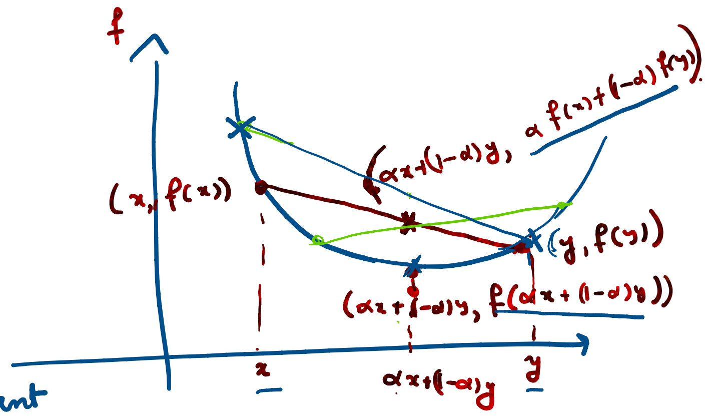
Lecture 5

Monday, 17 January 2022 1:52 PM

[Convexity]

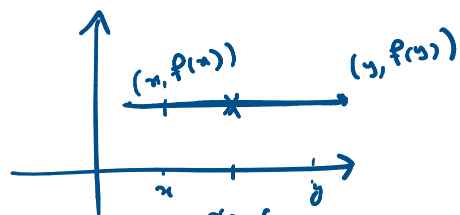


Recall: (Convexity)

$\underline{x}, \underline{y} \in \Omega$, f is convex;

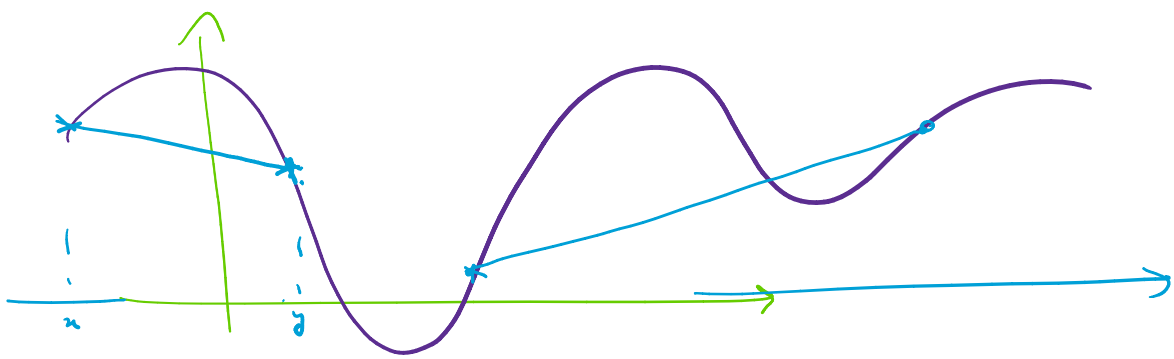
$$f(\alpha \underline{x} + (1-\alpha)\underline{y}) \le \alpha f(\underline{x}) + (1-\alpha) f(\underline{y})$$

[The function lies below its secant line (every secant line).



(not strictly convex)

Non-convex Not every secant line lies above the function; some lie below, some lie above, some lie above and below.



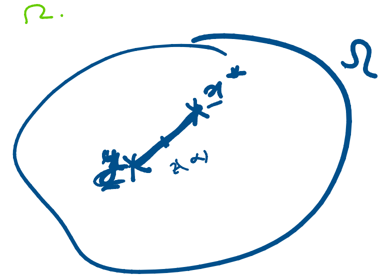
Convex Optimization Problems

Result 1: Let "Ω be a convex set. If "f is convex", then any local minimum of f in Ω is also a global minimum. [in Ω]

Ω & f are convex.
x^* is a local minimum of f in Ω
$\Leftrightarrow$ x^* is a global minimum of f in Ω.

Pf. Let f be convex, $\underline{x}^*$ be a local minimum of f in Ω.

$\Rightarrow$ $f(\underline{x}^*) \le f(\underline{x})$ for all $\underline{x} \in$ nbd of $\underline{x}^*$



Assume the contrary. (x^* is not a global minimum.) Let $\exists\ \underline{y} \in \Omega$ such that $\boxed{f(\underline{y}) < f(\underline{x}^*)}$.

that $\boxed{f(\underline{y}) < f(\underline{x}^*)}$.

Let $\underline{z}(\alpha) = \alpha \underline{x}^* + (1-\alpha)\underline{y}$ for $\alpha \in (0,1)$
$\in \Omega$ as Ω is a convex set.

f is $\boxed{\text{convex}}$ $\Rightarrow f(\underline{z}(\alpha)) \leq \alpha f(\underline{x}^*) + (1-\alpha) f(\underline{y})$
$< \alpha f(\underline{x}^*) + (1-\alpha) f(\underline{x}^*)$

$$\boxed{f(\underline{z}(\alpha)) < f(\underline{x}^*)} \quad \forall \alpha \in (0,1)$$

Hence $\exists$ points that are arbitrarily close to $\underline{x}^*$ and having a lower objective function values, thus contradicting the fact that $\underline{x}^*$ is a local minimum.

Justify: Choose $\alpha = 1 - \frac{1}{n}$; $\underline{z}(\alpha) = \left(1 - \frac{1}{n}\right)\underline{x}^* + \frac{1}{n}\underline{y}$

As $n \to \infty$ $\underline{z}(\alpha) \to \underline{x}^*$

and $f(\underline{z}(\alpha)) < f(\underline{x}^*)$,

Contradicting $\underline{x}^*$ is a local minimum.

$\therefore$ $\underline{x}^*$ is indeed a global minimum of f.

Result 2: [Let Ω be a convex set. If f is strictly convex in Ω then there exists at most one local minimum of f in Ω.] Consequently, if it exists, it is a unique global minimum of f in Ω.

Pf. Second part follows from Result 1.

We assume the contrary; let $\tilde{\underline{x}}$ and $\underline{x}^*$ be two local minima with $\tilde{\underline{x}} \neq \underline{x}^*$.

f is strictly convex $\Rightarrow$ f is convex $\Rightarrow$ $\tilde{\underline{x}}$ and $\underline{x}^*$ are global minima (from Result 1)

f is strictly convex, ...
global minima (from Result 1)

$$\Rightarrow \boxed{f(\tilde{x}) = f(x^*)}$$

Ω is convex ; $z(\alpha) = \alpha x^* + (1-\alpha)\tilde{x} \in \Omega$.

Since f is "strictly" convex,

$$f(z(\alpha)) < \alpha \underbrace{f(x^*)}_{f(\tilde{x})} + (1-\alpha) f(\tilde{x})$$

$$f(z(\alpha)) < f(\tilde{x})$$

This is a contradiction to the fact that $\tilde{x}$ is a global minimum. $\Rightarrow$ At most one local minimum exists for a strictly convex function.

Example: A quadratic form $f : \Omega \to \mathbb{R}$, $\Omega \subseteq \mathbb{R}^n$ is given by $f(x) = x^T Q x$; $x \in \mathbb{R}^n$; $\boxed{Q = Q^T}$ $Q \in \mathbb{R}^{n \times n}$ is convex on $\Omega \iff$ for all $x, y \in \Omega$, $(x-y)^T Q (x-y) \geq 0$

(Exercise) Use this result to establish $\boxed{f(x) = x_1 x_2 \text{ is convex}}$ in $\Omega = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1, x_2 \geq 0 \right\}$.

Solution Hint : $f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$

$$\iff \boxed{\alpha f(x) + (1-\alpha) f(y) - f(\alpha x + (1-\alpha) y) \geq 0}$$

Substitute $f(x) = x^T Q x$

Algebraic

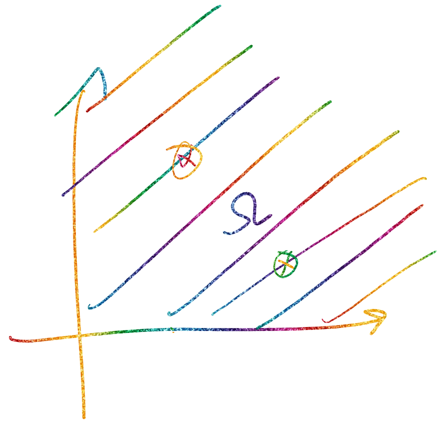
$$(x-y)^T Q (x-y) \geq 0$$

$\boxed{\text{Second part}}$

$$f(x) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (= x^T Q x)$$

$f(\underline{x}) = [\ x_2\ ]\ [\ 1/2 \quad 0\ ]\ [\ x_2\ ] \quad (= \underline{x} \ \ldots)$

$$(\underline{x}-\underline{y})^T Q (\underline{x}-\underline{y}) = \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \end{bmatrix} \begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \end{bmatrix}$$



$$= (x_1 - y_1)(x_2 - y_2) \overset{?}{\geq} 0$$

Choose $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

$\rightarrow (1-2)(2-1) < 0$.

f is <u>not convex</u>.

Verify f is not convex using defn of <u>convexity</u>.

Concave $\rightarrow$ if $(-f)$ is convex.

Can we have characterizations for <u>differentiable convex functions</u>?

$\boxed{\text{Result 3}}$ Let $f : \Omega \to \mathbb{R}$, $f \in C^{(1)}$ be defined on an $\boxed{\text{open}}$ convex set $\Omega \subset \mathbb{R}^n$. Then f is convex on Ω

$\Longleftrightarrow$ for all $\underline{x}, \underline{y} \in \Omega$, $\quad f(\underline{y}) \geq f(\underline{x}) + Df(\underline{x})(\underline{y}-\underline{x})$.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (\underline{y}-\underline{x})^T \nabla f(\underline{x})$

Proof: $\Rightarrow$ $\boxed{f : \Omega \to \mathbb{R}, \ f \in C^{(1)}, \ f \text{ is convex.}}$

$$f(\alpha \underline{y} + (1-\alpha)\underline{x}) \leq \alpha f(\underline{y}) + (1-\alpha) f(\underline{x})$$

$$\rightarrow f(\underline{x} + \alpha(\underline{y}-\underline{x})) \leq \alpha f(\underline{y}) + \underbrace{(1-\alpha) f(\underline{x})}$$

$$f(\underline{x} + \alpha(\underline{y}-\underline{x})) - f(\underline{x}) \leq \alpha [f(\underline{y}) - f(\underline{x})]$$

$$\frac{f(y) - f(x)}{y - x} \geq f'(x)$$

Convex functions lie above the tangent ...

$\Rightarrow f(\underline{x} + \alpha(\underline{y}-\underline{x})) = \ldots$

$$f(\underline{x} + \alpha(\underline{y}-\underline{x})) - f(\underline{x}) \leq \alpha\,[f(\underline{y}) - f(\underline{x})]$$

Convex functions lie above the tangents ~~at~~ the point on the curve.

$$\frac{f(\underline{x} + \alpha(\underline{y}-\underline{x})) - f(\underline{x})}{\alpha} \leq f(\underline{y}) - f(\underline{x})$$

Take limits as $\alpha \to 0$; $Df(\underline{x})(\underline{y}-\underline{x}) \leq f(\underline{y}) - f(\underline{x})$

$$\boxed{f(\underline{y}) \geq f(\underline{x}) + Df(\underline{x})(\underline{y}-\underline{x})}$$

$\Leftarrow$ $\boxed{\Omega \text{ is convex, } f \in C^{(1)},\ f(\underline{y}) \geq f(\underline{x}) + Df(\underline{x})(\underline{y}-\underline{x})}$ ✓

$\underline{u}, \underline{v}, \underline{w}$

S.T. f is convex.

Let $\underline{u}, \underline{v} \in \Omega$; Ω is convex $\Rightarrow$ $\boxed{\underline{w} = \alpha\underline{u} + (1-\alpha)\underline{v}} \in \Omega$

$\left.\begin{matrix} y = u \\ x = w \end{matrix}\right]$ $f(\underline{u}) \geq f(\underline{w}) + Df(\underline{w})(\underline{u}-\underline{w})$ $\times\ \alpha$

$\left.\begin{matrix} y = v \\ x = w \end{matrix}\right]$ $f(\underline{v}) \geq f(\underline{w}) + Df(\underline{w})(\underline{v}-\underline{w})$ $\times\ (1-\alpha)$

$\alpha \in (0,1)$

$$\alpha f(\underline{u}) + (1-\alpha) f(\underline{v}) \geq [\cancel{\alpha f(\underline{w})} + (1-\cancel{\alpha}) f(\underline{w})]$$
$$+\ Df(\underline{w})(\alpha\underline{u} - \alpha\underline{w} + (1-\alpha)(\underline{v}-\underline{w}))$$

$$\geq f(\underline{w}) + Df(\underline{w})[\alpha\underline{u} - \cancel{\alpha\underline{w}} + \underline{v} - \underline{w} - \alpha\underline{v} + \cancel{\alpha\underline{w}}]$$

$$\underbrace{\cancel{\alpha\underline{u}} + \cancel{(1-\alpha)\underline{v}} - \cancel{\alpha\underline{u}} - \cancel{(1-\alpha)\underline{v}}}_{0}$$

$\Rightarrow$ ...

$$\Rightarrow \boxed{\alpha f(u) + (1-\alpha) f(\underline{v}) \geq f(\alpha \underline{u} + (1-\alpha)\underline{v})}$$

$\Rightarrow f$ is **convex**.

$(f \in C^1)$

Convex $\Leftrightarrow f(\underline{y}) \geq f(\underline{x}) + Df(\underline{x})(\underline{y}-\underline{x})$ Check for convexity

Result 4 $\quad f \in C^2 \rightarrow$ Characterization for convexity

Let $f : \Omega \rightarrow \mathbb{R}$, $f \in C^2$ be defined on an open convex set $\Omega \subset \mathbb{R}^n$. Then f is convex on $\Omega \Leftrightarrow$ for each $\underline{x} \in \Omega$, the Hessian $F(\underline{x})$ of f at $\underline{x}$ is a positive semi-definite matrix.

$$F \geq 0 \qquad \nabla^2 f,\ D^2 f \geq 0.$$

Ex. $\quad f(\underline{x}) = x_1 x_2$

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

$$F = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \ngeq 0.$$

Revise:

Positive definite, Negative def.

Positive semi-definite, Negative semi-definite

Relevance to convex optimization problem

$\rightarrow$ global minimizer.

Relevance

FONC

Sufficient.

If $\underline{x}^*$ is a local minimizer, global minimizer.

$\Updownarrow$

$\underline{d}^T \nabla f(\underline{x}^*) \geq 0$

Proof in next class