Week 2

Nicolo Michelusi

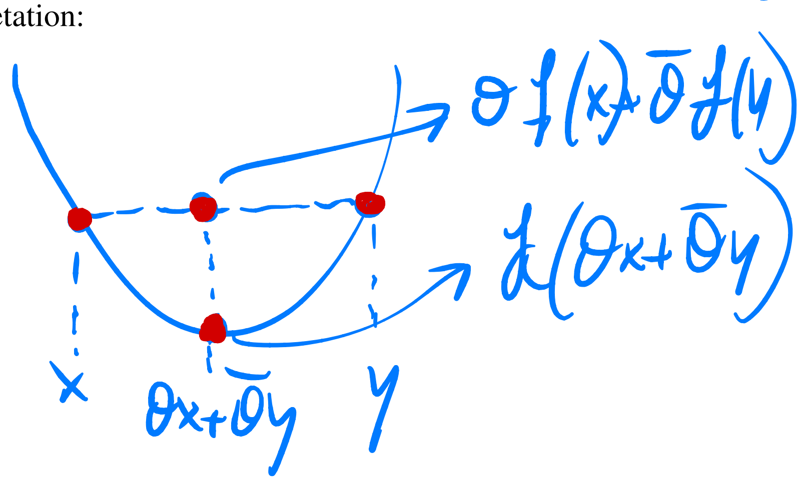
I. Convex functions

- Definition:

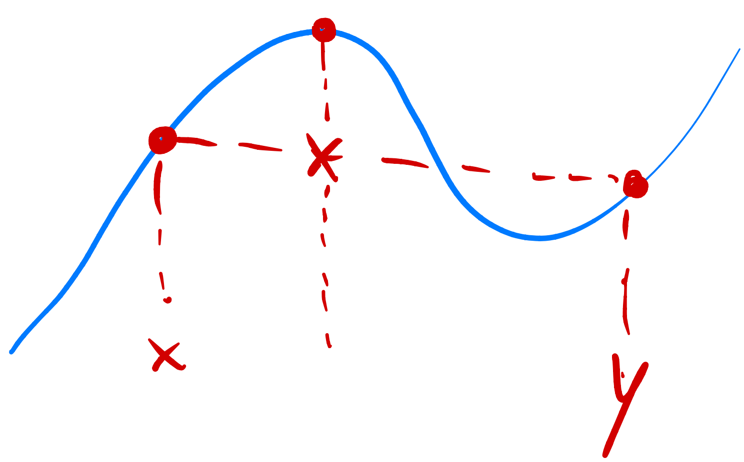
f convex $\iff$ dom f convex and

$$f(\theta x + \bar{\theta} y) \leq \theta f(x) + \bar{\theta} f(y); \ \forall x, y \in \mathrm{dom} f, \ \theta \in [0,1]$$

- Geometric interpretation:

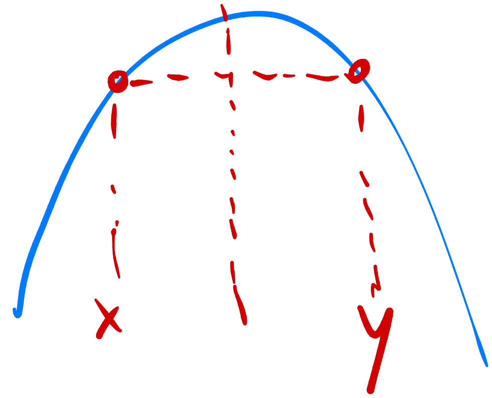


- Counter-example



- Concave functions

$$f(\theta x+\bar{\theta} y) \geq \theta f(x)+\bar{\theta} f(y)$$



- Strictly convex (concave) functions

$$f(\theta x+\bar{\theta} y) < (>)\ \theta f(x)+\bar{\theta} f(y) \quad \forall\, x,y \in \operatorname{dom} f$$

$$x \neq y,\ \theta \in (0,1)$$

(strict inequality

- Jensen's inequality: if f convex, then $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y),\ \forall \theta \in [0,1],\ \forall x, y, \in \operatorname{dom}(f)$

can be extended to infinite sums, integrals, expectations:

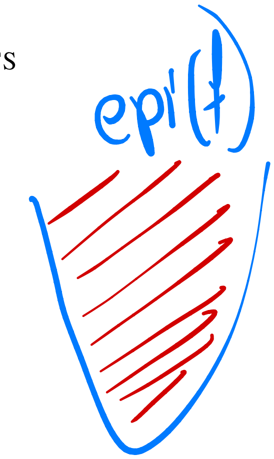
Let $p(x)$ with $\int p(x)\,dx = 1$, $p(x) \geq 0\ \forall x$

$$\Rightarrow f\left(\int p(x)\, x\, dx\right) \leq \int p(x) f(x)\, dx$$

if p is a probability distribution, $f(\mathbb{E}(x)) \leq \mathbb{E}(f(x))$

II. Relationship between convex functions and convex sets

- Epigraph of function, $\mathrm{epi}(f)$

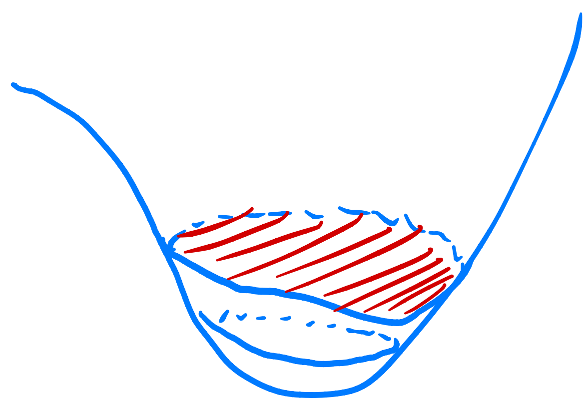


$$\mathrm{epi}(f) \equiv \{(x,t) : t \geq f(x),\ x \in \mathrm{dom}\, f\}$$

- f convex iff $\mathrm{epi}(f)$ convex

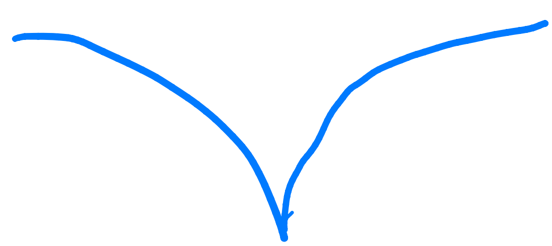
- sub-level set C_α

$$C_\alpha = \{x : f(x) \leq \alpha\}$$

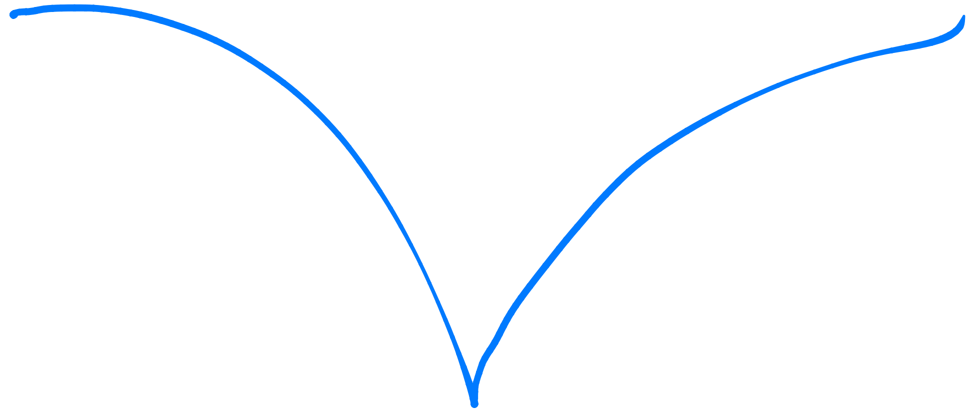


- f convex $\Rightarrow C_\alpha$ convex for all α

However, converse may not be true:



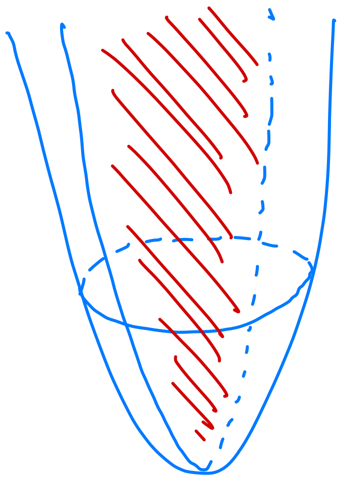
Q: is the converse true?



all sub-level sets convex but
f is NOT convex

III. Restriction to a line

- f convex iff convex when restricted to any line intersecting its domain



f convex $\iff$ dom f convex

and $g(t) \triangleq f(x+tv)$ convex

$\forall\, t$: $x+tv \in$ dom f

$\Longrightarrow$ convexity becomes a one dimensional problem

Not surprising since, to check convexity, we only need to check straight lines.

- Useful because we can reduce the problem of checking convexity of any function to a one-dimensional problem

- Example: is $f(x) = x_1 x_2$ convex?

choose $g(t) = f\left(0 + t\begin{pmatrix}1\\1\end{pmatrix}\right) = t^2$ convex

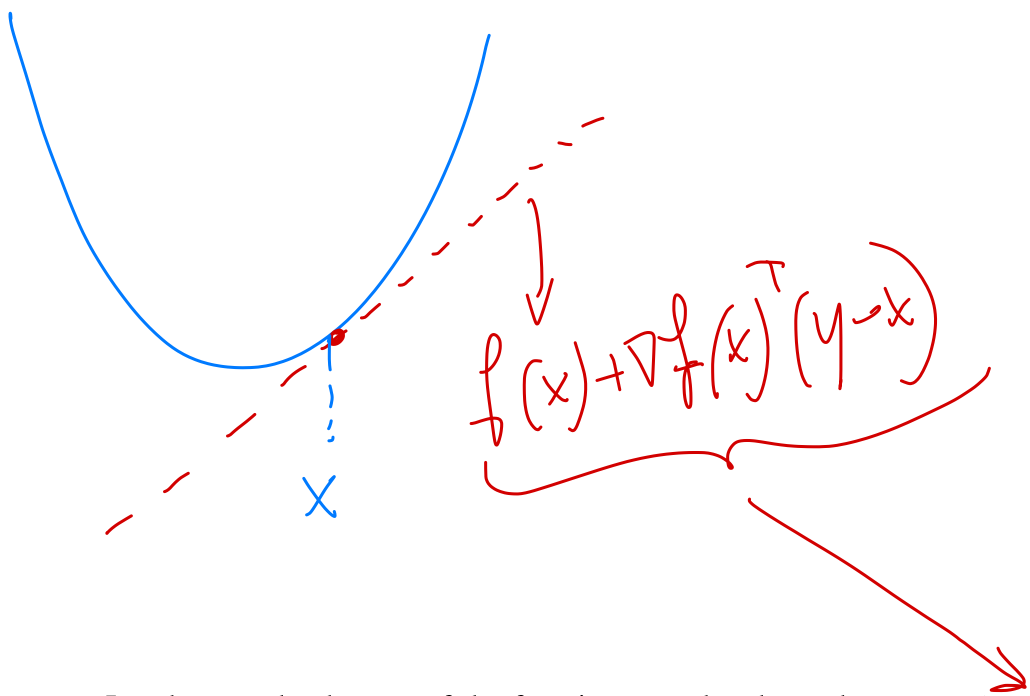
however, choose $g(t) = f\left(0 + t\begin{pmatrix}1\\-1\end{pmatrix}\right) = -t^2$ concave!

$\Longrightarrow$ f not concave nor convex!

IV. Conditions for convexity

- Often, not easy to check convexity. The following will be useful conditions:
  - First order conditions, ∇f
  - Second order conditions, $\nabla^2 f$ (Hessian)
  - Compositions that preserve convexity
- First order condition (f differentiable):

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



In other words, the rest of the function must be above the *supporting hyperplane*

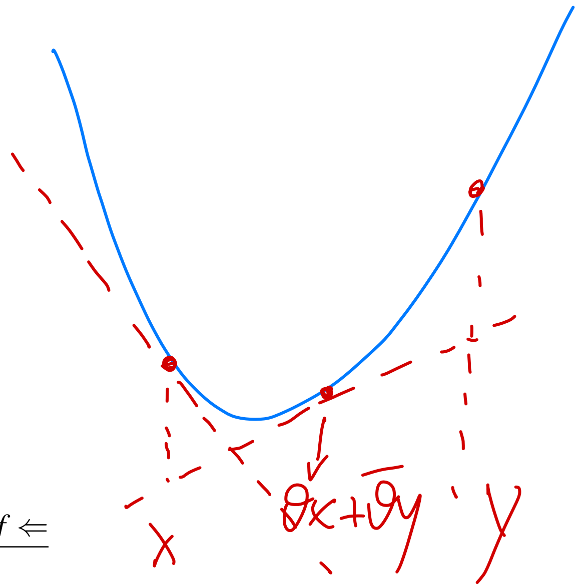
Proof of $\Rightarrow$

Assume f convex $\Rightarrow$ $g(t) = f(x+tv)$ is convex, $x \in \mathrm{dom} f$

and $g(\theta t_1) = g(\theta t_1 + \bar{\theta} 0) \leq \theta g(t_1) + \bar{\theta} g(0)$

$\Rightarrow t_1 \dfrac{g(\theta t_1) - g(0)}{\theta t_1} \leq g(t_1) - g(0)$ $\Rightarrow$ in the limit $\theta \to 0$: $t_1 g'(0) + g(0) \leq g(t_1)$

or equivalently: $f(\overbrace{x + t_1 v}^{y}) \geq f(x) + \nabla f^T(x) \cdot \underbrace{v t_1}_{y-x}$



Proof of $\Leftarrow$

Assume that $f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall y, x$

Let $x, y \in \mathrm{dom} f$ and $\theta \in [0,1]$

$$\Rightarrow \quad f(x) \geq f(\theta x + \bar{\theta} y)^T \underbrace{\left[\theta x + \bar{\theta} y - x\right]}_{\bar{\theta}(y-x)} + f(\theta x + \bar{\theta} y)$$

$$f(y) \geq f(\theta x + \bar{\theta} y)^T \underbrace{\left[\theta x + \bar{\theta} y - y\right]}_{\theta(x-y)} + f(\theta x + \bar{\theta} y)$$

$$\Rightarrow \theta f(x) + \bar{\theta} f(y) \geq f(\theta x + \bar{\theta} y) \Rightarrow \text{convex}$$

- Second order condition (f twice differentiable):

  f convex iff $\mathrm{dom}(f)$ is convex and $\nabla^2 f(x) \succeq 0, \forall x$ (strictly if $\succ 0$)

Special case, $f : \mathbb{R} \mapsto \mathbb{R}$

$$f''(x) \geq 0 \quad \forall x$$

More in general, $f : \mathbb{R}^n \mapsto \mathbb{R}$

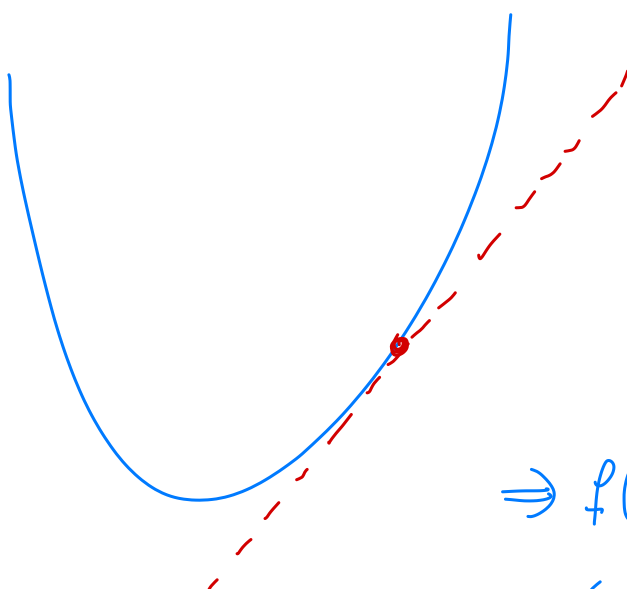
Restrict to a line: $g(t) = f(x+tv) \Rightarrow g'(t) = \nabla f(x+tv)^T \cdot v$

and $g''(t) = v^T H_f(x+tv) \cdot v \geq 0, \forall x, t, v \Rightarrow H_f(x) \succeq 0 \quad \forall x$

Hessian

Proof of $\Rightarrow$

Assume $f : \mathbb{R} \to \mathbb{R}$



f convex $\Rightarrow$ first-order condition

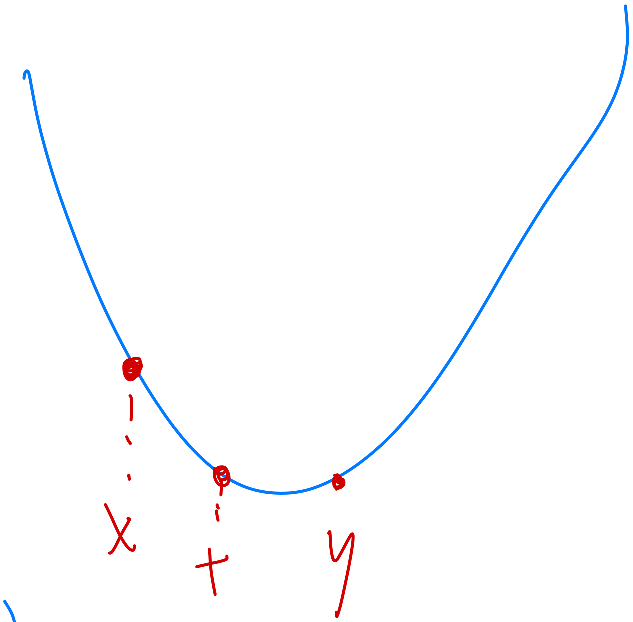
$$f(y) \geq f(x) + f'(x)(y-x)$$

$$f(x) \geq f(y) + f'(y)(x-y)$$

$$\Rightarrow f(x) + f(y) \geq f(y) + f(x) + (x-y)(f'(y) - f'(x))$$

$$\Leftrightarrow \frac{(x-y)(f'(y)-f'(x))}{(x-y)^2} \leq 0 \Leftrightarrow \frac{f'(y)-f'(x)}{y-x} \geq 0$$

$\Rightarrow$ in the limit $y \to x$: $f''(x) \geq 0$



Proof of $\Leftarrow$

Assume $f''(x) \geq 0$

$\Rightarrow$ mean-value theorem: $\exists\, t \in [x,y]$ s.t. $f(y) = f(x) + f'(x)(y-x) + \frac{1}{2}\underbrace{f''(t)(y-x)^2}_{\geq 0}$

$\Rightarrow$ $f(y) \geq f(x) + f'(x)(y-x)$

$\Rightarrow$ Result follows from first-order condition $\Rightarrow$ f convex!

V. Operations that preserve convexity

- non-negative weighted sums (also, infinite sums and integrals):

f_i convex $\Rightarrow \sum_i \alpha_i f_i(x)$ with $\alpha_i \geq 0$ is convex

$\mathrm{dom}\left(\sum_i \alpha_i f_i\right) = \bigcap_i \mathrm{dom}\, f_i$, convex

Q: what if some weights are negative? $\rightarrow$ might not be convex

- affine mapping of the argument: f convex, $g(x) = f(Ax + b)$

- Example: $\log(\sum_i e^{x_i})$ is convex

$\Rightarrow \log\left(e^{2x_1+x_2} + e^{x_1-x_2}\right) + 2\log\left(e^{x_1-x_2} + 3e^{x_2}\right)$ is convex

$3 \rightarrow e^{\ln 3}$

- Example: x^2 convex

$\Rightarrow (x_1 - x_2)^2 + 3(2x_1 + x_2)^2$ is convex

- **Q:** if $f(x)$ convex, is $f(-x)$ convex? Is $-f(x)$ convex?
- Using the above two properties, whenever we check for convexity, we can ignore the affine change of variables and non-negative weights, and focus on the simplest function form.

VI. Examples of convex functions

- To use convex optimization, it is very important to be able to quickly identify/convert to convex functions. For some problems, the key to success is to find/identify convex functions. Here is a list of commonly used convex functions with applications.
- Exponent $e^{\alpha x}$

$f'(x) = \alpha e^{\alpha x}$, $f''(x) = \alpha^2 e^{\alpha x} \geq 0 \Rightarrow$ convex

- negative log $-\ln(x)$

$f'(x) = -\frac{1}{x}$, $f''(x) = \frac{1}{x^2} > 0 \Rightarrow$ convex in $\mathrm{dom} f = \{x : x > 0\}$

- Example: Shannon capacity: $C = W \log_2(1 + P/N)$

Concave in P

Capacity increment due to increased power diminishes as P increases; *Principle of diminishing returns*, often assumed for utility functions

- Example: Power-rate: $P = N(2^{C/W} - 1)$

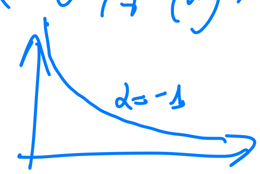
convex in C

To get the same increment in capacity, the required increase in power grows exponentially; in power systems, the cost of generation is often assumed to be convex.

- Powers, x^α, $x > 0$

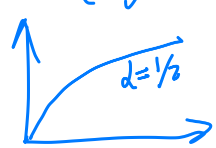
$f'(x) = \alpha x^{\alpha-1}$; $f''(x) = \alpha(\alpha-1)x^{\alpha-2}$

$\Rightarrow$ • for $\alpha < 0$, $f''(x) > 0 \Rightarrow$ convex



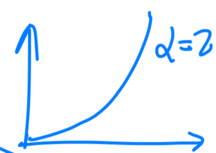
• for $\alpha = 0$, $x^\alpha = 1 \Rightarrow$ affine

• for $\alpha \in (0,1)$, $f''(x) < 0 \Rightarrow$ concave

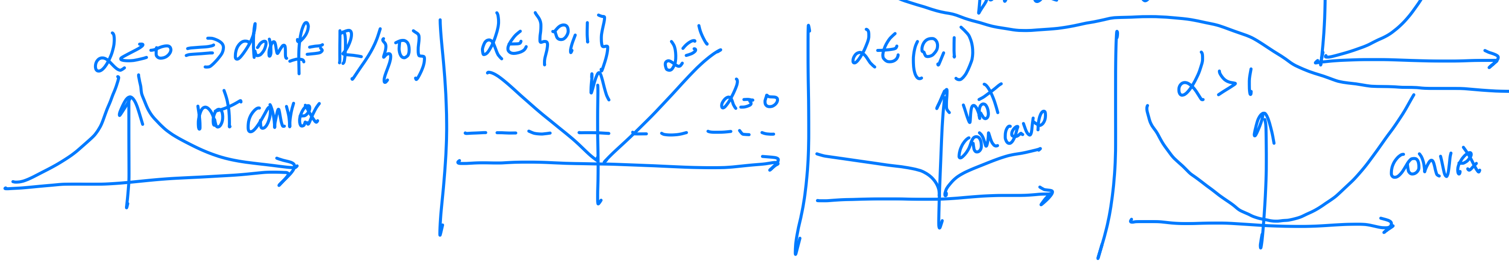


• for $\alpha = 1$, $x \Rightarrow$ linear

• for $\alpha > 1 \Rightarrow$ convex



- Power of absolute values, $|x|^p$, $p \geq 1$



- Norms $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \ p \geq 1$

use triangular inequality:

$\|\theta x + \bar{\theta} y\|_p \leq \|\theta x\|_p + \|\bar{\theta} y\|_p = \theta \|x\|_p + \bar{\theta} \|y\|_p$

- Ellipsoid $(x - x_0)^T P (x - x_0), \ P \succeq 0$ $\quad \|x\|_2^2$ is convex

Let $y = \sqrt{P}(x - x_0) \Rightarrow f(x) = \|\sqrt{P}(x - x_0)\|_2^2 \Rightarrow$ affine transformation, convex

- Example: mean squared errors, regression; given $(x_i, y_i), i = 1, \ldots, N$, find a, b such that $y = ax + b$ has the smallest error, defined as

$$\sum_i |y_i - (ax_i + b)|^p, \ p \geq 1$$

- Example: detection and likelihood functions: estimate x from $y = x + w,\ w \sim \mathcal{CN}(0, \sigma^2)$ via maximum likelihood

concave

- Example: entropy $x \log_2 x$, $x > 0$; if a source generates symbols S according to distribution $p_i = \mathbb{P}(S = s_i),\ i = 1, \ldots, N$, then entropy is

$$H(p) = -\sum_i p_i \log_2(p_i) \quad \text{where } \sum_i p_i = 1,\ p_i \geq 0\ \forall i$$

$\Rightarrow H(p)$ is concave wrt p (affine transform)

- let p_i be i-th permutation of entries of p, $i=1 \ldots n!$

$\Rightarrow H(p_i) = H(p)\ \forall i \Rightarrow H(p) = \sum_i \frac{1}{n!} H(p_i) \underset{\text{concavity}}{\leq} H\left(\sum_i \frac{1}{n!} p_i\right) = H\left(\frac{1}{n}\right) = -\log_2(n) \Rightarrow$ max entropy under uniform distribution

Entropy measures the uncertainty of the source, i.e. the amount of information generated by the source.

- Max: $\max_i x_i$ (non-differentiable!) (proof with epigraph)

$$\mathrm{epi}(f) = \{(x,t) : t \geq \max_i x_i\} \equiv \{(x,t) : t \geq x_i\ \forall i\} = \bigcap_i \{(x,t) : t \geq x_i\}$$

$\Rightarrow$ intersection of halfspaces $\Rightarrow$ epi(f) convex

$\Rightarrow \max_i x_i$ convex

- log-sum-exp: $\ln(\sum_i e^{x_i})$ (restrict to a line and use 2nd order condition)

Outline:

1) $g(t) = f(x + tv)$.

2) compute $g''(t)$

3) Use Schwarz inequality $|x^T z|^2 \le \|x\|^2 \cdot \|z\|^2$ for proper choice of x, z

- Example: log-moment generating function: $f(s) = \ln \mathbb{E}[e^{sX}]$ (restrict to a line and use 2nd order condition)

$f(s) = \ln \sum_i p_i e^{sX_i} = \ln \sum_i e^{\ln p_i + sX_i} \Rightarrow$ convex

- Example: high-SNR approximation of Shannon capacity

$$C_i = W \log_2 \left(1 + \frac{P_i}{\sum_{j \neq i} P_j + N} \right) \ge W \log_2 \left(\frac{P_i}{\sum_{j \neq i} P_j + N} \right)$$

(not convex nor concave; use change of variables $P_i = e^{x_i}$)

$$W \cdot \log_2 \left(\frac{P_i}{\sum_{j \neq i} e^{x_j} + N} \right) = \underbrace{W \log_2 (P)}_{\text{concave}} \underbrace{- W \underbrace{\log_2 \left(\sum_{j \neq i} e^{x_j} + e^{\ln N} \right)}_{\text{convex}}}_{\text{concave}}$$

- Low-SNR approximation,

$$C_i = W\log_2\left(1+\frac{P_i}{\sum_{j\neq i}P_j+N}\right) \leq \frac{W}{\ln(2)}\frac{P_i}{\sum_{j\neq i}P_j+N}$$

(not convex nor concave; take $y_i = \ln(C_i)$ and change of variables $P_i = e^{x_i}$)

$$\ln(C_i) \leq \ln\left(\frac{W}{\ln 2}\right) + \ln\left(\frac{e^{x_i}}{\sum_{j\neq i} e^{x_j} + N}\right)$$

$$= \ln\left(\frac{W}{\ln 2}\right) + x_i - \ln\left(\sum_{j\neq i} e^{x_j} + e^{\ln N}\right) \Rightarrow \text{concave}$$

- Geometric mean: $\left(\prod_i x_i\right)^{1/N}$ is concave (see Boyd page 74)
- Quadratic over linear: $x^2/y,\ y>0$ (check 2nd order condition)

VII. More operations that preserve convexity

- In addition to non-negative weighted sums and affine mappings of arguments
- Pointwise maximum and supremum of convex functions: $\max_i f_i(x)$ or $\sup_y f(x,y)$ (use epigraph)

$$\text{epi}(f) = \left\{(x,t) : t \geq \max_i f_i(x)\right\} = \left\{(x,t) : t \geq f_i(x)\ \forall i\right\}$$

$$= \bigcap_i \left\{(x,t) : t \geq f_i(x)\right\} = \bigcap_i \text{epi}(f_i) \Rightarrow \text{convex}$$

- Pointwise minimum and infimum of concave functions: $\min_i f_i(x)$ or $\inf_y f(x,y)$
- Examples:

  $\max_i x_i$

  $\max_i a_i^T x_i + b_i$ } max of affine $\Rightarrow$ convex

  $\min_i x_i$

  $\min_i a_i^T x_i + b_i$ } min of affine $\Rightarrow$ concave
- Sum of the largest r components of $x \in \mathbb{R}^n$

Let $p_i \in \{0,1\}^n$ be all permutations such that

$1^T p_i = r \Rightarrow f(x) = \max_i p_i^T \cdot x$

$\Rightarrow$ max of affine $\Rightarrow$ convex

- Distance to the farthest point: let $x \in \mathbb{R}^n$, $\mathcal{C} \subseteq \mathbb{R}^n$, $f(x) = \sup_{y \in \mathcal{C}} \underbrace{\|x-y\|_2}_{\text{convex}}$

$\Rightarrow f(x)$ is convex

- What about the infimum of convex functions? $f(x) = \inf_{y\in\mathcal{C}} g(x,y)$? Not always!

Special case: $g(x,y)$ convex wrt (x,y) and C convex

$\Rightarrow f$ is convex

Proof: let $y_x = \arg\min_{y\in C} g(x,y) \Rightarrow f(\theta x+\bar\theta z) = g(\theta x+\bar\theta z, y_{\theta x+\bar\theta z}) \le g(\theta x+\bar\theta z, y)\ \forall y\in C$

and in particular for $y = \theta y_x + \bar\theta y_z$: $f(\theta x+\bar\theta z) \le g(\theta x+\bar\theta z, \theta y_x+\bar\theta y_z) \le$

$\overset{\text{(convexity)}}{\le} \theta g(x,y_x) + \bar\theta g(z,y_z) = \theta f(x) + \bar\theta f(z)$

- Example: distance to a set, $\text{dist}(x,\underbrace{\mathcal{C}}_{\text{convex}}) = \inf_{y\in\mathcal{C}} \|x-y\|_2$

convex $\|x\|_2$ is convex $\Rightarrow \|x-y\|_2$ is convex

$\Rightarrow \text{dist}(x,C)$ is convex!

- Perspective of a function $g(x,t) = tf(x/t)$, $t>0$

If f convex $\Rightarrow g$ convex

Proof: $\text{epi}(g) = \{(x,t,z) : z \ge g(x,t)\} \equiv \{(x,t,z) : \frac{z}{t} \ge f(\frac{x}{t})\}$

$\equiv \{(x,t,z) : (\frac{x}{t}, \frac{z}{t}) \in \underbrace{\text{epi}(f)}_{\text{convex}}\}$

$\Rightarrow \text{epi}(g)$ is the inverse mapping of a convex set

$\Rightarrow \text{epi}(g)$ is convex $\Rightarrow g(x,t)$ is convex

- Example: find channel capacity one a user is served a portion $t < 1$ of the time; is it convex?

$$g(P,t) = t\ln\left(1+\frac{P}{t}\right) \text{ and } f(P) = -\ln(1+P) \text{ (convex)}$$

$$\Rightarrow g(P,t) = -t f\left(\frac{P}{t}\right) \quad \underline{\text{concave}}$$

- Example: $g(x) = (c^T x + d) f\left(\frac{Ax+b}{c^T x+d}\right)$

obtained by combining perspective function and affine transformation $\Rightarrow$ convex

- Compositions: $f(x) = h(g(x))$

$$f'(x) = h'(g(x)) \cdot g'(x)$$

$$f''(x) = h''(g(x))[g'(x)]^2 + h'(g(x))\, g''(x) \geq 0$$

Many special cases:

1) h convex non-decreasing, g convex

2) h convex non increasing; g concave

- Examples:

1) g convex, what about $e^{g(x)}$?

2) g convex, what about $g(x)^2$?

3) g convex and non-negative, what about $g(x)^2$?

4) g concave, what about $\ln g(x)$?

5) g concave, what about $1/g(x)$?

1) $h(y) = e^y$ convex increasing $\Rightarrow$ convex
2) $h(y) = y^2$ convex but not monotone $\Rightarrow$ NO!
3) $h(y) = y^2$ convex and increasing for $y \geq 0 \Rightarrow$ convex
4) $h(y) = \ln y$ concave increasing $\Rightarrow$ concave
5) $h(y) = 1/y$, but dom h not convex $\Rightarrow$ NO!

- Vector composition: $f(x) = h(g(x))$ with $h : \mathbb{R}^k \mapsto \mathbb{R}$ and $g : \mathbb{R} \mapsto \mathbb{R}^k$, what are the conditions? (assume differentiable)

$$f'(x) = \nabla h^T(g(x))\, g'(x)$$

$$f''(x) = g'(x)^T \cdot H_h(g(x)) \cdot g'(x) + \nabla h^T(g(x))\, g''(x)$$

$\Rightarrow$ many cases:

1) $H_h \succeq 0$, h non-decreasing in all components, g convex in all components

2) $H_h \succeq 0$, h non-increasing in all components, g concave in all components