446 Section 05

TA: Varun Ananth

Plans for today!

1. This
2. Reminders
3. Subgradients
4. Convexity
5. Midterm review

Reminders

- HW2 due next Wednesday 2/12, 11:59 PM
  - Are you keeping track of late days? Use them!
- <u>Midterm</u> TOMORROW.

K-fold CV + LASSO

Code on the website for parts 1 & 2

- Check it out to see how k-fold cross validation is done in numpy
- Also a manual implementation of LASSO!

Subgradients

Why subgradients?

You can have convex functions that are not differentiable everywhere

- GD still useful to find minima
- But we may need to find subgradients

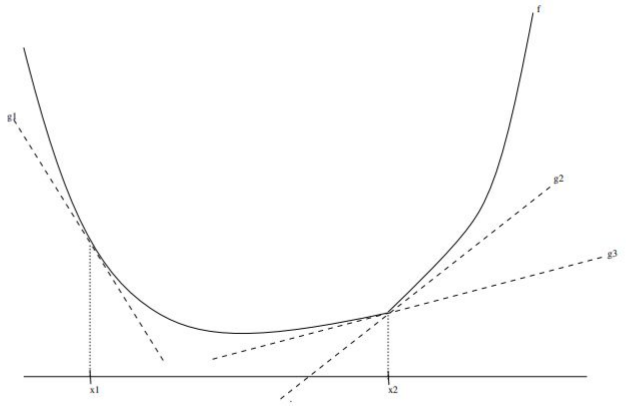
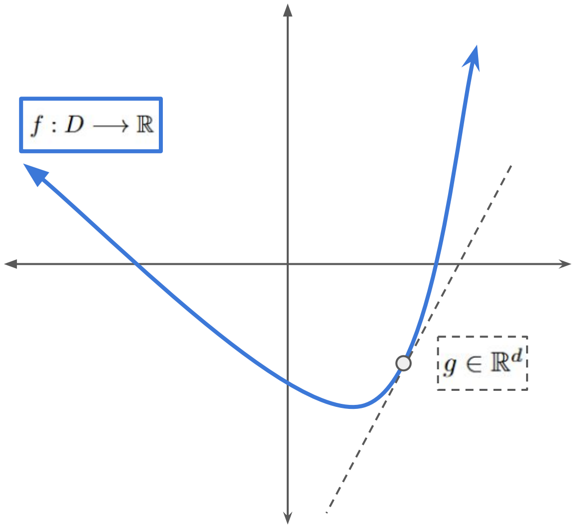


Figure 1: At x_1, the convex function f is differentiable, and g_1 (which is the derivative of f at x_1) is the unique subgradient at x_1. At the point x_2, f is not differentiable. At this point, f has many subgradients: two subgradients, g_2 and g_3, are shown.

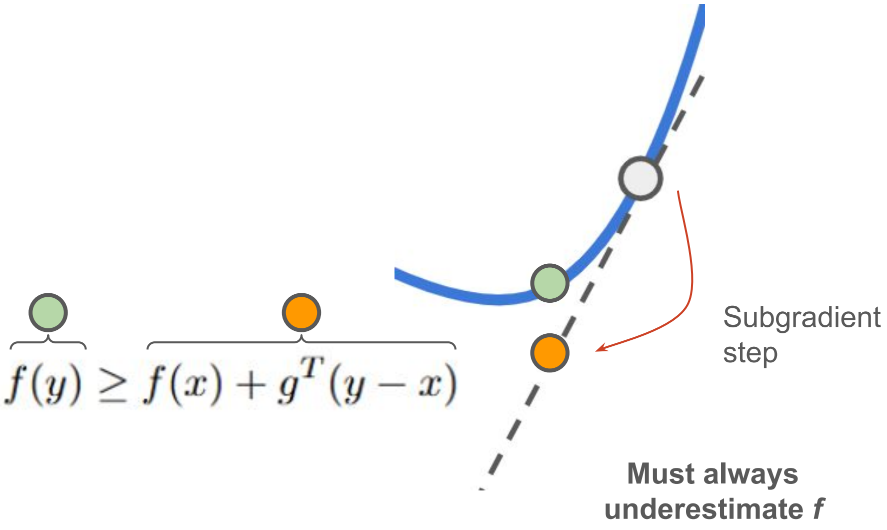
Subgradients visualized

Definition 1 (subgradients). *A vector $g \in \mathbb{R}^d$ is a subgradient of a convex function $f : D \longrightarrow \mathbb{R}$ at $x \in D \subseteq \mathbb{R}^d$ if*

$$f(y) \geq f(x) + g^T(y - x) \quad \textit{for all } y \in D.$$



Zoom in

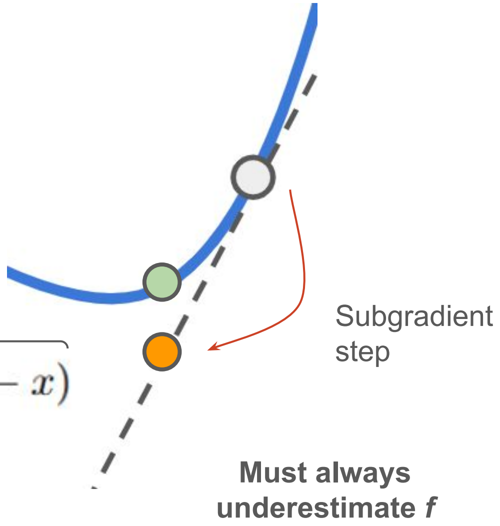


Subgradients visualized

Definition 1 (subgradients). *A vector $g \in \mathbb{R}^d$ is a subgradient of a convex function $f : D \longrightarrow \mathbb{R}$ at $x \in D \subseteq \mathbb{R}^d$ if*

$$f(y) \geq f(x) + g^T(y - x) \quad \textit{for all } y \in D.$$

$$\overbrace{f(y)} \geq \overbrace{f(x) + g^T(y - x)}$$



<u>Note:</u>

- You can have many subgradients at a point *x*
- Subgradients can exist even at non-differentiable points *x*

<u>Cool fact:</u>

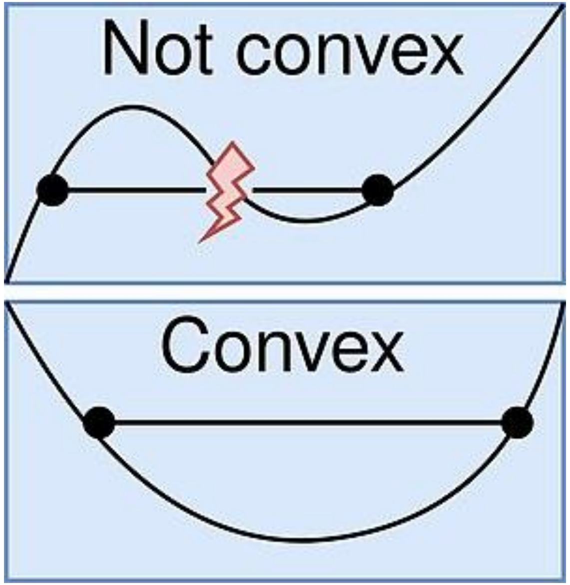
- If *f* **is** differentiable at *x*, then the gradient of *f* at *x* is also a subgradient of *f* at *x*

Convexity

Convexity in functions

Definition 2 (convex functions). *A function $f : \mathbb{R}^d \to \mathbb{R}$ is **convex** on a set A if for all $x, y \in A$ and $\lambda \in [0,1]$:*

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



The function between x and y

Must be less than or equal to

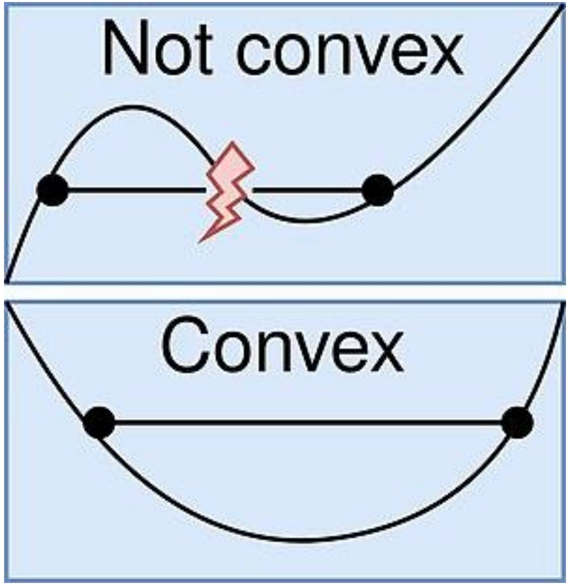
A straight line between x and y

Note: The sum of convex functions is convex

Convexity in functions

Definition 2 (convex functions). *A function* $f : \mathbb{R}^d \to \mathbb{R}$ *is* **convex** *on a set* A *if for all* $x, y \in A$ *and* $\lambda \in [0,1]$*:*

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



Guarantees that any local minimum we find will be as low as the global minimum

If you perform GD with a small step size on a convex loss function, you **will** reach the best possible performance!

Solutions

Midterm Review!

Questions/Chat Time!