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:

\[
\begin{align*}
\text{minimize} & \quad c^T x, \\
\text{subject to} & \quad Ax = b, \\
& \quad x \geq 0,
\end{align*}
\]

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 \( \Omega \) in terms of inequalities only:

\[
\Omega = \{ x \in \mathbb{R}^n \mid \tilde{A}x \leq \tilde{b} \},
\]

where

\[
\tilde{A} = \begin{pmatrix} A \\ -A \\ -I \end{pmatrix}, \quad \tilde{b} = \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}.
\]

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 \( \tilde{A}_x \) and \( \tilde{b}_x \) 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 \( \tilde{A}_x \) and \( \tilde{b}_x \). Thus \( \tilde{A}_x x = \tilde{b}_x \) and \( \tilde{A}_x x < \tilde{b}_x \).

Consider all points \( v_i \in \Omega \) such that rank \( \tilde{A}_{v_i} = n \); thus \( v_i = \tilde{A}_{v_i}^{-1}\tilde{b}_{v_i} \). Note that the number of such points is not larger than the number of ways of selecting \( n \) rows out of \( \ell \) 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

\[
P = \left\{ \sum_{i=1}^{N} \lambda_i v_i \mid \lambda_i \geq 0, \sum_{i=1}^{N} \lambda_i = 1 \right\},
\]

\[
C = \{ d \in \mathbb{R}^n \mid \tilde{A}d \leq 0 \}.
\]

* 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 \( \tilde{A} \in \mathbb{R}^{t \times n} \) and a vector \( \tilde{b} \in \mathbb{R}^t \) defining the set (2), and the sets \( P \) and \( C \) defined in (4). Suppose that \( \text{rank} \tilde{A} = n \). 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 by induction in \( \text{rank} \tilde{A}_x, x \in \Omega \).

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

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

Let \( x \in \Omega \) be such a point. Since \( \text{rank} \tilde{A}_x < n \) there is \( 0 \neq z \in \text{null} \tilde{A}_x \).

Consider a perturbed point \( x + \lambda z, \lambda \in \mathbb{R} \). Since \( \tilde{A}_x x < \tilde{b}_x \) and \( \tilde{A}_x z = 0 \), it holds that \( x + \lambda z \in \Omega \) for all small \( \lambda \).

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 \( 0 \neq z \in C \cap [-C] = \text{null} \tilde{A} \), which contradicts the assumption \( \text{rank} \tilde{A} = n \).

Case 2: Suppose \( \lambda^+ < \infty \) but \( \lambda^- = +\infty \). Consider the point \( x^+ = x + \lambda^+ z \).

Then \( x^+ \in \Omega \) since \( \Omega \) is closed. We claim that \( \text{rank} \tilde{A}_{x^+} \geq k \). Indeed, \( \tilde{A}_{x^+} \) is a submatrix of \( \tilde{A}_{x^+} \) (recall, \( \tilde{A}_x z = 0 \)) and thus \( \text{rank} \tilde{A}_{x^+} \geq k - 1 \). If \( \text{rank} \tilde{A}_{x^+} = k - 1 \) then the additional rows in \( \tilde{A}_{x^+} \) (in relation to \( \tilde{A}_x \)) may be expressed as linear combinations of rows in \( \tilde{A}_x \). Therefore, \( z \in \text{null} \tilde{A}_{x^+} \) and \( x^+ + \lambda z \in \Omega \), for all small \( \lambda \). This contradicts the selection of \( \lambda^+ \), 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 \( x \in P + (C + \lambda^+ \Omega) = P + C \), 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 \( P, C \). \( \square \)

Proposition 1 (Existence of extreme points; see Theorem 13.2 in N&W). Suppose that \( \Omega \) given by (2) is non-empty and \( \text{rank} \tilde{A} = n \). Then the set \( P \) defined in (4) is non-empty.
Proof. Take any $x \in \Omega \neq \emptyset$. If rank $\bar{A}_x = n$ we are done; otherwise we proceed as in the proof of Theorem 1 and define $\lambda^+, \lambda^-$. If $\lambda^+ < \infty$ we then go to the point $x^+$; otherwise $\lambda^- < \infty$ and then we go to the point $x^-$. In any case, rank $\bar{A}_{x^+} > \text{rank } \bar{A}_x$ or rank $\bar{A}_{x^-} > \text{rank } \bar{A}_x$. Repeating this procedure, we eventually reach a point $x \in \Omega$ where rank $\bar{A}_x = n$. $\square$