Bertsimas, Tsitsiklis: *Introduction to Linear Optimization*

Danny Nygård Hansen

15th December 2023

2 The geometry of linear programming

REMARK 2.1. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a map. Then the following properties are equivalent:

(a) There exists a matrix $\boldsymbol{A}$ and a vector $\boldsymbol{b}$ such that $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

(b) For all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $t \in \mathbb{R}$, $f((1-t)\boldsymbol{x} + t\boldsymbol{y}) = (1-t)f(\boldsymbol{x}) + tf(\boldsymbol{y})$.

Such a function is called *affine*. The first property clearly entails the second, so assume that f has the second property. First assume that $f(\boldsymbol{0}) = \boldsymbol{0}$. For $\beta \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^n$ be thus have

$$f(\beta\boldsymbol{x}) = f(\beta\boldsymbol{x} + (1-\beta)\boldsymbol{0}) = \beta f(\boldsymbol{x}) + (1-\beta)f(\boldsymbol{0}) = \beta f(\boldsymbol{x}),$$

so f is homogeneous. If also $\boldsymbol{y} \in \mathbb{R}^n$, then

$$\tfrac{1}{2}f(\boldsymbol{x}+\boldsymbol{y}) = f\left(\tfrac{1}{2}\boldsymbol{x} + \tfrac{1}{2}\boldsymbol{y}\right) = \tfrac{1}{2}f(\boldsymbol{x}) + \tfrac{1}{2}f(\boldsymbol{y}),$$

so f is also additive, hence linear. For general f, simply replace f with $f - f(\boldsymbol{0})$. ⌟

REMARK 2.2. If $f : \mathbb{R}^n \to \mathbb{R}^n$ preserves the Euclidean metric, then it is affine. Replacing f with $f - f(\mathbf{0})$ we may assume that $f(\mathbf{0}) = \mathbf{0}$, and so that f also preserves the inner product. For all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ we have

$$\begin{aligned}
\langle f(\beta\boldsymbol{x} + \boldsymbol{y}), f(\boldsymbol{z})\rangle &= \langle \beta\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}\rangle \\
&= \beta\langle \boldsymbol{x}, \boldsymbol{z}\rangle + \langle \boldsymbol{y}, \boldsymbol{z}\rangle \\
&= \beta\langle f(\boldsymbol{x}), f(\boldsymbol{z})\rangle + \langle f(\boldsymbol{y}), f(\boldsymbol{z})\rangle \\
&= \langle \beta f(\boldsymbol{x}) + f(\boldsymbol{y}), f(\boldsymbol{z})\rangle.
\end{aligned}$$

Since f is surjective, the claim follows. ⌟

REMARK 2.3: *Adjacent basic solutions and adjacent bases.*
Let $P = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \mathbf{0}\}$ be a polyhedron in standard form.

> *If $\boldsymbol{x}$ and $\boldsymbol{y}$ are adjacent basic solutions, then they can be obtained from two adjacent bases.*

Assume that $x_i = 0$ for $i \neq B(1), \ldots, B(m)$. Each of the m equality constraints are active at both $\boldsymbol{x}$ and $\boldsymbol{y}$, so there must be exactly one k such that $y_{B(k)} \neq 0$. Then $\boldsymbol{x}$ can be obtained from the basis with basic indices $\{B(1), \ldots, B(m)\}$ and $\boldsymbol{y}$ from a basis with basic indices $\{B(1), \ldots, B(k-1), j, B(k+1), \ldots, B(m)\}$ where $j \neq B(k)$. Hence these two bases are adjacent as claimed.

> *If two adjacent bases lead to distinct basic solutions $\boldsymbol{x}$ and $\boldsymbol{y}$, then these are adjacent.*

Similarly, there are (at least) $n-1$ constraints that are active at both $\boldsymbol{x}$ and $\boldsymbol{y}$ (the equality constraints along with all but one of the constraints given by the $B(k)$). Since $\boldsymbol{x} \neq \boldsymbol{y}$ there must be some k such that $x_{B(k)} = 0$ and $y_{B(k)} \neq 0$, so there are exactly $n-1$ constraints that are active at both $\boldsymbol{x}$ and $\boldsymbol{y}$. ⌟

EXERCISE 2.5

(Extreme points of isomorphic polyhedra) A mapping f is called ***affine*** if it is of the form $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$, where $\boldsymbol{A}$ is a matrix and $\boldsymbol{b}$ is a vector. Let P and Q be polyhedra in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. We say that P and Q are ***isomorphic*** if there exist affine mappings $f : P \to Q$ and $g : Q \to P$ such that $g(f(\boldsymbol{x})) = \boldsymbol{x}$ for all $\boldsymbol{x} \in P$, and $f(g(\boldsymbol{y})) = \boldsymbol{y}$ for all $\boldsymbol{y} \in Q$. (Intuitively, isomorphic polyhedra have the same shape.)

(a) If P and Q are isomorphic, show that there exists a one-to-one correspondence between their extreme points. In particular, if f and g are as above, show that $\boldsymbol{x}$ is an extreme point of P if and only if $f(\boldsymbol{x})$ is an extreme point of Q.

(b) (Introducing slack variables leads to an isomorphic polyhedron) Let $P = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$, where $\boldsymbol{A}$ is a matrix of dimensions $k \times n$. Let $Q = \{(\boldsymbol{x}, \boldsymbol{z}) \in \mathbb{R}^{n+k} \mid \boldsymbol{A}\boldsymbol{x} - \boldsymbol{z} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \geq \boldsymbol{0}\}$. Show that P and Q are isomorphic.

Solution. (a) For $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in P$ with $\boldsymbol{x} = \lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{z}$, then since f is affine we have $f(\boldsymbol{x}) = \lambda f(\boldsymbol{y}) + (1-\lambda) f(\boldsymbol{z})$. Hence if $\boldsymbol{x}$ is *not* an extreme point, then neither is $f(\boldsymbol{x})$. Since f is bijective with an affine inverse, this proves the claim.

(b) Simply note that the map $f \colon P \to Q$ given by $f(\boldsymbol{x}) = (\boldsymbol{x}, \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})$ is an isomorphism. □

Exercise 2.7

Suppose that $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{a}_i'\boldsymbol{x} \geq b_i, i = 1, \dots, m\}$ and $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{g}_i'\boldsymbol{x} \geq h_i, i = 1, \dots, k\}$ are representations of the same nonempty polyhedron P. Suppose that the vectors $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$ span $\mathbb{R}^n$. Show that the same must be true for the vectors $\boldsymbol{g}_1, \dots, \boldsymbol{g}_k$.

Solution. If $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$ span $\mathbb{R}^n$, then they must contain a collection of n linearly independent vectors. Then Theorem 2.6 implies that P has an extreme point. But this is a geometric property, so it does not depend on the representation. The same theorem then implies that the collection $\boldsymbol{g}_1, \dots, \boldsymbol{g}_k$ contains n linearly independent vectors. But then they span $\mathbb{R}^n$. □

Exercise 2.12

Consider a nonempty polyhedron P and suppose that for each variable x_i we have either the constraint $x_i \geq 0$ or the contraint $x_i \geq 0$. Is it true that P has at least one basic feasible solution?

Solution. Let I be the set of indices i such that we have the contraint $x_i \leq 0$, and let $\boldsymbol{A}$ be a diagonal matrix with a 1 as the ith entry on the diagonal if $i \notin I$, and -1 if $i \in I$. The map $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ is then an isomorphism between P and the polyhedron Q which is defined by the same contraints as P, except that the contraints on the form $x_i \leq 0$ are replaced by contraints $x_i \geq 0$. In particular, P and Q have the same extreme points by [TODO ref] Exercise 2.4, hence the same basic feasible solutions by Theorem 2.3.

Next introduce slack variables to Q, yielding a polyhedron R in standard form, which is isomorphic to Q by [TODO ref] Exercise 2.5. But since R is nonempty (since it is isomorphic to P) it has a basic feasible solution by Corollary 2.2. □

TODO 2.14, 2.15

EXERCISE 2.21

Suppose that Fourier–Motzkin elimination is used in the manner described at the end of Section 2.8 to find the optimal cost in a linear programming problem. Show how this approach can be augmented to obtain an optimal solution as well.

SOLUTION. Let $P \subseteq \mathbb{R}^n$ be a polyhedron, and assume that we wish to minimise the quantity $\boldsymbol{c}'\boldsymbol{x}$ subject to $\boldsymbol{x} \in P$. Define

$$P_n = \{(x_0, \boldsymbol{x}) \in \mathbb{R}^{n+1} \mid x_0 = \boldsymbol{c}'\boldsymbol{x} \text{ and } \boldsymbol{x} \in P\},$$

which is clearly a polyhedron in $\mathbb{R}^{n+1}$. To minimise $\boldsymbol{c}'\boldsymbol{x}$ with $\boldsymbol{x} \in P$ we thus find the minimal value of x_0 among elements $(x_0, \boldsymbol{x}) \in P_n$. Letting $P_{n-k} = \Pi_k(P_n)$ for $k = 1, \ldots, n$ we have $Q := P_0 = \Pi_n(P_n) = \{x_0 \in \mathbb{R} \mid \exists \boldsymbol{x} \in P\colon (x_0, \boldsymbol{x}) \in P_n\}$, so this is equivalent to finding the minimum of Q. Given the minimum of Q we wish to find a (not necessarily unique) $\boldsymbol{x} \in P$ such that $(\min Q, \boldsymbol{x}) \in P_n$.

We do this by recursively finding a suitable value for the variable we eliminate in each step of the algorithm. In the notation of the Elimination algorithm on page 72, let $P \subseteq \mathbb{R}^n$ be a polyhedron, and let $Q = \Pi_{n-1}(P)$ be the polyhedron obtained when eliminating the variable x_n. Let $\overline{\boldsymbol{x}}$ be an element of Q, i.e. a vector satisfying the inequalities

$$\begin{aligned} d_j + \boldsymbol{f}_j'\overline{\boldsymbol{x}} &\geq d_i + \boldsymbol{f}_i'\overline{\boldsymbol{x}}, && \text{if } a_{in} > 0 \text{ and } a_{jn} < 0, \\ 0 &\geq d_k + \boldsymbol{f}_k'\overline{\boldsymbol{x}}, && \text{if } a_{kn} = 0, \end{aligned}$$

for $i, j, k = 1, \ldots, n$. Now choose any $x_n \in \mathbb{R}$ that lies in all the intervals $[d_i + \boldsymbol{f}_i'\overline{\boldsymbol{x}}, d_j + \boldsymbol{f}_j'\overline{\boldsymbol{x}}]$. Then the vector $(\overline{\boldsymbol{x}}, x_n)$ satisfies the inequalities

$$\begin{aligned} x_n &\geq d_i + \boldsymbol{f}_i'\overline{\boldsymbol{x}}, && \text{if } a_{in} > 0, \\ d_j + \boldsymbol{f}_j'\overline{\boldsymbol{x}} &\geq x_n, && \text{if } a_{jn} < 0, \\ 0 &\geq d_k + \boldsymbol{f}_k'\overline{\boldsymbol{x}}, && \text{if } a_{kn} = 0. \end{aligned}$$

Such an x_n exists since these inequalities are a representation of P, and the former inequalities are a representation of Q, which is the projection of P by Theorem 2.10. Thus $(\overline{\boldsymbol{x}}, x_n) \in P$ as desired. □

EXERCISE 2.22

Let P and Q be polyhedra in $\mathbb{R}^n$.

(a) Show that $P + Q$ is a polyhedron.

(b) Show that every extreme point of $P + Q$ is the sum of an extreme point of P and an extreme point of Q.

Solution. (a) The Cartesian product $P \times Q$ is clearly a polyhedron. Its image under the linear map $(\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{x} + \boldsymbol{y}$ is exactly $P + Q$, and this is a polyhedron by Corollary 2.5.

(b) Let $\boldsymbol{z} \in P + Q$ and write $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$ for $\boldsymbol{x} \in P$ and $\boldsymbol{y} \in Q$. Assume that $\boldsymbol{x}$ is not an extreme point in P, so that there exist $\boldsymbol{x}_1, \boldsymbol{x}_2 \in P \setminus \{\boldsymbol{x}\}$ and a $\lambda \in [0,1]$ such that $\boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$. Letting $\boldsymbol{z}_i := \boldsymbol{x}_i + \boldsymbol{y}$ we easily find that $\boldsymbol{z} = \lambda \boldsymbol{z}_1 + (1-\lambda)\boldsymbol{z}_2$, so $\boldsymbol{z}$ is not an extreme point. $\square$