Convex Functions

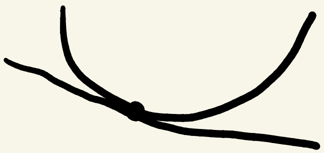
- Examples of 2nd order characterization.
- Operations preserving convexity
- Levelset and epigraph

Last time: 1st and 2nd order cond

1st order condition: $f: S \to \mathbb{R}$ (differentiable) with convex S

is convex iff

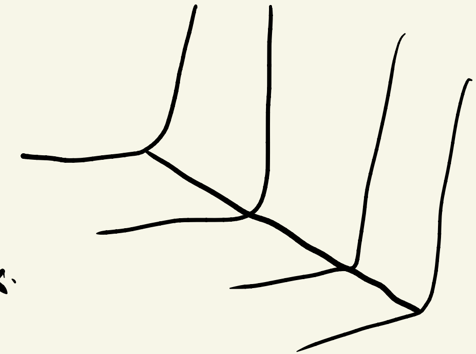
$$\boxed{f(x) + \nabla f(x)^T (y-x) \leq f(y), \quad \forall\, y, x \in S}$$



2nd order condition: For twice differentiable $f: S \to \mathbb{R}$ with convex S, f is convex iff

$$\boxed{\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in S.}$$

Example



1. $f(x) = x^\alpha$ for $x \geq 0$.

$$f''(x) = \alpha(\alpha-1)\, x^{\alpha-2}$$

$f''(x) \geq 0$ if $\alpha \leq 0$, $\alpha \geq 1$. $\Rightarrow$ convex.

$f''(x) \leq 0$ if $\alpha \in [0,1] \Rightarrow$ concave

2. Quadratic-over-line: $f(x,y) = x^2/y$ over $C = \{(x,y) \mid x \in \mathbb{R}, y > 0\}$

$$\nabla^2 f(x,y) = \frac{2}{y^3} z z^T \succeq 0, \qquad z = \begin{bmatrix} y \\ -x \end{bmatrix}.$$

Least squares objective: $f(x) = \frac{1}{2} \| Ax - b \|_2^2$

$$\nabla f(x) = A^T(Ax-b).$$

$$\nabla^2 f(x) = A^T A \succeq 0$$

always convex.

Example contd

3. Entropy function : $f(x) = -\sum_{i=1}^{n} x_i \log(x_i)$ on simplex.

$$\Delta_n : \left\{ x \in \mathbb{R}^n \mid x_i \geq 0 , \sum_{i=1}^{n} x_i = 1 \right\}$$

$$\frac{\partial f}{\partial x_i}(x) = -\log(x_i) - 1 \Rightarrow \frac{\partial^2 f}{\partial x_i^2}(x) = -\frac{1}{x_i}$$

Hessian is diagonal with $(\nabla^2 f(x))_{ii} = -\frac{1}{x_i}$

$\Rightarrow$ Entropy function is concave.

4. Log Sum Exponential: $f(x) = \log\left(\sum_{i=1}^{m} e^{a_i^T x}\right)$, $y_i = e^{x_i}$

$$\nabla^2 f(x) = \frac{1}{e^T y} \operatorname{diag}(y) - \frac{1}{(e^T y)^2} y y^T$$

$A = [a_1, \cdots a_m]$

verify $\nabla^2 f(x) \succcurlyeq 0$ for all x.

Operations that preserve convexity.

Verify convexity of function:

① Using definition : $f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)$

② 1st and 2nd order cond.

③ Operations persaving convexity:

- non-negative multiple
- sum
- composition with affine function
- composition with non-decreasing convex function
- pointwise maximum of convex functions.
- minimization.

Non-negative multiple, sum and affine composition

- f convex over a convex set $S \subseteq \mathbb{R}^n$, $\alpha \geq 0$
  $\Rightarrow$ αf is convex on S.

- $f_1, \dots, f_k$ convex over a convex set $S \subseteq \mathbb{R}^n$;
  $\Rightarrow$ $f_1 + \dots + f_k$ is convex over C.

- f convex over a convex set $S \subseteq \mathbb{R}^n$, $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$.
  $\Rightarrow$ $g(y) = f(Ay+b)$ over $D = \{y \mid Ay+b \in S\}$.

Example

$f(x) = x^2/y$

quadratic-over-linear: $h(y,t) = \|y\|^2/t$ is convex over

$C = \left\{ \begin{pmatrix} y \\ t \end{pmatrix} \in \mathbb{R}^{m+1} \;\middle|\; y \in \mathbb{R}^m,\; t > 0 \right\}.$

$$\underline{h(y,t)} = \frac{\|y\|^2}{t} = \sum_{i=1}^{m} y_i^2/t \leftarrow$$ sum of convex function.

generalized quadratic over linear:

$A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c = \mathbb{R}^n \setminus \{0\}$, $d \in \mathbb{R}$.

$g(x) = \|Ax+b\|^2/(c^T x + d)$ is convex over

$D = \{x \in \mathbb{R}^n \mid c^T x + d > 0\}.$

because $g(x) = h(Ax+b,\, c^T x + d)$ is a linear change of variable of $h \Rightarrow g$ is convex.

Composition with a non-decreasing convex function.

- $f : C \to \mathbb{R}$ is convex over convex set $C \subseteq \mathbb{R}^n$.

  $g : I \to \mathbb{R}$ is a one dimensional non-decreasing convex func.

  Assume $f(C) \subseteq I$. (image of f is contained in I).

  $$h(x) = g(f(x)) \text{ is convex over } C.$$

Example: $h(x) = e^{\|x\|^2}$ is convex because

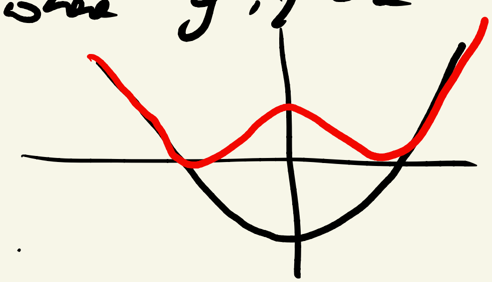
$g(t) = e^t$ is non-decreasing convex function.

and $f(x) = \|x\|^2$ is convex.

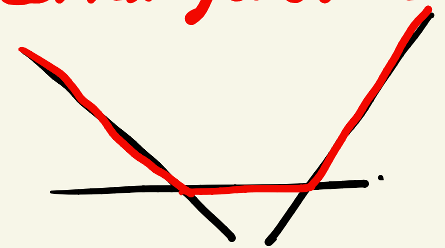
Non-example: Is there a non-convex h where g, f are convex?

$g(x) = x^2$, $f(x) = x^2 - 4$

$h(x) = (x^2 - 4)^2$ is not convex.



Pointwise maximum of convex functions.



- $f_1, \dots, f_k$ are convex functions over convex set $C \subseteq \mathbb{R}^n$

$$\Rightarrow f(x) = \max_{i \in [k]} f_i(x) \text{ is convex over } C.$$

$[k] = \{1, \dots, k\}$.

$$\max_i (a_i + b_i) \le \max_i a_i + \max_i b_i$$

Example: $f(x) = \max\{x_1, \dots, x_n\}$, $x \in \mathbb{R}^n$ is convex.

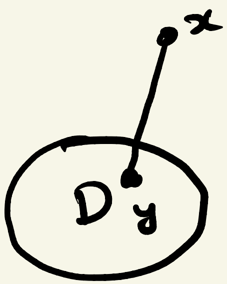
Example: $f(x) = x_{[1]} + \dots + x_{[k]}$, where $x_{[k]}$ is the k^{th} largest component of x. is convex because

$$f(x) = \max\{x_{i_1} + \dots + x_{i_k} \mid i_k \in \{1, \dots, n\} \text{ are different}\}$$

Minimization

- $f: C\times D \to \mathbb{R}$ convex defined over the set $C\times D$ where $C\subseteq \mathbb{R}^m$, $D\subseteq \mathbb{R}^n$ are convex.

$$g(x) = \min_{y\in D} f(x,y) \text{ is convex on } C.$$



Example: $C\subseteq \mathbb{R}^n$ be a convex set

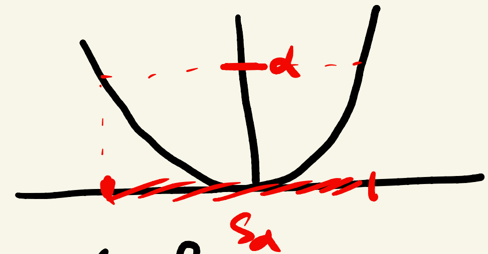
The distance function:

$$d(x,C) = \min_{y\in C} \|y-x\|$$

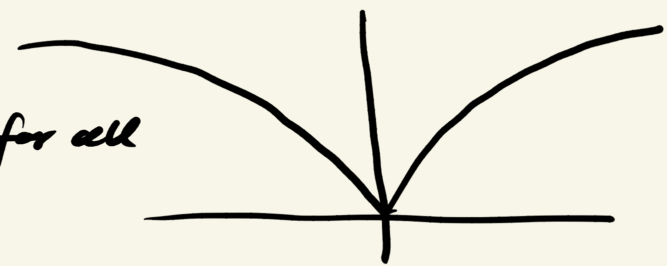
is convex over $\mathbb{R}^n$.

Level set

- $f: S \to \mathbb{R}$ defined over a set $S \subseteq \mathbb{R}^n$. α-level set of f is: $S_\alpha = \{x \in S \mid f(x) \leq \alpha\}$.

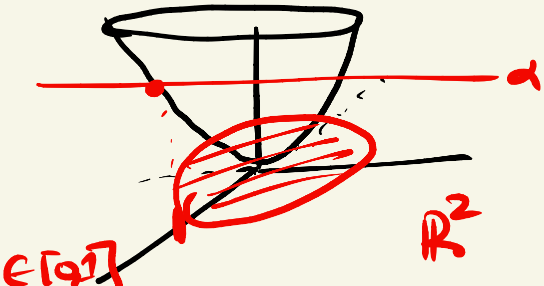


- $f: S \to \mathbb{R}$ convex defined over a convex set $S \subseteq \mathbb{R}^n$. Then every level set of f is a convex set.



- $f: S \to \mathbb{R}$ is quasi convex if for all $\alpha \in \mathbb{R}$, S_α is convex.

$S_\alpha = \{ x \in \mathbb{R}^2 \mid f(x) \le \alpha \}$



$x, y \in S_\alpha$

show $\lambda x + (1-\lambda) y \in S_\alpha$ $\quad \lambda \in [0,1]$

show $f(\lambda x + (1-\lambda) y) \le \alpha$. , $\quad f(x) \le \alpha$

$f(y) \le \alpha$