MATH464 – Lecture 4 (01/19/2023)

Today: * convex functions
* piecewise linear (PL) convex functions
* PL convex functions in LP

We first finish the formulation instance from last lecture.

3. Currency Exchange LP (continued..)

$$\max \sum_{i=1}^{n-1} x_{in} \qquad \text{(total \# units of currency } n)$$

$$\text{s.t.} \quad \sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij} \le u_i, \quad i=1,\dots,n-1 \qquad \text{(max exchange of currency } i)$$

$$\sum_{j=2}^{n} x_{1j} \le b \qquad \text{(amt of currency 1 at start)}$$

same!

$$\sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij} \le \underbrace{\sum_{\substack{k=1\\ k\neq i}}^{n-1} r_{ki}x_{ki}}_{\text{max \# currency } i \text{ we can get by converting any other currency to currency } i.} \quad i=2,\dots,n-1 \qquad \left(\begin{array}{l}\text{max amt of currency } i\\ \text{that can be exchanged}\end{array}\right)$$

$$x_{ij} \ge 0 \quad \forall i,j \qquad \text{(non-negativity).}$$

The limits of b and u_1 are **Both** applied to the total amount of Currency 1 that can be exchanged. The two bounds are independently imposed. The third set of constraints is the corresponding bound for currencies 2 to $n-1$ (similar to the bound b on currency 1).

Piecewise linear (PL) Convex Functions

→ a class of otherwise nonlinear functions that could still be modeled using linear functions!

We first define convex functions.

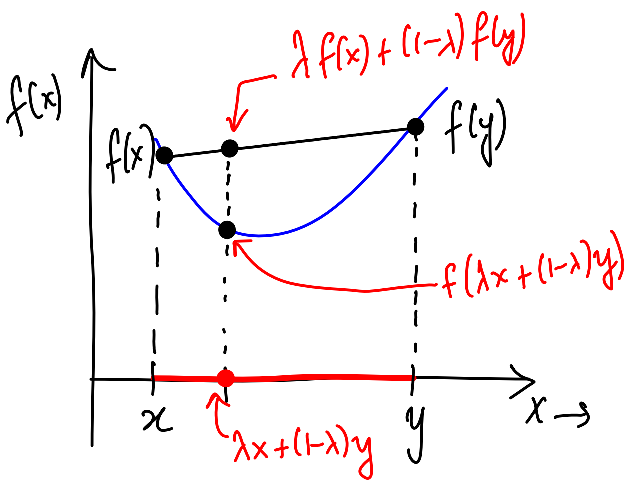
Def A function $f: \mathbb{R}^n \to \mathbb{R}$ is **convex** if $\forall \bar{x}, \bar{y} \in \mathbb{R}^n$ and $\forall\ \lambda \in [0,1]$

$$f(\lambda \bar{x} + (1-\lambda)\bar{y}) \leq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y}) \qquad (*)$$

$\geq$ → f is concave

$\lambda \bar{x} + (1-\lambda)\bar{y}$ for $\lambda \in [0,1]$ is any point in the line segment joining $\bar{x}$ and $\bar{y}$.

Illustration in 1D:



The graph of $f(\bar{x})$ lies at or below ($\leq$) the line segment connecting $f(\bar{x})$ and $f(\bar{y})$.

If $f(\bar{x})$ is linear, we get $=$ in $(*)$, i.e., a linear function is both convex and concave.

in place of $\leq$

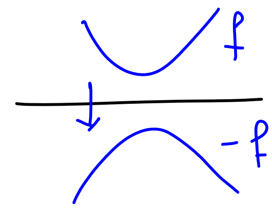
Def A **linear function** $f: \mathbb{R}^n \to \mathbb{R}$ is specified as

$$f(\bar{x}) = \bar{a}^T \bar{x} = \sum_{i=1}^{n} a_i x_i, \quad a_i \in \mathbb{R}\ \forall i \ (\text{or } \bar{a} \in \mathbb{R}^n).$$

Def An **affine function** $f : \mathbb{R}^n \to \mathbb{R}$ is specified as

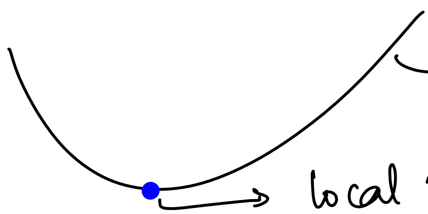
$$f(\bar{x}) = a_0 + \bar{a}^T \bar{x} = a_0 + \sum_{i=1}^{n} a_i x_i \,, \quad a_i \in \mathbb{R} \ \forall i.$$

Note f is convex iff $-f$ is concave. Multiplying $(*)$ by -1 flips the $\leq$ to $\geq$, giving that $-f(\cdot)$ satisfies the condition for a function being concave. → Intuitively, $-f$ is obtained by the mirror reflection of f across the x-axis.



Why study convex functions?

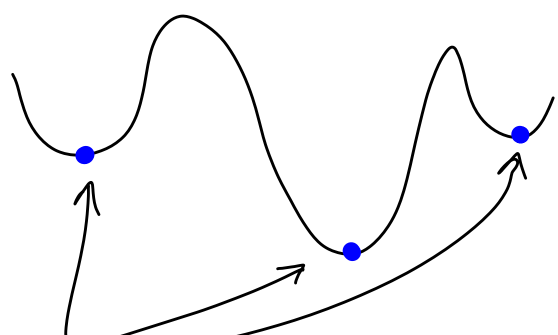
Convex functions have unique global minima!



convex function

local minimum is global minimum

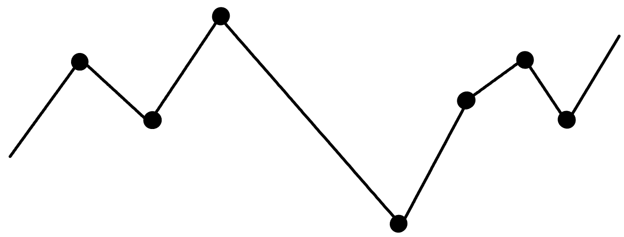
neither convex nor concave



local minima

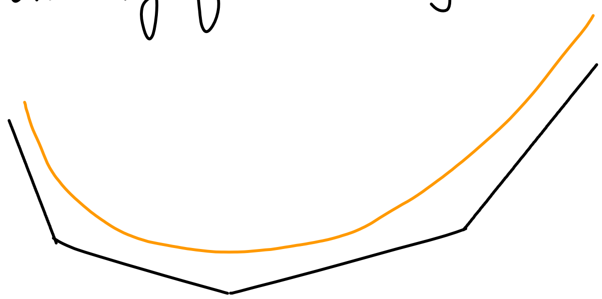
But then again, we are studying *linear programming*, and all the curves drawn so far look non-linear!?

We want to study piecewise linear (PL) functions, which consist of several pieces, each of which is linear.



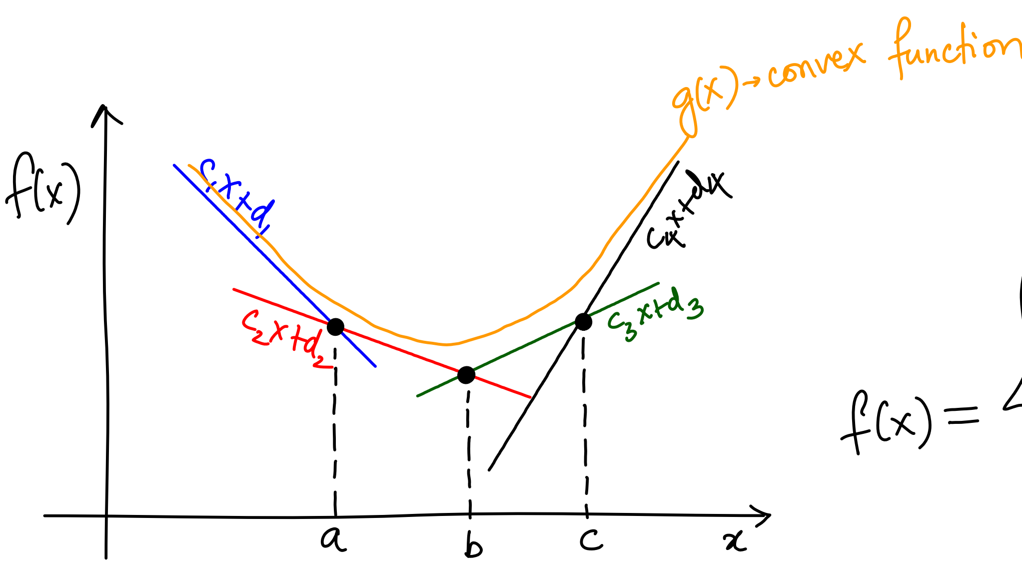
$\longrightarrow$ not convex or concave!

We will study piecewise linear convex functions, which have the following form. They could also be used to approximate nonlinear convex functions, as shown in orange.



We can explicitly define a PL convex function in 1D by specifying the expression for each linear piece as follows:

Note: the general such function could have m pieces, not necessarily 4 all the time!



$$f(x) = \begin{cases} c_1x+d_1, & x \le a \\ c_2x+d_2, & a \le x \le b \\ c_3x+d_3, & b \le x \le c \\ c_4x+d_4, & c \le x \end{cases}$$

A more compact way of stating $f(x)$ here is as follows:

$$f(x) = \max_{i=1,\dots,4} \{c_i x + d_i\}$$

replacing max with min will not give a concave function here! But,

$$h(x) = \min_{i=1,\dots,4} \{-(c_i x + d_i)\}$$ is a PL concave function.

We could extend this specification to higher dimensions to define PL convex functions in general as follows:

Def $f: \mathbb{R}^n \to \mathbb{R}$ is a piecewise-linear (PL) convex function if

$$f(\bar{x}) = \max_{i=1,\dots,m} (\bar{c}_i^T \bar{x} + d_i)$$ for $\bar{c}_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$ $\forall i$.

We can prove that $f(\bar{x})$ is indeed convex. Notice that each piece, i.e., $\bar{c}_i^T \bar{x} + d_i$ is affine, and hence convex. We prove a more general result — the max of a set of convex functions is convex.

Recall, $f(\bar{x})$ is convex if $f(\lambda \bar{x} + (1-\lambda)\bar{y}) \leq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$ $\forall \bar{x}, \bar{y} \in \mathbb{R}^n$, $\lambda \in [0,1]$.

This will be the first proof-type problem for us. We will see several such proofs — but they should be easier than many proofs in analysis, for instance!

Theorem 1.1 (BT-ILO) Let $f_1, \dots, f_m : \mathbb{R}^n \to \mathbb{R}$ be convex functions.
Then $f(\bar{x}) = \max_{i=1,\dots,m} f_i(\bar{x})$ is convex.

Proof Let $\bar{x}, \bar{y} \in \mathbb{R}^n$ and $\lambda \in [0,1]$.
We need to show $f(\lambda\bar{x} + (1-\lambda)\bar{y}) \leq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$.

$$f(\lambda\bar{x} + (1-\lambda)\bar{y}) = \max_{i=1,\dots,m} f_i\left(\lambda\bar{x} + (1-\lambda)\bar{y}\right)$$

$$\leq \max_{i=1,\dots,m} \left(\lambda f_i(\bar{x}) + (1-\lambda) f_i(\bar{y})\right), \text{ as each } f_i \text{ is convex}$$

$$\leq \max_{i=1,\dots,m} \lambda f_i(\bar{x}) + \max_{i=1,\dots,m} (1-\lambda) f_i(\bar{y})$$

$\rightarrow$ as $\max_i (a_i + b_i) \leq \max_i (a_i) + \max_i (b_i)$

$$= \lambda \max_{i=1,\dots,m} f_i(\bar{x}) + (1-\lambda) \max_{i=1,\dots,m} f_i(\bar{y})$$

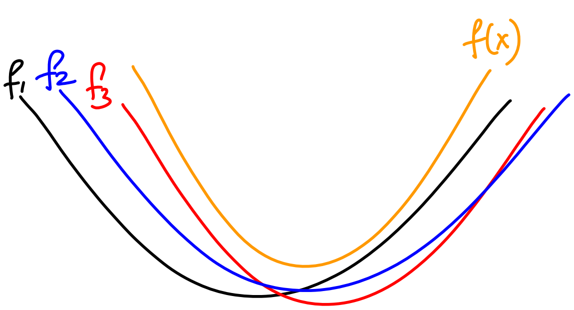
as both $\lambda, 1-\lambda \geq 0$ $\downarrow$ this is a crucial requirement!

$$= \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$$

i.e., $f(\lambda\bar{x} + (1-\lambda)\bar{y}) \leq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$.

QED or $\square$

The result works for nonlinear convex functions as well:



QED: quod erat demonstrandum: Latin for "that which has to be demonstrated". Specifies the end of proof. Some journals & books use $\blacksquare$ or $\square$ instead.

PL convex functions in LP

Consider the following generalization of an LP:

$$\min \left\{ \max_{i=1,\dots,m} \left(\bar{c}_i^T \bar{x} + d_i \right) \right\} \longrightarrow \text{PL convex function}$$

$$\text{s.t.} \quad A\bar{x} \geq \bar{b} \longrightarrow \text{sign restrictions are included here.}$$

We can model this problem as an LP:

$$\boxed{\begin{aligned} \min \quad & z \\ \text{s.t.} \quad & z \geq \bar{c}_i^T \bar{x} + d_i, \quad i=1,\dots,m \\ & A\bar{x} \geq \bar{b} \end{aligned}} \text{ — (LP)}$$

z is a single urs variable. We specify that z is $\geq$ each piece. Then when we minimize z, we get the desired model!

Note that (the optimal) z could indeed be <0 here, depending on how the m pieces sit!

What about PL convex functions in constraints?

If we have $f(\bar{x}) \leq h$ where $f(\bar{x}) = \max_{i=1,\dots,m} (\bar{f}_i^T \bar{x} + g_i)$,

then we can replace this constraint by m linear inequalities:

$$\bar{f}_i^T \bar{x} + g_i \leq h, \quad \text{for } i=1,\dots,m.$$

Once again, the logic is similar. To specify that the largest of m functions is $\leq h$, we instead specify each of the m pieces is $\leq h$. Then the largest of them will also be $\leq h$.