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: $x_1, \ldots, x_n$
  - Recall: a feasible solution to an optimization model is a choice of values for <u>all</u> decision variables that satisfies all constraints
- Easier to refer to a feasible solution as a <u>vector</u>: $\mathbf{x} = (x_1, \ldots, x_n)$
- 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

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

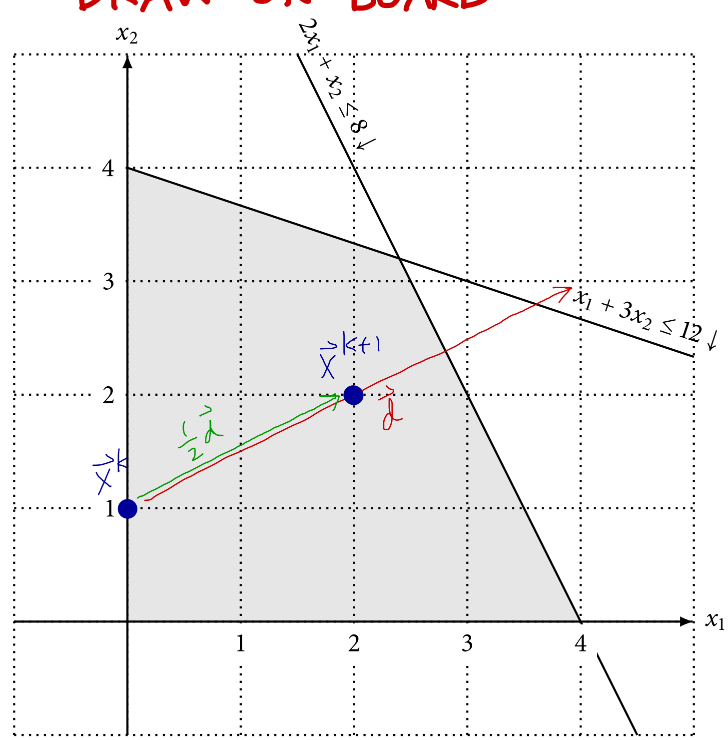
- <u>Linear programs fit into this framework</u>

DRAW ON BOARD

Example 1.

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

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



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 $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ (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 not locally optimal **do**
4   Determine a new feasible solution $\mathbf{x}^{k+1}$ that improves the objective value at $\mathbf{x}^k$
5   Set $k = k + 1$
6 **end while**

- Generates sequence of feasible solutions $\mathbf{x}^0, \mathbf{x}^1, \mathbf{x}^2, \ldots$
- 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?

$$\text{next solution} \longrightarrow \mathbf{x}^{k+1} = \underset{\substack{\uparrow \\ \text{current solution}}}{\mathbf{x}^k} + \underset{\substack{\uparrow \\ \text{step size}}}{\lambda}\mathbf{d} \longleftarrow \text{direction}$$

- In Example 1:

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

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

6 Improving directions

- 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 $\vec{x}^k$ is

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

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

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

$\nabla f(\vec{x}^k)^T \vec{d} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}^T \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 16 + 4 = 20 > 0$

$\Rightarrow$ Yes, $\vec{d} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$ is improving at $\vec{x}^k = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

- 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 **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 $\vec{x}^k + \lambda\vec{d} = \begin{pmatrix} 4\lambda \\ 1+2\lambda \end{pmatrix}$ still feasible?

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

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

(3) $x_1 \geq 0$
$\Leftrightarrow 4\lambda \geq 0$
$\Leftrightarrow \lambda \geq 0$

(4) $x_2 \geq 0$
$\Leftrightarrow 1+2\lambda \geq 0$
$\Leftrightarrow \lambda \geq -\frac{1}{2}$

$\Rightarrow$ $\vec{x}^k + \lambda\vec{d}$ is feasible when $0 \leq \lambda \leq \frac{7}{10}$

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

8 Feasible directions

- Some improving directions don't lead to any new feasible solutions
- **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 **x** if it is satisfied with equality
- For linear programs:
  - We have constraints of the form:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq b$$
$$a_1x_1 + a_2x_2 + \cdots + a_nx_n \geq b$$
$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

  - We can rewrite these constraints using vector notation:

Let $\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ $\Rightarrow$ $\vec{a}^T\vec{x} \leq b$, $\vec{a}^T\vec{x} \geq b$, $\vec{a}^T\vec{x} = b$

- ○ **d** is a feasible direction at **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:

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

Active constraints at $\vec{x}^k$: $x_1 \geq 0$

$\Rightarrow$ We need to check if $d_1 \geq 0$ – Yes!

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

9 Detecting unboundedness

- Suppose **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 **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