Lesson 13. Improving Search: Finding Better Solutions

1 A general optimization model

- For now, we will consider a general optimization model
- Decision variables: $x_1, \ldots, x_n$
    - 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: $\overset{\vec{x}}{\mathbf{x}} = (x_1, \ldots, x_n)$
- Let $f(\mathbf{x})$ and $g_i(\mathbf{x})$ for $i \in \{1, \ldots, m\}$ be multivariable functions in $\mathbf{x}$, not necessarily linear
- Let b_i for $i \in \{1, \ldots, m\}$ be constant scalars

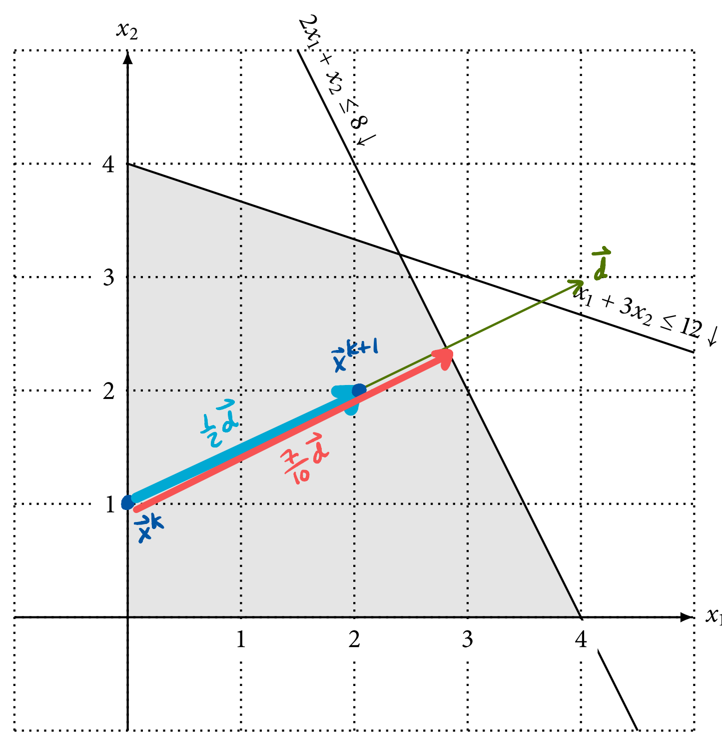
$$
\begin{aligned}
\text{minimize/maximize} \quad & f(\mathbf{x}) \\
\text{subject to} \quad & g_i(\mathbf{x}) \begin{Bmatrix} \le \\ \ge \\ = \end{Bmatrix} b_i \quad \text{for } i \in \{1, \ldots, m\}
\end{aligned}
\tag{$*$}
$$

- Linear programs fit into this framework

Example 1.

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$
\begin{aligned}
\text{maximize} \quad & \overbrace{4x_1 + 2x_2}^{f(\vec{x})} \\
\text{subject to} \quad & \underbrace{x_1 + 3x_2}_{g_1(\vec{x})} \le \underbrace{12}_{b_1} && (1) \\
& \underbrace{2x_1 + x_2}_{g_2(\vec{x})} \le \underbrace{8}_{b_2} && (2) \\
& \underbrace{x_1}_{g_3(\vec{x})} \ge \underbrace{0}_{b_3} && (3) \\
& \underbrace{x_2}_{g_4(\vec{x})} \ge \underbrace{0}_{b_4} && (4)
\end{aligned}
$$

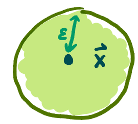


2 Improving search algorithms, informally

- 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 $\varepsilon > 0$):



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

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

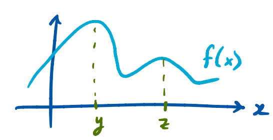
- A feasible solution $\mathbf{x}$ to optimization model ($*$) is **locally optimal** if for some value of $\varepsilon > 0$:

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

- 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

$\vec{x}^{k+1} = \vec{x}^k + \lambda \vec{d}$ ← improving + feasible

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?

$$\underbrace{\mathbf{x}^{k+1}}_{\text{new solution}} = \underbrace{\mathbf{x}^k}_{\text{current solution}} + \underbrace{\lambda}_{\text{step size}} \underbrace{\mathbf{d}}_{\text{direction}}$$

- In Example 1:

Let $\vec{x}^k = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ $\quad \lambda = \frac{1}{2}$ $\quad \vec{d} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$

$$\Rightarrow \vec{x}^{k+1} = \vec{x}^k + \lambda \vec{d} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

6 Improving directions

$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ $\quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$

$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2$

$\vec{a}^T \vec{b} = a_1 b_1 + a_2 b_2$

- 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

$$\frac{\nabla f(\vec{x}^k)^T \vec{d}}{\|\vec{d}\|} = \text{rate of change in } f \text{ at } \vec{x}^k \text{ in the direction } \vec{d}$$

- Maximizing f: $\mathbf{d}$ is an improving direction at $\mathbf{x}^k$ if $\boxed{\nabla f(\vec{x}^k)^T \vec{d} > 0}$
- Minimizing f: $\mathbf{d}$ is an improving direction at $\mathbf{x}^k$ if $\boxed{\nabla f(\vec{x}^k)^T \vec{d} < 0}$
- In Example 1:

Is $\vec{d} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$ improving at $\vec{x}^k = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$?

$f(\vec{x}) = 4x_1 + 2x_2 \Rightarrow \nabla f(\vec{x}) = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$ (for any $\vec{x}$!)

$\Rightarrow \nabla f(\vec{x}^k)^T \vec{d} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}^T \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 20 > 0$

$\begin{pmatrix} 4 \\ 2 \end{pmatrix}^T \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \underbrace{\begin{pmatrix} 4 & 2 \end{pmatrix}}_{1\times 2} \underbrace{\begin{pmatrix} 4 \\ 2 \end{pmatrix}}_{2\times 1}$

$= 4(4) + 2(2)$

$= 20$

$\Rightarrow$ Yes, $\vec{d}$ is improving at $\vec{x}^k$

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

You can go in the direction $\vec{d}$ from $\vec{x}^k$ as far as you want and still improve

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:

For what values of λ is $\vec{x}^k + \lambda \vec{d} = \begin{pmatrix} 4\lambda \\ 1+2\lambda \end{pmatrix}$ feasible?

our example $\rightarrow$ $\vec{x}^k = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\vec{d} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$

$$
\begin{array}{llll}
\text{maximize} & \overbrace{4x_1 + 2x_2}^{f(\vec{x})} & & \\
\text{subject to} & \underbrace{x_1 + 3x_2}_{g_1(\vec{x})} \le \underbrace{12}_{b_1} & & (1) \\
& \underbrace{2x_1 + x_2}_{g_2(\vec{x})} \le \underbrace{8}_{b_2} & & (2) \\
& \underbrace{x_1}_{g_3(\vec{x})} \ge \underbrace{0}_{b_3} & & (3) \\
& \underbrace{x_2}_{g_4(\vec{x})} \ge \underbrace{0}_{b_4} & & (4)
\end{array}
$$

Check all the constraints:

(1) $x_1 + 3x_2 \le 12$
$\iff 4\lambda + 3(1+2\lambda) \le 12$
$\iff 10\lambda + 3 \le 12$
$\iff 10\lambda \le 9$
$\iff \lambda \le \frac{9}{10}$

(2) $2x_1 + x_2 \le 8$
$\iff 2(4\lambda) + (1+2\lambda) \le 8$
$\iff 10\lambda + 1 \le 8$
$\iff 10\lambda \le 7$
$\iff \lambda \le \frac{7}{10}$

(3) $x_1 \ge 0$
$\iff 4\lambda \ge 0$
$\iff \lambda \ge 0$

(4) $x_2 \ge 0$
$\iff 1 + 2\lambda \ge 0$
$\iff \lambda \ge -\frac{1}{2}$

$\Rightarrow$ $\vec{x}^k + \lambda\vec{d}$ is feasible when $\lambda \in \left[0, \frac{7}{10}\right]$

$\Rightarrow$ maximum step size: $\lambda = \frac{7}{10}$

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:

$$
\begin{aligned}
a_1x_1 + a_2x_2 + \cdots + a_nx_n &\le b \\
a_1x_1 + a_2x_2 + \cdots + a_nx_n &\ge b \\
a_1x_1 + a_2x_2 + \cdots + a_nx_n &= b
\end{aligned}
$$

  - We can rewrite these constraints using vector notation:

Let: $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ $\quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ $\quad \Rightarrow \quad \begin{aligned} \vec{a}^T\vec{x} &\le b \\ \vec{a}^T\vec{x} &\ge b \\ \vec{a}^T\vec{x} &= b \end{aligned}$

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

If there are no active constraints at a feasible solution $\vec{x}$, then any direction is feasible.

- In Example 1:

$$
\begin{aligned}
\text{maximize} \quad & \overbrace{4x_1 + 2x_2}^{f(\vec{x})} \\
\text{subject to} \quad & \underbrace{x_1 + 3x_2}_{g_1(\vec{x})} \le \underbrace{12}_{b_1} \quad (1) \\
& \underbrace{2x_1 + x_2}_{g_2(\vec{x})} \le \underbrace{8}_{b_2} \quad (2) \\
& \underbrace{x_1}_{g_3(\vec{x})} \ge \underbrace{0}_{b_3} \quad (3) \\
& \underbrace{x_2}_{g_4(\vec{x})} \ge \underbrace{0}_{b_4} \quad (4)
\end{aligned}
$$

Is $\vec{d} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$ a feasible direction at $\vec{x}^k = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$?

Active constraints?

(1) $0 + 3(1) \le 12$ not active

(2) $2(0) + (1) \le 8$ not active

(3) $0 \ge 0$ active

(4) $1 \ge 0$ not active

(3) is of the form $\underbrace{\vec{a}^T\vec{x}}_{\begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} \ge 0$ $\Rightarrow$ need to check if $\underbrace{\vec{a}^T\vec{d}}_{\begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} 4 \\ 2 \end{pmatrix}} \ge 0$ $\Rightarrow$ Yes!

$\Rightarrow \vec{d}$ is feasible at $\vec{x}^k$.

9 Detecting unboundedness → improve obj. fn. value infinitely.

- 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?

LP is unbounded : $f(\vec{x}^k + \lambda\vec{d})$ improves and $\vec{x}^k + \lambda\vec{d}$ remains feasible as $\lambda \to \infty$

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