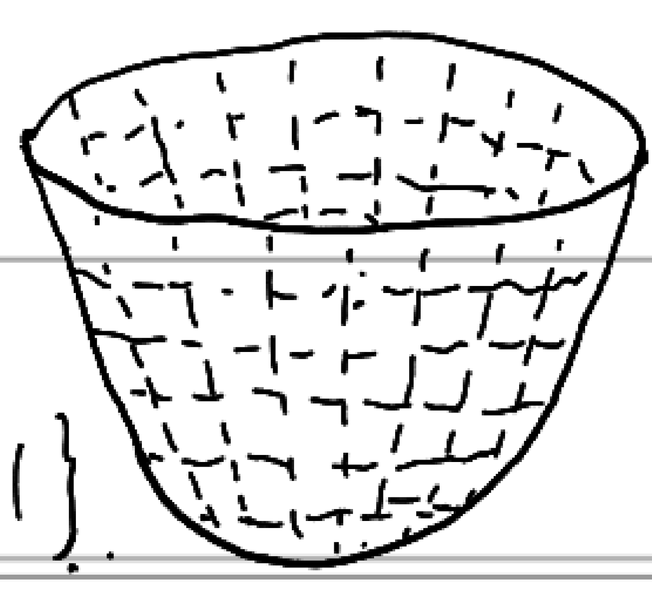
Remark: $\nabla^2 f(x) \succ 0$ not necessary in general. but necessary for quadratic.

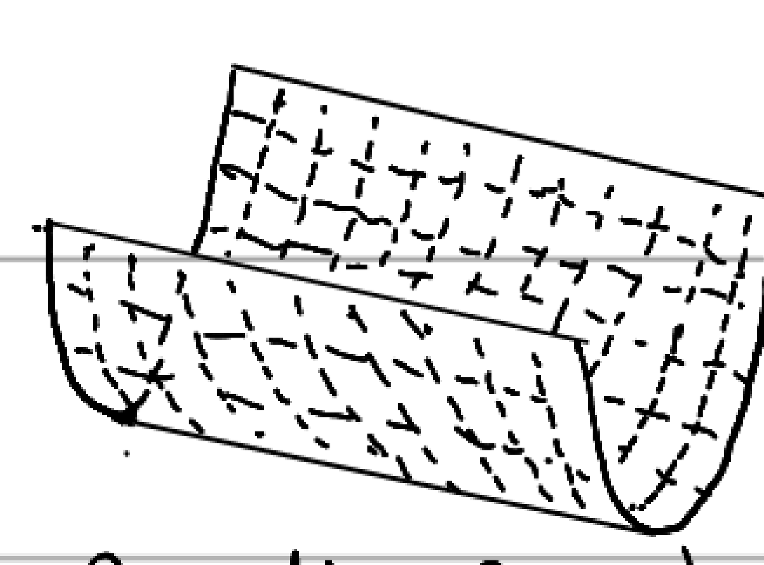
Proof: "$\Rightarrow$" part of strictly convex: note $\nabla f(x) = Qx + w$.

$$f(x+d) = f(x) + \nabla f(x)^T d + \frac{1}{2} d^T Q d. \quad \forall d \neq 0.$$

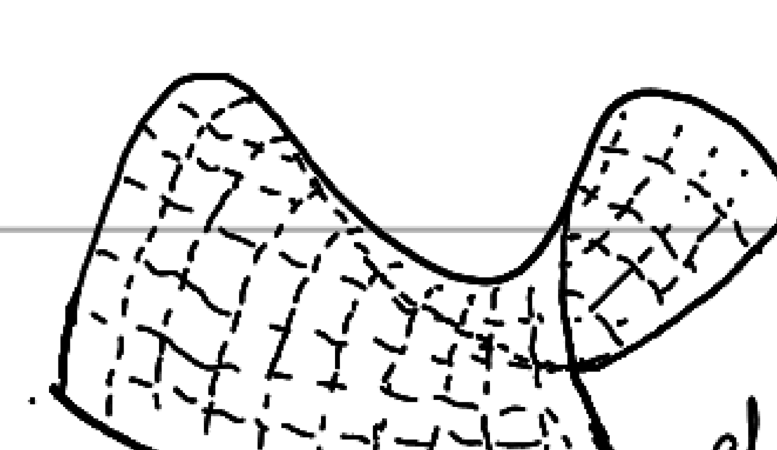
By first-order condition. $\frac{1}{2} d^T Q d > 0 \Rightarrow Q \succ 0$. $\square$

$Q = \mathrm{diag}\{1, 1\}$.





$Q = \mathrm{diag}\{0, 1\}$



$Q = \mathrm{diag}\{1, -1\}$.