MATH11111: Fundamentals of Optimization **Fall 2022**

Lecture 5 Introduction to Convex Optimization

Lecturer: E. Alper Yıldırım *Week: 2*

5.1 Outline

- Convex Optimization
- Connection with Linear Programming
- Properties of Convex Optimization
- Review Problems

5.2 Quick Review of Lecture 4

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and let $\alpha \in \mathbb{R}$.

- $\mathcal{L}_\alpha(f) = \{x \in \mathbb{R}^n : f(x) = \alpha\}$ is the level set of f.
- $\mathcal{L}_\alpha^-(f) = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$ is the sublevel set of f.
- $\mathcal{L}_\alpha^+(f) = \{x \in \mathbb{R}^n : f(x) \geq \alpha\}$ is the superlevel set of f.
- Sublevel sets of convex functions and superlevel sets of concave functions are convex sets.
- A linear function is both convex and concave.
- Level sets of linear functions (hyperplanes) and sublevel and superlevel sets of linear functions (halfspaces) are convex sets.

5.3 Relation with Constrained Optimization

Recall our generic constrained optimization problem:

$$
\begin{array}{llrcll}
\text{(P)} & \min & f(x) & & & \\
 & \text{s.t.} & & & & \\
 & & g_i(x) & \geq & b_i, & i \in M_1, \\
 & & \ell_i(x) & \leq & b_i, & i \in M_2, \\
 & & h_i(x) & = & b_i, & i \in M_3,
\end{array}
$$

where M_1, M_2, and M_3 are finite index sets; $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function; $g_i : \mathbb{R}^n \to \mathbb{R},\ i \in M_1$; $\ell_i : \mathbb{R}^n \to \mathbb{R},\ i \in M_2$; and $h_i : \mathbb{R}^n \to \mathbb{R},\ i \in M_3$. Each of the functional relations $g_i(x) \geq b_i,\ i \in M_1$; $\ell_i(x) \leq b_i,\ i \in M_2$; $h_i(x) = b_i,\ i \in M_3$ is a constraint.

Remark 5.1. *The feasible region $\mathcal{S} \subseteq \mathbb{R}^n$ of (P) is given by*

$$\mathcal{S} = \{x \in \mathbb{R}^n : g_i(x) \geq b_i,\ i \in M_1; \quad \ell_i(x) \leq b_i,\ i \in M_2; \quad h_i(x) = b_i,\ i \in M_3\}.$$

Therefore, $\mathcal{S} \subseteq \mathbb{R}^n$ is given by the intersection of each of (i) the superlevel set $\mathcal{L}^+_{b_i}(g_i)$ for each $i \in M_1$; (ii) the sublevel set $\mathcal{L}^-_{b_i}(\ell_i)$ for each $i \in M_2$; and (iii) the level set $\mathcal{L}_{b_i}(h_i)$ for each $i \in M_3$, i.e.,

$$\mathcal{S} = \left(\bigcap_{i \in M_1} \mathcal{L}^+_{b_i}(g_i)\right) \bigcap \left(\bigcap_{i \in M_2} \mathcal{L}^-_{b_i}(g_i)\right) \bigcap \left(\bigcap_{i \in M_3} \mathcal{L}_{b_i}(h_i)\right).$$

Therefore, understanding of superlevel, sublevel, and level sets of real-valued functions is fundamental in understanding the geometry of the feasible region $\mathcal{S} \subseteq \mathbb{R}^n$.

5.4 Convex Optimization

Recall our generic constrained optimization problem:

$$\begin{array}{llrcll}
\text{(P)} & \min & f(x) & & & \\
& \text{s.t.} & & & & \\
& & g_i(x) & \geq & b_i, & i \in M_1, \\
& & \ell_i(x) & \leq & b_i, & i \in M_2, \\
& & h_i(x) & = & b_i, & i \in M_3,
\end{array}$$

where M_1, M_2, and M_3 are finite index sets; $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function; $g_i : \mathbb{R}^n \to \mathbb{R},\ i \in M_1$; $\ell_i : \mathbb{R}^n \to \mathbb{R},\ i \in M_2$; and $h_i : \mathbb{R}^n \to \mathbb{R},\ i \in M_3$.

Definition 5.1. *(P) is called a* convex optimization problem *if (i) f is a convex function; (ii) g_i is a concave function for each $i \in M_1$; (iii) ℓ_i is a convex function for each $i \in M_2$; and (iv) h_i is a linear function for each $i \in M_3$. Otherwise, (P) is called a* nonconvex optimization problem.

Remark 5.2. *Recall that (P) is a linear programming problem if each of f; $g_i,\ i \in M_1$; $\ell_i,\ i \in M_2$; and $h_i,\ i \in M_3$ is a linear function. Since every linear function is both convex and concave by Proposition 4.3, it follows that a linear programming problem is a convex optimization problem.*

5.4.1 Properties of Convex Optimization Problems

Proposition 5.1. *Let (P) be a convex optimization problem. Then, each of the feasible region $\mathcal{S} \subseteq \mathbb{R}^n$ and the set of optimal solutions $\mathcal{S}^* \subseteq \mathbb{R}^n$ is a convex set.*

Proof. If $\mathcal{S} = \emptyset$, then it is convex. Otherwise, each of the level sets of linear functions, sublevel sets of convex functions, and superlevel sets of concave functions is a convex set by Proposition 4.1, Proposition 4.2, and Corollary 4.5, respectively. Since convexity is preserved under taking intersections (see Remark 3 in Section 3.2), $\mathcal{S}$ is a convex set.

If $\mathcal{S}^* = \emptyset$, then it is convex. Otherwise, let $z^* \in \mathbb{R}$ denote the optimal value. For any $x^1 \in \mathcal{S}^*$, $x^2 \in \mathcal{S}^*$, and any $\lambda \in [0,1]$, note that $\lambda x^1 + (1-\lambda)x^2 \in \mathcal{S}$ since $\mathcal{S}$ is a convex set. Furthermore, by definition of the optimal value and the convexity of f,

$$z^* \leq f\left(\lambda x^1 + (1-\lambda)x^2\right) \leq \lambda f(x^1) + (1-\lambda) f(x^2) = \lambda z^* + (1-\lambda) z^* = z^*.$$

Therefore, $\lambda x^1 + (1-\lambda)x^2 \in \mathcal{S}^*$ and $\mathcal{S}^*$ is a convex set. $\square$

Remarks

1. Convex optimization problems possess very nice geometric and theoretical properties.
2. A very large class of convex optimization problems can be efficiently solved by powerful algorithms.
3. For every convex optimization problem, each of the feasible region $\mathcal{S}$ and the set of optimal solutions $\mathcal{S}^*$ is a convex set.
4. Linear programming is a very special class of convex optimization with further additional desirable properties.
5. Henceforth, we will mostly focus on linear programming in this course.

Exercises

Question 5.1. *Determine whether the following optimization problem is a convex optimization problem:*

$$\min\{|x| : x \geq -3, \quad x^2 \leq 4\}$$

Question 5.2. *Determine whether the following optimization problem is a convex optimization problem:*

$$\min\{|x| : x \geq -3, \quad x^2 \leq 4, \quad x^3 = 1\}$$

MATH11111: Fundamentals of Optimization **Fall 2022**

Lecture 6 Vertices of Convex Sets and Introduction to Polyhedra

Lecturer: E. Alper Yıldırım *Week: 2*

6.1 Outline

- Vertices of Convex Sets
- Polyhedra and Polytopes
- Review Problems

6.2 Vertices of Convex Sets

Recall that a set $\mathcal{C} \subseteq \mathbb{R}^n$ is a convex set if, for every $x \in \mathcal{C}$ and for every $y \in \mathcal{C}$, all the vectors on the line segment that joins x and y also belong to $\mathcal{C}$.

Definition 6.1. *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a convex set. A vector $\hat{x} \in \mathcal{C}$ is called a* vertex *of $\mathcal{C}$ if there exists a hyperplane $\mathcal{H} = \{x \in \mathbb{R}^n : a^T x = \alpha\}$, where $a \in \mathbb{R}^n \backslash \{\mathbf{0}\}$, and a corresponding halfspace $\mathcal{H}^+ = \{x \in \mathbb{R}^n : a^T x \geq \alpha\}$ such that*

(i) $\mathcal{C} \cap \mathcal{H} = \{\hat{x}\}$*; and*

(ii) $\mathcal{C} \subseteq \mathcal{H}^+$.

Such a hyperplane $\mathcal{H}$ is called a supporting hyperplane *of $\mathcal{C}$.*

Remarks

1. If $n = 1$, then a hyperplane is a point on the real line and a halfspace is a half line. Therefore, if $\mathcal{C} \subseteq \mathbb{R}$ is a convex set, then $\hat{x} \in \mathcal{C}$ is a vertex of $\mathcal{C}$ if and only if it is an end-point of $\mathcal{C}$. For instance, the convex set $\mathcal{C} = [0, 1]$ has two vertices given by 0 and 1. The convex set $\mathcal{C} = [0, \infty)$ has only one vertex given by 0. The convex set $\mathcal{C} = (0, \infty)$ has no vertices (why not?).

2. If $n = 2$, then a hyperplane is a line in two dimensions and a halfspace is either side of such a line. Therefore, if $\mathcal{C} \subseteq \mathbb{R}$ is a convex set, then $\hat{x} \in \mathcal{C}$ is a vertex of $\mathcal{C}$ if and only if there is a line that intersects $\mathcal{C}$ exactly at the point $\hat{x} \in \mathcal{C}$.

   (a) Let
   $$\mathcal{C} = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1, \quad x_2 \geq 0\},$$
   i.e., it is the set of all points in the upper semicircle of the unit circle centred at the origin including the boundary points. Then, $\mathcal{C}$ has an infinite number of vertices and the set of all vertices of $\mathcal{C}$ is given by
   $$\mathcal{V} = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1, \quad x_2 \geq 0\},$$
   i.e., set of all points on the boundary of the upper semicircle.

(b) Let

$$\mathcal{C} = \{x \in \mathbb{R}^2 : |x_1| \leq 1, \quad |x_2| \leq 1\},$$

i.e., it is the set of all points in the square centred at the origin and four corner points at $[\pm 1, \pm 1]$. There are only four vertices which are precisely given by the four corner points.

(c) Let

$$\mathcal{C} = \{x \in \mathbb{R}^2 : x_1 + x_2 \leq 1\},$$

i.e., it is a halfspace since it is given by the sublevel set of a linear function (see Definition 4.6). You can easily verify that this convex set has no vertices.

6.3 Vertices and Optimization of Linear Objective Functions

Proposition 6.1. *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a convex set. A vector $\hat{x} \in \mathcal{C}$ is a vertex of $\mathcal{C}$ if and only if there exists a linear function $\ell(x) = a^T x$, where $a \in \mathbb{R}^n \backslash \{\mathbf{0}\}$, such that $\hat{x}$ is the* **unique** *optimal solution of the optimization problem*

$$\begin{array}{lll} (P) & \min & a^T x \\ & s.t. & \\ & & x \in \mathcal{C}. \end{array}$$

Proof. $\Rightarrow$: Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a convex set and let $\hat{x} \in \mathcal{C}$ be a vertex of $\mathcal{C}$. Then, there exists a hyperplane $\mathcal{H} = \{x \in \mathbb{R}^n : a^T x = \alpha\}$, where $a \in \mathbb{R}^n \backslash \{\mathbf{0}\}$, and the corresponding halfspace $\mathcal{H}^+ = \{x \in \mathbb{R}^n : a^T x \geq \alpha\}$ such that $\mathcal{C} \cap \mathcal{H} = \{\hat{x}\}$ and $\mathcal{C} \subseteq \mathcal{H}^+$. Since $\mathcal{C} \subseteq \mathcal{H}^+$, we have $a^T x \geq \alpha$ for each $x \in \mathcal{C}$. Since $\mathcal{C} \cap \mathcal{H} = \{\hat{x}\}$, it follows that $a^T x > \alpha$ for each $x \in \mathcal{C} \backslash \{\hat{x}\}$. Therefore, $\hat{x}$ is the unique optimal solution of (P).

$\Leftarrow$: Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a convex set and let $\hat{x} \in \mathcal{C}$ be the unique optimal solution of (P). Let $\alpha \in \mathbb{R}$ denote the optimal value of (P). Then, $a^T \hat{x} = \alpha$. Since $\hat{x} \in \mathcal{C}$ is the unique optimal solution of (P), we have $a^T x > \alpha$ for each $x \in \mathcal{C} \backslash \{\hat{x}\}$. Therefore, if we define the hyperplane $\mathcal{H} = \{x \in \mathbb{R}^n : a^T x = \alpha\}$, where $a \in \mathbb{R}^n \backslash \{\mathbf{0}\}$, and the corresponding halfspace $\mathcal{H}^+ = \{x \in \mathbb{R}^n : a^T x \geq \alpha\}$, we obtain $\mathcal{C} \cap \mathcal{H} = \{\hat{x}\}$ and $\mathcal{C} \subseteq \mathcal{H}^+$. Therefore, $\hat{x}$ is a vertex of $\mathcal{C}$. □

Remark 6.1. *Vertices of a convex set play an important role in the minimization of a linear function over that convex set, i.e., each vertex of a convex set is the unique optimal solution for the optimization problem of minimizing some linear function over that set.*

6.4 Polyhedra

Recall that each level set of a linear function is a hyperplane (see Definition 4.2) and that each sublevel or superlevel set of a linear function is a halfspace (see Definition 4.6).

Definition 6.2. *A set $\mathcal{P} \subseteq \mathbb{R}^n$ is called a* polyhedron *if it is given by the intersection of a finite number of hyperplanes and a finite number of halfspaces.*

Remark 6.2. *Every polyhedron is a convex set since every hyperplane and every halfspace is a convex set and convexity is preserved under taking intersections (see Remark 2 in Section 3.2).*

6.4.1 Linear Programming and Polyhedra

Recall our generic constrained optimization problem:

$$\begin{array}{lllll} \text{(P)} & \min & f(x) & & \\ & \text{subject to (s.t.)} & & & \\ & & g_i(x) & \geq & b_i, \quad i \in M_1, \\ & & \ell_i(x) & \leq & b_i, \quad i \in M_2, \\ & & h_i(x) & = & b_i, \quad i \in M_3, \end{array}$$

Remark 6.3. *Recall that (P) is a linear programming problem if each of f; $g_i,\ i \in M_1$; $\ell_i,\ i \in M_2$; and $h_i,\ i \in M_3$ is a linear function. The feasible region of every linear programming problem is a polyhedron since it is given by the intersection of a finite number of hyperplanes and a finite number of halfspaces.*

6.4.2 Bounded Sets and Polytopes

Definition 6.3. *A set $\mathcal{S} \subseteq \mathbb{R}^n$ is* bounded *if there exists a real number $K \in \mathbb{R}$ such that*

$$x \in \mathcal{S} \Rightarrow |x_j| \leq K, \quad j = 1, \ldots, n.$$

Definition 6.4. *A bounded polyhedron is called a* polytope.

Exercises

Question 6.1. *In $\mathbb{R}$, does there exist a convex set with no vertices? One vertex? Two vertices? Three vertices?*

Question 6.2. *In $\mathbb{R}^2$, for any $k = 0, 1, \ldots$, show that you can construct a convex set with exactly k vertices.*

MATH11111: Fundamentals of Optimization	**Fall 2022**
Lecture 7	Characterisation of Vertices of Polyhedra
Lecturer: E. Alper Yıldırım	*Week: 2*

7.1 Outline

- Active (Binding) Constraints
- Basic Solutions and Basic Feasible Solutions
- Connection with Vertices
- Review Problems

7.2 Quick Review

- Given a convex set $\mathcal{C} \subseteq \mathbb{R}^n$, $\hat{x} \in \mathcal{C}$ is called a *vertex* of $\mathcal{C}$ if there exists a hyperplane $\mathcal{H} = \{x \in \mathbb{R}^n : a^T x = \alpha\}$, where $a \in \mathbb{R}^n \backslash \{\mathbf{0}\}$, and a corresponding halfspace $\mathcal{H}^+ = \{x \in \mathbb{R}^n : a^T x \geq \alpha\}$ such that
  - $\mathcal{C} \cap \mathcal{H} = \{\hat{x}\}$;
  - $\mathcal{C} \subseteq \mathcal{H}^+$.
- A set $\mathcal{P} \subseteq \mathbb{R}^n$ is called a *polyhedron* if it is given by the intersection of a finite number of hyperplanes and a finite number of halfspaces.
- Every polyhedron is a convex set.
- The feasible region of every linear programming problem is a polyhedron.

7.3 Active (Binding) Constraints

Consider a polyhedron $\mathcal{P} \subseteq \mathbb{R}^n$ given by

$$\mathcal{P} = \left\{ x \in \mathbb{R}^n : \begin{array}{ll} (a^i)^T x \geq b_i, & i \in M_1, \\ (a^i)^T x \leq b_i, & i \in M_2, \\ (a^i)^T x = b_i, & i \in M_3 \end{array} \right\},$$

where M_1, M_2, and M_3 are finite sets, and $a^i \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ for each $i \in M_1 \cup M_2 \cup M_3$.

Definition 7.1. *Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a polyhedron and let $\hat{x} \in \mathbb{R}^n$. The set of indices of* active *(or* binding*) constraints at $\hat{x}$ is given by*

$$I(\hat{x}) = \left\{ i \in M_1 \cup M_2 \cup M_3 : (a^i)^T \hat{x} = b_i \right\}.$$

Note that $M_3 \subseteq I(\hat{x})$ for each $\hat{x} \in \mathcal{P}$.

7.4 Basic Solutions and Basic Feasible Solutions

Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a polyhedron given by

$$\mathcal{P} = \left\{ x \in \mathbb{R}^n : \begin{array}{ll} (a^i)^T x \geq b_i, & i \in M_1, \\ (a^i)^T x \leq b_i, & i \in M_2, \\ (a^i)^T x = b_i, & i \in M_3 \end{array} \right\}.$$

Definition 7.2. *Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a polyhedron and let $\hat{x} \in \mathbb{R}^n$.*

(i) $\hat{x}$ is a basic solution *if all of the equality constraints are active at $\hat{x}$ (i.e., $M_3 \subseteq I(\hat{x})$) and the set $\{a^i : i \in I(\hat{x})\} \subset \mathbb{R}^n$ contains n linearly independent vectors (i.e., the set $\{a^i : i \in I(\hat{x})\}$ spans $\mathbb{R}^n$).*

(ii) $\hat{x}$ is a basic feasible solution *if $\hat{x}$ is a* basic solution *and $\hat{x}$ is feasible (i.e., $\hat{x} \in \mathcal{P}$).*

7.4.1 Basic Feasible Solutions and Vertices

Proposition 7.1. *Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a polyhedron and let $\hat{x} \in \mathcal{P}$. Then, $\hat{x}$ is a basic feasible solution of $\mathcal{P}$ if and only if $\hat{x}$ is a vertex of $\mathcal{P}$.*

Proof. $\Rightarrow$: Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a polyhedron and let $\hat{x} \in \mathcal{P}$ be a basic feasible solution of $\mathcal{P}$. Let $I(\hat{x}) = \big\{i \in M_1 \cup M_2 \cup M_3 : (a^i)^T \hat{x} = b_i\big\}$. We need to construct a vector $a \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ and a real number $\alpha \in \mathbb{R}$ such that $a^T \hat{x} = \alpha$ and $a^T x > \alpha$ for each $x \in \mathcal{P} \backslash \{\hat{x}\}$. Let $I \subseteq I(\hat{x})$ be such that the set $\{a^i : i \in I\}$ is linearly independent and spans $\mathbb{R}^n$ (i.e., it is a basis for $\mathbb{R}^n$). Let $a = \sum\limits_{i \in I \cap M_1} (a^i) + \sum\limits_{i \in I \cap M_2} (-a^i) + \sum\limits_{i \in I \cap M_3} (a^i)$. Note that $a \neq \mathbf{0}$ since the set $\{a^i : i \in I\}$ is linearly independent. Let $\alpha = a^T \hat{x} = \sum\limits_{i \in I \cap M_1} (a^i)^T \hat{x} + \sum\limits_{i \in I \cap M_2} (-a^i)^T \hat{x} + \sum\limits_{i \in I \cap M_3} (a^i)^T \hat{x} = \sum\limits_{i \in I \cap M_1} b_i + \sum\limits_{i \in I \cap M_2} (-b_i) + \sum\limits_{i \in I \cap M_3} b_i$. Let $\mathcal{H} = \{x \in \mathbb{R}^n : a^T x = \alpha\}$ and $\mathcal{H}^+ = \{x \in \mathbb{R}^n : a^T x \geq \alpha\}$. Then, $\hat{x} \in \mathcal{P} \cap \mathcal{H}$. For any $x \in \mathcal{P}$, we have $a^T x = \sum\limits_{i \in I \cap M_1} (a^i)^T x + \sum\limits_{i \in I \cap M_2} (-a^i)^T x + \sum\limits_{i \in I \cap M_3} (a^i)^T x \geq \sum\limits_{i \in I \cap M_1} b_i + \sum\limits_{i \in I \cap M_2} (-b_i) + \sum\limits_{i \in I \cap M_3} b_i = \alpha$. Therefore, $\mathcal{P} \subseteq \mathcal{H}^+$. Finally, for any $x \in \mathcal{P} \cap \mathcal{H}$, we have $(a^i)^T x = b_i$ for each $i \in I$, which implies that $(a^i)^T (x - \hat{x}) = 0$ for each $i \in I$. Since the set $\{a^i : i \in I\}$ is a basis for $\mathbb{R}^n$, it follows that $x - \hat{x} = \mathbf{0}$, which implies that $\mathcal{P} \cap \mathcal{H} = \{\hat{x}\}$. Therefore, $\hat{x}$ is a vertex of $\mathcal{P}$.

$\Leftarrow$: Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a polyhedron and let $\hat{x} \in \mathcal{P}$ be a vertex of $\mathcal{P}$. Then, there exists a vector $a \in \mathbb{R}^n \backslash \{\mathbf{0}\}$, a real number $\alpha \in \mathbb{R}$, a hyperplane $\mathcal{H} = \{x \in \mathbb{R}^n : a^T x = \alpha\}$ and a corresponding halfspace $\mathcal{H}^+ = \{x \in \mathbb{R}^n : a^T x \geq \alpha\}$ such that $\mathcal{P} \subseteq \mathcal{H}^+$ and $\mathcal{P} \cap \mathcal{H} = \{\hat{x}\}$. Let $I(\hat{x}) = \big\{i \in M_1 \cup M_2 \cup M_3 : (a^i)^T \hat{x} = b_i\big\}$. Suppose, for a contradiction, that $\hat{x}$ is not a basic feasible solution. Since $\hat{x}$ is feasible, it is then not a basic solution. Therefore, the set $\{a^i : i \in I(\hat{x})\}$ does not contain n linearly independent vectors. Then, there exists a vector $d \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ such that $(a^i)^T d = 0$ for each $i \in I(\hat{x})$. Let $\epsilon > 0$ be a real number. Consider $\hat{x} + \epsilon d$ and $\hat{x} - \epsilon d$. Note that $M_3 \subseteq I(\hat{x})$ since $\hat{x} \in \mathcal{P}$. Since $(a^i)^T d = 0$ for each $i \in I(\hat{x})$, we have $(a^i)^T (\hat{x} + \epsilon d) = b_i$ and $(a^i)^T (\hat{x} - \epsilon d) = b_i$ for each $\epsilon > 0$ and each $i \in I(\hat{x})$. For each $i \in M_1 \backslash I(\hat{x})$, we have $(a^i)^T \hat{x} > b_i$, which implies that $(a^i)^T (\hat{x} + \epsilon d) \geq b_i$ and $(a^i)^T (\hat{x} - \epsilon d) \geq b_i$ if ϵ is sufficiently small but positive. Similarly, for each $i \in M_2 \backslash I(\hat{x})$, we have $(a^i)^T \hat{x} < b_i$, which implies that $(a^i)^T (\hat{x} + \epsilon d) \leq b_i$ and $(a^i)^T (\hat{x} - \epsilon d) \leq b_i$ if ϵ is sufficiently small but positive. Therefore, there exists $\epsilon^* > 0$ such that $\hat{x} - \epsilon^* d \in \mathcal{P}$ and $\hat{x} + \epsilon^* d \in \mathcal{P}$. Since $a^T x > \alpha$ for each $x \in \mathcal{P} \backslash \{\hat{x}\}$, we have $a^T (\hat{x} - \epsilon^* d) = \alpha - \epsilon^* a^T d > \alpha$ and $a^T (\hat{x} + \epsilon^*) = \alpha + \epsilon^* a^T d > \alpha$. Hence, $a^T d < 0$ and $a^T d > 0$, which is a contradiction. Therefore, $\hat{x}$ is a basic feasible solution.

$\square$

Remarks

1. For a polyhedron $\mathcal{P} \subseteq \mathbb{R}^n$, there is a one-to-one correspondence between vertices and basic feasible solutions.
2. The definition of a vertex is geometric and is therefore not very useful in an algorithmic framework.
3. On the other hand, the definition of a basic feasible solution is algebraic, i.e., for a given polyhedron $\mathcal{P}$ and a given vector $\hat{x}$, one can check if $\hat{x}$ is a basic feasible solution by simply using tools from linear algebra.

Exercises

Question 7.1. *Consider the following polyhedron:*

$$\mathcal{P} = \{x \in \mathbb{R}^2 : x_1 \geq 1,\ x_1 - x_2 \geq 0,\ x_1 \geq 0,\ x_2 \leq 1,\ x_1 + x_2 = 2\}$$

For each of the following vectors in $\mathbb{R}^2$, determine whether it is a basic solution, basic feasible solution, both, or neither.

(i) $x^1 = [1, 1]^T$

(ii) $x^2 = [3/2, 1/2]^T$

(iii) $x^3 = [2, 0]^T$

(iv) $x^4 = [0, 2]^T$

(v) $x^5 = [2, 2]^T$

MATH11111: Fundamentals of Optimization | **Fall 2022**

Lecture 8 | Existence and Finiteness of Basic Feasible Solutions

Lecturer: E. Alper Yıldırım | *Week: 2*

8.1 Outline

- Existence of Basic Feasible Solutions
- Finiteness of Basic Feasible Solutions
- Review Problems

8.2 Quick Review

Let

$$\mathcal{P} = \{x \in \mathbb{R}^n : (a^i)^T x \geq b_i, \ i \in M_1; \ (a^i)^T x \leq b_i, \ i \in M_2; (a^i)^T x = b_i, \ i \in M_3\}$$

be a polyhedron and let $\hat{x} \in \mathbb{R}^n$.

- A constraint is active (or binding) at $\hat{x}$ if it is satisfied with equality.
- Let $I(\hat{x})$ denote the set of indices of all active constraints at $\hat{x}$.
- $\hat{x}$ is a *basic solution* if all of the equality constraints are active at $\hat{x}$ (i.e., $M_3 \subseteq I(\hat{x})$) and the set $\{a^i : i \in I(\hat{x})\} \subset \mathbb{R}^n$ contains n linearly independent vectors (i.e., the set $\{a^i : i \in I(\hat{x})\}$ spans $\mathbb{R}^n$).
- $\hat{x}$ is a *basic feasible solution* if $\hat{x}$ is a *basic solution* and $\hat{x}$ is feasible (i.e., $\hat{x} \in \mathcal{P}$).
- $\hat{x}$ is a basic feasible solution of $\mathcal{P}$ if and only if $\hat{x}$ is a vertex of $\mathcal{P}$.

8.3 Existence of Vertices

Question 1. *Does every nonempty polyhedron necessarily have at least one vertex?*

Definition 8.1. *Let $\mathcal{P} = \{x \in \mathbb{R}^n : (a^i)^T x \geq b_i, \ i \in M_1; \ (a^i)^T x \leq b_i, \ i \in M_2; \ (a^i)^T x = b_i, \ i \in M_3\}$ be a polyhedron. $\mathcal{P}$* contains a line *if there exists a vector $\tilde{x} \in \mathcal{P}$ and a nonzero vector $d \in \mathbb{R}^n$ such that $\tilde{x} + \lambda d \in \mathcal{P}$ for every real number λ.*

Consider the line in $\mathbb{R}^2$ given by $x_1 + x_2 = 2$. Alternatively, the same line can be represented by a point on the line and the direction of the line. For instance, $\tilde{x} = [1, 1]^T$ is a vector on this line and starting from this point, one can move in the direction $d = [1, -1]^T$ or in its opposite direction and will always remain on this line. Therefore,

$$\{x \in \mathbb{R}^n : x_1 + x_2 = 2\} = \{\tilde{x} + \lambda d : \lambda \in \mathbb{R}\}.$$

The latter representation holds for any line in $\mathbb{R}^n$. Note that Definition 8.1 uses the second representation.

The next proposition gives a complete characterisation of polyhedra that contain at least one vertex.

Proposition 8.1. *Let* $\mathcal{P} = \{x \in \mathbb{R}^n : (a^i)^T x \geq b_i,\ i \in M_1;\ (a^i)^T x \leq b_i,\ i \in M_2;\ (a^i)^T x = b_i,\ i \in M_3\}$ *be a nonempty polyhedron.* $\mathcal{P}$ *has at least one vertex if and only if it does not contain a line.*

Proof. $\Rightarrow$: We will use proof by contrapositive, i.e., we will show that if a nonempty polyhedron contains a line, then it does not have any vertices. Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a nonempty polyhedron that contains a line. Then, there exists a vector $\tilde{x} \in \mathcal{P}$ and a nonzero vector $d \in \mathbb{R}^n$ such that $\tilde{x} + \lambda d \in \mathcal{P}$ for every real number λ. Since $\tilde{x} \in \mathcal{P}$, we have $(a^i)^T \tilde{x} \geq b_i$ for each $i \in M_1$, $(a^i)^T \tilde{x} \leq b_i$ for each $i \in M_2$, and $(a^i)^T \tilde{x} = b_i$ for each $i \in M_3$. Since $\tilde{x} + \lambda d \in \mathcal{P}$ for every real number λ, we have $(a^i)^T(\tilde{x} + \lambda d) \geq b_i$ for each $i \in M_1$, $(a^i)^T(\tilde{x} + \lambda d) \leq b_i$ for each $i \in M_2$, and $(a^i)^T(\tilde{x} + \lambda d) = b_i$ for each $i \in M_3$. Therefore, we have $(a^i)^T d = 0$ for each $i \in M_1 \cup M_2 \cup M_3$. Therefore, for any $x \in \mathcal{P}$, since $I(x) \subseteq M_1 \cup M_2 \cup M_3$, it follows that $(a^i)^T d = 0$ for each $i \in I(x)$. Therefore, the set $\{a^i : i \in I(x)\}$ cannot contain n linearly independent vectors for any $x \in \mathcal{P}$. It follows that $\mathcal{P}$ does not have any vertices.

$\Leftarrow$: Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a nonempty polyhedron that does not contain a line. We need to show that $\mathcal{P}$ has at least one vertex. By the argument in the previous part of the proof, there does not exist a vector $d \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ such that $(a^i)^T d = 0$ for each $i \in M_1 \cup M_2 \cup M_3$. Let $\hat{x} \in \mathcal{P}$ be an arbitrary vector and let $I(\hat{x}) = \left\{i \in M_1 \cup M_2 \cup M_3 : (a^i)^T \hat{x} = b_i\right\}$. If the set $\{a^i : i \in I(\hat{x})\}$ contains n linearly independent vectors, then we are done since $\hat{x}$ is a vertex. Otherwise, there exists a vector $d \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ such that $(a^i)^T d = 0$ for each $i \in I(\hat{x})$. Since $M_3 \subseteq I(\hat{x})$, we have $(a^i)^T d = 0$ for each $i \in M_3$. Consider the line $\hat{x} + \lambda d$, where $\lambda \in \mathbb{R}$. Since $\mathcal{P}$ does not contain a line, there exist a nonzero $\lambda^* \in \mathbb{R}$ such that $x^* = \hat{x} + \lambda^* d \in \mathcal{P}$ and an index $i^* \in (M_1 \cup M_2) \backslash I(\hat{x})$ such that $(a^{i^*})^T x^* = b_{i^*}$. Therefore, $(a^{i^*})^T d \neq 0$ and $I(x^*) \supseteq I(\hat{x}) \cup \{i^*\}$. We claim that a^{i^*} is not a linear combination of the vectors in the set $\{a^i : i \in I(\hat{x})\}$. Otherwise, there would exist real numbers $\alpha_i,\ i \in I(\hat{x})$ such that $a^{i^*} = \sum\limits_{i \in I(\hat{x})} \alpha_i a^i$. Then, $(a^{i^*})^T d = \sum\limits_{i \in I(\hat{x})} \alpha_i (a^i)^T d = 0$, which contradicts with $(a^{i^*})^T d \neq 0$. Therefore, the number of linearly independent vectors indexed by $I(x^*)$ is at least one larger than that indexed by $I(\hat{x})$. By repeating this procedure as many times as needed, we obtain a feasible solution whose set of active constraints contains a subset of n linearly independent vectors. Therefore, $\mathcal{P}$ contains a vertex. ☐

8.3.1 Implications on Polytopes

Remark 8.1. *Recall that a bounded polyhedron is called a polytope.*

Corollary 8.2. *Let* $\mathcal{P} \subseteq \mathbb{R}^n$ *be a nonempty polytope. Then,* $\mathcal{P}$ *has at least one vertex.*

Proof. Since $\mathcal{P}$ is nonempty and bounded, it cannot contain a line. By Proposition 8.1, $\mathcal{P}$ has at least one vertex. ☐

8.4 Number of Vertices of a Polyhedron

Proposition 8.2. *Let* $\mathcal{P} = \{x \in \mathbb{R}^n : (a^i)^T x \geq b_i,\ i \in M_1;\ (a^i)^T x \leq b_i,\ i \in M_2;\ (a^i)^T x = b_i,\ i \in M_3\}$ *be a nonempty polyhedron. Then,* $\mathcal{P}$ *contains at most a finite number of vertices.*

Proof. Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a nonempty polyhedron. If $\mathcal{P}$ contains a line, then it has no vertices. Otherwise, for any $x \in \mathcal{P}$, let $I(x) = \left\{i \in M_1 \cup M_2 \cup M_3 : (a^i)^T x = b_i\right\}$. Note that $M_3 \subseteq I(x)$. Therefore, $I(x) = M_3 \cup J$, where $J \subseteq M_1 \cup M_2$. The number of different subsets of $M_1 \cup M_2$ is given by $2^{|M_1| + |M_2|}$, which is a finite number. Therefore, for any $x \in \mathcal{P}$, $I(x)$ can be equal to a finite number of different sets. Among those different sets, only a subset of them will satisfy the condition that the set $\{a^i : i \in I(x)\}$ contains n linearly independent vectors. Finally, if $x^1 \in \mathcal{P}$ and $x^2 \in \mathcal{P}$ are such that $I(x^1) = I(x^2) = I$ and the set $\{a^i : i \in I\}$ contains n linearly independent vectors, then $x^1 = x^2$ since $(a^i)^T(x^1 - x^2) = 0$ for each $i \in I$. Therefore, $\mathcal{P}$ contains at most a finite number of vertices. □

8.4.1 Number of Vertices of a General Convex Set

Consider the following convex set $\mathcal{C} \subseteq \mathbb{R}^2$:

$$\mathcal{C} = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}.$$

Note that $\mathcal{C}$ is the circle centred at the origin with radius 1. It is easy to show that $\mathcal{C} \subseteq \mathbb{R}^2$ is a convex set. However, $\mathcal{C}$ is not a polyhedron since it cannot be written as the intersection of a finite number of halfspaces and hyperplanes. Note that $\mathcal{C}$ has an **infinite** number of vertices and the set of all vertices of $\mathcal{C}$ is given by

$$\mathcal{V} = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\},$$

i.e., set of all points on the boundary of the circle. As illustrated by this example, a polyhedron is a very special kind of convex set.

Exercises

Question 8.1. *Let $\mathcal{P}^1 \subset \mathbb{R}^n$ be a polyhedron that contains at least one vertex and let $\mathcal{P}^2 \subset \mathbb{R}^n$ be an arbitrary nonempty polyhedron. Let $\mathcal{P} = \mathcal{P}^1 \cap \mathcal{P}^2$. Suppose that $\mathcal{P}$ is nonempty.*

(i) Show that $\mathcal{P}$ is a polyhedron.

(ii) Show that $\mathcal{P}$ contains at least one vertex.

Question 8.2. *Consider the following polyhedron:*

$$\mathcal{P} = \{x \in \mathbb{R}^n : x_j \geq 0,\ j = 1, \ldots, n;\ x_j \leq 1,\ j = 1, \ldots, n\}$$

How many vertices does $\mathcal{P}$ have?