1 Foundation

1.1 Introduction

⋆ (mathematical) optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le b_i, i = 1, \dots, m \end{array}$$

- $x = (x1, \dots, x_n)$: optimization variables
- $f_0 : \mathbf{R}^n \to \mathbf{R}$: objective function
- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \dots, m$: constraint functions

optimal solution x^* has smallest value of f_0 among all vectors that satisfy the constaints

EX: portfolio optimization

- variables: amounts invested in different assets
- constraints: budget, max./min. investment per asset, minimum return
- object: overall risk or return variance

EX: data fitting

- variables: model parameters
- constraints: prior information, parameter limits
- objective: measure of misfit or prediction error

⋆ Least-squares

$$\text{minimize} \quad ||Ax - b||_2^2$$

solving least-squares problems

- analytical solution: $x^* = (A^T A)^{-1} A^T b$
- reliable and efficient algorithms and software
- computation time proportional to $n^2 k$ ($A \in \mathbf{R}^{k \times n}$); less if structured
- a mature technology

using least-squares

- least-squares problems are easy to recognize
- a few standard techniques increase flexibility (e.g., including weights, adding regularization terms)

⋆ Linear programming

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i,\ i = 1, \ldots, m \end{array}$$

solving linear programs

- no analytical formula for solution
- reliable and efficient algorithms and software
- computation time proportional to $n^2 m$ if $m \geq n$; less with structure
- a mature technology

using linear programming

- not as easy to recognize as least-squares problems
- a few standard tricks used to convert problems into linear programs (e.g., problems involving l_1- or $l_\infty-$norms, piecewise-linear functions)

⋆ Convex optimization problem

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

- objective and constraint functions are convex

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y)$$

  if $\alpha + \beta = 1, \alpha \geq 0, \beta \geq 0$

- includes least-squares problems and linear programs as special cases

solving convex optimization problems

- no analytical solution
- reliable and efficient algorithms
- computation time (roughly) proportional to $\max\{n^3, n^2 m, F\}$, where F is cost of evaluating f_i's and their first and second derivatives
- almost a technology

using convex optimization

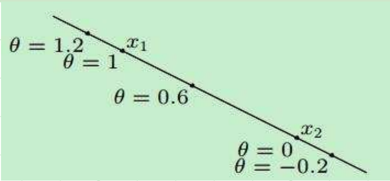
- often difficult to recognize
- many tricks for transforming problems into convex form
- surprisingly many problems can be solved via convex optimization

1.2 Convex sets

⋆ Affine set

line through x_1,x_2: all pints

$$x = \theta x_1 + (1-\theta)x_2 \quad (\theta \in \mathbf{R})$$



DEF: affine set

affine set: contains the line through any two distinct points in the set

⋆ Convex set

line segment between x_1 and x_2: all points

$$x = \theta x_1 + (1-\theta)x_2$$

with $0 \leq \theta \leq 1$

DEF: convex set

convex set: contains line segment between any two points in the set

$$x_1, x_2 \in C, 0 \leq \theta \leq 1 \Rightarrow \theta x_1 + (1-\theta)x_2 \in C$$

⋆ Convex combination and convex hull

DEF: convex combination

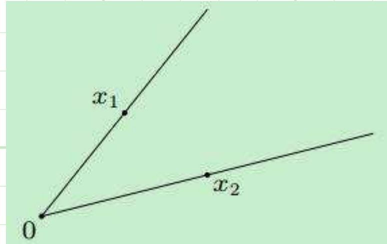
convex combination of $x_1, \ldots, x_k$: any point x of the form

$$x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k$$

with $\theta_1 + \cdots + \theta_k = 1, \theta_i | geq 0$

DEF: convex hull conv S

convex hull conv S: set of all combinations of points in S



⋆ Convex cone

DEF: conic (nonnegative) combination

conic (nonnegative) combination of x_1 and x_2: any point of the form

$$x = \theta_1 x_1 + \theta_2 x_2$$

withe $\theta_1 \geq 0, \theta_2 \geq 0$

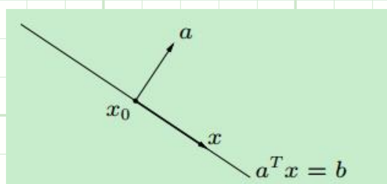
DEF: convex cone

convex cone: set that contains all conic combinations of points in the set

⋆ Hyperplanes and halfspaces

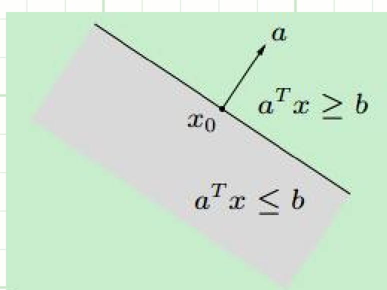
DEF: hyperplane

hyperplane: set of the form $\{x | a^T x = b\} \quad (a \neq 0)$



DEF: halfspace

halfspace: set of the form $\{x | a^T x \leq b\} \quad (a \neq 0)$



remark

- a is the normal vector
- hyperplanes are affine and convex; halfspaces are convex

section name goes here

* term definition

DEF: term - and it's definition

* an example

EX: example heading

* a system of equations

$$\begin{cases} 2x + 4y = 2 \\ 2x + 6y = 3 \end{cases}$$

★ working a multistep problem

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \xrightarrow{R_1+R_2} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \Rightarrow \begin{cases} x = 1 \\ y = 2 \\ z = 3 \end{cases}$$

★ a vector in R^3

$$v = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

★ a multi-step process

$$A \xrightarrow{\text{do stuff}} B \xrightarrow{\text{more stuff}} C$$

★ an enumerated list

1. this is the first item in an enumerated list
2. this is the second item in an enumerated list

★ manually broken lines

the first line
the second line
the third line

★ some math

$$\int_a^b f(x)\,dx \int f(x)\,dx \; \frac{\pi}{2} \; \sqrt{\theta} \; n = 1, 2, 3 \ldots 4$$