Intro to LO, Lecture 4

Martin Böhm
University of Wrocław, Winter 2023/2024

Refresh: Concepts from metric spaces

All of the definitions below work for general topological spaces, but we are only interested in the Euclidean space $\mathbb{R}^n$.

D(Open set): A set $S \subseteq \mathbb{R}^n$ is called open if for each point $p \in S$, there exists a radius r such that all points from a ball $B(p, r)$ are present in S.

D(Closed set): A set S is closed if its complement $(\mathbb{R}^n \setminus S)$ is open.

D(Bounded set): A set S is bounded if S fits into a ball of some finite diameter d.

D(Compact set): A set S is compact if it is closed and bounded.

T: Consider a function $f : \mathbb{R}^n \to \mathbb{R}$. If S is a compact set and f is continuous on S, then there exists points in S where f attains its infimum and supremum value over S.

Refresh: Convexity

D: A set $K \subseteq \mathbb{R}^d$ is a *convex set*, if $\forall x, y \in K, \forall t \in [0,1] : tx+(1-t)y \in K$. In other words, if you take two points inside the convex set K, the entire line segment between those two points must belong to K.

D: A vector x is a *convex combination* of a set of vectors $a_1, a_2, \ldots a_n$ if $x = \sum_{i=1}^n \alpha_i a_i$, where α_i are real numbers satisfying $\sum_{i=1}^n \alpha_i = 1$ and also $\forall i : \alpha_i \in [0,1]$.

A set of vectors/points $V \subseteq \mathbb{R}^d$ is *in a convex position*, if it holds that no vector $v \in V$ is a convex combination of the rest.

O: Let Y be the set of convex combinations of points from X. Then every convex combination of points of Y is a convex combination of points of X.

O: If all points $\boldsymbol{x} \in X$ satisfy the inequality $\boldsymbol{a}^T\boldsymbol{x} \le b$, then any convex combination of points from X satisfies this inequality.

D: As with linearity and affinity, for convexity we also define a span/hull:

If we have a set of vectors $V \subseteq \mathbb{R}^d$, its *convex hull* is a set of all vectors C, which are convex combinations of any finite subset of the vectors in V.

Here, we really need to consider any finite subset of V, because convex sets in general do not have a finite basis.

T: The convex hull of $X, \mathrm{conv}(X)$, is equal to the set of all convex combinations

$$Y = \left\{ \alpha_0 \boldsymbol{a}_0 + \ldots + \alpha_k \boldsymbol{a}_k \mid k \in \mathbb{N}, \boldsymbol{a}_i \in X, \alpha_i \ge 0, \sum_{i=0}^k \alpha_i = 1 \right\}.$$

T(Caratheodory): Let $X \subseteq \mathbb{R}^n, \dim(X) = d$. Then,

$$\mathrm{conv}(X) = \left\{ \alpha_0 \boldsymbol{a}_0 + \ldots + \alpha_d \boldsymbol{a}_d \mid \boldsymbol{a}_i \in X, \alpha_i \ge 0, \sum_{i=0}^d \alpha_i = 1 \right\}.$$

D: A *hyperplane* is any affine space in $\mathbb{R}^d$ of dimension $d-1$. Thus, on a 2D plane, any line is a hyperplane. In the 3D space, any plane is a hyperplane, and so on.

A hyperplane splits the space $\mathbb{R}^d$ into two *halfspaces*. We count the hyperplane itself as a part of both halfspaces.

D: A *convex polyhedron* (sometimes also called H-polytope) is any object in $\mathbb{R}^d$ that is an intersection of finitely many halfspaces. Alternatively, we can say that a convex polytope is any set of points of the form $\{x | Ax \le b\}$ for some real matrix A and some real vector b.

Two observations from high school

O: Consider the angle ϕ at b for the triangle abc. Then, $(b-a)^T(c-b) = ||b-a|| \cdot ||c-b|| \cdot \cos\phi$. Since the two norms are always non-negative, the sign of the scalar product depends only on $\cos\phi$.

O: Consider a triangle abc and again ϕ angle at b. Then:

- If the angle is a right angle, then b is the closest point to c on the entire line generated by ab.
- If the angle is acute (less than $\pi/2$), then there is some point other than a and b that is closest to c on the line segment ab.
- If the angle is obtuse, then the point closest to c on the line segment ab is b (and the closest point on the *line* ab is outside the line segment ab).

Separation theorem

T(Separation theorem): Let $C, D \subseteq \mathbb{R}^n$ be nonempty, closed, convex and disjoint. and let C be bounded. Then there is a hyperplane $\{\boldsymbol{x} \mid \boldsymbol{a}^T\boldsymbol{x} = b\}$, which strongly separates C and D, i.e., one that $C \subseteq \{\boldsymbol{x} \mid \boldsymbol{a}^T\boldsymbol{x} < b\}$ a $D \subseteq \{\boldsymbol{x} \mid \boldsymbol{a}^T\boldsymbol{x} > b\}$.

Proof steps.

1. If both $C \times D$ are bounded, they are also compact and so is the product $C \times D \subseteq \mathbb{R}^{2n}$, and since the Euclidean norm is continuous, we can find minimizers $c \in C$ and $d \in D$.
2. If D is unbounded, we restrict ourselves to a set $D' \subseteq D$ that is bounded as follows: Let the max distance in C be α. We sample one $c' \in C$ and $d' \in D$ and let their distance be β. Now, restrict D to $D' = D \cap B(c', \alpha + \beta)$. A simple observation shows that all points in D that have a chance to be closer than β to C live in D'.
3. Find the closest points in C and D, call them c, d, respectively. Let $a = c - d$ be the direction between them. The separating hyperplane will be orthogonal to $c - d$ and will touch the point $(c+d)/2$.
4. $a^T d - a^T c = ||a||^2 > 0$.
5. If the difference of two numbers is a non-negative real, then their average is strictly between them, in other words: if $b = a^T(c+d)/2$, then $a^T c < b < a^T d$.
6. Finally, we observe that for all other points $c'' \in C$ holds that $a^T c'' < a^T c$. We observe this geometrically, using our observations from high school.

Vertices and basic feasible solutions

D(Polyhedron): A set of points is called a *polyhedron* (or an H-polytope) if it the set is the intersection of finitely many halfspaces.

D(Polytope): A set of points is called a *polytope* (or a V-polytope) if it is a convex hull of a finite number of points.

D(Vertex): Let P be a convex set. A point $z \in P$ is a *vertex* if z cannot be written as a convex combination of any other two points in P.

D(Basic feasible solution): Let $P = \{x \mid Ax \le b\}$ be a polyhedron and let $z \in P$. Then A_z is the submatrix of A consisting of those rows a_i of A for which $a_i z = b_i$. We say a point $z \in P$ is an *basic feasible solution* if $\mathrm{rank}\,(A_z) = n$.

T: Let $P = \{x \mid Ax \le b\}$ be a polyhedron in $\mathbb{R}^n$ and let $z \in P$. Then z is a vertex of P if and only if z is a basic feasible solution.

P: $\Rightarrow$: A_z not full rank $\to$ find $c : A_z c = 0$. z is then a conv. combination of $z + \delta c, z - \delta c$ for some $\delta > 0$.

We can find this δ since every equation outside A_z held with a strict inequality, so even though c might be the direction where the inequality breaks eventually, it will hold at least for some small δ.

$\Leftarrow$: If z not a vertex $\leftarrow$ find two elements which combine to it; $A_v(x-y) = 0$.

A polyhedron is a convex hull of its vertices

T: Let P be a bounded polyhedron, with vertices $x_1, \ldots, x_t$. Then $P = \mathrm{conv.hull}\,\{x_1, \ldots, x_t\}$.

P: Clearly

$$\mathrm{conv.hull}\,\{x_1, \ldots, x_t\} \subseteq P$$

since $x_1, \ldots, x_t$ belong to P and since P is convex. The reverse inclusion amounts to:

$$\text{if } z \in P \text{ then } z \in \mathrm{conv.hull}\,\{x_1, \ldots, x_t\}.$$

We prove it by induction on $n - \mathrm{rank}\,(A_z)$. Informally, we take a point $z \in P$ and we find two points in P of higher rank of A_z which have z within their line segment. To find these, we take z and move in opposite directions until we reach the boundary of P, since the boundary points will satisfy more equalities than z.

Formally, we find the respective increases

$$\mu_0 := \max\{\mu \mid z + \mu c \in P\},$$
$$\nu_0 := \max\{\nu \mid z - \nu c \in P\}.$$

And define $x := z + \mu_0 c$ and $y := z - \nu_0 c$ to be the two boundary points.

For the computation, we need to observe a "form of duality": For the μ_0 defined above, we have

$$\mu_0 = \min\left\{ \frac{b_i - a_i z}{a_i c} \mid a_i \text{ is a row of } A; a_i c > 0 \right\}.$$

C: Every bounded polyhedron is a polytope.

T(Without proof): Every polytope is a bounded polyhedron.

Exercises

Exercise one. Find all the vertices of the following polytope: $P = \{(x,y,z) \mid x+y \le 2, y+z \le 4, x+z \le 3, -2x-y \le 3, -y-2z \le 3, -2x-z \le 2\}$.

Exercise two. Prove the following statements that give insight into where optimal solutions lie for a linear program. We have already mentioned some of them at the lecture.

1. If the polyhedron of feasible solutions of a linear program is *bounded*, and it has at least one feasible solution, it has an optimal solution.
2. If a linear program has an optimal solution, it also has one on the boundary of the polyhedron.
3. If a bounded linear program has an optimal solution, it also has one in a *vertex* of the polyhedron. (Induction might be useful for this one.)

Exercise three. Check if the point $v = (1,1,1,1)$ is a vertex of a polytope P defined as the following set of inequalities:

$$\begin{pmatrix} -1 & -6 & 1 & 3 \\ -1 & -2 & 7 & 1 \\ 0 & 3 & -10 & -1 \\ -6 & -11 & -2 & -12 \\ 1 & 6 & -1 & -3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \le \begin{pmatrix} -3 \\ 5 \\ -8 \\ -7 \\ 4 \end{pmatrix}$$

The next two exercises introduce a bit more naming for the boundary objects of polyhedra, which we call *faces*.

D: Let P be some convex polyhedron in $\mathbb{R}^d$. We say that a hyperplane H is a *supporting hyperplane* if it touches the polyhedron, but does not cut it.

In other words, if the hyperplane H is defined as $\{x \in \mathbb{R}^d | c^T x = t\}$, then we say H is supporting if and only if it holds that (i) $P \cap H$ is nonempty; and (ii) $\forall y \in P : \{c^T y \le t\}$ or it holds that $\forall y \in P : \{c^T y \ge t\}$.

D: A *face* F of a polytope P is any set of the form $F = P \cap H$ for any supporting hyperplane H.

D: A *vertex* of a polytope P is a face of dimension 0 (a single point). An *edge* is any face of dimension 1 (a line segment, half-line or a line). On the other side of the spectrum, a *facet* of P is a face of dimension $d-1$.

Exercise four. We need to establish the equivalence of the vertex definition above with the one from the lecture. In other words, prove that for any point v in a bounded convex polyhedron P:

The point v is a basic feasible solution if and only if there exists a supporting hyperplane H such that $P \cap H = \{v\}$.

Exercise five. Prove that any bounded convex polyhedron of dimension d in $\mathbb{R}^d$ has at least $d+1$ vertices and at least $d+1$ facets.

The next two exercises deal with *convex cones*. Citing from Schrijver's lecture notes:

Convex cones are special cases of convex sets. A subset C of $\mathbb{R}^n$ is called a *convex cone* if for any $x, y \in C$ and any $\lambda, \mu \ge 0$ one has $\lambda x + \mu y \in C$.

For any $X \subseteq \mathbb{R}^n$, cone (X) is the smallest cone containing X. One easily checks:

$$\text{cone}(X) = \{\lambda_1 x_1 + \cdots \lambda_t x_t \mid x_1, \ldots, x_t \in X; \lambda_1, \ldots, \lambda_t \ge 0\}$$

Exercise six. Let $C \subseteq \mathbb{R}^n$. Then C is a closed convex cone if and only if $C = \bigcap \mathcal{F}$ for some collection $\mathcal{F}$ of linear halfspaces. (Notice the halfspaces are linear, i.e., going through 0.)

Exercise seven. For any subset X of $\mathbb{R}^n$, define

$$X^* := \left\{ y \in \mathbb{R}^n \mid x^T y \le 1 \text{ for each } x \in X \right\}$$

1. Show that for each convex cone C, C^* is a closed convex cone.
2. Show that for each closed convex cone $C, (C^*)^* = C$.