Non-linear optimization modeling

15.093: Optimization

Dimitris Bertsimas
Alexandre Jacquillat

Sloan School of Management
Massachusetts Institute of Technology

Sample problems and formulations

The approximation problem

Formulation (The approximation problem)

$$\min_{\boldsymbol{x}} \quad \{\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\| : \boldsymbol{x} \in \mathcal{X}\}$$

- Non-linearities from norm-based approximation:

$$\text{Least-square approximation:} \quad \min \sum_{i=1}^{m} \left(\boldsymbol{a}_i^\top \boldsymbol{x} - b_j\right)^2$$

$$\text{Minimax approximation:} \quad \min \max_{i=1,\cdots,m} \left|\boldsymbol{a}_i^\top \boldsymbol{x} - b_j\right|$$

$$\text{Absolute-residual approximation:} \quad \min \sum_{i=1}^{m} \left|\boldsymbol{a}_i^\top \boldsymbol{x} - b_j\right|$$

- Examples of applications across disciplines
  - Model fitting: approximating a response vector $\boldsymbol{b}$ via $\boldsymbol{A}\boldsymbol{x}$
  - Geometry: projecting vector $\boldsymbol{b}$ onto the subspace spanned by $\boldsymbol{A}$
  - Engineering: design decision $\boldsymbol{x}$ with minimal error from target $\boldsymbol{b}$

Approximation: application to linear regression

- Features: independent variables x_{ij}, for $i = 1, \cdots, n$, $j = 1, \cdots, m$
- Target: dependent variable y_i, for $i = 1, \cdots, n$
- Linear regression objectives:
  - Approximation: finding coefficients $\beta_1, \cdots, \beta_m$ such that

$$\sum_{j=1}^{m} \beta_j x_{ij} \approx y_i \text{ for all } i = 1, \cdots, n$$

  - Regularization: finding "well-behaved" coefficients $\beta_1, \cdots, \beta_m$

$$\text{Example of a quadratic penalty: } R(\boldsymbol{\beta}) = \sum_{j=1}^{m} \beta_j^2$$

  - Robustness: lower weight to less "important" data points (e.g., outliers)

$\rightarrow$ Weighted and regularized linear regression

$$\min_{\boldsymbol{\beta}} \quad \sum_{i=1}^{n} w_i \left(y_i - \sum_{j=1}^{m} x_{ij}\beta_j \right)^2 + \lambda R(\boldsymbol{\beta})$$

Approximation: application to logistic regression

- Features: independent variables x_{ij}, for $i = 1, \cdots, n$, $j = 1, \cdots, m$
- Target: dependent variable $y_i \in \{0, 1\}$, for $i = 1, \cdots, n$
- Logistic regression model:

$$p_i = \mathbb{P}\left(y = 1 | x_1, \cdots, x_m\right) = \frac{e^{\beta_1 x_1 + \cdots + \beta_m x_m}}{1 + e^{\beta_1 x_1 + \cdots + \beta_m x_m}}$$

- Approximation problem: finding coefficients $\beta_1, \cdots, \beta_m$ such that

$$\frac{e^{\beta_1 x_{i1} + \cdots + \beta_m x_{im}}}{1 + e^{\beta_1 x_{i1} + \cdots + \beta_m x_{im}}} \text{ is high } \iff y_i = 1$$

$\rightarrow$ Regularized logistic regression based on the log-likelihood function

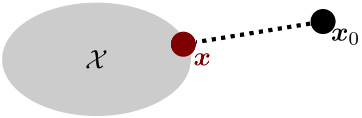
$$\max_{\boldsymbol{\beta}} \quad \sum_{i=1}^{n} \left(\log p_i \mathbf{1}(y_i = 1) + \log(1 - p_i) \mathbf{1}(y_i = 0)\right) + \lambda R(\boldsymbol{\beta})$$

$$\max_{\boldsymbol{\beta}} \quad \sum_{i=1}^{n} \sum_{j=1}^{m} y_i \beta_j x_{ij} - \sum_{i=1}^{n} \log\left(1 + e^{\beta_1 x_{i1} + \cdots + \beta_m x_{im}}\right) + \lambda R(\boldsymbol{\beta})$$

Approximation: geometric application

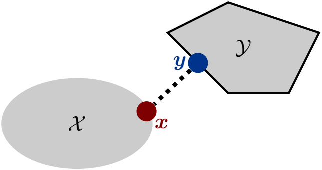
Projecting x_0 onto $\mathcal{X}$

$$\min_{\boldsymbol{x}} \{\|\boldsymbol{x} - \boldsymbol{x}_0\| : \boldsymbol{x} \in \mathcal{X}\}$$



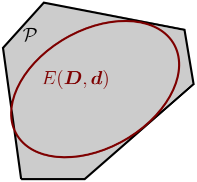
Distance between $\mathcal{X}$ and $\mathcal{Y}$

$$\min_{\boldsymbol{x},\boldsymbol{y}} \{\|\boldsymbol{x} - \boldsymbol{y}\| : \boldsymbol{x} \in \mathcal{X}, \boldsymbol{y} \in \mathcal{Y}\}$$



Maximum inscribed ellipsoid

$$\begin{aligned} \max\ & \log(|\det(\boldsymbol{D}|)) \\ \text{s.t. } & E(\boldsymbol{D}, \boldsymbol{d}) \subseteq \mathcal{P} \end{aligned}$$



$$E(\boldsymbol{D}, \boldsymbol{d}) = \left\{\boldsymbol{x} \in \mathbb{R}^n : (\boldsymbol{x} - \boldsymbol{d})^\top \boldsymbol{D}^{-2} (\boldsymbol{x} - \boldsymbol{d}) \le 1\right\}$$

Portfolio optimization: statement

- Portfolio optimization problem: investing in stocks to manage risk and maximize reward
- Data
  - n stocks, indexed by $i, j = 1, \cdots, n$
  - Budget B, target β on expected portfolio reward
  - The reward r_j of stock j is a random variable, but we can estimate its mean μ_i, its variance $\mathsf{Var}(r_j)$, and the covariance σ_{ij} between rewards:

$$\mu_j = \mathbb{E}(r_j)$$
$$\sigma_{ij} = \mathbb{E}[(r_j - \mu_j)(r_i - \mu_i)] = Cov(r_i, r_j)$$
$$\sigma_{jj} = \mathsf{Var}(r_j)$$

- Objective: Minimize total portfolio variance (risk)
- Constraints:
  1. Expected reward of total portfolio is above target β
  2. Total amount invested stay within our budget
  3. No short sales: investments are non-negative

Portfolio optimization: formulation

- Decision variable

$$x_j : \text{ amount invested in stock } j = 1, \cdots, n$$

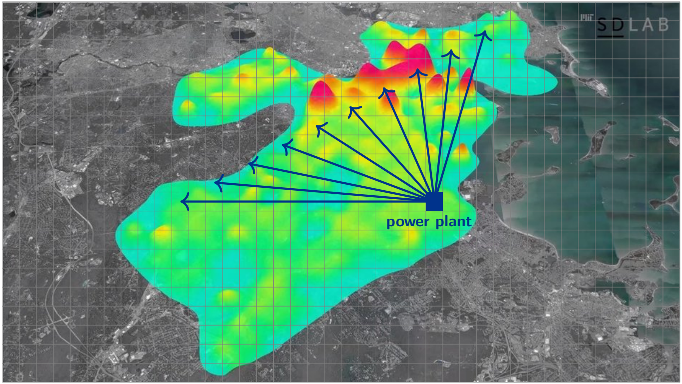
Formulation (Portfolio optimization)

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{n} x_i^2 \mathit{Var}(x_i) + \sum_{i \neq j} 2 x_i x_j \sigma_{ij} \\
\textit{s.t.} \quad & \sum_{i=1}^{n} x_i \leq B \\
& \sum_{i=1}^{n} \mu_i x_i \geq \beta \\
& x_1, \cdots, x_n \geq 0
\end{aligned}
$$

$\rightarrow$ Non-linearities from risk management objective (variance-covariance)

Positioning a power plant

- Data: map of electricity demand in the Boston area
  - m demand nodes, indexed by $i = 1, \cdots, m$
  - Location of each demand node: $\boldsymbol{c}_1, \cdots, \boldsymbol{c}_m \in \mathbb{R}^n$
  - Demand in each location: $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m \in \mathbb{R}^n$
- Where to locate a new power plant to best serve customer demand?
- Non-linearities from norm-based placement



The Fermat-Weber problem

- Decision variable $\boldsymbol{x} \in \mathbb{R}^n$: location of new power plant
- Objectives:
  - Fermat-Weber problem: minimizing total weighted distance
  - Ball circumscription problem: minimizing maximum distance

Formulation (Fermat-Weber problem)

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \sum_{i=1}^{m} \boldsymbol{w}_i \|\boldsymbol{x} - \boldsymbol{c}_i\| \\ \textit{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \quad \textit{feasible sites} \end{aligned}$$

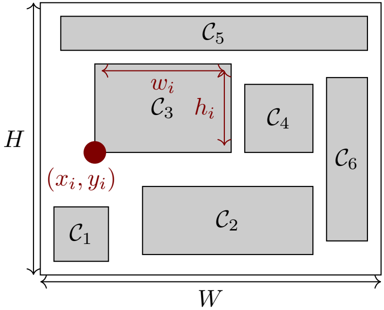
Formulation (Ball circumscription problem)

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \Delta \\ \textit{s.t.} \quad & \|\boldsymbol{x} - \boldsymbol{c}_i\| \leq \Delta, \ \forall i = 1, \cdots, m \\ & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \quad \textit{feasible sites} \end{aligned}$$

Floor planning problem: statement

- Locating n rectangular objects in an $W \times H$ bounding area
  - Objective: minimizing size (area, perimeter) of the bounding area
  - Constraints: object dimensions, no overlap, practical considerations
- Decisions:
  - placement of each object via lower left corner: (x_i, y_i)
  - configuration of each object: width w_i, heigth h_i
- Applications: manufacturing planning, computational design, etc.

- To simplify, we assume that the relative positioning is given
- $(i,j) \in \mathcal{H}$ if object $\mathcal{C}_i$ must be located left of object $\mathcal{C}_j$
  - Example: $(1,2) \in \mathcal{H}$
- $(i,j) \in \mathcal{V}$ if object $\mathcal{C}_i$ must be located below object $\mathcal{C}_j$
  - Example: $(6,5) \in \mathcal{V}$



Floor planning problem: sample formulation

- Objective: minimizing the bounding perimeter $\min\ 2(W+H)$
- Each box lies within the bounding area:

$$x_i \geq 0,\ y_i \geq 0,\ x_i + w_i \leq W,\ y_i + h_i \leq H,\ \forall i = 1, \cdots, n$$

- Relative positioning constraints, with a buffer $\rho \geq 0$:

$$x_i + w_i + \rho \leq x_j,\ \forall (i,j) \in \mathcal{H}$$
$$y_i + h_i + \rho \leq x_j,\ \forall (i,j) \in \mathcal{V}$$

- Minimum area and aspect ratio for each object:

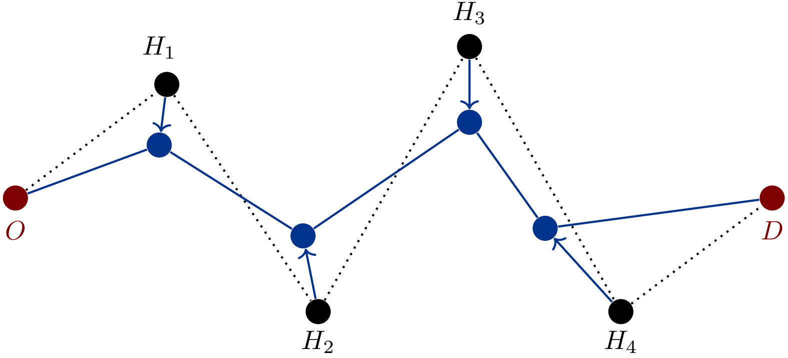
$$w_i h_i \geq A_i,\ \forall i = 1, \cdots, n$$
$$l_i \leq h_i / w_i \leq u_i,\ \forall i = 1, \cdots, n$$

- Restricting the distance between objects

$$x_i \leq u_i \leq x_i + w_i,\ \forall i = 1, \cdots, n$$
$$y_i \leq v_i \leq y_i + h_i,\ \forall i = 1, \cdots, n$$
$$\|(u_i, v_i) - (u_j, v_j)\| \leq D_{ij},\ \forall i,j = 1, \cdots, n$$

Shortest path with vehicle-customer coordination: statement

- A vehicle travels from an origin O to a destination D, at speed $\overline{S}$
- Each customer $i = 1, \cdots, n$ starts in a home location H_i, and can walk up to W_i to the pickup location at speed S_i
- Where and when to pick up each customer to minimize travel times?
- Application: routing engine of "walking products" in ride-sharing



Shortest path with vehicle-customer coordination: formulation

- Decision variables:
  - M_i: location of stop $i = 0, \cdots, n+1$
  - v_i: time of stop $i = 0, \cdots, n+1$
- Objective: minimizing the arrival time at destination
- Constraints
  - Start at the origin, end at the destination
  - Coordination requirements between vehicle and customers
  - Maximum walking distance for customers

$$
\begin{aligned}
\min\ & v_{K+1} \\
\text{s.t. } & v_0 = \overline{v},\ M_0 = O,\ M_{n+1} = D \\
& v_i \geq v_{i-1} + \frac{\|M_i - M_{i-1}\|}{\overline{S}} \quad \forall i = 1, \cdots, n+1 \\
& v_i \geq \frac{\|M_i - H_i\|}{S_i} \quad \forall i = 1, \cdots, n \\
& \|M_i - H_i\| \leq W_i \quad \forall i = 1, \cdots, n \\
& v_i \geq 0,\ M_i \in \mathbb{R}^2 \quad \forall i = 0, \cdots, n+1
\end{aligned}
$$

The general problem

Non-linear optimization (NLO)

- $f(\boldsymbol{x}) \colon \mathbb{R}^n \to \mathbb{R}$ is a continuous (usually differentiable) function
- $g_i(\boldsymbol{x}) \colon \mathbb{R}^n \mapsto \mathbb{R}, i = 1, \cdots, m$
- $h_j(\boldsymbol{x}) \colon \mathbb{R}^n \mapsto \mathbb{R}, j = 1, \cdots, l$

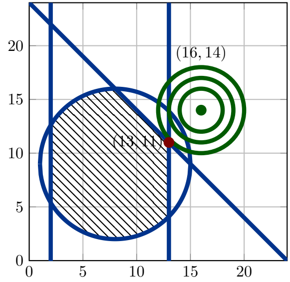
Formulation (Non-linear optimization)

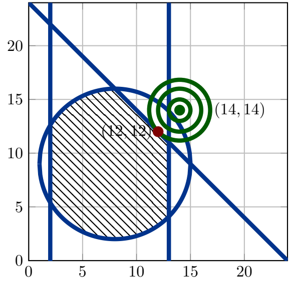
$$
(NLO) \qquad \min \quad f(\boldsymbol{x})
$$
$$
\textit{s.t.} \quad \left.\begin{array}{rcl}
g_1(\boldsymbol{x}) & \leq & 0 \\
\vdots & & \\
g_m(\boldsymbol{x}) & \leq & 0 \\
h_1(\boldsymbol{x}) & = & 0 \\
\vdots & & \\
h_\ell(\boldsymbol{x}) & = & 0
\end{array}\right\} \mathcal{F} \textit{ feasible region}
$$

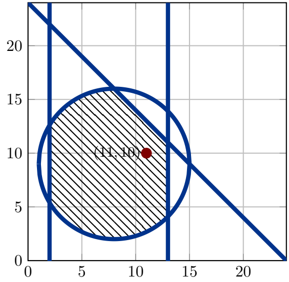
Difficulty 1: optima may not lie in corners of the feasible region

$$\min \quad (x-a)^2 + (y-b)^2$$
$$(x-8)^2 + (y-9)^2 \leq 49$$
$$2 \leq x \leq 13, \quad x + y \leq 24$$

Solution in corner	**Solution on boundary**	**Solution in interior**
$(a,b) = (16,14)$	$(a,b) = (14,14)$	$(a,b) = (11,10)$







Difficulty 2: a local optimum may not be a global optimum

Definition (Ball)

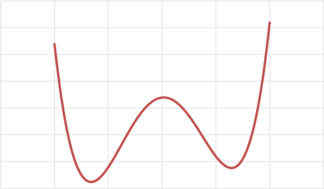
Ball centered at $\overline{\boldsymbol{x}}$ with radius ε: $\mathcal{B}(\overline{\boldsymbol{x}}, \varepsilon) := \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \overline{\boldsymbol{x}}\| \leq \varepsilon\}$

Definition (Local optimum)

$\boldsymbol{x} \in \mathcal{F}$ is a **local minimum** of NLO if there exists $\varepsilon > 0$ such that $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ for all $\boldsymbol{y} \in B(\boldsymbol{x}, \varepsilon) \cap \mathcal{F}$

Definition (Global optimum)

$\boldsymbol{x} \in \mathcal{F}$ is a **global minimum** of NLO if $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ for all $\boldsymbol{y} \in \mathcal{F}$



The role of convexity

Convex sets

Definition (Convex set)

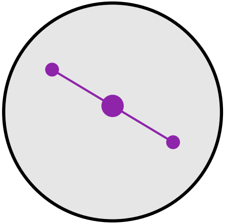
A subset $S \subseteq \mathbb{R}^n$ is a **convex set** if

$$\boldsymbol{x}, \boldsymbol{y} \in S \implies \theta \boldsymbol{x} + (1-\theta)\boldsymbol{y} \in S \qquad \forall \theta \in [0,1]$$

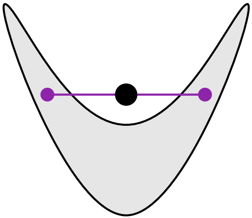
Proposition

The intersection of convex sets is a convex set.

Convex set



Non-convex set



Convex functions

Definition (Convex function)

$f : \mathbb{R}^n \mapsto \mathbb{R}$ is a **convex function** if

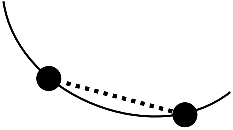
$$f(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y}), \ \forall \boldsymbol{x}, \boldsymbol{y}, \ \forall \theta \in [0,1]$$

Definition (Concave function)

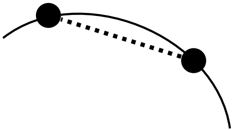
$f : \mathbb{R}^n \mapsto \mathbb{R}$ is a **concave function** if

$$f(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y}) \geq \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y}), \ \forall \boldsymbol{x}, \boldsymbol{y}, \ \forall \theta \in [0,1]$$

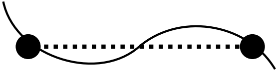
A convex function



A concave function



Neither



Properties of convex functions

Proposition

$f_1(\cdot)$, $f_2(\cdot)$ *convex;* $\alpha, \beta \geq 0 \implies f(\boldsymbol{x}) := \alpha f_1(\boldsymbol{x}) + \beta f_2(\boldsymbol{x})$ *convex*

Proposition

$f(\cdot) \implies g(\boldsymbol{y}) := f(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b})$ *convex.*

Proposition

$f_1(\cdot), \cdots, f_n(\cdot)$ *convex* $\implies g(\boldsymbol{x}) := \max\{f_1(\boldsymbol{x}), \cdots, f_n(\boldsymbol{x})\}$ *convex.*

Proposition

$h(\cdot)$ *convex and non-decreasing;* $g(\cdot)$ *convex* $\implies f(\boldsymbol{x}) := h(g(x))$ *convex.*

Proposition

$\|\cdot\|$ *norm* $\implies f(\boldsymbol{x}) := \|\boldsymbol{x}\|$ *convex.*

Differentiability in n dimensions

Definition (Gradient)

$f : \mathbb{R}^n \mapsto \mathbb{R}$ is differentiable at $\bar{\boldsymbol{x}}$ if there exists $\nabla f(\bar{\boldsymbol{x}}) \in \mathbb{R}^n$ and $\rho : \mathbb{R}^n \mapsto \mathbb{R}$ with $\rho(\boldsymbol{x}) \to 0$ as $\boldsymbol{x} \to \bar{\boldsymbol{x}}$, such that

$$f(\boldsymbol{x}) = f(\bar{\boldsymbol{x}}) + \nabla f(\bar{\boldsymbol{x}})^\top (\boldsymbol{x} - \bar{\boldsymbol{x}}) + R(\boldsymbol{x}) \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|$$

Gradient: vector of partial derivatives: $\nabla f(\overline{\boldsymbol{x}}) = \left(\frac{\partial f(\overline{\boldsymbol{x}})}{\partial x_1}, \cdots, \frac{\partial f(\overline{\boldsymbol{x}})}{\partial x_n} \right)^\top$

Definition (Hessian)

$f : \mathbb{R}^n \mapsto \mathbb{R}$ is twice differentiable at $\overline{\boldsymbol{x}}$ if there exists $\nabla f(\overline{\boldsymbol{x}}) \in \mathbb{R}^n$ and $\rho : \mathbb{R}^n \mapsto \mathbb{R}$ with $\rho(\boldsymbol{x}) \to 0$ as $\boldsymbol{x} \to \overline{\boldsymbol{x}}$, such that

$$f(\boldsymbol{x}) = f(\overline{\boldsymbol{x}}) + \nabla f(\overline{\boldsymbol{x}})^\top (\boldsymbol{x} - \overline{\boldsymbol{x}}) + \frac{1}{2} (\boldsymbol{x} - \overline{\boldsymbol{x}})^\top H(\overline{\boldsymbol{x}}) (\boldsymbol{x} - \overline{\boldsymbol{x}}) + \rho(\boldsymbol{x}) \|\boldsymbol{x} - \overline{\boldsymbol{x}}\|^2$$

Hessian: matrix of second partial derivatives: $H(\overline{\boldsymbol{x}})_{ij} = \frac{\partial^2 f(\overline{\boldsymbol{x}})}{\partial x_i \partial x_j}$

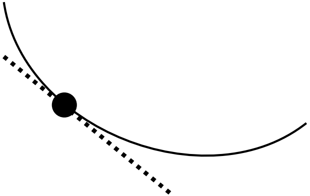
Recognizing convex functions: first-order conditions

Proposition (First-order condition)

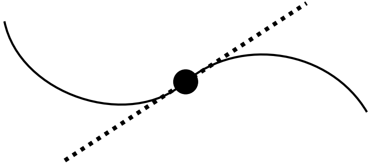
Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a differentiable function. It is convex if and only:

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^\top (\boldsymbol{y} - \boldsymbol{x}), \ \forall \boldsymbol{x}, \boldsymbol{y}$$

A convex function



A non-convex function



- Interpretation: A function is convex if and only if its first-order Taylor approximation is an underestimator of the function across its domain

$\rightarrow$ Global information from local information

Recognizing convex functions: first-order conditions (proof)

Proposition (First-order condition)

Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a differentiable function. It is convex if and only:

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^\top(\boldsymbol{y} - \boldsymbol{x}),\ \forall \boldsymbol{x}, \boldsymbol{y}$$

$\Longrightarrow$ Consider $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. By convexity, we have:

$$\begin{aligned} & f(\theta\boldsymbol{y} + (1-\theta)\boldsymbol{x}) \leq \theta f(\boldsymbol{y}) + (1-\theta) f(\boldsymbol{x}) \\ & \frac{f(\boldsymbol{x} + \theta(\boldsymbol{y} - \boldsymbol{x})) - f(\boldsymbol{x})}{\theta} \leq f(\boldsymbol{y}) - f(\boldsymbol{x}) \\ & \theta \to 0 : \quad \Longrightarrow \quad \nabla f(\boldsymbol{x})^\top(\boldsymbol{y} - \boldsymbol{x}) \leq f(\boldsymbol{y}) - f(\boldsymbol{x}) \end{aligned}$$

$\Longleftarrow$ Consider $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $\theta \in [0, 1]$ and define $\boldsymbol{z} = \theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}$

$$\left.\begin{array}{l} f(\boldsymbol{x}) \geq f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^\top(\boldsymbol{x} - \boldsymbol{z}) \\ f(\boldsymbol{y}) \geq f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^\top(\boldsymbol{y} - \boldsymbol{z}) \end{array}\right\} \Longrightarrow \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y}) \geq f(\boldsymbol{z})$$

Recognizing convex functions: second-order conditions

Definition (Positive semi-definite matrix)

A symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is positive semi-definite, written $\boldsymbol{A} \succeq \mathbf{0}$, if $\boldsymbol{u}^\top \boldsymbol{A} \boldsymbol{u} \geq 0$ for all $\boldsymbol{u} \in \mathbb{R}^n$. Equivalently, $\boldsymbol{A} \succeq \mathbf{0}$ if and only if all its eigenvalues are nonnegative.

Proposition (Second-order condition)

Let $f \cdot)$ be a twice continuously differentiable function.
f is a convex function if and only if $H(\boldsymbol{x}) \succeq \mathbf{0}$ for all $\boldsymbol{x} \in \mathbb{R}^n$

- Interpretation: A function is convex if and only if it has an upward curvature across its domain
- Example of a quadratic function
  - Definition: $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P} \boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} + r$
  - Gradient: $\nabla f(\boldsymbol{x}) = \boldsymbol{P}\boldsymbol{x} + \boldsymbol{q}$
  - Hessian: $H(\boldsymbol{x}) = \boldsymbol{P}$

  $\rightarrow$ The function is convex if and only if $\boldsymbol{P}$ is positive semi-definite

Convex optimization

Convex optimization formulation

- A convex optimization problem has three features
  - A convex objective function, to be minimized
  - Convex inequality constraints
  - Linear equality constraints

Formulation (Convex optimization problem)

Consider convex functions $f : \mathbb{R}^n \to \mathbb{R}$, $g_1, \cdots, g_m : \mathbb{R}^n \to \mathbb{R}$. The convex optimization problem is defined as:

$$\begin{aligned} \min \quad & f(\boldsymbol{x}) \\ \textit{s.t.} \quad & g_1(\boldsymbol{x}) \le 0 \\ & \dots \\ & g_m(\boldsymbol{x}) \le 0 \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{aligned}$$

Proposition

The feasible region $\mathcal{F}$ of a convex optimization problem is a convex set

Examples

- A linear optimization problem is a convex optimization problem
- The Fermat-Weber and ball circumscription problems are convex optimization problems

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \sum_{i=1}^m \boldsymbol{w}_i \|\boldsymbol{x}-\boldsymbol{c}_i\| \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \end{aligned}$$

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \Delta \\ \text{s.t.} \quad & \|\boldsymbol{x}-\boldsymbol{c}_i\| \le \delta,\ \forall i = 1,\cdots,m \\ & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \end{aligned}$$

- Quadratically constrained problems are convex optimization problems

$$\begin{aligned} \min \quad & (\boldsymbol{A}_0\boldsymbol{x}+\boldsymbol{b}_0)^\top(\boldsymbol{A}_0\boldsymbol{x}+\boldsymbol{b}_0) - \boldsymbol{c}_0^\top\boldsymbol{x} - d_0 \\ \text{s.t.} \quad & (\boldsymbol{A}_i\boldsymbol{x}+\boldsymbol{b}_i)^\top(\boldsymbol{A}_i\boldsymbol{x}+\boldsymbol{b}_i) - \boldsymbol{c}_i^\top\boldsymbol{x} - d_i \le 0,\ \forall i = 1,\cdots,m \end{aligned}$$

- The portfolio optimization is not a convex optimization problem if the variance-covariance matrix is not positive semi-definite

$$\min\ \boldsymbol{x}^\top\boldsymbol{\Sigma}\boldsymbol{x} \quad \text{or} \quad \boldsymbol{x}^\top\boldsymbol{\Sigma}\boldsymbol{x} \le \sigma$$

Local and global minima in convex optimization

Theorem

Let $\mathcal{F}$ be a convex set and $f : \mathcal{F} \mapsto \mathbb{R}$ be a convex function. Consider the following optimization problem:

$$\min_{\boldsymbol{x} \in \mathcal{F}} f(\boldsymbol{x})$$

If $\boldsymbol{x}^*$ is a local minimum, then $\boldsymbol{x}^*$ is a global minimum.

- Proof:
  - Assume by contradiction that $\exists \boldsymbol{y} \in \mathcal{F} : f(\boldsymbol{y}) < f(\boldsymbol{x}^*)$
  - By convexity of $f(\cdot)$:

$$f(\theta \boldsymbol{x}^* + (1-\theta)\boldsymbol{y}) \leq \theta f(\boldsymbol{x}^*) + (1-\theta) f(\boldsymbol{y}) < f(\boldsymbol{x}^*) \text{ for all } \theta \in (0,1)$$

  - At the same time, $\theta \boldsymbol{x}^* + (1-\theta)\boldsymbol{y} \in \mathcal{F}$ by convexity of the set $\mathcal{F}$
  - → As $\theta \to 0$, there exists a point in any ball centered in $\boldsymbol{x}^*$ that is feasible and achieves a strictly smaller objective value than $\boldsymbol{x}^*$
  - That is, $\boldsymbol{x}^*$ is not a local minimum, resulting in a contradiction

Conclusion

Summary

Takeaway

Convex and non-linear optimization are powerful modeling frameworks.

Takeaway

Non-linear optimization encompasses a vast range of problem classes, with important implications in terms of modeling and algorithms.

Takeaway

Convex optimization problems can be solved efficiently, thanks to the global optimality of local optima.

Takeaway

Still, some convex optimization problems are easier than others.

Takeaway

Non-convex optimization problems are much harder.