CME 307 / MS&E 311 / OIT 676: Optimization

Introduction

Professor Udell

Management Science and Engineering
Stanford

September 25, 2024

Announcements

announcements:

- website: https://stanford-cme-307.github.io/web
- instructors: Madeleine Udell and Dan Iancu
- TAs: Zach Frangella and Pratik Rathore
- Ed for discussion and announcements
- fill out course survey (see website)
- talk to instructors after class and/or at office hours (see website)
- class attendance is required. will post slides, generally no recordings

before class starts: find someone you haven't met and introduce yourselves.

- name, major, year
- something fun you did this summer
- why are you interested in optimization?
- what are you hoping to learn?

Agenda for today

- Understand course objectives and expectations
- Identify several types of optimization problem
- Meet someone you've not met before
- Discuss challenges in a real-world optimization problem
- Review basic linear algebra

Outline

What is an optimization problem?

Course goals and expectations

Linear Algebra Review

(Integer) linear optimization problem

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax = b \\
& Cx \leq d \\
& \ell \leq x \leq u \\
\text{variable} & x \in \mathbf{Z}^{n_1} \times \mathbf{R}^{n_2}
\end{array}
$$

- objective $c^T x$
- equality constraints $Ax = b$
- inequality constraints $Cx \leq d$
- lower and upper bounds $\ell \leq x \leq u$
- integer variables if $n_1 > 0$

problem data:

- $c \in \mathbf{R}^n$, $n = n_1 + n_2$ total variables
- $\ell \in \mathbf{R}^n$, $u \in \mathbf{R}^n$
- $A \in \mathbf{R}^{m_1 \times n}$, $b \in \mathbf{R}_1^m$, $C \in \mathbf{R}^{m_2 \times n}$, $d \in \mathbf{R}^{m_2}$ $m = m_1 + m_2$ total constraints

LP example: diet problem

I am designing a diet to be as cheap as possible but adequately nutritious. We have a list of essential nutrients and how much an active person needs of each. We also know the cost to procure each food, and how much of each nutrient is in each food.

- x_j servings of food j, $j = 1, \ldots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \ldots, m$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

LP example: diet problem

I am designing a diet to be as cheap as possible but adequately nutritious. We have a list of essential nutrients and how much an active person needs of each. We also know the cost to procure each food, and how much of each nutrient is in each food.

- x_j servings of food j, $j = 1, \ldots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \ldots, m$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

extensions:

- foods come from recipes?

LP example: diet problem

I am designing a diet to be as cheap as possible but adequately nutritious. We have a list of essential nutrients and how much an active person needs of each. We also know the cost to procure each food, and how much of each nutrient is in each food.

- x_j servings of food j, $j = 1, \ldots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \ldots, m$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

extensions:

- foods come from recipes? $x = By$

LP example: diet problem

I am designing a diet to be as cheap as possible but adequately nutritious. We have a list of essential nutrients and how much an active person needs of each. We also know the cost to procure each food, and how much of each nutrient is in each food.

- x_j servings of food j, $j = 1, \ldots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \ldots, m$

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax = b \\
& x \geq 0
\end{array}
$$

extensions:

- foods come from recipes? $x = By$
- ensure diversity in diet?

LP example: diet problem

I am designing a diet to be as cheap as possible but adequately nutritious. We have a list of essential nutrients and how much an active person needs of each. We also know the cost to procure each food, and how much of each nutrient is in each food.

- x_j servings of food j, $j = 1, \ldots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \ldots, m$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

extensions:

- foods come from recipes? $x = By$
- ensure diversity in diet? $y \leq u$

LP example: diet problem

I am designing a diet to be as cheap as possible but adequately nutritious. We have a list of essential nutrients and how much an active person needs of each. We also know the cost to procure each food, and how much of each nutrient is in each food.

- x_j servings of food j, $j = 1, \ldots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \ldots, m$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

extensions:

- foods come from recipes? $x = By$
- ensure diversity in diet? $y \leq u$
- ranges of nutrients?

LP example: diet problem

I am designing a diet to be as cheap as possible but adequately nutritious. We have a list of essential nutrients and how much an active person needs of each. We also know the cost to procure each food, and how much of each nutrient is in each food.

- x_j servings of food j, $j = 1, \dots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \dots, m$

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax = b \\
& x \geq 0
\end{array}
$$

extensions:

- foods come from recipes? $x = By$
- ensure diversity in diet? $y \leq u$
- ranges of nutrients? $Ax + s = b, \ \ l \leq s \leq u$

Nonlinear optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m_1 \\ & h_i(x) = 0, \quad i = 1, \ldots, m_2 \\ \text{variable} & x \in \mathbf{R}^n \end{array}$$

- objective f_0
- inequality constraints f_i
- equality constraints h_i

problem data:

- (blackbox) code to evaluate f_i and h_i for any $x \in \mathbf{R}^n$
- (first order) and to compute gradients
- (second order) and to compute Hessians

Example: process control

You are the process engineer for a desalination plant that produces drinking water. The plant has a variety of knobs, collected in vector x, that you can turn to control the process. These control, *e.g.*, how much water is pumped into the plant, how much pressure is used to force the water through filters, and how much of each chemical is added to the water.

- $f_0(x)$: cost of water produced
- $f_i(x)$: level of each measured impurity in the water
- b_i: maximum allowable level of each impurity

Given a setting of the knobs, you can observe the cost of water produced and the levels of impurities.

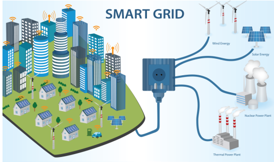
What is the optimal setting of the knobs?

Why optimization?

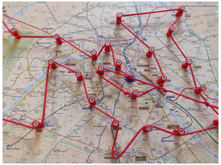
declarative programming:

- model: specify what you require and what you prefer
- solve: then figure out how to get it

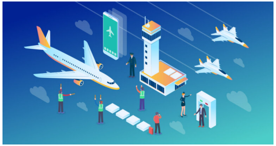
Optimization in operations



Energy



Routing



Scheduling

Supply chain

- Optimization improves efficiency throughout the economy
- $\Longrightarrow$ more productivity, less waste, lower costs, lower carbon, more utility

Where is optimization used?

- statistical estimation and machine learning
- controls (robotics, finance)
- operations (supply chain, logistics, routing, scheduling)
- ...

characteristics of these problems differ:

- discrete vs continuous variables
- constrained vs unconstrained
- linear vs nonlinear
- estimated vs known problem data

Optimization problems

important optimization problem classes:

- linear
- integer
- nonlinear (with linear or nonlinear constraints)
- quadratic
- unconstrained
- finite-sum
- conic
- convex
- black-box with (0, 1, or 2)-order oracle

Modularity in optimization

how to optimize:

1. model problem as a mathematical optimization problem
2. identify the properties of the problem
3. use an appropriate solver (or write a new one)

... and iterate:

- approximate the problem to make it easier
- solve a sequence of approximated problems that converge to solve the original problem
- or initialize ("warm-start") a solver for the original problem with a solution to the approximated problem

Outline

Course goals

look at goals, materials, and grading on course website:
https://stanford-cme-307.github.io/web/

- Which goals sound exciting?
- Which goals don't make sense?
- What else do you hope to accomplish?
- Do expectations make sense given course goals?

Outline

What is an optimization problem?

Course goals and expectations

Linear Algebra Review

Linear independence

- vectors $a_1, \ldots, a_k \in \mathbf{R}^n$ are **linearly dependent** if there exist scalars $\lambda, \ldots, \lambda$ not all zero so that

$$\lambda_1 a_1 + \ldots + \lambda_k a_k = 0$$

- otherwise, they are **linearly independent**
- the **span** of $a_1, \ldots, a_k$ is the set of all linear combinations $\lambda_1 a_1 + \ldots + \lambda_k a_k$
- a **linear subspace** L is a set closed under addition and scalar multiplication

$$v,\ w \in L \implies v + w \in L,\ \lambda v \in L\ \forall \lambda \in \mathbf{R}$$

- an **affine subspace** A is a set of the form $x_0 + L$ where $x_0 \in \mathbf{R}^n$ and L is a linear subspace.
- fact: A is affine if and only if

$$v,\ w \in A \implies \lambda v + (1 - \lambda) w \in A\ \forall \lambda \in \mathbf{R}$$

Span and nullspace

matrix $A \in \mathbf{R}^{m \times n}$ with columns $a_1, \ldots, a_n$. define

- span of A
- nullspace of A
- rank of A

proof: on board

Span and nullspace

matrix $A \in \mathbf{R}^{m \times n}$ with columns $a_1, \ldots, a_n$. define

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- rank of A: $\mathrm{Rank}(A) = \dim(\mathrm{span}(A))$

proof: on board

Span and nullspace

matrix $A \in \mathbf{R}^{m \times n}$ with columns $a_1, \ldots, a_n$. define

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\textbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- rank of A: $\mathrm{Rank}(A) = \dim(\mathrm{span}(A))$

geometry? what is the relationship between these?

proof: on board

Span and nullspace

matrix $A \in \mathbf{R}^{m \times n}$ with columns $a_1, \ldots, a_n$. define

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- rank of A: $\mathrm{Rank}(A) = \dim(\mathrm{span}(A))$

geometry? what is the relationship between these?

rank nullity Theorem:

$$\mathrm{Rank}(A) + \dim(\mathbf{nullspace}(A)) = n$$

proof: on board

Span and nullspace

matrix $A \in \mathbf{R}^{m \times n}$ with columns $a_1, \ldots, a_n$. define

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\textbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- rank of A: $\mathrm{Rank}(A) = \dim(\mathrm{span}(A))$

geometry? what is the relationship between these?

rank nullity Theorem:

$$\mathrm{Rank}(A) + \dim(\textbf{nullspace}(A)) = n$$

proof: on board

Solution of linear system

matrix $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $m \leq n$ ("fat matrix"). define

- solution set of linear system $\{x : Ax = b\}$

Solution of linear system

matrix $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $m \leq n$ ("fat matrix"). define

- solution set of linear system $\{x : Ax = b\}$

Q: when does a solution exist? what is the dimension of the solution set? when is the solution unique?

Solution of linear system

matrix $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $m \leq n$ ("fat matrix"). define

- solution set of linear system $\{x : Ax = b\}$

Q: when does a solution exist? what is the dimension of the solution set? when is the solution unique?

- solution exists if $b \in \mathrm{span}(A)$. define x_0 so that $Ax_0 = b$.

Solution of linear system

matrix $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $m \leq n$ ("fat matrix"). define

- solution set of linear system $\{x : Ax = b\}$

Q: when does a solution exist? what is the dimension of the solution set? when is the solution unique?

- solution exists if $b \in \mathrm{span}(A)$. define x_0 so that $Ax_0 = b$.
- solution set is an affine subspace of dimension $n - \mathrm{Rank}(A)$.

Solution of linear system

matrix $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $m \leq n$ ("fat matrix"). define

- solution set of linear system $\{x : Ax = b\}$

Q: when does a solution exist? what is the dimension of the solution set? when is the solution unique?

- solution exists if $b \in \mathrm{span}(A)$. define x_0 so that $Ax_0 = b$.
- solution set is an affine subspace of dimension $n - \mathrm{Rank}(A)$. proof:
  - **nullspace**(A), is a linear subspace of dimension $n - \mathrm{Rank}(A)$ by rank-nullity theorem
  - solution set is $\{x : Ax = b\} = \{x_0 + v : v \in \textbf{nullspace}(A)\}$

Solution of linear system

matrix $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $m \leq n$ ("fat matrix"). define

- solution set of linear system $\{x : Ax = b\}$

Q: when does a solution exist? what is the dimension of the solution set? when is the solution unique?

- solution exists if $b \in \mathrm{span}(A)$. define x_0 so that $Ax_0 = b$.
- solution set is an affine subspace of dimension $n - \mathrm{Rank}(A)$. proof:
  - **nullspace**(A), is a linear subspace of dimension $n - \mathrm{Rank}(A)$ by rank-nullity theorem
  - solution set is $\{x : Ax = b\} = \{x_0 + v : v \in \textbf{nullspace}(A)\}$
- solution is unique if **nullspace**$A = 0$, *e.g.*, if $m = n$ and A is full rank. in that case, we say A is **invertible** and write $x_0 = A^{-1}b$

Solution of linear system

matrix $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $m \leq n$ ("fat matrix"). define

- solution set of linear system $\{x : Ax = b\}$

Q: when does a solution exist? what is the dimension of the solution set? when is the solution unique?

- solution exists if $b \in \mathrm{span}(A)$. define x_0 so that $Ax_0 = b$.
- solution set is an affine subspace of dimension $n - \mathrm{Rank}(A)$. proof:
  - **nullspace**(A), is a linear subspace of dimension $n - \mathrm{Rank}(A)$ by rank-nullity theorem
  - solution set is $\{x : Ax = b\} = \{x_0 + v : v \in \textbf{nullspace}(A)\}$
- solution is unique if $\textbf{nullspace}A = 0$, *e.g.*, if $m = n$ and A is full rank. in that case, we say A is **invertible** and write $x_0 = A^{-1}b$

if these are confusing:

- review linear algebra and prove them all!
- come to office hours and/or review session this Friday

What next?

- website: https://stanford-cme-307.github.io/web
- Ed for discussion and announcements
- fill out course survey (linked on website)
- talk to instructors after class and during office hours (see website)
- class attendance is required. will post some slides, generally no recordings