Solving Linear Programs: The Basics

- Introduction to convex analysis: convexity, convex combination, hyperplanes, half-spaces, extreme points
- Convex and concave functions, the gradient, local and global optima, the fundamental theorem of convex programming
- Convex geometry: polyhedra, the Minkowski-Weyl theorem (the Representation Theorem)
- Solving linear programs using the Minkowski-Weyl theorem, the relation of optimal feasible solutions and extreme points
- Solving simple linear programs with the graphical method
- The feasible region (bounded, unbounded, empty) and optimal solutions (unique, alternative, unbounded)

Convex Sets

• For each $0 \le \lambda \le 1$, the expression

$$\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$$

is called the convex combination of vectors $m{x}_1$ and $m{x}_2$

• Geometrically, the convex combinations of x_1 and x_2 span the **line segment** between point x_1 and x_2



Convex Sets

• A set $X \subset \mathbb{R}^n$ is **convex** if for each points x_1 and x_2 in X it holds that

$$\forall \lambda \in [0,1] : \lambda \boldsymbol{x}_1 + (1-\lambda) \boldsymbol{x}_2 \in X$$

• In other words, X is convex if it contains all convex combinations of each of its points



Convex Sets: Examples

• The convex combinations of k points $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$:

$$X = \left\{ \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i : \sum_{i=1}^{k} \lambda_i = 1, \forall i \in \{1, \dots, k\} : \lambda_i \ge 0 \right\}$$
$$X = \operatorname{conv} \{ \boldsymbol{x}_i : 1 \le i \le k \}$$

- The 3-sphere: $X = \{ [x, y, z] : x^2 + y^2 + x^2 \le 1 \}$
- Vector space: $X = \{ x : Ax = 0 \}$
- Affine space (translated vector space): $X = \{ x : Ax = b \}$
- Feasible region of a linear program:

$$X = \{ \boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0} \}$$
$$X = \{ \boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \}$$

Convex and Concave Functions

• A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is **convex** on a convex set $X \subseteq \mathbb{R}^n$ if for each x_1 and x_2 in X:

 $f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \le \lambda f(\boldsymbol{x}_1) + (1-\lambda)f(\boldsymbol{x}_2) \quad \forall \lambda \in [0,1]$

- The line segment between any two points $f(x_1)$ and $f(x_2)$ on the graph of the function lies above or on the graph
- Function f is **concave** if (-f) is convex



Convex and Concave Functions

• The set of points "above" the graph of some function $f: \mathbb{R}^n \mapsto \mathbb{R}$ is called the **epigraph** of f

 $epi(f) = \{(\boldsymbol{x}, y) : \boldsymbol{x} \in \mathbb{R}^n, y \in \mathbb{R}, y \ge f(\boldsymbol{x})\}$

• Simple convexity check: f is convex on set \mathbb{R}^n if and only if ${\rm epi}(f)$ is convex



Optimization on a Convex Set

- Given function $f : \mathbb{R}^n \mapsto \mathbb{R}$ and set X, solve the generic optimization problem $\max f(x) : x \in X$
- Some $\bar{x} \in X$ is a global optimal solution (or global optimum) if for each $x \in X$: $f(\bar{x}) \ge f(x)$
- An $\bar{x} \in X$ is a local optimum if there is a neighborhood $N_{\epsilon}(\bar{x})$ of \bar{x} (an open ball of radius $\epsilon > 0$ with centre \bar{x}) so that $\forall x \in N_{\epsilon}(\bar{x}) \cap X$: $f(\bar{x}) \ge f(x)$

point *A* is a local optimum and point *B* is a global optimum on the closed interval $[x_1, x_2]$



Optimization on a Convex Set

- Theorem: Let X be a nonempty convex set in ℝⁿ and let f: ℝⁿ → ℝ be a concave function on X. Consider the optimization problem max f(x) : x ∈ X. Then, if x̄ ∈ X is a local optimal solution then it is also a global optimum
- **Proof:** $\bar{x} \in X$ is a local optimum, therefore there is nonempty neighborhood $N_{\epsilon}(\bar{x})$, $\epsilon > 0$ so that

$$\forall \boldsymbol{x} \in N_{\epsilon}(\bar{\boldsymbol{x}}) \cap X : f(\bar{\boldsymbol{x}}) \ge f(\boldsymbol{x})$$

- Suppose that \bar{x} is not a global optimal solution and therefore there exists $\hat{x} \in X : f(\hat{x}) > f(\bar{x})$
- Since X is convex it contains all convex combinations of \hat{x} and \bar{x} (i.e., the line segment between \hat{x} and \bar{x})
- Also, since $N_{\epsilon}(\bar{x})$ is nonempty it contains points different from \bar{x} that are also on the line segment between \hat{x} and \bar{x}

Optimization on a Convex Set

- Since f is concave, the graph of f lies above, or on the line segment joining \hat{x} and \bar{x}

$$\forall \lambda \in (0,1] : f(\lambda \hat{\boldsymbol{x}} + (1-\lambda)\bar{\boldsymbol{x}}) \ge \lambda f(\hat{\boldsymbol{x}}) + (1-\lambda)f(\bar{\boldsymbol{x}})$$

$$\overset{*}{>} \lambda f(\bar{\boldsymbol{x}}) + (1-\lambda)f(\bar{\boldsymbol{x}}) = f(\bar{\boldsymbol{x}})$$

- The strict inequality (marked with * above) comes from the assumption that \bar{x} is not a global optimum: $f(\hat{x}) > f(\bar{x})$
- Consequently, all neighborhoods of \bar{x} contain feasible points $x \neq \bar{x}$ for which $f(x) > f(\bar{x})$, which contradicts our assumption that \bar{x} is a local optimal solution
- Corollary: Let X be a nonempty convex set in \mathbb{R}^n , let $f: \mathbb{R}^n \to \mathbb{R}$ be a convex function on X, and let $\bar{x} \in X$ be a local optimal solution to the optimization problem $\min f(x) : x \in X$. Then, \bar{x} is also a global optimal solution

The Gradient

Given a set X ⊆ ℝⁿ and function f : ℝⁿ → ℝ. We call f differentiable at a point x̄ ∈ X if there exits ∇f(x̄) gradient vector and function α : ℝⁿ → ℝ so that for all x ∈ X:

$$f(\boldsymbol{x}) = f(\bar{\boldsymbol{x}}) + \nabla f(\bar{\boldsymbol{x}})^T (\boldsymbol{x} - \bar{\boldsymbol{x}}) + \|\boldsymbol{x} - \bar{\boldsymbol{x}}\| \alpha(\bar{\boldsymbol{x}}, \boldsymbol{x} - \bar{\boldsymbol{x}}) ,$$

where $\lim_{\boldsymbol{x}\to\bar{\boldsymbol{x}}}\alpha(\bar{\boldsymbol{x}},\boldsymbol{x}-\bar{\boldsymbol{x}})=0$

- "Linearization" by approximating with the first-order Taylor series: $f(\boldsymbol{x}) \approx f(\bar{\boldsymbol{x}}) + \nabla f(\bar{\boldsymbol{x}})^T (\boldsymbol{x} \bar{\boldsymbol{x}})$
- If *f* is differentiable then the gradient is well-defined and can be written as the vector of the partial derivatives:

$$\nabla f(\boldsymbol{x})^T = \begin{bmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} & \frac{\partial f(\boldsymbol{x})}{\partial x_2} & \dots & \frac{\partial f(\boldsymbol{x})}{\partial x_n} \end{bmatrix}$$

Hyperplanes and Half-spaces

- Hyperplane: all $x \in \mathbb{R}^n$ satisfying the equation $a^T x = b$ for some a^T row *n*-vector (the **normal** vector) and scalar *b*
- The hyperplane $X = \{ \boldsymbol{x} : \boldsymbol{a}^T \boldsymbol{x} = b \}$ divides the space \mathbb{R}^n into two half-spaces
 - "lower" half-space: $\{ \boldsymbol{x} : \boldsymbol{a}^T \boldsymbol{x} \leq b \}$
 - "upper" half-space: $\{ \boldsymbol{x} : \boldsymbol{a}^T \boldsymbol{x} > b \}$
- Hyperplanes and half-spaces are convex



Extreme Points

Given a convex set X, a point x ∈ X is called an extreme point of X if x cannot be obtained as the convex combination of two points in X different from x:

$$oldsymbol{x} = \lambda oldsymbol{x}_1 + (1-\lambda) oldsymbol{x}_2$$
 and $0 \leq \lambda \leq 1 \Rightarrow oldsymbol{x_1} = oldsymbol{x_2} = oldsymbol{x}$



- x_1 and x_4 are extreme points, x_2 and x_3 are not
- extreme points correspond to the "corner points" of a convex set

Polyhedra

• A polyhedron is a geometric object with "flat" sides



• By the word "polyhedron" we will usually mean a "convex polyhedron"

Convex Polyhedra

• **Definition 1:** the intersection of finitely many (closed) half-spaces

$$X = \{ \boldsymbol{x} : \boldsymbol{a}_i \boldsymbol{x} \le b_i, i \in \{1, \dots, m\} \} = \{ \boldsymbol{x} : \boldsymbol{A} \boldsymbol{x} \le \boldsymbol{b} \}$$

- Corollary: the feasible region of a linear program forms a convex polyhedron
 - \circ canonical form: $\max\{ \boldsymbol{c}^T \boldsymbol{x} : \boldsymbol{A} \boldsymbol{x} \leq \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0} \}$
 - \circ standard form: $\max\{ \boldsymbol{c}^T \boldsymbol{x} : \boldsymbol{A} \boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0} \}$
- **Definition 2:** convex combinations of finitely many points

$$X = \left\{ \sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i : \sum_{i=1}^{n} \lambda_i = 1, \forall i \in \{1, \dots, n\} : \lambda_i \ge 0 \right\}$$

The Minkowski-Weyl Theorem

- The Representation Theorem of Bounded Polyhedra: the two definitions are equivalent
- The Strong Minkowski-Weyl Theorem: if the intersection of finitely many half-spaces is bounded then it can be written as the convex combination of finitely many extreme points

$$P = \{ \boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \} \Leftrightarrow P = \operatorname{conv} \{ \boldsymbol{x}_j : 1 \le j \le k \}$$



The Minkowski-Weyl Theorem: Example



• Half-space intersection:

$$X = \{ x_1, x_2 : x_1 + x_2 \leq 6 \\ x_2 \leq 3 \\ x_1, x_2 \geq 0 \}$$

 Convex combination of extreme points:

$$X = \operatorname{conv}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}0\\3\end{bmatrix}, \begin{bmatrix}3\\3\end{bmatrix}, \begin{bmatrix}6\\0\end{bmatrix}\right)$$

Linear Programs and Extreme Points

- The Fundamental Theorem of Linear Programming: if the feasible region of a linear program is bounded then the at least one optimal solution is guaranteed to occur at an extreme point of the feasible region
- Proof: given the below linear program

$$\begin{array}{ll} \max & \boldsymbol{c}^T \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A} \boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{array}$$

• May also be cast in the form $\max\{ \boldsymbol{c}^T \boldsymbol{x} : \boldsymbol{A}' \boldsymbol{x} \leq \boldsymbol{b}' \}$

$$\begin{array}{ll} \max & \boldsymbol{c}^T \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A} \boldsymbol{x} \leq \boldsymbol{b} \\ & -\boldsymbol{A} \boldsymbol{x} \leq -\boldsymbol{b} \\ & -\boldsymbol{x} \leq \boldsymbol{0} \end{array}$$

Lineáris Programs and Extreme Points

- The feasible region $X = \{ x : Ax = b, x \ge 0 \}$ of any linear program is a polyhedron
- Suppose now that X is bounded and let the extreme points of X be x_1, x_2, \ldots, x_k
- Using the strong Minkowski-Weyl theorem, X can be written as the convex combination of its extreme points

Lineáris Programs and Extreme Points

- Substitute the above \boldsymbol{x} into the linear program
- Observe that *x* automatically satisfies the constraints!
- The objective function: $\max c^T (\sum_{j=1}^k \lambda_j x_j) = \max \sum_{j=1}^k (c^T x_j) \lambda_j$
- We have obtained another linear program in which the variables are now the scalars λ_j

$$\max \sum_{j=1}^{k} (\boldsymbol{c}^{T} \boldsymbol{x}_{j}) \lambda_{j}$$

s.t.
$$\sum_{j=1}^{k} \lambda_{j} = 1$$
$$\lambda_{j} \ge 0 \qquad \forall j \in \{1, \dots, k\}$$

Lineáris Programs and Extreme Points

- Our goal is to obtain the maximum of the linear program $\sum_{j=1}^{k} (\boldsymbol{c}^T \boldsymbol{x}_j) \lambda_j : \sum_{j=1}^{k} \lambda_j = 1, \lambda_j \ge 0$
- It is easy to see that it is enough to find the extreme point for which $c^T x_j$ is maximal, let this extreme point be x_p
- Then, $\lambda_p = 1$, $\lambda_j = 0$: $j \neq p$ solves the new linear program and the optimal objective function value is $c^T x_p$.
- We have proved only for the bounded case, but the theorem also holds for the linear programs with unbounded feasible region
- In general a linear program cannot be written in the equivalent using extreme points (which would then be easy to solve), since the number of extreme points is exponential
- In lower dimensions it can: graphical method

Extreme points: Example



max	x_1	+	$2x_2$		
s.t.	x_1	+	x_2	\leq	6
			x_2	\leq	3
	$x_1,$		x_2	\geq	0

• Extreme points: $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \end{bmatrix}$

Extreme points: Example



• Compute the objective function value $c^T x_j$ for each extreme point x_j :

 $\boldsymbol{c}^T \boldsymbol{x}_1 = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{vmatrix} 0 \\ 0 \end{vmatrix} = 0$ $\boldsymbol{c}^T \boldsymbol{x}_2 = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{vmatrix} 0 \\ 3 \end{vmatrix} = 6$ $\boldsymbol{c}^T \boldsymbol{x}_3 = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{vmatrix} 3 \\ 3 \end{vmatrix} = \begin{bmatrix} 9 \end{bmatrix}$ $\boldsymbol{c}^T \boldsymbol{x}_4 = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{vmatrix} 6 \\ 0 \end{vmatrix} = 6$

Optimal Resource Allocation Revisited

- **Exercise:** a paper mill manufactures two types of paper, standard and deluxe
 - $\circ~\frac{1}{2}~m^3$ of wood is needed to manufacture 1 m^2 of paper (both standard or deluxe)
 - producing 1 m² of standard paper takes 1 man-hour, whereas 1 m² of deluxe paper requires 2 man-hours
 - every week 40 m³ wood and 100 man-hours of workforce is available
 - the profit is 3 thousand USD per 1 m² of standard paper and 4 thousand USD per 1 m² of deluxe paper
- **Question:** how much standard and how much deluxe paper should be produced by the paper mill per week to maximize profits?













A Problem of Logistics

- Transshipment problem:
 - \circ from point A to point C
 - 15 tonnes of goods
 - ship capacities are 10 tonnes each
 - cost ([thousand USD per tonne]): $c_{AB} = 2$, $c_{BC} = 4$, $c_{AC} = 3$
 - objective: minimize costs



A Problem of Logistics X_{AC} 15 🕂 min $2x_{AB} + 4x_{BC} + 3x_{AC}$ $-x_{AB} - x_{AC} = -15$ s.t. 10 $-x_{AB} + x_{BC} = 0$ $x_{BC} + x_{AC} = 15$ $0 \le x_{AB} \le 10$ 5 $0 \le x_{BC} \le 10$ $0 \le x_{AC} \le 10$ X BC 5 10 15 5 10 5 AB – p. 31











– p. 36

• The furthest point along the inverse(!) of the normal vector $\begin{bmatrix} 2 & 4 & 3 \end{bmatrix}^T$ (inverse as the the optimization is minimization)

 $x_{AB} = 5$ $x_{BC} = 5$ $x_{AC} = 10$



The Feasible Region



The Feasible Region



$$X = \{ [x_1, x_2] : -x_1 + 2x_2 \le 8 \\ x_1, x_2 \ge 0 \}$$

Unbounded: there is point $x \in X$ and vector d so that the **ray** from point x along direction d lies entirely within the feasible region:

 $\forall \lambda > 0 : \boldsymbol{x} + \lambda \boldsymbol{d} \in X$

The Feasible Region





max	$3x_1 + x_2$
s.t.	$x_1 + x_2 \le 6$
	$-x_1 + 2x_2 \le 8$
	$x_1, x_2 \ge 0$

Unique optimal solution on **bounded** feasible region: the objective is maximized at a unique point



 $\begin{array}{ll} \max & -x_1 + x_2 \\ \text{s.t.} & -x_1 + 2x_2 \leq 8 \\ & x_1, x_2 \geq 0 \end{array}$

Unique optimal solution on an **unbounded** feasible region



max	$-x_1 + 2x_2$
s.t.	$x_1 + x_2 \le 6$
	$-x_1 + 2x_2 \le 8$
	$x_1, x_2 \ge 0$

Alternative optimal solutions on a **bounded** feasible region: the objective function attains its maximum at more than one distinct points



$$\begin{array}{ll} \max & -x_1 + 2x_2 \\ \text{s.t.} & -x_1 + 2x_2 \le 8 \\ & x_1, x_2 \ge 0 \end{array}$$

Alternative optimal solutions on an **unbounded** feasible region



$$\begin{array}{ll} \max & x_1 + x_2 \\ \text{s.t.} & -x_1 + 2x_2 \le 8 \\ & x_1, x_2 \ge 0 \end{array}$$

Unbounded optimal solution: the objective function value can be increased without limit inside the feasible region