CME 252: Introduction to Optimization

AJ Friend
ICME, Stanford University

Introduction

Outline

Course Goals

- get students using CVXPY to solve real (convex) optimization problems as quickly as possible
- teach just enough theory to do so
- focus on **modeling**: **what** you want to solve instead of **how** to solve it
- convex optimization as a starting point to consider more general optimization problems

Audience

- anyone interested in using optimization in their work
- no background in optimization is necessary
- do need to be comfortable with
  - linear algebra
  - basic programming (any language)

Logistics

- course website at ajfriend.github.io/cme252
  - announcements
  - homework
  - lecture materials
- schedule
  - MW 3:30-4:50 in McCullough 115
  - 8 sessions
  - office hours TBD (Thursday?)
- Piazza `http://piazza.com/stanford/fall2015/cme252`

Pyhton/CVXPY

- we'll use Python and CVXPY to solve optimization problems
- need a working Python distribution with
  - numpy
  - scipy
  - matplotlib
  - CVXPY
- example code given in Jupyter (IPython) notebooks
- HW0 to get you set up (out before Wednesday)
- additional help Wednesday, office hours, Piazza

Homework

- 1 assignment per week, due on Friday
- submit python script solving a few optimization problems
- HW1 released by this Friday, due next Friday
- HW0 (not graded)
    - Python/CVXPY setup
    - Jupyter (IPython) notebooks
    - homework submission

Topic Outline (tentative)

- types of optimization
- convex sets and functions
- convex optimization and modeling
- regression, least-squares, curve-fitting and variants
- in-depth examples from various fields (SVM, logistic regression,...)
- basics of gradient descent
- non-convex problems (and convex approaches)

vary topics based on student interest

Won't Cover

- other optimization classes (global optimization, integer programming,. . .)
- within convex optimization:
  - optimality conditions (maybe a *little*)
  - duality and Lagrange multipliers
  - in-depth algorithms
  - fancy convex sets and functions (SDP, perspective functions,. . .)

Questions?

Optimization Overview

Outline

Optimization

Optimization given an objective function, finding a best (or good enough) choice among a set of (possibly constrained) options

Mathematical optimization

Mathematical optimization problem has form

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in C \end{array}$$

- $x \in \mathbf{R}^n$ is **decision variable** (to be found)
- C is a set describing **acceptable** points
- f is objective function (choose best acceptable point)
- problem data are hidden inside f and C
- variations: different ways to represent problem, maximize a utility function, satisfaction (feasibility), optimal trade off, and more

The good news

Everything is an optimization problem

- *choose parameters* in model to fit data (minimize misfit or error on observed data)
- *optimize actions* (minimize cost or maximize profit)
- *allocate resources* over time (minimize cost, power; maximize utility)
- *engineering design* (trade off weight, power, speed, performance, lifetime)

The bad news

In full generality, optimization problems can be quite difficult

- generally NP-hard
- local vs. global minimizers
- heuristics required, hand-tuning, luck, babysitting

The bad news

In full generality, optimization problems can be quite difficult

- generally NP-hard
- local vs. global minimizers
- heuristics required, hand-tuning, luck, babysitting

But...

- we can do a lot by restricting to convex models
- local minimizers are global
- we have good computational tools
  - modeling languages (CVX, CVXPY, JuMP, AMPL, GAMS) to write problems down
  - good solver software to obtain solutions

Optimization in one variable

$$\text{minimize} \quad f(x) \in C : \mathbf{R} \to \mathbf{R}$$

Optimization in one variable

$$\text{minimize} \quad f(x) \in C : \mathbf{R} \to \mathbf{R}$$

- x is a real variable

Optimization in one variable

$$\text{minimize} \quad f(x) \in C : \mathbf{R} \to \mathbf{R}$$

- x is a real variable
- $f(x)$ is the objective function, which returns a single real number

Optimization in one variable

$$\text{minimize} \quad f(x) \in C : \mathbf{R} \to \mathbf{R}$$

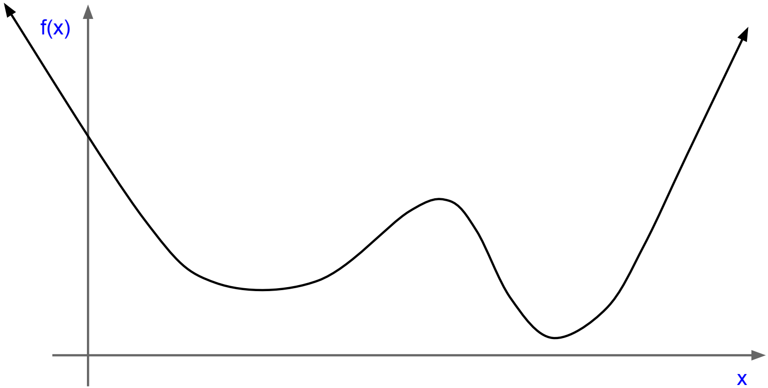
- x is a real variable
- $f(x)$ is the objective function, which returns a single real number
- Local optimization: look for a point x^* such that $f(x^*) \leq f(x)$ for all points x near x^*

Optimization in one variable

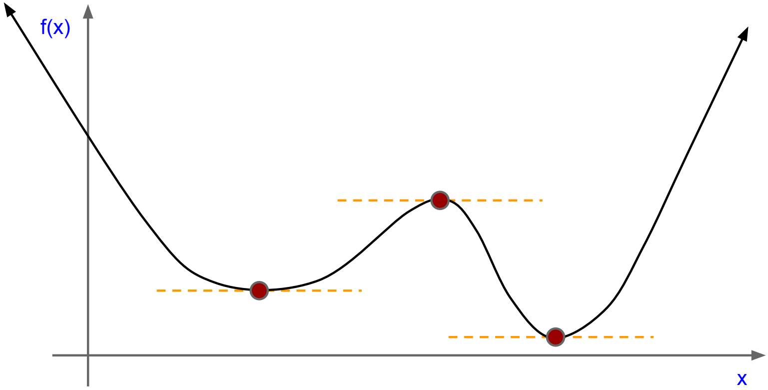
$$\text{minimize} \quad f(x) \in C : \mathbf{R} \to \mathbf{R}$$

- x is a real variable
- $f(x)$ is the objective function, which returns a single real number
- Local optimization: look for a point x^* such that $f(x^*) \leq f(x)$ for all points x near x^*
- Global optimization: look for a point x^* such that $f(x^*) \leq f(x)$ for all points x in domain of interest

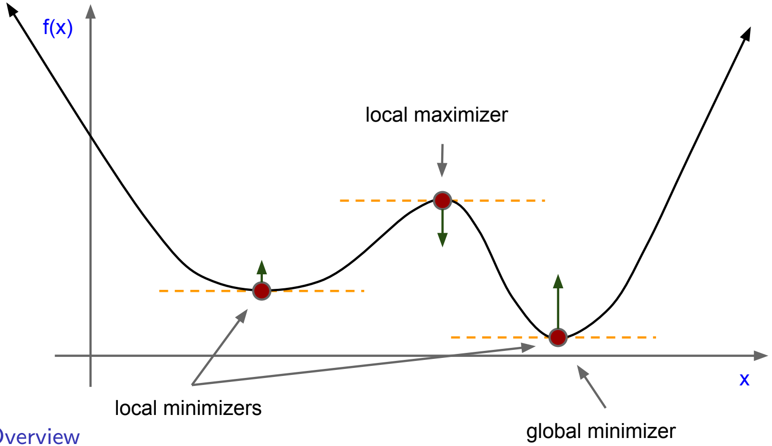
Optimization in one variable: example objective function



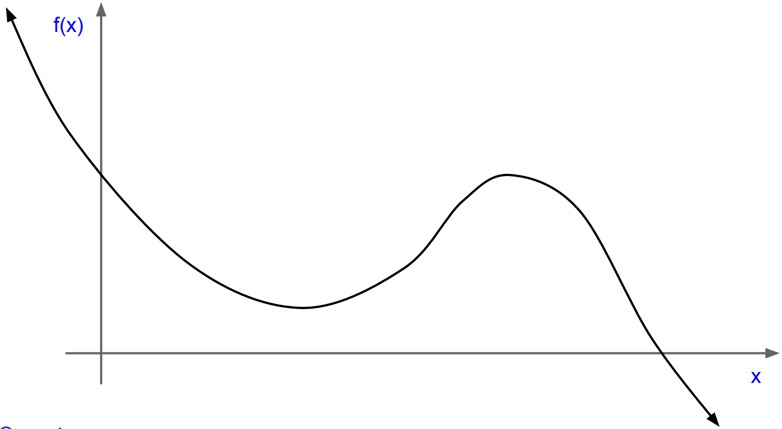
Optimization in one variable: critical points, $f'(x) = 0$



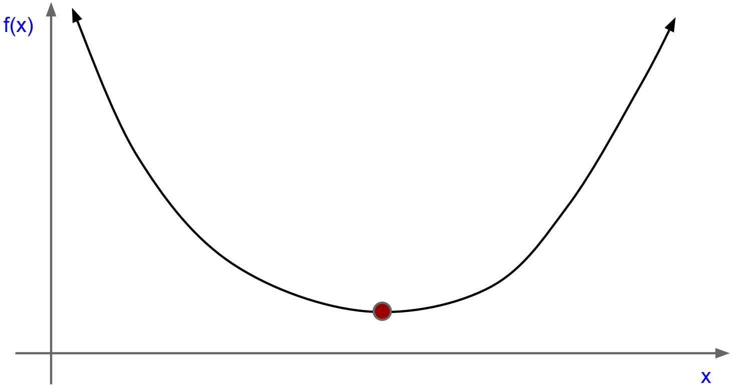
Optimization in one variable: local optima, $f''(x) = ?$



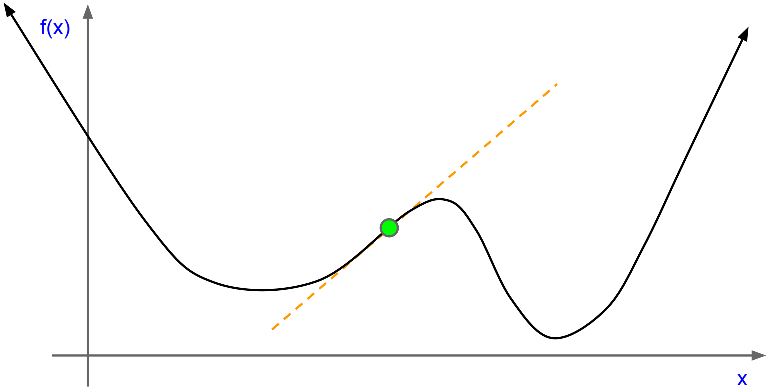
Optimization in one variable: unbounded below



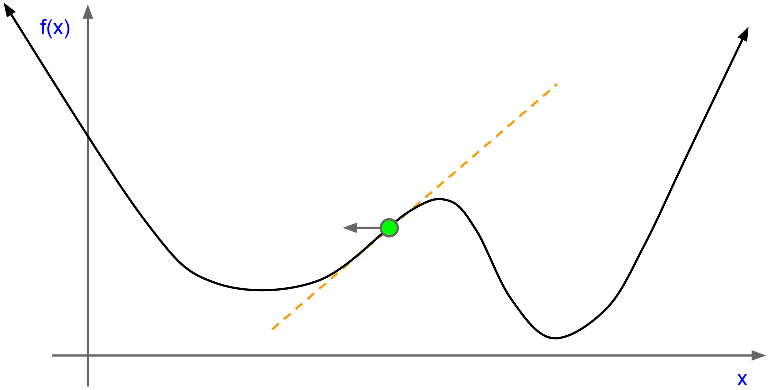
Optimization in one variable: convex objective



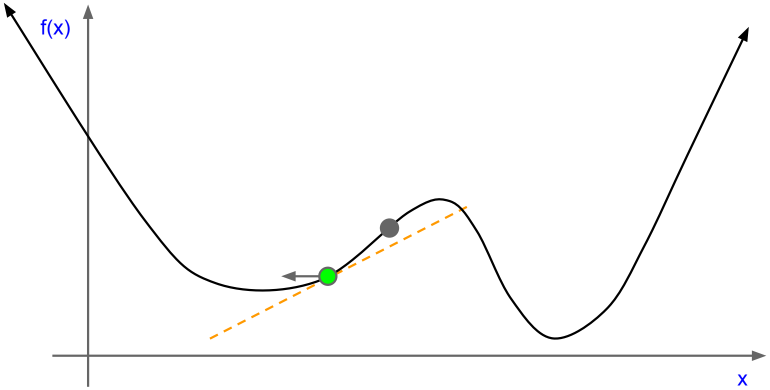
Optimization in one variable: algorithm basics



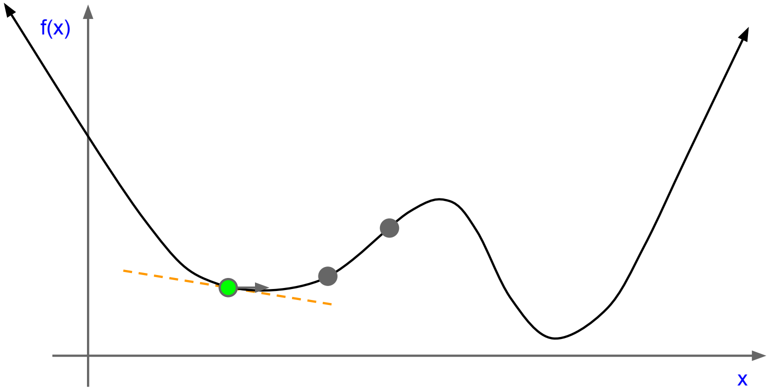
Optimization in one variable: algorithm basics



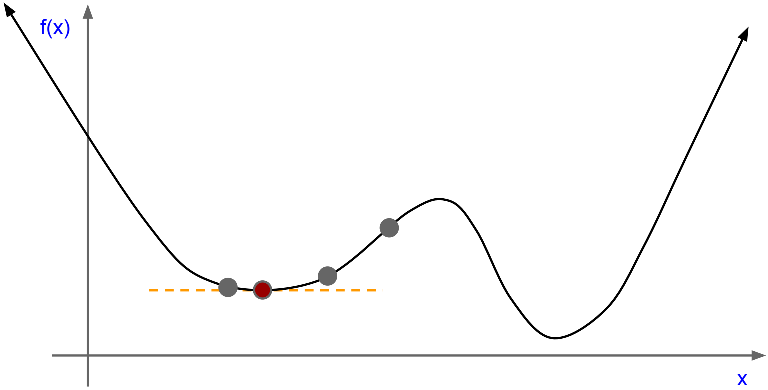
Optimization in one variable: algorithm basics



Optimization in one variable: algorithm basics



Optimization in one variable: algorithm basics



Optimization in one variable: algorithm basics

Optimization in one variable: algorithm basics

- Start with an initial guess x_0

Optimization in one variable: algorithm basics

- Start with an initial guess x_0

- Goal: generate sequence that converges to solution

$$x_0, x_1, x_2, x_3, \cdots \to x^*$$

Optimization in one variable: algorithm basics

- Start with an initial guess x_0
- Goal: generate sequence that converges to solution

$$x_0, x_1, x_2, x_3, \cdots \to x^*$$

- Notation for sequence and convergence: $\{x_k\} \to x^*$

Optimization in one variable: algorithm basics

- Start with an initial guess x_0
- Goal: generate sequence that converges to solution

$$x_0, x_1, x_2, x_3, \cdots \to x^*$$

- Notation for sequence and convergence: $\{x_k\} \to x^*$
- Key algorithm property: ***descent condition***

$$f(x_{k+1}) < f(x_k)$$

- the simplest algorithm works! (for local minimizers)

Optimization in one variable: algorithm basics

- Start with an initial guess x_0
- Goal: generate sequence that converges to solution

$$x_0, x_1, x_2, x_3, \cdots \to x^*$$

- Notation for sequence and convergence: $\{x_k\} \to x^*$
- Key algorithm property: ***descent condition***

$$f(x_{k+1}) < f(x_k)$$

- the simplest algorithm works! (for local minimizers)

Why Convexity?

Convex optimization:

- local minimizers are global
- useful theory of convexity
- effective algorithms and available software that provide: global solutions, polynomial complexity, and algorithms that scale
- convenient **language** to discuss problems
- expressive: **lots of applications**

Non-convex optimization:

- in general, no guarantee that minimizers are global
- solvers often use convex optimization as a sub-routine
- modeling tools are more difficult to use
- solution process may require expert guidance or tweaking

Classes of mathematical optimization problems

- There are many (!) classes of mathematical optimization problems (and associated solvers)
- The primary problem features are:
  - Variable type: {continuous, discrete}
  - Domain: {unconstrained, constrained}
  - Model: {convex, non-convex}

Variables

Continuous variables take real numbers as values (within limits):

$$x \in \mathbf{R}$$

Discrete variables typically take integers as values:

$$x \in \{0, 1, 2, 3, \ldots\}$$

Boolean or binary variables are a special case of this:

$$x \in \{0, 1\}$$

Problems with discrete variables are generally harder than those with continuous variables.

Variables

Example of continuous variables:

- maximum likelihood estimate of the mean
- parameters in a linear model
- asset allocation in mean-variance portfolio optimization
- position in a standard coordinate system
- speed (in, say, a model to minimize fuel consumption)

Example of discrete variables:

- A $\{0, 1\}$ selector for facility location. Say variable $x_{ij} = 1$ if and only if resource i is placed in location j and zero otherwise.
- An integer representing the number of people allocated to a task. It would be unwise and perhaps illegal to allocate half a person.

Domain

Unconstrained mathematical optimization problems only require an objective function:

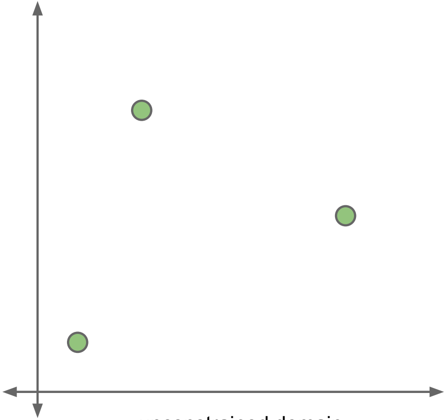
$$\text{minimize} \quad f_0(x)$$

Constrained optimization problems limit the domain with equations or inequalities:

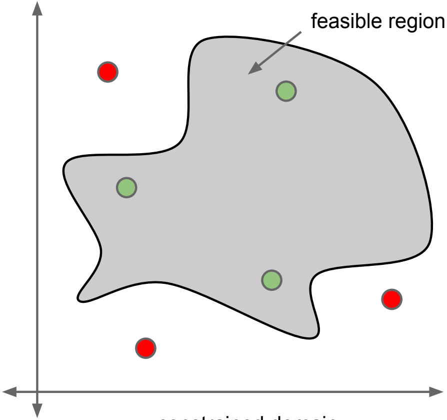
$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \end{array}$$

Solvers for constrained optimization typically rely on a solver for unconstrained optimization.

Domain



unconstrained domain
(all points considered acceptable)



constrained domain
(only green points acceptable)

This class

This class will primarily cover the **bold** topics:

- Variable type: {**continuous**, discrete}
- Domain: {**unconstrained**, **constrained**}
- Model: {**convex**, non-convex}

Summary

- Mathematical optimization is an important and useful tool in science, engineering, and industry
- The optimization community has produced a large set of good tools to solve problems
  - there are a mix of open-source and commercial packages
- Art: mapping your problem into a mathematical model that can be attacked using an existing tool
- Next class: jump into theory and examples of convex sets and functions

CVXPY/Jupyter Example

Outline

CVXPY/Jupyter Example