Lesson 17. Improving Search: Finding Better Solutions

1 A general optimization model

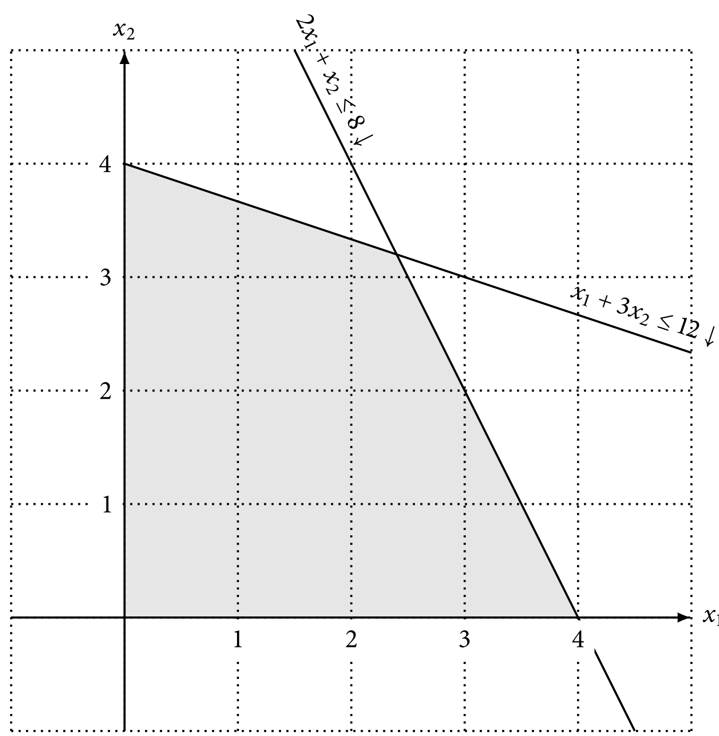
- For the next few lessons, we will consider a general optimization model
- Decision variables: $x_1, \dots, x_n$
  - Recall: a feasible solution to an optimization model is a choice of values for all decision variables that satisfies all constraints
- Easier to refer to a feasible solution as a vector: $\mathbf{x} = (x_1, \dots, x_n)$
- Let $f(\mathbf{x})$ and $g_i(\mathbf{x})$ for $i \in \{1, \dots, m\}$ be multivariable functions in $\mathbf{x}$, not necessarily linear
- Let b_i for $i \in \{1, \dots, m\}$ be constant scalars

$$\begin{aligned} \text{minimize/maximize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \begin{Bmatrix} \leq \\ \geq \\ = \end{Bmatrix} b_i \quad \text{for } i \in \{1, \dots, m\} \end{aligned} \tag{$*$}$$

- Linear programs fit into this framework

Example 1.

$$\begin{aligned} \text{maximize} \quad & 4x_1 + 2x_2 \\ \text{subject to} \quad & x_1 + 3x_2 \leq 12 \quad & (1) \\ & 2x_1 + x_2 \leq 8 \quad & (2) \\ & x_1 \geq 0 \quad & (3) \\ & x_2 \geq 0 \quad & (4) \end{aligned}$$



2 Improving search algorithms, informally

- Idea:
  - Start at a feasible solution
  - Repeatedly move to a "close" feasible solution with better objective function value
- The **neighborhood** of a feasible solution is the set of all feasible solutions "close" to it
  - We can define "close" in various ways to design different types of algorithms
- Let's start formalizing these ideas

3 Locally and globally optimal solutions

- ε-**neighborhood** $N_\varepsilon(\mathbf{x})$ of a solution $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ (where $\varepsilon > 0$):

$$N_\varepsilon(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{y}) \leq \varepsilon\}$$

  where $d(\mathbf{x}, \mathbf{y})$ is the distance between solution $\mathbf{x}$ and $\mathbf{y}$

- A feasible solution $\mathbf{x}$ to optimization model $(*)$ is **locally optimal** if for some value of $\varepsilon > 0$:

$$f(\mathbf{x}) \quad \text{is better than} \quad f(\mathbf{y}) \qquad \text{for all feasible solutions } \mathbf{y} \in N_\varepsilon(\mathbf{x})$$

- A feasible solution $\mathbf{x}$ to optimization model $(*)$ is **globally optimal** if:

$$f(\mathbf{x}) \quad \text{is better than} \quad f(\mathbf{y}) \qquad \text{for all feasible solutions } \mathbf{y}$$

  - Also known simply as an **optimal solution**
- Global optimal solutions are locally optimal, but not vice versa
- In general: harder to check for global optimality, easier to check for local optimality

4 The improving search algorithm

1 Find an initial feasible solution $\mathbf{x}^0$
2 Set $k = 0$
3 **while** $\mathbf{x}^k$ is <u>not</u> locally optimal **do**
4   Determine a new <u>feasible</u> solution $\mathbf{x}^{k+1}$ that <u>improves</u> the objective value at $\mathbf{x}^k$
5   Set $k = k + 1$
6 **end while**

- Generates sequence of feasible solutions $\mathbf{x}^0, \mathbf{x}^1, \mathbf{x}^2, \dots$
- In general, improving search converges to a local optimal solution, not a global optimal solution
- Let's concentrate on line 4 – finding better feasible solutions

5 Moving between solutions

- How do we move from one solution to the next?

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda \mathbf{d}$$

- In Example 1:

6 Improving directions

- We want to choose $\mathbf{d}$ so that $\mathbf{x}^{k+1}$ has a better value than $\mathbf{x}^k$
- $\mathbf{d}$ is an **improving direction** at solution $\mathbf{x}^k$ if

$$f(\mathbf{x}^k + \lambda \mathbf{d}) \quad \text{is better than} \quad f(\mathbf{x}^k) \qquad \text{for all positive } \lambda \text{ “close” to } 0$$

- How do we find an improving direction?
- The **directional derivative** of f in the direction $\mathbf{d}$ at solution $\mathbf{x}^k$ is

- Maximizing f: $\mathbf{d}$ is an improving direction at $\mathbf{x}^k$ if
- Minimizing f: $\mathbf{d}$ is an improving direction at $\mathbf{x}^k$ if
- In Example 1:

- For linear programs in general: if $\mathbf{d}$ is an improving direction at $\mathbf{x}^k$, then $f(\mathbf{x}^k + \lambda \mathbf{d})$ improves as $\lambda \to \infty$

7 Step size

- We have an improving direction $\mathbf{d}$ – now how far do we go?
- One idea: find maximum value of λ so that $\mathbf{x}^k + \lambda\mathbf{d}$ is still feasible
- Graphically, we can eyeball this
- Algebraically, we can compute this – in Example 1:

8 Feasible directions

- Some improving directions don't lead to any new feasible solutions
- $\mathbf{d}$ is a **feasible direction** at feasible solution $\mathbf{x}^k$ if $\mathbf{x}^k + \lambda\mathbf{d}$ is feasible for all positive λ "close" to 0
- Again, graphically, we can eyeball this
- A constraint is **active** at feasible solution $\mathbf{x}$ if it is satisfied with equality
- For linear programs:
  - We have constraints of the form:

$$
\begin{aligned}
a_1x_1 + a_2x_2 + \cdots + a_nx_n &\leq b \\
a_1x_1 + a_2x_2 + \cdots + a_nx_n &\geq b \\
a_1x_1 + a_2x_2 + \cdots + a_nx_n &= b
\end{aligned}
$$

  - We can rewrite these constraints using vector notation:

- $\mathbf{d}$ is a feasible direction at $\mathbf{x}$ if
    - $\mathbf{a}^\top \mathbf{d} \leq 0$ for each active constraint of the form $\mathbf{a}^\top \mathbf{x} \leq b$
    - $\mathbf{a}^\top \mathbf{d} \geq 0$ for each active constraint of the form $\mathbf{a}^\top \mathbf{x} \geq b$
    - $\mathbf{a}^\top \mathbf{d} = 0$ for each active constraint of the form $\mathbf{a}^\top \mathbf{x} = b$

- In Example 1:

9 Detecting unboundedness

- Suppose $\mathbf{d}$ is an improving direction at feasible solution $\mathbf{x}^k$ to a linear program
- Also, suppose $\mathbf{x}^k + \lambda \mathbf{d}$ is feasible for all $\lambda \geq 0$
- What can you conclude?

10 Summary

- Line 4 boils down to finding an improving and feasible direction $\mathbf{d}$ and an accompanying step size λ
- We discussed conditions on whether a direction is improving and feasible
- We don't know how to systematically find such directions... yet