# A New Character for Polyhedrons in $\mathbb{R}^n_+$

Evan O'Dea

January 19, 2007

# 1 Notation and Terminology

We begin by introducing the notation used throughout the paper. First, a subscript on a vector refers to a specific component of that vector. A number as a subscript refers to a single component, such as $x_1, x_2, x_3, x_i$, or $x_n$. A set as a subscript refers to a new vector composed of the indicated components. For example, if $S = \{1, 2, 4\}$, then $x_S = (x_1, x_2, x_4)^T$. Similar to how a subscript on a vector refers to certain components, we use a subscript on a matrix to refer to a subset of the columns of the matrix. A number as a subscript refers to a single column of a matrix. A set as a subscript refers to the indicated collection of columns, so using $S$ from above, $A_S = [A_1|A_2|A_4]$. Such labeling of vectors and matrices complement each other so that matrix multiplication works and we get $A_S x_S = A_1x_1 + A_2x_2 + A_4x_4$. We use $e$ to represent a vector of 1's. The length of $e$ varies to accommodate context. For example, given the matrix multiplication $Ae$, we infer that the length of $e$ is the same as the number of columns in $A$. We use the notation $<_L$ to mean lexicographically less than, and $>_L$ to mean lexicographically greater than. We say that $x$ is lexicographically less than $y$ if the minimal index, $i$, such that $x_i \neq y_i$ satisfies $x_i < y_i$. We write $x <_L y$.

For a permutation $\sigma$, we write $x_\sigma$ to denote the vector $x$ after its components are reordered according to $\sigma$. For example, given $x = (8, 5, 9, 7)^T$ and $\sigma = (3, 2, 4, 1)$, we have $x_\sigma = (x_3, x_2, x_4, x_1)^T = (9, 5, 7, 8)^T$. We use the function "sort" as a function that rearranges the components of a vector in a non-increasing order. For example, using $x = (8, 5, 9, 7)^T$ from above, we get $\text{sort}(x) = (9, 8, 7, 5)^T$. The "sort-order" of a vector $x$ is the order of the components of $x$ from greatest to smallest. We formalize this notion of sort-order by defining the set $\Sigma(x) = \{\sigma : x_\sigma = \text{sort}(x)\}$. We see that for $x = (8, 5, 9, 7)^T$, we have $\Sigma(x) = \{(3, 1, 4, 2)\}$. This is a clear and succinct way to denote that $x_3 \geq x_1 \geq x_4 \geq x_2$. We should point out that when a vector $x$ contains repeated values, there are multiple permutations on the components of $x$ that produce $\text{sort}(x)$, and so in general $|\Sigma(x)| \geq 1$. Consider $x = (5, 7, 6, 6)^T$. Using bold font to show the effect of the repeated 6, we have

$$\text{sort}(x) = (7, 6, 6, 5)^T = (x_2, x_{\mathbf{3}}, x_{\mathbf{4}}, x_1)^T = (x_2, x_{\mathbf{4}}, x_{\mathbf{3}}, x_1)^T,$$

so $\Sigma(x) = \{(2, \mathbf{3}, \mathbf{4}, 1), (2, \mathbf{4}, \mathbf{3}, 1)\}$. We say that two vectors $x$ and $y$ are sort-similar if

$$\Sigma(x) \cap \Sigma(y) \neq \emptyset.$$

We should point out that sort-similarity is not transitive. For example, consider $x = (4, 5, 6)^T, y = (4, 6, 6)^T$, and $z = (4, 6, 5)^T$. We get

$$\Sigma(x) = \{(3, 2, 1)\}, \quad \Sigma(y) = \{(3, 2, 1), (2, 3, 1)\}, \quad \Sigma(z) = \{(2, 3, 1)\}.$$

Notice that $x$ and $y$ are sort-similar, and $y$ and $z$ are sort-similar, but $x$ and $z$ aren't sort-similar. This example illustrates that two vectors $x$ and $z$ are not sort-similar if and only if there exist indices $i$ and $j$ such that $x_i > x_j$ and $z_i < z_j$. This fact is fundamental to some of the proofs included in this paper.

# 2 Introduction

Our work is motivated by the beam selection problem in radiotherapy, as presented in [1]. We will briefly review the essential parts of their treatment design process, but for an in-depth look please refer to the paper. The method in [1] is based on a probability distribution that assigns a probability to each angle from which radiation might be delivered. The probability assignment is based on the lexmin-sort vector, which we call $z^*$ and will define shortly. In [1], the authors present an iterative algorithm that computes the lexmin-sort $z^*$, and we expect our audience to have some familiarity with this, but we will review the algorithm and introduce our own notation to better suit our purposes.

In the first step of treatment design, a physician details a *prescription*. The prescription defines the amount of energy sought to be delivered to the target, as well as the limitations on the energy that can safely be delivered to the healthy anatomy. We model the prescription as a vector $\hat{b}$, which describes the bounds for the various anatomical structures. A *treatment* refers to one possible way to deliver radiation to the patient, that is, a particular treatment assigns an exposure time to each angle. We model a treatment as a vector $x$, whose components represent the exposure time for each angle. Finally, since the physical interactions of the beams and the energy that they deliver follow linear relationships, we can create a linear transformation that accurately maps a treatment to the energy that it delivers throughout the patient's anatomy. We model this as a matrix $A$ so that the function $x \mapsto Ax$ maps the treatment $x$ to the energy that it delivers throughout the anatomy. This forms the basis for the system from which an ideal treatment $x$ will be selected. We consider the polyhedron

$$\mathcal{P} = \{x : Ax = b, x \geq 0\},$$

where all $x \in \mathcal{P}$ are considered optimal. Assume in all cases that $A \in \mathbb{R}^{m \times n}$.

In the beam selection method presented in [1], a probability distribution assigns value to each $i \in \{1, 2, 3, \ldots, n\}$. Since each index of $x \in \mathcal{P}$ represents one angle from which radiation might be emitted, this effectively assigns probability to each angle. The probability distribution is based on the lexmin-sort vector of $\mathcal{P}$, defined as

$$z^* = \text{lexmin}\{\text{sort}(x) : x \in \mathcal{P}\}.$$

Additionally, we define

$$\mathcal{P}^* = \text{arglexmin}\{\text{sort}(x) : x \in \mathcal{P}\} = \{x \in \mathcal{P} : \text{sort}(x) = z^*\}.$$

The probability distribution that the authors of [1] have designed assigns probability as

$$p(i) = \frac{1}{\|x^*\|} x^*.$$

This method effectively assigns high probability values, meaning that it is useful in beam selection, but for the purpose of reducing the size of the problem we are interested in angles that are assigned a value of zero, i.e. $p(i) = 0$. If we can isolate angles such that $x_i = 0$ for all $x \in \mathcal{P}$, then we can remove them from consideration and thus reduce the size of the problem. So for the purpose of radiotherapy, our question becomes Does $p(i) = 0$ imply that $x_i = 0$ for all $x \in \mathcal{P}$? (The reverse of this is clearly true.) Notice, though, that if $x_i = 0$ for all $x \in \mathcal{P}$, then in our system $\mathcal{P}$ we have an implied equality for $x_i \geq 0$. So our question in radiotherapy leads us to the broader task of studying if (and if so, when) this probability distribution can identify implied equalities in linear systems.

We pose our question using two characters defined on $\{1, 2, 3, \ldots, n\}$. We define the character $(B|N)$ by

$$N = \{i : x_i = 0 \ \ \forall\, x \in \mathcal{P}\}, \quad \text{and} \quad B = \{i : \exists\, x \in \mathcal{P} \text{ such that } x_i > 0\},$$

and $(\beta|\varepsilon)$ by

$$\eta = \{i : x_i^* = 0\}, \ \text{and} \ \ \beta = \{i : x_i^* > 0\}.$$

So we see that we have $p(i) = 0 \iff x_i = 0$ for all $x \in \mathcal{P}$ if and only if we have $(B|N) = (\beta|\eta)$. It is simple to conjure examples that show that this is not always the case, so our work is aimed at identifying conditions that could guarantee that $(B|N) = (\beta|\eta)$.

## 3 Results and Conjectures

Now we move into our results, the first of which requires a lengthy lemma.

**Lemma 3.1** *Let $h, k$ be such that $h, k \in \mathbb{R}^n$, $h, k \geq 0$, $h \neq k$, and sort$(h) =$ sort$(k)$. Let $h_a$ and $h_b$ be two components of $h$ such that $h_a > h_b$ and $h_a - h_b = \min\{h_i - h_j : h_i \neq h_j\}$. (Notice that there must exist two such elements because if $h_i = h_j$ for all $i, j$ then $h = k$, contradicting our hypothesis.) Let $x(\alpha) = (1-\alpha)h + \alpha k$. If*

$$\alpha \in \left(0, \frac{h_a - h_b}{h_{\max} + h_a - h_{\min} - h_b}\right),$$

*then $x(\alpha)$ and $h$ are sort-similar, and sort$(x(\alpha)) <_L$ sort$(h)$.*

**Proof:** In the process of forming $x(\alpha)$ from $h$ and $k$, we say that $h_i$ "combines" with $k_i$ to form $x_i(\alpha)$. Suppose that $x(\alpha)$ and $h$ are not sort-similar and suppose that $h_a$ combines with $k_{\min}$ and $h_b$ combines with $k_{\max}$. Since $x(\alpha)$ and $h$ are not sort-similar, then there exist $i, j$ (not necessarily distinct from $a, b$) such that $h_i > h_j$ and $x_i(\alpha) < x_j(\alpha)$, so we have

$$\begin{array}{rcl} x(\alpha')_i & < & x(\alpha')_j \\ (1-\alpha')h_i + \alpha' k_i & < & (1-\alpha')h_j + \alpha' k_j \\ (1-\alpha')(h_i - h_j) & < & \alpha'(k_j - k_i) \\ (1-\alpha')(h_a - h_b) \leq (1-\alpha')(h_i - h_j) & < & \alpha'(k_j - k_i) \leq \alpha'(k_{\max} - k_{\min}) \\ (1-\alpha')h_a + \alpha' k_{\min} & < & (1-\alpha')h_b + \alpha' k_{\max} \\ x(\alpha')_a & < & x(\alpha')_b. \end{array}$$

Now we show that our constraints on $\alpha$ guarantee $x_a(\alpha) > x_b(\alpha)$, regardless of which components of $k$ $h_a$ and $h_b$ combine with.

Let $\alpha' \in \left(0, \frac{h_a - h_b}{h_{\max} + h_a - h_{\min} - h_b}\right)$. Then

$$\begin{array}{rcl} \frac{h_a - h_b}{h_{\max} + h_a - h_{\min} - h_b} & > & \alpha' \\ h_a - h_b & > & \alpha'(h_{\max} + h_a - h_{\min} - h_b) \\ & > & \alpha' h_{\max} + \alpha' h_a - \alpha' h_{\min} - \alpha' h_b \\ h_a - \alpha' h_a + \alpha' h_{\min} & > & h_b - \alpha' h_b + \alpha' h_{\max} \\ (1-\alpha')h_a + \alpha' h_{\min} & > & (1-\alpha')h_b + \alpha' h_{\max}. \end{array}$$

Since $h_{\min} = k_{\min}$ and $h_{\max} = k_{\max}$, for all $i$ we have

$$\begin{array}{rcl} (1-\alpha')h_a + \alpha' k_{\min} & > & (1-\alpha')h_b + \alpha' k_{\max} \\ (1-\alpha')h_a + \alpha' h_i \leq (1-\alpha')h_a + \alpha' h_{\min} & > & (1-\alpha')h_b + \alpha' h_{\max} \geq (1-\alpha')h_b + \alpha' h_i. \end{array}$$

So $x_a(\alpha') > x_b(\alpha')$. Therefore given our constraint on $\alpha$, we are assured that if $h_i > h_j$, then we have $x_i(\alpha) > x_j(\alpha)$. Thus $x(\alpha)$ and $h$ are sort-similar.

Now we show that sort$(x(\alpha)) <_L$ sort$(h)$. From the definition of $x(\alpha)$ we see that

$$\begin{array}{rcl} h_i = k_i & \Leftrightarrow & h_i = x(\alpha)_i \\ h_i > k_i & \Leftrightarrow & h_i > x(\alpha)_i. \end{array}$$

Since $h \neq k$ and sort$(h) =$ sort$(k)$, we know that $h$ and $k$ are different permutations of the same values, so for some $i$ we might have $h_i = k_i$. Since $h \neq k$, there exists some $i$ such that $h_i \neq k_i$. Let $t$ be the index such that $h_t \neq k_t$ and $h_t \geq h_j$ for all $h_j \neq k_j$. So $h_t$ is the maximal component of $h$ that doesn't combine with an equivalent component of $k$. For each $h_i > h_t$, we have $h_i = k_i$, so $h_i = x_i(\alpha)$. Since $h$ and $k$ are permutations of the same values, each $k_i > h_t$ has already combined with $h_i > h_t$. Thus since $h_t \neq k_t$, we have $h_t > k_t$, and so $h_t > x_t(\alpha)$. Since $x(\alpha)$ and $h$ are sort-similar, we have that the first discrepancy encountered when lexicographically comparing sort$(x(\alpha))$

and sort($h$) is $x_t(\alpha) < h_t$. Therefore sort($x(\alpha)$) $<_L$ sort($h$).
■

Now we use this lemma to prove our first theorem.

**Theorem 3.2** $|\mathcal{P}^*| = 1$.

**Proof:** Suppose $h, k \in \mathcal{P}^*$ with $h \neq k$. Then $h, k \in \mathcal{P}$, so $h, k \geq 0$. Letting $x(\alpha) = (1-\alpha)h + \alpha k$ as defined in Lemma 3.1, we have $x(\alpha) \in \mathcal{P}$ for $\alpha \in (0,1)$ because it's a convex combination of $h$ and $k$. Let us constrain

$$0 < \alpha < \frac{h_1 - h_2}{h_{\max} + h_1 - h_{\min} - h_2}.$$

Since $h, k \in \mathcal{P}^*$ we know that sort($h$) = sort($k$) = $z^*$. So by Lemma 3.1, $x(\alpha)$ and $h$ are sort-similar and sort($x(\alpha)$) $<_L$ sort($h$), contradicting that $h \in \mathcal{P}^*$. Thus there cannot be two elements in $\mathcal{P}^*$.
■

We continue by discussing the lexmin-sort algorithm presented in [1] and the relevance of the lexmin-sort vector. We expect some familiarity with the lexmin-sort algorithm, but we adapt the notation to the $(B|N)$ and $(\beta|\eta)$ partitions to better suit our presentation.

We begin by defining $\eta_0 = \{1, 2, \ldots, n\}$ and $\beta_0 = \emptyset$. As the algorithm progresses, we fix certain components of $x \in \mathcal{P}$, and in doing so remove them from $\eta_t$ and assign them to $\beta_t$. The intuition behind these steps and the selected notation is that in the $t$-th iteration of the algorithm, $\eta_t$ describes the set of all components of $x$ that are not yet fixed, so $\eta_t$ is the set of indices that could possibly be in $\eta$. Contrarily, $\beta_t$ is the set of indices that we have so far determined are in $\beta$. To be consistent with our notation from the start, we think of $Ax = b$ as $A_{\eta_0} x_{\eta_0} = b$, and we think of $\mathcal{P}$ as $\mathcal{P}_0$. We then define

$$z_1^* = \min\{\|x_{\eta_0}\|_\infty : x_{\eta_0} \in \mathcal{P}_0\},$$

$$\beta_1 = \{i : (x_{\eta_0})_i \text{ must be } z_1^* \text{ when } \|x_{\eta_0}\|_\infty \text{ is minimized to } z_1^*\}, \text{ and } \eta_1 = \eta_0 \setminus \beta_1.$$

Next we define

$$\mathcal{P}_1 = \{x : x \in \mathcal{P}_0,\ x_{\beta_1} = z_1^*\},$$

and then begin the process again, this time working with $\mathcal{P}_1$. We define

$$z_2^* = \min\{\|x_{\eta_1}\|_\infty : x \in \mathcal{P}_1\},$$

$$\beta_2 = \{i : (x_{\eta_1})_i \text{ must be } z_2^* \text{ when } \|x_{\eta_1}\|_\infty \text{ is minimized to } z_2^*\}, \text{ and } \eta_2 = \eta_1 \setminus \beta_2,$$

and then define $\mathcal{P}_2 = \{x : x \in \mathcal{P}_1,\ x_{\beta_2} = z_2^*\}$ to repeat. The algorithm terminates in the $i$-th iteration when $\eta_i = \emptyset$, and the result is that $x^*$ is the only element contained in $\mathcal{P}_i$.

We attain the following generalized form for the algorithm.

- Define $\eta_0 = \{1, 2, \ldots, n\}$, $\beta_0 = \emptyset$, and $\mathcal{P}_0 = \mathcal{P} = \{x : Ax = b, x \geq 0\}$ as above; and initialize $t = 1$.

- Do until $\eta_t = \emptyset$:

Let

$$\begin{array}{rcl} z_t^* & = & \min\{\|x_{\eta_{(t-1)}}\|_\infty : x \in \mathcal{P}_{t-1}\} \\ \beta_t & = & \{i : (x_{\eta_{(t-1)}})_i \text{ must be } z_t^* \text{ when } |x_{\eta_{(t-1)}}|_\infty = z_t^*\} \\ \eta_t & = & \eta_{(t-1)} \setminus \beta_t \\ \mathcal{P}_t & = & \{x : x \in \mathcal{P}_{t-1},\ x_{\beta_t} = z_t^*\} \\ t & = & t+1 \end{array}$$

- When $\eta_t = \emptyset$, we have $|\mathcal{P}_t| = 1$ with $\mathcal{P}_t = \{x^*\}$.

From $x^*$ we can then find $(\beta|\eta)$. If in the last iteration $z_t^* = 0$, then $\eta = \eta_{t-1}$. Now we present some corollaries that follow directly from the algorithm.

**Corollary 3.3** *$\mathcal{P}_t \subseteq \mathcal{P}_{t-1}$ for all $t$, so for $i > j$ we have $\mathcal{P}_i \subseteq \mathcal{P}_j$.*

**Corollary 3.4** *$z_t < z_{t-1}$ for all $t$, so for $i > j$ we have $z_i < z_j$.*

Now we take a more in-depth look at the steps of the algorithm to help illustrate the process. Consider the first step of the algorithm. To make $\mathcal{P}_1$ we fix $x_{\beta_1} = z_1^*$, but consider how this affects the equation $A_{\eta_0} x_{\eta_0} = b$. By definition $\beta_1$ and $\eta_1$ form a partition of $\eta_0$, so we get

$$A_{\eta_0} x_{\eta_0} = [A_{\eta_1}|A_{\beta_1}]x = A_{\eta_1} x_{\eta_1} + A_{\beta_1} x_{\beta_1} = b.$$

Fixing $x_{\beta_1} = z_1^*$ then gives

$$A_{\eta_1} x_{\eta_1} + z_1^* A_{\beta_1} e = b,$$

$$\text{so} \quad A_{\eta_1} x_{\eta_1} = b - z_1^* A_{\beta_1} e.$$

This provides an alternative way to represent $\mathcal{P}_1$:

$$\mathcal{P}_1 = \{x : x \in \mathcal{P}_0, x_{\beta_1} = z_1^*\} = \{x : A_{\eta_1} x_{\eta_1} = b - z_1^* A_{\beta_1} e,\ x \geq 0\}.$$

We can carry out this process again in the next step by separating $A_{\eta_1}$ according to $\beta_2$ and $\eta_2$, and then fixing $x_{\beta_2} = z_2^*$:

$$\begin{array}{rcl} A_{\eta_1} x_{\eta_1} & = & b - z_1^* A_{\beta_1} e \\ [A_{\eta_2}|A_{\beta_2}] x_{\eta_1} & = & b - z_1^* A_{\beta_1} e \\ A_{\eta_2} x_{\eta_2} + A_{\beta_2} x_{\beta_2} & = & b - z_1^* A_{\beta_1} e \\ A_{\eta_2} x_{\eta_2} + z_2^* A_{\beta_2} e & = & b - z_1^* A_{\beta_1} e \\ A_{\eta_2} x_{\eta_2} & = & b - z_1^* A_{\beta_1} e - z_2^* A_{\beta_2} e. \end{array}$$

Again we have the alternative form $\mathcal{P}_2 = \{x : A_{\eta_2} x_{\eta_2} = b - z_1^* A_{\beta_1} e - z_2^* A_{\beta_2} e,\ x \geq 0\}$. In general, we have

$$\mathcal{P}_t = \{x : x \in \mathcal{P}_{t-1},\ x_{\beta_{(t-1)}} = z_{t-1}^*\} = \left\{ x_{\eta_t} : A_{\eta_t} x_{\eta_t} = b - \sum_{j=1}^{t} z_j^* A_{\beta_j} e,\ x_{\eta_t} \geq 0 \right\}.$$

Before we conclude our discussion of the algorithm, let us define $u$ to be the last iteration of the algorithm so that $\eta_u = \emptyset$, $\mathcal{P}_u = \{x^*\}$, and $z_u^* = \min\{x_i^* : i = 1, 2, \ldots, n\}$. Notice that we have $z_u^* > 0$ if and only if $\eta = \emptyset$.

Now we introduce some theorems and proofs that expand our understanding of the relationships between a linear system and how $(B|N)$ and $(\beta|\eta)$ compare to each other. But first, we provide the following corollary that follows directly from the definitions of these partitions.

**Observation 3.5** *We have $(B|N) \neq (\beta|\eta)$ if and only if $|N| < |\eta|$. So if $|\eta| = 0$, then $(B|N) = (\beta|\eta)$.*

**Theorem 3.6** *Let $A^+ \in \mathbb{R}^{n\times m}$ be the generalized inverse of $A \in \mathbb{R}^{m\times n}$. Given $Ax = b$, if either $A^+b > 0$ or if there exists $y \in Null(A)$ such that $y > 0$, then there exists $x > 0$ such that $Ax = b$.*

**Proof:** From the properties of the generalized inverse, we know that solutions to $Ax = b$ are of the form

$$x = A^+b + (I - A^+A)q, \quad \text{for any } q \in \mathbb{R}^n.$$

From the following algebra, we see that $(I - A^+A)q \in \text{Null}(A)$:

$$A[(I - A^+A)q] = (A - AA^+A)q = (A - A)q = (0)q = 0.$$

Further, by letting $q \in \text{Null}(A)$, we see that the elements of $\text{Null}(A)$ are of the form $(I - A^+A)q$. Thus $x = A^+b + y$, where $y$ is any vector in $\text{Null}(A)$.

If $A^+b > 0$, then we let $y = 0$ and we have $x = A^+b > 0$.

If $A^+b \not> 0$ and $y > 0$, then we let

$$\delta > \max\left\{\frac{(A^+b)_i}{y_i} \; i = 1, 2, \ldots, n\right\},$$

so $A^+b + \delta y > 0$. Since $y \in \text{Null}(A)$, we have that $\delta y \in \text{Null}(A)$. Thus we have $x = A^+b + \delta y > 0$.
■

**Theorem 3.7** *There exists $x \in \mathcal{P}_t$ with $x_{\eta_t} > 0$ for all $t$ if and only if there exists $x \in \mathcal{P}_u$ with $x_{\eta_u} > 0$.*

**Proof:** The fact that $x_{\eta_t} > 0$ for all $t \Longrightarrow x_{\eta_u}$ follows directly from the definitions of $\eta_t$ and $\eta_u$. Let $x' \in \mathcal{P}_u$ with $x'_{\eta_u} > 0$. Since $u$ is the last iteration of the algorithm, we have $\eta_u \cup \left(\bigcup_{j=1}^{u} \beta_j\right) = \{1, 2, \ldots, n\}$. Suppose $i \in \eta_u$. Then $x'_i > 0$. Now suppose $i \notin \eta_u$. Then we have $i \in \beta_t$ for some $t$, so $x'_i = z^*_t > 0$. So for all $i$ we have $x'_i > 0$, so $x' > 0$. Since $\mathcal{P}_u \subseteq \mathcal{P}_t$ for all $t$, we have that $x' \in \mathcal{P}_t$ for all $t$.
■

**Theorem 3.8** *If $(\beta|\eta) = (B|N)$ with $\eta = N = \emptyset$, then there exists $x \in \mathcal{P}_{u-1}$ with $x_{\eta_{u-1}} > 0$.*

**Proof:** Suppose that there does not exist $x \in \mathcal{P}_{u-1}$ with $x_{\eta_{u-1}} > 0$. Then there exists $i$ such that $x_i = 0$ for all $x \in \mathcal{P}_{u-1}$. Thus by the definition of $u$ as the last iteration, $i \in \eta$, and so $\eta \neq \emptyset$, and the theorem is proven by contrapositive.
■

The reverse of Theorem 3.8 is true only given that the columns of $A_{\eta_{u-1}}$ are linearly independent, and we present this theorem followed by an example illustrating the necessity of the additional restriction.

**Theorem 3.9** *If the columns of $A_{\eta_{u-1}}$ are linearly independent and if $\exists\, x \in \mathcal{P}_{u-1}$ with $x_{\eta_{u-1}} > 0$, then $(\beta|\eta) = (B|N)$ with $\eta = N = \emptyset$.*

**Proof:** Assume the hypothesis. Since the columns of $A_{\eta_{u-1}}$ are linearly independent, the linear transformation $x_{\eta_{u-1}} \mapsto A_{\eta_{u-1}} x_{\eta_{u-1}}$ is 1:1. So

$$A_{\eta_{u-1}} x_{\eta_{u-1}} = b - \sum_{j=1}^{u-1} z_j^* A_{\beta_j} e$$

has a unique solution, $\hat{x}$, with $\hat{x}_{\eta_{u-1}} > 0$ by hypothesis. So $z_u^* > 0$, $\eta = \emptyset$, and $(B|N) = (\beta|\eta)$.
■

**Example 3.10** *We demonstrate that the condition that the columns of $A_{\eta_{u-1}}$ are linearly independent is necessary for Theorem 3.9.*

*Let*

$$A_0 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \textit{and} \;\; b = (2,2)^T,$$

$$\textit{so} \;\; \mathcal{P}_0 = \{(1,2,x_3)^T : x_3 \geq 0\}.$$

*We see that $(1,2,1)^T \in \mathcal{P}_0$, so $B = \{1,2,3\}$ and $N = \emptyset$. Proceeding with the algorithm we have $z_1^* = 2$ and $\beta_1 = \{2\}$, and we get*

$$\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} x_{\eta_1} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} (2) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

*with $\mathcal{P}_1 = \{(1,x_3)^T : x_3 \geq 0\}$. Proceeding again we have $z_2^* = 1$ and $\beta_2 = \{1\}$, which gives*

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} x_{\eta_2} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} - \begin{bmatrix} 2 \\ 0 \end{bmatrix} (1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

*with $\mathcal{P}_2 = \{(x_3) \;\; \text{for any} \;\; x_3 \geq 0\}$. We see that $\exists\, x_{\eta_2} \in \mathcal{P}_2$ with $x_{\eta_2} > 0$, but $z_3^* = 0$, so $3 \in \eta$. Thus $\eta \neq N$ and $(B|N) \neq (\beta|\eta)$.*

In some cases when $(B|N) \neq (\beta|\eta)$, negligible adjustment of $b$ can result in $(B|N) = (\beta|\eta)$ as desired.

**Example 3.11** *Let $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$ and $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Then $\mathcal{P} = \left\{ \begin{pmatrix} 1 - x_3 \\ 1 + x_3 \\ x_3 \end{pmatrix} : x_3 \geq 0 \right\}$, so $N = \emptyset$. But $z^* = (1,1,0)^T$, so $\eta = \{3\} \neq N$. However, if we change $b$ to $b' = \begin{pmatrix} 1+\varepsilon \\ 1 \end{pmatrix}$ for arbitrarily small $\varepsilon > 0$, then the results are different. We get that $\mathcal{P} = \left\{ \begin{pmatrix} 1 + \varepsilon - x_3 \\ 1 + x_3 \\ x_3 \end{pmatrix} : x_3 \geq 0 \right\}$, so again we have $N = \emptyset$, but this time $z^* = \begin{pmatrix} 1 + \frac{\varepsilon}{2}, & 1 + \frac{\varepsilon}{2}, & \frac{\varepsilon}{2} \end{pmatrix}^T$, so $\eta = \emptyset = N$.*

In the beam selection problem, the prescription is based on medical approximation, so tiny adjustments would not have adverse effects on the patient. Though we have observed this possibility, we have not yet had an opportunity to investigate how we might identify the direction(s) in which we could shift $b$ to achieve $(B|N) = (\beta|\eta)$. However, further inspection verified that in some cases no adjustment of $b$ provides $(B|N) = (\beta|\eta)$ as desired.

**Example 3.12** *Let* $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$ *and* $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ *be arbitrary. Then* $\mathcal{P} = \left\{ \begin{pmatrix} b_1 + x_3 \\ b_2 + x_3 \\ x_3 \end{pmatrix} : x_3 \geq 0 \right\}$, *so* $N = \emptyset$. *But* $z^* = (b_1, b_2, 0)^T$, *so* $\eta = \{3\} \neq N$. *This shows that* $3 \in \eta$ *for all* $b$, *highlighting that we can't always perturb* $b$ *to get our result.*

**Theorem 3.13**

$$\beta = B \quad \Longrightarrow \quad Null(A_\eta) \cap \left(\mathbb{R}_+^{|\eta|} \setminus \{0\}\right) = \emptyset.$$

**Proof:** By contrapositive: Suppose

$$y \in \text{Null}(A_\eta) \cap \left(\mathbb{R}_+^{|\eta|} \setminus \{0\}\right).$$

Let $y' \in \mathbb{R}^n$ be such that $y'_\beta = 0, y'_\eta = y$. Since $y'_\beta = 0$, $x^*_\eta = 0$, and $y'_\eta \in \text{Null}(A_\eta)$, we have

$$\begin{aligned} A(x^* + y') &= A_\beta(x^* + y')_\beta + A_\eta(x^* + y')_\eta \\ &= A_\beta x^*_\beta + A_\beta y'_\beta + A_\eta x^*_\eta + A_\eta y'_\eta \\ &= A_\beta x^*_\beta + 0 + 0 + 0 \\ &= A_\beta x^*_\beta. \end{aligned}$$

Similarly,

$$b = Ax^* = A_\beta x^*_\beta + A_\eta x^*_\eta = A_\beta x^*_\beta + 0,$$

so

$$A(x^* + y') = A_\beta x^*_\beta = b.$$

Also, we know $x^* \geq 0$ and $y' \geq 0$, therefore $(x^* + y') \in \mathcal{P}$. But $\exists\ i \in \eta$ such that $y'_i > 0$, so $(x^*_i + y_i) > 0$. Thus $i \in B$ and $i \notin N$, and so $\eta \neq N$ and $\beta \neq B$.
■

The reverse of Theorem 3.13 is not true, which we show with the following counterexample.

**Example 3.14** *Let*

$$A = \begin{bmatrix} 1 & 0 & 2 & 1 & 0 \\ 0 & 1 & 3 & 0 & 1 \end{bmatrix} \quad \text{and} \quad b = \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

*Then the vector* $x^* = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \in \mathcal{P}^*$. *So* $\eta = \{2, 5\}$, *and* $A_\eta = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$. *Thus*

$$Null(A_\eta) = span\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}, \quad so \quad Null(A_\eta) \cap \left(\mathbb{R}_+^{|\eta|} \setminus \{0\}\right) = \emptyset.$$

*Thus the hypothesis is true, but* $x' = \begin{pmatrix} 1 \\ 1 \\ \frac{1}{2} \\ 2 \\ \frac{1}{2} \end{pmatrix} \in \mathcal{P}$, *so* $B = \{1, 2, 3, 4, 5\} \neq \{1, 3, 4\} = \beta$.

To close our paper, we consider the environment(s) in which $z^*$ exists. For $\mathcal{P} = \{x : Ax = b, x \geq b\}$, let $\Sigma_j$ be one of the $n!$ permutation matrices on the columns of $A$. (So $A\Sigma_j$ is a reordering of the columns of $A$.) We define the following:

$$\begin{array}{l} \mathcal{P}_j = \{x : A\Sigma_j x = b, x \geq 0\}, \\ \mathcal{C} = \{x \in \mathbb{R}^n : x_1 \geq x_2 \geq x_3 \geq \cdots \geq x_n\}, \text{ and} \\ \bar{\mathcal{P}} = \mathcal{C} \cap \text{conhull}\left\{\bigcup \mathcal{P}_j\right\}. \end{array}$$

Let us represent $\bar{\mathcal{P}}$ as $\{x : \bar{A}x = \bar{b}, x \geq 0\}$. Notice that for any $x \in \mathcal{P}$, the union $\bigcup \mathcal{P}_j$ contains all permutations of the components of $x$. Therefore $\text{sort}(x) \in \bar{\mathcal{P}}$ for all $x \in \mathcal{P}$, and so $z^* \in \bar{\mathcal{P}}$.

**Conjecture 3.15** *$z^*$ is a vertex of $\bar{\mathcal{P}}$.*

We do not have a formal proof for this idea, but we have established the conjecture with an additional condition. This proof relies on the following generalization of Lemma 3.1.

**Lemma 3.16** *For any $x \in \mathbb{R}^n$, there exists $\delta > 0$ such that for all $q$ with $d(x,q) < \delta$, we have that $q$ and $x$ are sort-similar.*

**Proof:** Let $x_a$ and $x_b$ be components of $x$ such that $x_a > x_b$ and $x_a - x_b = \min\{x_i - x_j : x_i - x_j > 0\}$. Let $\delta < \frac{1}{2}(x_a - x_b)$. Then for all $q \in N(x, \delta)$ we have

$$\begin{array}{rcl} q_a > x_a - \delta & > & x_a - \frac{1}{2}(x_a - x_b) \\ & = & \frac{1}{2}x_a + \frac{1}{2}x_b \\ & = & x_b + \frac{1}{2}(x_a - x_b) > x_b + \delta > q_b. \end{array}$$

So $q_a > q_b$, and so $q$ and $x$ are sort-similar.
■

**Theorem 3.17** *If $z^*$ contains no repeated values, then $z^*$ is a vertex of $\bar{\mathcal{P}}$.*

**Proof:** Suppose $z^*$ is not a vertex of $\bar{\mathcal{P}}$. Then $z^*$ is not a basic feasible solution of $\bar{\mathcal{P}}$, and so the columns of $A_\beta$ are linearly dependent. Therefore there exists $y \neq 0$ such that $\bar{A}_\beta y = 0$. So for any scalar $\alpha$, we have

$$\bar{A}(z^* + \alpha y) = \bar{A}z^* + \bar{A}(\alpha y) = \bar{A}z^* + 0 = \bar{b}.$$

By selecting $\alpha$ sufficiently close to zero, we get that $(z^* + \alpha y) \geq 0$, so $(z^* + \alpha y) \in \bar{\mathcal{P}}$. Let $t$ be an index such that $y_t \neq 0$ and $z_t^* > z_i^*$ for all $i$ with $y_i \neq 0$. (We can assert strict inequality here because $z^*$ contains no repeated values.) From Lemma 3.16, let $\delta > 0$ be such that for all $q \in N(z^*, \delta)$, $q$ and $x$ are sort-similar. Select $\alpha' \in (-\delta, \delta) \setminus \{0\}$ sufficiently close to zero such that $(z^* - \alpha' y) \in N(z^*, \delta)$. Then $(z^* - \alpha' y)$ and $z^*$ are sort-similar. Also select $\alpha'$ with appropriate sign so that $(z^* - \alpha' y)_t = (z_t^* - \alpha' y) < z_t^*$. Then for $z_i^* > z_t^*$, since $y_i = 0$ by definition of $t$, we have that $(z^* - \alpha' y)_i = z_i^*$. Therefore $\text{sort}(z^* - \alpha' y) <_L \text{sort}(z^*)$, contradicting the definition of $z^*$. Thus $z^*$ is a vertex of $\bar{\mathcal{P}}$.
■

# 4 Conclusion

We've only begun to understand the relationships between $(\beta|\eta)$ and $(B|N)$. This area of study is new, so we think it can be a fruitful field for further research, and hopefully we can accumulate more results and better understand the relationships at hand.

## 5 Supplementary Examples

To end, we provide several examples of simple linear systems, pointing out how they compare to the theorems and counterexamples previously established.

**1.** Follows Theorem 6:

$$[A|b] = \left[\begin{array}{cccc|c} 2 & 0 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 & 8 \\ 2 & 0 & 0 & 1 & 2 \end{array}\right], \quad \mathcal{P} = \left\{ \begin{pmatrix} 1 \\ 8 - x_3 \\ x_3 \\ 0 \end{pmatrix} \right\}, \quad x^* = \begin{pmatrix} 1 \\ 4 \\ 4 \\ 0 \end{pmatrix}, \quad N = \{4\}, \ \ \eta = \{4\},$$

$$\text{Null}(A) = \text{span}\left\{ \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad \text{Null}(A_\eta) = \{ \begin{pmatrix} 0 \end{pmatrix} \}, \ \ \text{Null}(A_\eta) \cap \left(\mathbb{R}_+^{|\eta|} \setminus \{0\}\right) = \emptyset.$$

**2.** Counterexample to the reverse of Theorem 6:

$$[A|b] = \left[\begin{array}{cccc|c} 2 & 0 & 0 & 0 & 2 \\ 0 & -1 & 1 & 0 & 4 \\ 0 & -2 & 2 & 0 & 8 \\ 2 & 0 & 0 & 1 & 2 \end{array}\right], \quad \mathcal{P} = \left\{ \begin{pmatrix} 1 \\ x_3 - 4 \\ x_3 \\ 0 \end{pmatrix} \right\}, \quad x^* = \begin{pmatrix} 1 \\ 0 \\ 4 \\ 0 \end{pmatrix}, \quad N = \{4\}, \ \ \eta = \{2, 4\},$$

$$\text{Null}(A) = \text{span}\left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad \text{Null}(A_\eta) = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}, \ \ \text{Null}(A_\eta) \cap \left(\mathbb{R}_+^{|\eta|} \setminus \{0\}\right) = \emptyset.$$

**3.** Counterexample to the reverse of Theorem 6:

$$[A|b] = \left[\begin{array}{cccc|c} 2 & 0 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0 & 4 \\ 0 & 2 & -2 & 0 & 8 \\ 2 & 0 & 0 & 1 & 2 \end{array}\right], \quad \mathcal{P} = \left\{ \begin{pmatrix} 1 \\ 4 + x_3 \\ x_3 \\ 0 \end{pmatrix} \right\}, \quad x^* = \begin{pmatrix} 1 \\ 4 \\ 0 \\ 0 \end{pmatrix}, \quad N = \{4\}, \ \ \eta = \{3, 4\},$$

$$\text{Null}(A) = \text{span}\left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad \text{Null}(A_\eta) = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}, \ \ \text{Null}(A_\eta) \cap \left(\mathbb{R}_+^{|\eta|} \setminus \{0\}\right) = \emptyset.$$

**4.** Follows Theorem 6:

$$[A|b] = \left[\begin{array}{cccc|c} 2 & 0 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 & 4 \\ 0 & 2 & 2 & 0 & 8 \\ 2 & 0 & 0 & 1 & 2 \end{array}\right], \quad \mathcal{P} = \left\{ \begin{pmatrix} 1 \\ 4 - x_3 \\ x_3 \\ 0 \end{pmatrix} \right\}, \quad x^* = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 0 \end{pmatrix}, \quad N = \{4\}, \ \ \eta = \{4\},$$

$$\text{Null}(A) = \text{span}\left\{ \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad \text{Null}(A_\eta) = \{ \begin{pmatrix} 0 \end{pmatrix} \}, \ \ \text{Null}(A_\eta) \cap \left(\mathbb{R}_+^{|\eta|} \setminus \{0\}\right) = \emptyset.$$

**5.** Counterexample to the reverse of Theorem 6, and example showing that reducing $\|x\|_\infty$ does not force all other components to increase.

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 1 & 2 & 18 \\ 1 & 2 & 3 & 24 \end{array}\right], \quad \mathcal{P} = \left\{ \begin{pmatrix} 12 - x_3 \\ 6 - x_3 \\ x_3 \end{pmatrix} \right\}, \quad x^* = \begin{pmatrix} 6 \\ 0 \\ 6 \end{pmatrix}, \quad N = \emptyset, \ \ \eta = \{2\},$$

$$\text{Null}(A) = \text{span}\left\{\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}\right\}, \quad \text{Null}(A_\eta) = \{\begin{pmatrix} 0 \end{pmatrix}\}, \quad \text{Null}(A_\eta) \cap \left(\mathbb{R}_+^{|\eta|} \setminus \{0\}\right) = \emptyset.$$

Some example elements of $\mathcal{P}$ with decreasing $\|x\|_\infty$ follow. Notice that $x_2$ decreases instead of increasing.

$$\begin{pmatrix} 12 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 11 \\ 5 \\ 1 \end{pmatrix}, \begin{pmatrix} 10 \\ 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 9 \\ 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 8 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 7 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 6 \\ 0 \\ 6 \end{pmatrix}$$

**6.** There exists $y \in \text{Null}(A)$ with $y > 0$, but $(B|N) \neq (\beta|\eta)$ because there does not exist such $y$ at each step of the lexmin-sort algorithm, so there does not exist $y \in \text{Null}(A_\eta)$ with $y > 0$

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 0 & -1 & 12 \\ 0 & 1 & -1 & -6 \end{array}\right], \quad \mathcal{P} = \left\{\begin{pmatrix} 12 + x_3 \\ -6 + x_3 \\ x_3 \end{pmatrix}\right\}, \quad x^* = \begin{pmatrix} 18 \\ 0 \\ 6 \end{pmatrix}, \quad N = \emptyset, \quad \eta = \{2\},$$

$$\text{Null}(A) = \text{span}\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}, \quad \text{Null}(A_\eta) = \{\begin{pmatrix} 0 \end{pmatrix}\}, \quad \text{Null}(A_\eta) \cap \left(\mathbb{R}_+^{|\eta|} \setminus \{0\}\right) = \emptyset.$$

# References


[1] M. Ehrgott, A. Holder, and J. Reese., *Beam Selection in Radiotherapy Design.*, Trinity University, Mathematics Technical Report #95, 2005.