Representation theorem for polyhedral sets[⋆]

Anton Evgrafov

Department of Mathematical Sciences, NTNU anton.evgrafov@math.ntnu.no

Consider the following linear programming problem in the standard form:

$$\begin{aligned} &\underset{x\in\mathbb{R}^n}{\text{minimize}} && c^T x, \\ &\text{subject to} && Ax = b, \\ & && x \geq 0, \end{aligned} \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$, $c \in \mathbb{R}^N$, $b \in \mathbb{R}^m$. The existence of solutions for a feasible and bounded problem (1) relies upon the representation of the feasible set $\Omega = \{\, x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \,\}$ as a sum $\Omega = P + C$, P is a convex, closed, and bounded set and C is a closed convex cone.

Before we begin, we reformulate Ω in terms of inequalities only:

$$\Omega = \{\, x \in \mathbb{R}^n \mid \tilde{A}x \leq \tilde{b} \,\}, \tag{2}$$

where

$$\tilde{A} = \begin{pmatrix} A \\ -A \\ -I \end{pmatrix}, \qquad \tilde{b} = \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}. \tag{3}$$

Note that the matrix $\tilde{A} \in \mathbb{R}^{(2m+n)\times n}$ always has rank n due to the presence of the identity matrix in the last block-row. The representation theorem applies to all matrices $\tilde{A} \in \mathbb{R}^{\ell\times n}$ with rank n (full column rank in particular $\ell \geq n$), not only matrices of the form (3).

For every $x \in \Omega$ we will write $\overset{=}{A}_x$ and $\overset{=}{b}_x$ to denote those rows of $\tilde{A}$ and the corresponding components of $\tilde{b}$, where the inequalities are active (binding) at x. The rest of the rows of $\tilde{A}$/components of $\tilde{b}$ will be denoted with $\overset{<}{A}_x$ and $\overset{<}{b}_x$. Thus $\overset{=}{A}_x x = \overset{=}{b}_x$ and $\overset{<}{A}_x x < \overset{<}{b}_x$.

Consider all points $v_i \in \Omega$ such that $\operatorname{rank} \overset{=}{A}_{v_i} = n$; thus $v_i = \overset{=}{A}{}_{v_i}^{-1} \overset{=}{b}_{v_i}$. Note that the number of such points is not larger than the number of ways of selecting n rows out of ℓ possibilities, that is $\ell!/(n!(\ell-n)!)$, and in principle could be 0. For a given $\tilde{A}$ and $\tilde{b}$ we will denote this number with N. Let

$$\begin{aligned} P &= \left\{ \sum_{i=1}^{N} \lambda_i v_i \mid \lambda_i \geq 0, \sum_{i=1}^{N} \lambda_i = 1 \right\}, \\ C &= \{\, d \in \mathbb{R}^n \mid \tilde{A}d \leq 0 \,\}. \end{aligned} \tag{4}$$

[⋆] Based on Section 3.2.3 in "Introduction to continuous optimization" by N. Andréasson, AE, M. Patriksson, E. Gustavsson, M. Önnheim: Studentlitteratur (2013), 2nd ed.

Theorem 1 (Representation theorem). *Consider a matrix* $\tilde{A} \in \mathbb{R}^{\ell \times n}$ *and a vector* $\tilde{b} \in \mathbb{R}^{\ell}$ *defining the set* (2)*, and the sets* P *and* C *defined in* (4)*. Suppose that* $\operatorname{rank} \tilde{A} = n$*. If* P *is non-empty then* $\Omega = P + C$*.*

Proof. The inclusion $P + C \subset \Omega$ is easy to verify. The other inclusion is proved is by induction in $\operatorname{rank} \bar{A}_x$, $x \in \Omega$.

First, consider the points in $x \in \Omega$ with $\operatorname{rank} \bar{\bar{A}}_x = n$. These are precisely the points v_i defining the non-empty set P. Thus $x = v_i + 0$, for some $i = 1, \dots, N$. Note that $0 \in C$, thus $x \in P + C$.

Now assume that the representation holds for all $x \in \Omega$ such that $k \leq \operatorname{rank} \bar{\bar{A}}_x \leq n$. We will show that the representation holds also for points $x \in \Omega$ with $\operatorname{rank} \bar{\bar{A}}_x = k - 1$.

Let $x \in \Omega$ be such a point. Since $\operatorname{rank} \bar{\bar{A}}_x < n$ there is $0 \neq z \in \operatorname{null} \bar{\bar{A}}_x$. Consider a perturbed point $x + \lambda z$, $\lambda \in \mathbb{R}$. Since $\overset{<}{A}_x x < \overset{<}{b}_x$ and $\bar{\bar{A}}_x z = 0$, it holds that $x + \lambda z \in \Omega$ for all small λ.

Let $\lambda^+ = \sup\{\, \lambda \in \mathbb{R} : x + \lambda z \in \Omega \,\}$ and $\lambda^- = \sup\{\, \lambda \in \mathbb{R} : x - \lambda z \in \Omega \,\}$. If $\lambda^+ = +\infty$ then

$$\tilde{A} z = \lim_{\lambda \to +\infty} \lambda^{-1} \tilde{A}[x + \lambda z] \leq \lim_{\lambda \to +\infty} \lambda^{-1} \tilde{b} = 0.$$

and therefore $z \in C$. Similarly, if $\lambda^- = +\infty$ then $-z \in C$.

Case 1: Suppose that $\lambda^- = \lambda^+ = +\infty$; then $0 \neq z \in C \cap [-C] = \operatorname{null} \tilde{A}$, which contradicts the assumption $\operatorname{rank} \tilde{A} = n$.

Case 2: Suppose $\lambda^+ < \infty$ but $\lambda^- = +\infty$. Consider the point $x^+ = x + \lambda^+ z$. Then $x^+ \in \Omega$ since Ω is closed. We claim that $\operatorname{rank} \bar{\bar{A}}_{x^+} \geq k$. Indeed, $\bar{\bar{A}}_x$ is a submatrix of $\bar{\bar{A}}_{x^+}$ (recall, $\bar{\bar{A}}_x z = 0$) and thus $\operatorname{rank} \bar{\bar{A}}_{x^+} \geq k-1$. If $\operatorname{rank} \bar{\bar{A}}_{x^+} = k-1$ then the additional rows in $\bar{\bar{A}}_{x^+}$ (in relation to $\bar{\bar{A}}_x$) may be expressed as linear combinations of rows in $\bar{\bar{A}}_x$. Therefore, $z \in \operatorname{null} \bar{\bar{A}}_{x^+}$ and $x^+ + \lambda z \in \Omega$, for all small λ. This contradicts the selection of λ^+, which was such that $x + \lambda z \notin \Omega$, $\lambda > \lambda^+$. It remains to utilize the induction hypothesis for x^+, that is $x^+ = x + \lambda z \in P + C$, and as a result $x \in P + (C + \lambda^+(-z)) = P + C$, since in this case $-z \in C$.

Case 3: Suppose $\lambda^+ = +\infty$ but $\lambda^- < \infty$. This case is completely symmetric with *Case 2*.

Case 4: Suppose that $\lambda^+ < \infty$ and $\lambda^- < \infty$. In this case the induction hypothesis applies to both x^+ and x^-. Therefore

$$x = \frac{\lambda^+}{\lambda^+ + \lambda^-} x^- + \frac{\lambda^-}{\lambda^+ + \lambda^-} x^+ \in \frac{\lambda^+}{\lambda^+ + \lambda^-}(P + C) + \frac{\lambda^-}{\lambda^+ + \lambda^-}(P + C) \subset P + C,$$

where the last inclusion is owing to the convexity of P, C. □

Proposition 1 (Existence of extreme poitns; see Theorem 13.2 in N&W). *Suppose that* Ω *given by* (2) *is non-empty and* $\operatorname{rank} \tilde{A} = n$*. Then the set* P *defined in* (4) *is non-empty.*

Proof. Take any $x \in \Omega \neq \emptyset$. If $\operatorname{rank} \bar{\bar{A}}_x = n$ we are done; otherwise we proceed as in the proof of Theorem 1 and define λ^+, λ^-. If $\lambda^+ < \infty$ we then go to the point x^+; otherwise $\lambda^- < \infty$ and then we go to the point x^-. In any case, $\operatorname{rank} \bar{\bar{A}}_{x^+} > \operatorname{rank} \bar{\bar{A}}_x$ or $\operatorname{rank} \bar{\bar{A}}_{x^-} > \operatorname{rank} \bar{\bar{A}}_x$. Repeating this procedure, we eventually reach a point $x \in \Omega$ where $\operatorname{rank} \bar{\bar{A}}_x = n$. □