Math 381 - Fall 2022

Jay Newby

University of Alberta

Week 12

Last Time

1. Motivating examples of optimization

Today

1. Continuous optimization problems
2. Convex sets
3. Convex functions
4. Critical points and classification of critical points

Continuous optimization problems

Continuous optimization problem

$$\min_x f(x) \quad \text{subject to} \quad g(x) = 0, \quad h(x) \leq 0.$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, and $h : \mathbb{R}^n \to \mathbb{R}^p$.

We can say something about uniqueness for convex problems

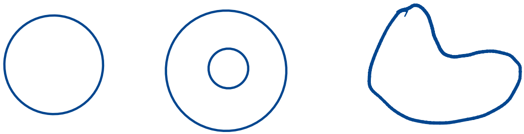
Definition: convex set

The set $S \subseteq \mathbb{R}^n$ is convex if

$$\{\alpha x + (1-\alpha)y : 0 \leq \alpha \leq 1\} \subseteq S$$

for all $x, y \in S$.

Examples:



We can say something about uniqueness for convex problems

Definition: Convex function

A function $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$, with S a convex set, is a convex function if

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y),$$

for all $\alpha \in [0,1]$ and all $x, y \in S$.

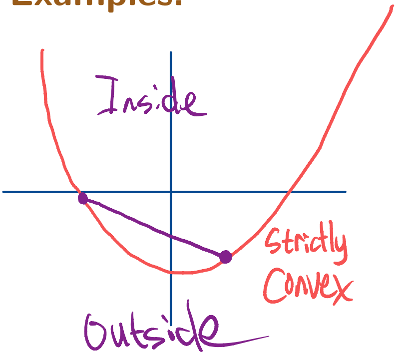
Definition: Strictly convex function

A convex function is strictly convex if

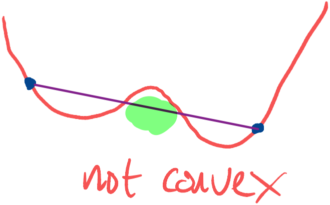
$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y),$$

for all $\alpha \in (0,1)$ and all $x, y \in S$.

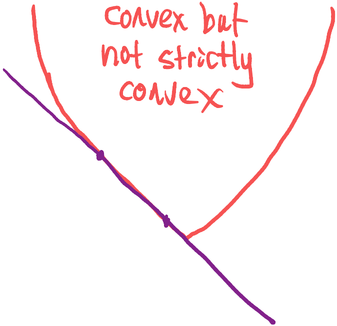
Examples:



Inside
Outside
strictly
Convex



not convex



convex but
not strictly
convex

Uniqueness of the global minimum for strictly convex function on convex sets

- Sublevel sets of a convex function are convex
- Any local minimum of a convex function f on a convex set S is a global minimum on S
- Any local minimum of a strictly convex function f on a convex set S is a unique global minimum on S

Unconstrained optimality conditions

How can we know if a given point $x \in S$ is a local minimum of f?

Gradient

Definition: gradient

$f : \mathbb{R}^n \to \mathbb{R}$ differentiable

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Critical points

Definition: Critical point

A point $\hat{x}$ such that $\nabla f(\hat{x}) = 0$ is a critical point.

Example: 1D problem

Suppose we have a smooth function $f(x)$. Wlog let $\hat{x} = 0$ so that $f'(0) = 0$. Expand around $\hat{x}$ with

$$f(x) \sim f(0) + \cancelto{0}{x f'(0)} + \frac{1}{2} x^2 f''(0) + \cdots$$

$$f(x) \sim f(0) + \frac{1}{2}x^2 f''(0)$$

We want to show that $f(x) \geq f(\hat{x})$ for all $x \in (-\delta, \delta)$ for some $\delta > 0$.

Example: $\overset{n}{2}$D problem

Suppose we have a smooth function $f(x)$. Wlog let $\hat{x} = 0$ so that ~~$f'(0) = 0$~~. Expand around $\hat{x}$ with

$\nabla f(0) = 0$

$$f(x) \sim f(0) + \frac{1}{2} x^T H(0) x$$

Definition: Hessian matrix

Given smooth $f : \mathbb{R}^n \to \mathbb{R}$. The $n \times n$ symmetric matrix with elements

$$H_{ij}(x) = \frac{\partial^2}{\partial x_i \partial x_j} f(x).$$

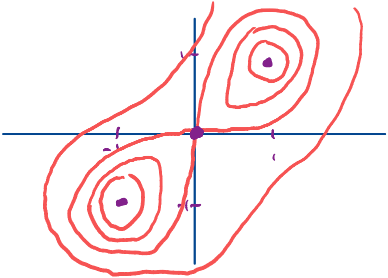
A sufficient condition for local minimum is $x^T H(\hat{x}) x > 0$ for all x in some neighborhood of $\hat{x}$.

Classifying critical points

- $H(\hat{x})$ positive definite (all eigenvalues have positive real part) implies that $\hat{x}$ is a local min
- $H(\hat{x})$ negative definite (all eigenvalues have negative real part) implies that $\hat{x}$ is a local max
- $H(\hat{x})$ has eigenvalues with mixed sign (but all nonzero) implies that $\hat{x}$ is a saddle point
- $H(\hat{x})$ has some zero eigenvalues: various abnormal situations can occur

Examples:

$f(x,y) = x^4 - 4xy + y^4$



$$\nabla f(x,y) = \begin{bmatrix} 4x^3 - 4y \\ -4x + 4y^3 \end{bmatrix}$$

$$H(x,y) = \begin{bmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{bmatrix}$$

Critical point $(1,1)$ or $(-1,-1)$

$$H(1,1) = \begin{bmatrix} 12 & -4 \\ -4 & 12 \end{bmatrix}$$

$$\lambda^2 - Tr(H)\lambda + |H| = 0$$

$$\lambda^2 - 24\lambda + 12^2 - 16 \qquad \lambda = 8, 16$$

critical $(0,0)$

$$H(0,0) = \begin{bmatrix} 0 & -4 \\ -4 & 0 \end{bmatrix}$$

$$\lambda^2 - 16 = 0 \qquad \lambda = \pm 4$$