2013

Abstracted primal-dual affine programming

Tien Chih

ABSTRACTED PRIMAL-DUAL AFFINE PROGRAMMING

By

Tien Chih

B.A. University of Hawaii at Hilo, 2007
M.A. The University of Montana, USA, 2009

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in Mathematics

The University of Montana
Missoula, MT

December 2013

Approved by:

Sandy Ross, Associate Dean of the Graduate School
Graduate School

Dr. George McRae, Chair
Mathematical Sciences

Dr. Kelly McKinnie
Mathematical Sciences

Dr. Jennifer McNulty
Mathematical Sciences

Dr. Thomas Tonev
Mathematical Sciences

Dr. Ronald Premuroso
Accounting

Chih, Tien Ph.D., December 2013 Mathematics

Abstracted Primal-Dual Affine Programming

Committee Chair: George McRae, Ph.D.

The classical study of linear (affine) programs, pioneered by George Dantzig and Albert Tucker, studies both the theory, and methods of solutions for a linear (affine) primal-dual maximization-minimization program, which may be described as follows:

"Given $A \in \mathbb{R}^{m\times n}, \vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n, d \in \mathbb{R}$, find $\vec{x} \in \mathbb{R}^n$ such that $A\vec{x} \leq \vec{b}$, and $\vec{x} \geq 0$, that maximizes the affine functional $f(\vec{x}) := \vec{c} \cdot \vec{x} - d$; and find $\vec{y} \in \mathbb{R}^m$ such that $A^\top \vec{y} \geq \vec{c}$, and $\vec{y} \geq 0$, that minimizes the affine functional $g(\vec{y}) := \vec{y} \cdot \vec{b} - d$."

In this classical setting, there are several canonical results dealing with the primal-dual aspect of affine programming. These include: I: Tucker's Key Equation, II: Weak Duality Theorem, III: Convexity of Solutions, IV: Fundamental Theorem of Linear (Affine) Programming, V: Farkas' Lemma,VI: Complementary Slackness Theorem, VII: Strong Duality Theorem, VIII: Existence-Duality Theorem, IX: Simplex Algorithm.

We note that although the classical setting involves finite dimensional real vector spaces, moreover the classical viewpoint of these problems, the key results, and the solutions are extremely coordinate and basis dependent. However, these problems may be stated in much greater generality. We can define a function-theoretic, rather than coordinate-centric, view of these problem statements. Moreover, we may change the underlying ring, or abstract to a potentially infinite dimensional setting. Integer programming is a well known example of such a generalization. It is natural to ask then, which of the classical facts hold in a general setting, and under what hypothesis would they hold?

We describe the various ways that one may generalize the statement of an affine program. Beginning with the most general case, we prove these facts using as few hypotheses as possible. Given each additional hypothesis, we prove all facts that may be proved in this setting, and provide counterexamples to the remaining facts, until we have successfully established all of our classical results.

Acknowledgements

I would like to give special thanks for my advisor, Dr. George McRae. His perspective, knowledge and experience has forever changed the way I think of and do Mathematics. His patience, guidance and support, coupled with his insight and expertise, has made completing this project possible, although his influence throughout my PhD work extends far beyond that.

I would also like to thank my committee: Dr. Kelly McKinnie, Dr. Jenny McNulty, Dr. Thomas Tonev, and Dr. Ron Premuroso, for going above and beyond to continuously supporting me, providing me with help and advice, and for their making available their knowledge and expertise. Their continued support throughout this process has been invaluable.

I would also like to thank the Department of Mathematical Sciences at the University of Montana. In particular I would like to thank my masters advisor Dr. Nikolaus Vonesson for continuing to be someone I could run to for help, even after the conclusion of our work, and Dr. Eric Chesebro for his advice and help critiquing this document. I would also like to thank the Associate Chair of Graduate Studies, Dr. Emily Stone for her continued diligence and support during this entire process.

I would also like to thank all of the graduate students for sharing my passion of mathematics, and giving me their unique insights into a shared field we all love. I would like to give special thanks to all the graduate students who assisted in editing this document. Additionally, I would like to thank former graduate student Dr. Demitri Plessas has being especially willing to act as a sounding board during my entire time as a graduate student.

I would also like to give special thanks for my family for their personal support through these years and all years prior.

Finally, I would like to thank Tara Ashley Clayton for her ongoing and continuous support through the entire schooling process, even as she attends to her own schooling as well. Her ability to understand and cope with the odd and sometimes aggravating lifestyle of a PhD student is immeasurable, and her emotional support during the more trying times of this process is invaluable. To her I give my love and everlasting gratitude.

Notation and Conventions

R	a ring, typically a (ordered) (division ring)	57 (58) (29)
X	a left R module (vector space), space of solutions	57 (29)
Y	a left R module (vector space), space of constraints	57 (29)
$\mathrm{Hom}_R(X,Y)$	collection of left homomorphisms from X to Y	57
X^*, Y^*	$\mathrm{Hom}_R(X,R), \mathrm{Hom}_R(Y,R)$	57
A	an element of $\mathrm{Hom}_R(X,Y)$	57
$\vec{b}$	element of Y, upper bound of primal problem	57
$\mathbf{c}$	element of $\mathrm{Hom}_r(X,R)$, primal objective	57
$\square_{\oplus}$	the non negatives of $\square$	67
$\hat{\mathcal{X}}$	collection of spanning maps when X a vector space	84
$\hat{\mathcal{Y}}$	collection of spanning maps when Y a vector space	84
$\mathcal{X}$	the image of 1_R under the induced inclusions (basis of X)	84
$\mathcal{Y}$	the image of 1_R under the induced inclusions (basis of Y)	84
α_i	the row-like projection map $\hat{y}_i \circ A, \hat{y}_i \in \hat{\mathcal{Y}}$	85
$\mathcal{A}$	collection of all row-like projection maps $\{\alpha_i\}$	85
$\mathcal{M}$	an oriented matroid defined on a set E	131
$\hat{\alpha}_i$	the map defined on $X \oplus R$ to simulate $b_i - \alpha_i$ in OM program	155
X	the $\hat{\alpha}_i$ that forms a vector(circuit)	155
Y	the $\hat{\alpha}_j$ induced by$\mathcal{A}\backslash X$	155
B	$B : Y \to \{+,-,0\}$, records coefficient of f for $\hat{\alpha}_j$	155
C	$C : X \to \{+,-,0\}$, records coefficient of $\hat{\alpha}_i$ as summand of g	155
$\mathbf{A}$	$\mathbf{A} : X \times Y \to \{+,-,0\}$, records coefficient of $\hat{\alpha}_i$ as summand of $\hat{\alpha}_j$	155

Contents

List of Figures

List of Tables

Chapter 1

Introduction: Primal-Dual Affine Programs, Classical and Abstract

1.1 The Goal

The goal of this dissertation is to develop generalizations and abstractions of the mathematical features and structures of the systems of traditional affine primal-dual programming situations over finite dimensional real vector spaces. The traditional descriptions are very coordinate and basis intensive descriptions where the goal is to develop a coordinate-free and basis-free approach.

The classical study of the theory and applications of the canonical primal maximization and the dual minimization affine optimization problems have been well studied over the last hundred years. This traditional study has been very basis and coordinate dependent, as is natural for the applications to real world settings.

In this proposed research we want to abstract and generalize these situations in several possible directions. In particular, with a focus on a "basis-free" and "coordinate free" approach, similar to the coordinate-free approach to geometry done by John von Neumann [vN98]. By doing so, we hope to reveal the underlying mathematical features. Some of our questions include:

¿ How much of the classical theory generalizes to linear and affine (possibly continuous) transformations for arbitrary dimensions (e.g. Banach Spaces)?

¿ How much of the classical theory generalizes to (ordered) ground rings of scalars (e.g. Integers, Rationals, Hyperreals)?

¿ How much of the Rockafeller-Bland oriented matroid approach can be generalized to these settings?

In this chapter, all classical results and definitions of linear or affine programming can be found in [NT93], unless specifically stated. Similarly, standard results and definitions about algebra may be found in [DF04], analysis results and definitions may be found in [Roy10], and results and definitions regarding oriented matroids may be found in [BLVS+99].

1.2 Classical Affine Primal-Dual Programs

1.2.1 An Example of a Classical Affine Primal-Dual Program

Consider the following classical example two affine optimization problems. One problem is a maximization (or primal) problem and the other (the dual) is a minimization problem.

Problem 1.2.1 (The Lumbermill Production Problem)**.** Suppose a lumber mill produces three products: 2×4 small lumber, 2×10 large lumber, and particle board. This mill uses two inputs: large logs and small logs. The small lumber sells for \$3 per unit, large dimensional sells for \$2 per unit and the particle board for \$4 per unit. The small lumber requires 1 and 2 large and small logs to produce a unit. The large lumber requires 3 large logs and 1 small log to produce a unit, and the particle board requires 2 large logs and 1 of the large and small logs respectively. The total materials consist of 10 large logs and 8 small logs. There is also an initial setup cost of \$5. **How much of each type of lumber should be produced to maximize revenue?**

We place this information in the following data table:

Table 1.1: Lumbermill Resources

Resource	Small lumber	Large lumber	Particle board	**Total Resource**
Large Logs	1	3	2	10
Small Logs	2	1	1	8
Revenue:	\$3	\$2	\$4	-\$5

With this table, given some quantity of small lumber, large lumber and particle board, we may easily compute both how much resources are used, and how much revenue is earned. The constant \$5 is the "fixed cost" of operating the mill. By allowing x_1, x_2, x_3 to be the quantities of small lumber, large lumber, and particle board respectively, we may phrase the problem as the following row systems of inequalities:

$$\begin{aligned}
\textbf{Maximize: } f(x_1, x_2, x_3) &:= 3x_1 + 2x_2 + 4x_4 - 5, \\
\text{subject to: } x_1 + 3x_2 + 2x_3 &\leq 10 \\
2x_1 + x_2 + x_3 &\leq 8 \\
x_1, x_2, x_3 &\geq 0.
\end{aligned}$$

Which has a natural linear algebraic formulation:

Let $A := \begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 3}, \vec{b} := [10,8]^\top \in \mathbb{R}^2, \vec{c} := [3,2,4]^\top \in \mathbb{R}^3, d := 5 \in \mathbb{R}$. Find $\vec{x} \in \mathbb{R}^3$ such that:

$$\begin{aligned} \textbf{Maximize: } f(\vec{x}) &:= \vec{c}^\top \vec{x} - d, \\ \text{subject to: } A\vec{x} &\leq \vec{b} \\ \vec{x} &\geq 0. \end{aligned}$$

One may take the information in this original problem, and formulate an entirely new problem:

Problem 1.2.2 (The Same Lumbermill - The Insurance Problem)**.** Consider the information given in Problem 1.2.1. The Vice President in charge of finance wishes to insure the raw materials used in the production of lumber. In order to do this she needs to assign a valuation to each of the raw materials used. Since the cost of insurance is proportional to the value of the insured materials, she wishes to minimize this valuation, and yet be adequately compensated should these materials be destroyed in some accident. **How much should the large and small logs be valued, to minimize cost, while making sure that the value of any destroyed materials meets or exceeds the value of the products they would have produced?**

We again place this information in a table:

With this "transposed" table, given some evaluation of large and small logs, we may easily compute both the evaluation of each product compared to their actual value, and the total evaluation of the material. By allowing y_1, y_2 to be the values of the large and small logs respectively, we may phrase the problem as follows:

Table 1.2: Lumbermill Values

Product	Large logs	Small logs	**Total Value**
Small lumber	1	2	3
Large lumber	3	1	2
Particle board	2	1	4
Value:	\$3	\$2	-\$5

$$\begin{aligned} \textbf{Minimize: } g(y_1, y_2) &:= 10y_1 + 8y_2 - 5, \\ \text{subject to: } y_1 + y_2 &\geq 3 \\ 3y_1 + y_2 &\geq 2 \\ 2y_1 + y_2 &\geq 4 \\ y_1, y_2 &\geq 0. \end{aligned}$$

It is convenient and efficient to consider this minimization system as a column system of inequalities. This system also has a has a natural linear algebraic formulation:

Let $A := \begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 3}, \vec{b} := [10, 8]^\top \in \mathbb{R}^2, \vec{c} := [3, 2, 4]^\top \in \mathbb{R}^3, d := 10 \in \mathbb{R}$. Find $\vec{y} \in \mathbb{R}^2$ such that:

$$\begin{aligned} \textbf{Minimize: } g(\vec{y}) &:= \vec{y}^\top \vec{b} - d, \\ \text{subject to: } \vec{y}^\top A &\geq \vec{c}^\top \\ \vec{y} &\geq 0. \end{aligned}$$

Thus given the information about total resource ($\vec{b}$) revenue per product ($\mathbf{c}$) a linear relation between goods and products (A) and an initial cost, we were able to formulate both a maximization problem and a minimization problem. The Lumbermill Production problem (the row system) is the *primal maximization problem* and the Insurance problem is the *dual minimization problem* (the column system).

1.2.2 Introduction to Classical Affine Programming

Here, we reiterate the basic definitions and goals of an primal-dual affine programming problem. The classical affine primal-dual optimization problem can be described as follows [NT93] :

Given the data $A, \vec{b}, \vec{c}, d$, where A is a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^m$, in other words, $A \in \mathbb{R}^{m\times n}$, together with its transpose (or adjoint), $A^\top : \mathbb{R}^m \to \mathbb{R}^n$, $\vec{b}, \vec{c}$ are two constant vectors, $\vec{b} \in \mathbb{R}^m$ and $\vec{c} \in \mathbb{R}^n$, and $d \in \mathbb{R}$ is a fixed scalar. We note that all algebra done here is with respect to standard matrix multiplication, and all vectors are assumed to be column vectors. With this data, define an affine function $f : \mathbb{R}^n \to \mathbb{R}$ (called the *primal objective function*) by the scalar equation:

(F) $\quad f(\vec{x}) = \vec{c}^\top \vec{x} - d$ for $\vec{x} \in \mathbb{R}^n$.

Then define a vector variable $\vec{t} \in \mathbb{R}^m$ (called the *primal slack variable*) by the vector equation (a "row" system):

(P) $\quad \vec{t} = \vec{b} - A\vec{x}$ for the vector variable $\vec{x} \in \mathbb{R}^n$.

Simultaneously, define the *dual objective function* as the affine functional $g : \mathbb{R}^m \to \mathbb{R}$ by the vector equation:

(G) $\quad g(\vec{y}) = \vec{y}^\top \vec{b} - d$ for $\vec{y} \in \mathbb{R}^m$.

Then define a vector variable $\vec{s} \in \mathbb{R}^n$ (called the *dual slack* or *surplus variable*)by the vector equation (a "column" system):

(D) $\quad \vec{s}^\top = \vec{y}^\top A - \vec{c}^\top$ for a vector variable $\vec{y} \in \mathbb{R}^m$.

Here, we might notice (as Albert Tucker pointed out in the 1960's) that the information in these four equations from the row system and the column system can be captured in what is called a *Tucker tableau* (but what Tucker called a "condensed" tableau):

Figure 1.1: A Tucker Tableau

$$\begin{array}{c|c|c|l}
 & \vec{x}^\top & \text{-1} & \\
\hline
\vec{y} & A & \vec{b} & = -\vec{t} \\
\hline
-1 & \vec{c}^\top & d & = f \\
\hline
 & \| & \| & \\
 & \vec{s}^\top & g &
\end{array}$$

Example 1.2.3. Recall the Lumbermill Problem 1.2.1. The primal maximization program can be entered into the following tableau:

$$
\begin{array}{c|ccc|c|l}
 & x_1 & x_2 & x_3 & \text{-1} & \\
\hline
y_1 & 1 & 3 & 2 & 10 & = -t_1 \\
y_2 & 2 & 1 & 1 & 8 & = -t_2 \\
\hline
-1 & 3 & 2 & 4 & 5 & = f \\
\hline
 & \parallel & \parallel & \parallel & \parallel & \\
 & s_1 & s_2 & s_3 & g &
\end{array}
$$

We can easily recover the original primal maximization program by looking at each row of this tableau, with the first 2 rows recording inequalities ($\leq$) and the last row recording the objective function. Thus the primal problem is sometimes referred to as the *row system*. Similarly, if we look at the first 3 columns as inequalities ($\geq$), and the last column as an objective function, we recover the dual minimization program. Thus, the dual program is sometimes referred to as the *column system*.

Also here, we might immediately observe that

(Tucker Duality Equation) $\quad g - f = \vec{s}^{\top}\vec{x} + \vec{y}^{\top}\vec{t}.$

In terms of the Tucker tableau, Tucker thought of this equation as: the inner product of the top and bottom marginal labels equals the inner product of the left and right marginal labels. He called this the "key" equation.

To get the classical optimization problems, we need to require non-negativity constraints. The classical primal maximization affine problem (or "program") is:

"FIND $\vec{x} \in \mathbb{R}^n$ that maximizes the primal objective affine functional $f = \vec{c}^{\top}\vec{x} - d$ subject to the constraints:

(P) $\qquad A\vec{x} - \vec{b} = -\vec{t}.$

and

(NN-P) $\quad \vec{x} \geq 0, \vec{t} \geq 0$ (i.e. $\vec{x}, \vec{t}$ are in the respective *non-negative cones*)."

The slack variable $\vec{t}$ measures the "difference" between the bound $\vec{b}$ and $A\vec{x}$, and so we see that $\vec{t} \geq 0$ implies each entry of $\vec{b} - A\vec{x}$ is non-negative. Thus, the above condition is often stated as: "FIND $\vec{x} \in \mathbb{R}^n$ that maximizes f subject to $A\vec{x} \leq \vec{b}$ and $\vec{x} \geq 0$.

Similarly, the classical dual minimization affine program is:

"FIND $\vec{y} \in \mathbb{R}^m$ that minimizes the dual objective affine functional $g = \vec{y}^{\top}\vec{b} - d$ subject to the constraints:

(D) $\qquad \vec{y}A - \vec{c}^{\top} = \vec{s}^{\top}.$

and

(NN-D) $\quad \vec{y} \geq 0, \vec{s} \geq 0$ (i. e. $\vec{y}, \vec{s}$ are in the respective *non-negative cones*)."

As in the primal case, we may suppress mention of the slack variable $\vec{s}$ and state this condition as: "FIND $\vec{y} \in \mathbb{R}^m$ that minimizes g subject to $\vec{y}^{\top}A \geq \vec{c}$ and $\vec{y} \geq 0$.

The two conditions (P) & (NN-P) say that the affine transformation in (P) takes the non-negative orthant of the domain to the non-negative orthant of the co-domain. Similarly, (D) & (NN-D) show that the affine transformation in (D) takes the non-negative cone of its domain to the non-negative cone of its co-domain.

1.2.3 Basic facts

Using only the definition of the objective functions and slack variables, as affine functionals or variables, we are able to prove two basic facts.

Proposition 1.2.4 (Classical Fact I: Tucker's Key Equation). *Given $A \in \mathbb{R}^{m\times n}, \vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n, d \in R$, and $f, g, \vec{s}, \vec{t}$ described above by only (F), (P), (G), (D) (and not necessarily (NN-P) or (NN-D)), we have that:*

$$\vec{x}^\top \vec{s} + (-1)g(\vec{y}) = \vec{y}^\top(-\vec{t}) + (-1)f(\vec{x}).$$

Proof. We note that:

$$\begin{aligned}
\vec{x}^\top \vec{s} + (-1)g(\vec{y}) &= \vec{x}^\top(A\vec{y} - \vec{c}) - \vec{b}^\top \vec{y} + d \\
&= \vec{x}^\top A\vec{y} - \vec{x}^\top \vec{c} - \vec{b}^\top \vec{y} + d. \\
&= \vec{y}^\top(A\vec{x} - \vec{b}) - \vec{c}^\top \vec{x} + d \\
&= \vec{y}^\top(-\vec{t}) + (-1)f(\vec{x})
\end{aligned}$$

□

Albert Tucker described what he called the "Key Equation" as: " The sum of the dot products of the upper and lower marginal labels, is equal to the sum of the dot products of the left and right marginal labels." [JGKR63]

As a corollary, we have:

Corollary 1.2.5 (Tucker's Duality Equation). *Given $A \in \mathbb{R}^{m\times n}, \vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n, d \in R$, and $f, g, \vec{s}, \vec{t}$ described above (using only (F), (P), (G), (D)), we have that:*

$$g(\vec{y}) - f(\vec{x}) = \vec{s}^\top \vec{x} + \vec{y}^\top \vec{t}.$$

Proof. By Theorem 1.2.4, we have:

$$\begin{aligned}\vec{x}^\top \vec{s} + (-1)g(\vec{y}) &= \vec{y}^\top(-\vec{t}) + (-1)f(\vec{x}) \\ g(\vec{y}) - f(\vec{x}) &= \vec{s}^\top \vec{x} + \vec{y}^\top \vec{t}\end{aligned}$$

□

1.2.4 Results regarding Order

Using the ordering of the real numbers, we can prove some additional results about affine primal dual programming.

Theorem 1.2.6 (Classical Fact II: Weak Duality)**.** *Given $A \in \mathbb{R}^{m\times n}, \vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n, d \in R$, and f, g described above (using (F), (P), (G), (D) and (NN-P), (NN-D), and that $\mathbb{R}$ is an ordered ring of scalars), we have that:*

$$g(\vec{y}) \geq f(\vec{x})$$

for feasible $\vec{x}, \vec{y}$.

Proof. Using that $\mathbb{R}$ is an ordered ring of scalars; and since $\vec{x}, \vec{y}$ are feasible, each entry $\vec{x}_i, \vec{y}_j \geq 0$ for $1 \leq i \leq n, 1 \leq j \leq m$. Moreover, the same is true for $\vec{s}_i, \vec{t}_j$. Thus

$$\vec{s}^\top \vec{x} = \sum_{i=1}^{n} \vec{s}_i \vec{x}_i \geq 0$$

and

$$\vec{y}^\top \vec{t} = \sum_{j=1}^{m} \vec{y}_j \vec{t}_j \geq 0.$$

Thus:

$$\begin{aligned} g(\vec{y}) - f(\vec{x}) &= \vec{s}^{\top}\vec{x} + \vec{y}^{\top}\vec{t} \geq 0 \\ g(\vec{y}) &\geq f(\vec{x}). \end{aligned}$$

□

Corollary 1.2.7. *Given the hypothesis above (Proposition 1.2.6), if there is a feasible pair of solutions $\vec{x} \in \mathbb{R}^n, \vec{y} \in \mathbb{R}^m$ such that $f(\vec{x}) = g(\vec{y})$, then $\vec{x}, \vec{y}$ are both optimal solutions.*

Proof. Suppose that $\vec{x}$ is not an optimal solution. It follows that there is a feasible $\vec{x}' \in \mathbb{R}^n$ such that $f(\vec{x}') > f(\vec{x})$. But then $f(\vec{x}') > g(\vec{y})$ which contradicts Corollary 1.2.5. Thus $\vec{x}$ is optimal. Similarly, $\vec{y}$ is optimal as well. □

This gives us a certificate of optimality.

Another notion that may be introduced along with the ordering of the real numbers is the notion of line segments and convexity in the vector spaces $\mathbb{R}^n, \mathbb{R}^m$ or $\mathbb{Q}^n, \mathbb{Q}^m$.

Definition 1.2.8. Let $R \in \{\mathbb{R}, \mathbb{Q}\}$. Given $\vec{v}, \vec{w} \in R^n$, the *line segment*, denoted $[\vec{v}, \vec{w}]$, with endpoints $\vec{v}, \vec{w}$ is the set:

$$[\vec{v}, \vec{w}] := \{\lambda \cdot \vec{v} + (1 - \lambda) \cdot \vec{w} : \lambda \in R, 0 \leq \lambda \leq 1\}.$$

The *open line segment*, with end points $\vec{w}, \vec{v}$ is the set:

$$]\vec{v}, \vec{w}[:= \{\lambda \cdot \vec{v} + (1 - \lambda) \cdot \vec{w} : \lambda \in R, 0 < \lambda < 1, \lambda \in R\}.$$

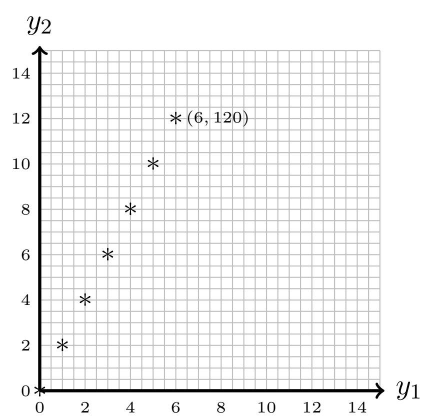
Here, we also describe a generalization of this concept to finite dimensional free modules over $\mathbb{Z}, \mathbb{Q}$ or $\mathbb{R}$. This generalization is due to the author. Note that we will later generalize this definition even further to all ordered rings.

Definition 1.2.9. Let $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$. Let $\vec{v}, \vec{w} \in R^n$. Let S be the collection of vectors $\vec{z}$ such that there is a $r_{\vec{z}} \in R, r_{\vec{z}} > 0$ where $\vec{w} = r_{\vec{z}} \cdot \vec{z} + \vec{v}$. Then the *generalized line segment* with endpoints $\vec{v}, \vec{w}$ is:

$$[[\vec{v}, \vec{w}]] := \bigcup_{\vec{z} \in S} \bigcup_{a \in R, 0 \leq a \leq r_{\vec{z}}} \vec{v} + a \cdot \vec{z}.$$

Example 1.2.10. Consider $(0,0), (6,12)$ in $\mathbb{Z}^2$. We note that $S = \{(6,12), (3,6), (2,4), (1,2)\}$, with values for $r_{\vec{z}} : 1, 2, 3, 6$. So the generalized line segment with endpoints $(0,0)$ to $(6,12)$ is $[[(0,0), (6,12)]] := \{(0,0), (1,2), (2,4), (3,6), (4,8), (5,10), (6,12)\}$.

Figure 1.2: Line segment in $\mathbb{Z}^2$



Proposition 1.2.11. *For $\mathbb{R}^n$, or $\mathbb{Q}^n$, and scalars from $\mathbb{R}$ or $\mathbb{Q}$ respectively, a line segment with endpoints $\vec{v}, \vec{w}$, and a generalized line segment with the same endpoints are the same.*

Proof. If $\vec{v} = \vec{w}$, then we are done. otherwise, since $\vec{w} - \vec{v}$ and 1 is a valid choice for $\vec{z}$ and $r_{\vec{z}}$, we have that $[\vec{v}, \vec{w}] \subseteq [[\vec{v}, \vec{w}]]$. Conversely. given any choice $\vec{z}, r_{\vec{z}}, a \in [0, r]$ the point $\vec{v} + a \cdot \vec{z} = \vec{v} + (\frac{a}{r_{\vec{z}}}) \cdot (r_{\vec{z}} \cdot \vec{z})$. Since $0 \leq \frac{a}{r_{\vec{z}}} \leq 1$, $\vec{v} + a \cdot \vec{z}$ is a point in the standard line segment. Thus $[\vec{v}, \vec{w}] = [[\vec{v}, \vec{w}]]$.

$\square$

Remark 1.2.12. We note that for these examples in $R \in \{\mathbb{Z}, \mathbb{R}, \mathbb{Q}\}$, it turns out that the line segment $[[\vec{v}, \vec{w}]]$ is the standard line segment $[\vec{v}, \vec{w}] \subseteq \mathbb{R}^n$, intersected with R^n. Thus it may seem unnecessary to define such a complex definition. However, we show in Example 2.4.16 that this is not always the case, thus requiring such a definition.

Definition 1.2.13. Given a set $C \subseteq \mathbb{R}^n$, we say that C is a *convex set* if given $\vec{w}, \vec{v} \in C$, the line segment with endpoints $\vec{v}, \vec{w}$ is contained in C.

With these notions, we may make some statements about the solutions of affine programs.

Lemma 1.2.14. *Let $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ and $A : R^n \to R^m$, be a linear transformation, $\vec{b} \in R^m, \vec{c} \in R^n$. Let $C_n \subseteq R^n$, $C_m \subseteq R^m$ be convex sets. Then the following hold:*

1. *$A^{-1}(C_m)$ is a convex set.*
2. *$C_n + \vec{c}, C_m + \vec{b}$ are convex sets.*
3. *The intersection of convex sets is convex.*

Proof. Let $\lambda \in [0, 1]$.

1. Let $\vec{w}, \vec{v} \in A^{-1}(C_m)$, and consider $\vec{x} := A(\vec{v})$ and $\vec{y} := A(\vec{w})$. Then $[[\vec{x}, \vec{y}]] \subseteq C_m$. So consider any $\vec{z}$ such that there is an $r \in R, r > 0$ where$\vec{w} = v + r \cdot \vec{z}$, and let $0 \leq a \leq r$. Then consider $\vec{v} + a \cdot \vec{z} \in [[\vec{v}, \vec{w}]]$, it follows that

$$A(\vec{v} + a \cdot \vec{z}) = \vec{x} + a \cdot A(\vec{z}).$$

   Since $\vec{x} + r \cdot A(\vec{z}) = A(\vec{v}) + A(r \cdot \vec{z}) = A(\vec{w}) = \vec{y}$, it follows that $A(\vec{v} + a \cdot \vec{z}) \in [\vec{x}, \vec{y}] \subseteq C_m$. Thus $\vec{v} + a \cdot \vec{z} \in A^{-1}(C_m)$.

2. Let $\vec{v}, \vec{w} \in C_m$, and consider $[[\vec{v} + \vec{c}, \vec{w} + \vec{c}]] \subseteq C_m$. Then consider $\vec{v} + \vec{c} + a \cdot \vec{z} \in [[\vec{v} + \vec{c}, \vec{w} + \vec{c}]]$,

where $\vec{z}$ is a vector such that there is a $r \in R, r > 0$ such that $r \cdot \vec{z} + \vec{v} + \vec{c} = \vec{w} + \vec{c}$, and $0 \leq a \leq r$. Then it is clear that $\vec{v} + a \cdot \vec{z} \in [[\vec{v}, \vec{w}]] \subseteq C_m$ and so $\vec{v} + a \cdot \vec{z} + \vec{c} \in C_m + \vec{c}$.

3. Let $D \subseteq \mathbb{R}^n$ be a convex set, then consider $D \cap C_n$. It follows that $[[\vec{v}, \vec{w}]] \subseteq C_n, D$ and so $[[\vec{v}, \vec{w}]] \subseteq C_n \cap D$. Thus $C_n \cap D$ is a convex set.

$\square$

Definition 1.2.15. Given $A \in \mathbb{R}^{m \times n}, \vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n, d \in \mathbb{R}$, $\vec{x} \in \mathbb{R}^n$ is *primal feasible* if $\vec{x} \geq_{\mathbb{R}^n} 0$ and $A\vec{x} \leq_{\mathbb{R}^m} \vec{b}$. Similarly a $\vec{y} \in \mathbb{R}^m$ is *dual feasible* if $\vec{y} \geq_{\mathbb{R}^m} 0$ and $\vec{y}^\top A \geq_{\mathbb{R}^n} \vec{c}^\top$.

Then, the *primal feasible region* is the collection of all primal feasible $\vec{x} \in \mathbb{R}^n$ and the *dual feasible region* is the collection of all dual feasible $\vec{y} \in \mathbb{R}^n$.

Corollary 1.2.16 (Classical Fact III Part 1: Convexity of feasible region)**.** *The feasible regions of a the primal and dual affine programming problems are convex sets.*

Proof. The feasible region of the primal problem is the set $\mathbb{R}^n_\oplus \cap A^{-1}(\vec{b} - \mathbb{R}^m_\oplus)$, which is the intersection of convex sets, and so is convex. (We use $\mathbb{R}^n_\oplus$ to denote the vectors with non-negative entries in $\mathbb{R}^n$).

The feasible region of the dual problem is the set $\mathbb{R}^m_\oplus \cap (A^\top)^{-1}(\vec{c} + \mathbb{R}^n_\oplus)$, which is also convex.

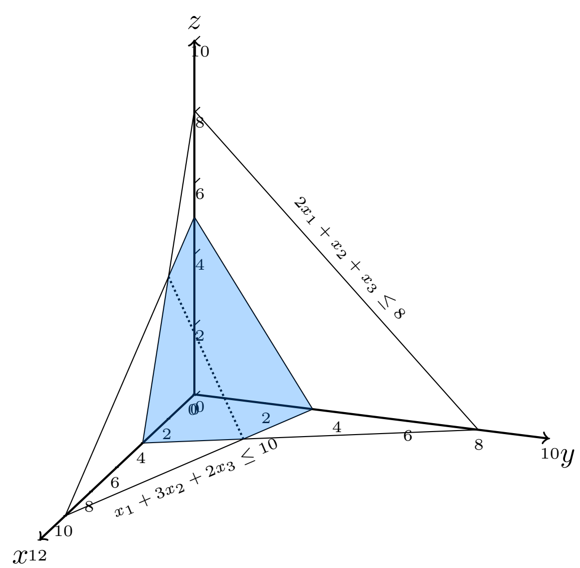
$\square$

Example 1.2.17. Recall the Lumbermill Problem 1.2.1. The feasible region of the primal maximization program is the collection of vectors that satisfy:

$$
\begin{aligned}
x_1 + 3x_2 + 2x_3 &\leq 10 \\
2x_1 + x_2 + x_3 &\leq 8 \\
x_1, x_2, x_3 &\geq 0
\end{aligned}
$$

which may be illustrated by Figure 1.3.

Figure 1.3: Lumbermill Primal-Feasible Region

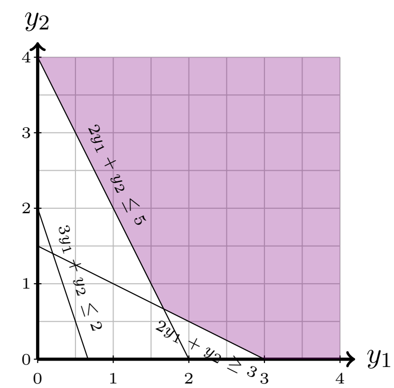


Similarly, the dual feasible region, bounded by the following inequalities:

$$
\begin{aligned}
2y_1 + y_2 &\geq 3 \\
3y_1 + y_2 &\geq 2 \\
2y_1 + y_2 &\geq 4 \\
y_1, y_2 &\geq 0
\end{aligned}
$$

can be show in Figure 1.4.

Figure 1.4: Lumbermill Dual-Feasible Region



Corollary 1.2.18 (Classical Fact III Part 2: Convexity of optimal solutions)**.** *The set of optimizers for an affine programming problem is convex.*

Proof. If the set of optimizers is empty, then it is vacuously convex. Otherwise, let $\vec{x} \in \mathbb{R}^n$ be an optimizer for the primal programming problem. Consider that $f(\vec{x})$ is a singleton and thus is convex. Then $f^{-1}(f(\vec{x}))$ is convex, and the intersection of this set with the feasible region is also convex.

Similarly the optimizers of the dual problem are convex. □

Here we introduce a new concept.

Definition 1.2.19. Given a convex set $C \subseteq \mathbb{R}^n$, an *extreme point* of C is a point that is not contained in any open line segment contained in C.

Definition 1.2.20. Given $\vec{x}, \vec{y} \in \mathbb{R}^n$, a point $\vec{z}$ is a *convex combination* of $\vec{x}, \vec{y}$ if it lies on a line segment with end points $\vec{x}, \vec{y}$. If it lies on an open line segment with end points $\vec{x}, \vec{y}$, it is called a *proper convex combination.*

Theorem 1.2.21 (Classical Fact IV: Fundamental Theorem of Affine Programming)**.** *Let P be the feasible region of a primal (or dual) affine program, and let $\ell : \mathbb{R}^n \to \mathbb{R}$ be a non-constant linear (or affine) functional. Then if there is a point $\vec{x} \in P$ that is a maximizer (or minimizer), then there is a point $\vec{x}'$ that is a maximizer (or minimizer) that is an extreme point of P*

Proof. It suffices to show that no interior point of P is an optimizer of ℓ. Let $\vec{x}$ be in the interior of P. Then there is an $\epsilon > 0$ such that $B(\vec{x}, \epsilon) \subset P$. Let e_i be a basis vector such that $\ell(e_i) \neq 0$, and without loss of generality, suppose $\ell(e_i) > 0$. Then $\ell(\vec{x} - \frac{\epsilon}{2}e_i)\vec{x}) < \ell(\vec{x}) < \ell(\vec{x} + \frac{\epsilon}{2}e_i)$. Thus $\vec{x}$ cannot be an optimizer of ℓ. □

Definition 1.2.22. Let $A : \mathbb{R}^n \to \mathbb{R}^m$, be a linear transformation, $\vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n$, $d \in \mathbb{R}$ define a primal-dual affine program. If for each $s \in \mathbb{R}$ there is a primal feasible $\vec{x} \in \mathbb{R}^n$ such that $f(\vec{x}) \geq s$, then the primal program said to be *unbounded*. Similarly if there is a dual feasible $\vec{y}$ such that $g(\vec{y}) \leq s$ then the dual program are said to be *unbounded*.

Theorem 1.2.23 (Small Existence-Duality Theorem)**.** *Let $A : \mathbb{R}^n \to \mathbb{R}^m$, be a linear transformation, $\vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n$, $d \in \mathbb{R}$ define a primal-dual affine program. Then the following hold:*

1. *If the solution values to the primal program are unbounded, then the dual program is infeasible.*
2. *If the solution values to the dual program are unbounded, then the primal program is infeasible.*

Proof. I the dual program is feasible, there is a feasible $\vec{y} \in \mathbb{R}^m$. Thus for any feasible $\vec{x} \in \mathbb{R}^n$, by Theorem 1.2.6, $f(\vec{x}) \leq g(\vec{y})$ and the solution values for the primal program are bounded. Similarly, if the primal program is feasible, the solution values to the dual program are bounded. □

1.2.5 Facts using $\mathbb{R}$

Using the properties of the real numbers, that it is an Archimedean, least upper bound closed field, we may prove the following classic result, the Farkas' Lemma. This fact is used to prove several powerful primal-dual programming results.

Theorem 1.2.24 (Classical Fact V: The Farkas' Lemma)**.** *Let A be a $m \times n$ matrix, and $\vec{c} \in \mathbb{R}^n$, then exactly one of the following is true:*

1. *There is a $\vec{x} \in \mathbb{R}^n$ such that $A\vec{x} \leq 0$ and $\vec{c}^\top \vec{x} > 0$.*
2. *There is a vector $\vec{y} \in \mathbb{R}^m, \geq 0$, that is, each entry non-negative, such that $A^\top \vec{y} = \vec{c}$.*

Proof. Suppose that (2) does not hold. Let S be the set of non-negative linear combinations of the rows of $A, a_1, a_2, \ldots a_m, S := \{A^\top \vec{y} : \vec{y} \in \mathbb{R}^m, \vec{y} \geq_{\mathbb{R}^m} 0\} \subseteq \mathbb{R}^n$. Since (2) does not hold, $\vec{c} \notin S$. Since S is the non-negative combination of a finite collection of vectors in $\mathbb{R}^n$, it may be expressed as the finite intersection of half-spaces in $\mathbb{R}^n$, and as such it is closed and convex. Thus, we may find a separating hyperplane H, defined by $a \in \mathbb{R}, \vec{v} \in \mathbb{R}^n$, where $H = \{\vec{w} \in \mathbb{R}^n : \vec{x}^\top \vec{w} = a\}$, and

$$\vec{v}^\top \vec{c} > a > \vec{v}^\top \vec{s}, \text{ for each } \vec{s} \in S.$$

Since $\vec{0} \in S$, we may assume $a > 0$. Thus for any $\vec{y} \in \mathbb{R}^m, \vec{y} \geq 0$, we have $(\vec{y}A)^\top \in S$, by definition of S. Thus for

$$a > (\vec{x}^\top (\vec{y}A)^\top)^\top = \vec{y}A\vec{x} = \sum_{j=1}^{m} \vec{x}_j (A\vec{v})_j.$$

It must be the case that each $(A\vec{x})_j \leq 0$, else we may select $\vec{w}$ with w_j sufficiently large, so that $\vec{y}_j(A\vec{x})_j > a$, and $w_i = 0$ when $i \neq j$. Thus for this $\vec{x}$, $A\vec{v} \leq 0$. Moreover $\vec{x}^\top \vec{c} > a > 0$.

$\square$

Example 1.2.25. Suppose $A := \begin{pmatrix} 3 \\ 5 \end{pmatrix}$, $\vec{c} = \begin{pmatrix} 10 \end{pmatrix}$. Then, by the Farkas' Lemma, either (a) there is a $x \in \mathbb{R}^1$ such that $\begin{pmatrix} 3 \\ 5 \end{pmatrix} x \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $10x > 0$, or (b) there is a $\vec{y} \in \mathbb{R}^2$ such that $\begin{pmatrix} 3 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 3y_1 + 5y_2 = 10$, $y_i \geq 0$.

Clearly $3x, 5x \leq 0$ if and only if $x \leq 0$. But for such x, $10x \leq 0$. Thus (a) does not hold. By the Farkas' Lemma, (b) does hold, and we see $\vec{y} = \begin{pmatrix} \frac{10}{3} \\ 0 \end{pmatrix}$ is a non-negative vector, such that $\begin{pmatrix} \frac{10}{3} \\ 0 \end{pmatrix} \begin{pmatrix} 3 & 5 \end{pmatrix} = 10$.

However, suppose we define $A' := \begin{pmatrix} -1 \\ -1 \end{pmatrix}$, $\vec{c} = \begin{pmatrix} 10 \end{pmatrix}$. Given any non-negative y_1, y_2, $\begin{pmatrix} -1 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = -y_1 - y_2 \neq 10$. Thus, (b) does not hold, and by the Farkas' Lemma, (a) holds. We note that if we let $x = 1$, $\begin{pmatrix} -1 \\ -1 \end{pmatrix} (1) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, but $10(1) > 0$.

We note that the proof of this result is highly dependent on the existence of a separating hyperplane, which is a property of the real numbers. If we wish to extend the Farkas' Lemma to a more general case, then one may need to provide an alternative proof to this result.

Corollary 1.2.26. *Let A be a $m \times n$ matrix, and $\vec{b} \in \mathbb{R}^m$, then the following are equivalent:*

1. *$A\vec{x} \leq_{\mathbb{R}^m} \vec{b}$ has no solution.*
2. *There exists a $\vec{y} \in \mathbb{R}^m$ such that $\vec{y} \geq_{\mathbb{R}^m} \vec{0}, \vec{y}^\top \vec{b} < 0$, and $A^\top \vec{y} \geq_{\mathbb{R}^n} \vec{0}$.*

Proof. Define $A' \in \mathbb{R}^{m \times n+1}$, where $A' = [A| - \vec{b}]$. Then $A\vec{x} \leq \vec{b}$ has a solution if and only if $A'\vec{x'} \leq 0$ has a solution where $\vec{x'}_{n+1} > 0$. Define $\vec{c'} \in \mathbb{R}^{n+1}$ where $\vec{c'}_i = \delta_{i,n+1}$. If $A\vec{x} \leq \vec{b}$ has no solution, then it must be the case that when $A'\vec{x}' \leq 0$, then $\vec{c'}^\top \vec{x'} \leq 0$. Thus condition (1) of the Farkas Lemma does not hold. It follows that condition 2 holds, and there is a $\vec{y} \in \mathbb{R}^m, \vec{y} \geq 0$ such that $\vec{y}^\top A' = \vec{c'}^\top$, and thus $\vec{y}^\top A = \vec{0}$, $\vec{y}^\top(-\vec{b}) = 1$ and $\vec{y}^\top \vec{b} < 0$.

□

Definition 1.2.27. Let $A : \mathbb{R}^n \to \mathbb{R}^m$, be a linear transformation, $\vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n$, $d \in \mathbb{R}$ define a primal-dual affine program. Then the feasible decision variables $\vec{x}, \vec{y}$ and induced slack variables $\vec{s}, \vec{t}$ are *complementary slack* if for each $i, 1 \leq i \leq n$, and each $j, 1 \leq j \leq m$:

1. $\vec{x}_i \neq 0 \implies \vec{s}_i = 0$.
2. $\vec{s}_i \neq 0 \implies \vec{x}_i = 0$.
3. $\vec{y}_j \neq 0 \implies \vec{t}_j = 0$.
4. $\vec{t}_j \neq 0 \implies \vec{y}_j = 0$.

Here, we state an another definition

Definition 1.2.28. Let $A : \mathbb{R}^n \to \mathbb{R}^m$, be a linear transformation, $\vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n$, $d \in \mathbb{R}$ define a primal-dual affine program. Then the feasible decision variables $\vec{x}, \vec{y}$ and induced slack variables $\vec{s}, \vec{t}$ are *complementary slack* if $\vec{s}^\top \vec{x} = \vec{y}^\top \vec{t} = 0$.

Proposition 1.2.29. *Let $A : \mathbb{R}^n \to \mathbb{R}^m$, be a linear transformation, $\vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n$, $d \in \mathbb{R}$ define a primal-dual affine program. Then the definitions given in Definition 1.2.27 and Definition 1.2.28 about decision variables $\vec{x}, \vec{y}$ and induced slack variables $\vec{s}, \vec{t}$ are complementary slack, are equivalent.*

Proof. We note that $\vec{s}^\top \vec{x} = \sum_{i=1}^{n} \vec{s}_i \vec{x}_i$. Since $\vec{x}, \vec{y}$ are feasible, each $\vec{x}_i, \vec{y}_j, \vec{s}_i, \vec{t}_j \geq 0$. Thus $\sum_{i=1}^{n} \vec{s}_i \vec{x}_i = 0$ if and only if each $\vec{s}_i \vec{x}_i = 0$, which is true if and only if:

- $\vec{x}_i \neq 0 \implies \vec{s}_i = 0$
- $\vec{s}_i \neq 0 \implies \vec{x}_i = 0$.

Similarly, $\vec{y}^\top \vec{t} = \sum_{j=1}^{m} \vec{y}_j \vec{t}_j$, which is 0 if and only if:

- $\vec{y}_j \neq 0 \implies \vec{t}_j = 0$
- $\vec{t}_j \neq 0 \implies \vec{y}_j = 0$.

□

We may then state the following characterization of optimal solutions:

Theorem 1.2.30 (Classical Fact VI: Complementary Slackness Theorem)**.** *Let $A : \mathbb{R}^n \to \mathbb{R}^m$, be a linear transformation, $\vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n$, $d \in \mathbb{R}$ define a primal-dual affine program. Then the feasible decision variables $\vec{x}, \vec{y}$ are optimal solutions if and only if $\vec{x}, \vec{y}$ and induced slack variables $\vec{s}, \vec{t}$ are complementary slack.*

Proof. Suppose $\vec{x}, \vec{y}, \vec{s}, \vec{t}$ are complementary slack. Recall that $g(\vec{y}) - f(\vec{x}) = \vec{s}^\top \vec{x} + \vec{y}^\top \vec{t}$ by Theorem 1.2.5. Then since these variables are complementary slack, $g(\vec{y}) - f(\vec{x}) = 0$. Thus $g(\vec{y}) = f(\vec{x})$ and by Theorem 1.2.6, both solutions are optimal.

Conversely, suppose that $\vec{x}^*, \vec{y}^*$, are optimal with slack variables $\vec{s}^*, \vec{t}^*$. By the Farkas' Lemma 1.5.2, either (2) we may find a $\vec{y}' \in \mathbb{R}^m$ such that $A^\top \vec{y}' = \vec{c}, \vec{y}' \geq 0$ or (1) there is a $\vec{x}' \in \mathbb{R}^n$ such that $\vec{x'}^\top A \leq 0$, $\vec{c}^\top \vec{x}' > 0$.

We notice here that $A\vec{x} \le \vec{b}, \vec{x} \ge 0$ if and only if $(A, -I_n)\vec{x} \le (\vec{b}, 0)$. Similarly $A^\top \vec{y}' \ge \vec{c}$ if and only if there is a $\vec{w} \in \mathbb{R}^n$ such that $A^\top \vec{y}' - I_n \vec{w} = \mathbf{c}$. Thus we may let $A' := [A| - I_n]$, and $\vec{b}' := [\vec{b}|0]^\top$.

If (2) holds, then given any feasible $\vec{y}$, and non-negative real number $r \in \mathbb{R}_\oplus$, $A(\vec{x} + r\vec{x}')^\top \le A\vec{x} \le \vec{b}$, since $A(r\vec{x}') \le 0$. Thus $\vec{y} + r\vec{y}'$ is feasible. But $f(\vec{x} + r\vec{x}') = \vec{c}^\top \vec{x} + r\vec{c}^\top \vec{x'} - d$. Since $\vec{c}^\top \vec{x'} > 0$, this can be made arbitrarily large by choice of r, and so by Theorem 1.2.23, the dual program is infeasible, a contradiction.

Thus, it follows that we may find a $\vec{y}' \in \mathbb{R}^{m+n}$ such that $A^\top \vec{y}' = \vec{c}, \vec{y'} \ge 0$. We claim that there is such a $\vec{y}'$, such that $\vec{y}'_j = 0$ whenever $\vec{t}^*_j \ne 0$.

Let $J := \{j : 1 \le j \le m, \vec{t}^*_j \ne 0\}$. We note that given any feasible $\vec{x} \in \mathbb{R}^n$, we may write $\vec{x} := \vec{x}^* + \vec{x}'$ where $\vec{x}' := \vec{x} - \vec{x}^*$. Let $\vec{b}' \in \mathbb{R}^m$ be the vector such that $\vec{b}'_j := \vec{b}_j - A\vec{x}^*_j$ for each $1 \le j \le m$. Let A_J be the matrix whose rows are $(A_J)_j = A_j, j \in J$, $\vec{0}^\top$ otherwise. Let A_I be the matrix consisting of the remaining rows of A, i.e. $(A_I)_i = A_i, i \notin J$, $\vec{0}^\top$ otherwise.

We want to show that $A_I \vec{x}' \le \vec{0} \implies \vec{c}^\top \vec{x}' \le 0$. Suppose this is not true, and that there is a $\vec{x}'$ such that $A_I \vec{x}' \le 0$, $\vec{c}^\top \vec{x}' > 0$. Then consider that for some $\epsilon > 0$, $A_J \epsilon \vec{x}' < \vec{b}'$, Since $A_I \vec{x}' \le 0$, $A_I \epsilon \vec{x}' \le 0$ as well, and $A\epsilon \vec{x}' \le \vec{b}'$. Thus $\vec{c}^\top \epsilon \vec{x}' \le 0$, else $\vec{x}^* + \vec{x}$" is an improved solution to the primal problem.

Thus $\vec{c}^\top \vec{x}' >\le 0$, contradicting the existence of such a $\vec{x}'$. It follows that $A_I \vec{x}' \le \vec{0} \implies \vec{c}^\top \vec{x}' \le 0$, and so by Theorem 1.5.2 $\vec{c} = A_I^\top \vec{y}'$ for some $\vec{y}' \ge 0$. WLOG, we may assume that $\vec{y'}_j = 0$ whenever $j \in J$. Then

$$\vec{c}^\top \vec{x} = \vec{y'}^\top A\vec{x} = \vec{y'}^\top \vec{b}.$$

So by Theorem 1.2.6, this $\vec{y}'$ is optimal, and $\vec{x}^*, \vec{y}', \vec{t}^*$ and $\vec{s}' := A^T \vec{y}' - \vec{c}$ are complimentary slack.

Since $\vec{y^*}$ is also optimal, it follows that $\vec{y^*}^\top \vec{b} = \vec{y'}^\top \vec{b}$, and thus $g(\vec{y^*}) - f(\vec{x^*})$ is also 0, and $\vec{x^*}, \vec{y^*}$ and induced slack variables $\vec{s^*}, \vec{t^*}$ are complementary slack.

$\square$

Theorem 1.2.31 (Classical Fact VII: Strong Duality)**.** *Let $\vec{x^*} \in \mathbb{R}^n, \vec{y^*} \in \mathbb{R}^m$ be a pair of feasible optimal solutions for the primal and dual programs respectively. Then*

$$f(\vec{x^*}) = g(\vec{y^*}).$$

Proof. Recall that by Theorem 1.2.5

$$g(\vec{y^*}) - f(\vec{x^*}) = \vec{s}^\top \vec{x} + \vec{y}^\top \vec{t}.$$

By Theorem 1.2.30, $\vec{s}^\top \vec{x} + \vec{y}^\top \vec{t} = 0$, so $f(\vec{x^*}) = g(\vec{y^*})$. $\square$

1.2.6 Simplex Algorithm

Given the initial data that defines a primal dual affine program $(A \in \mathbb{R}^{m\times n}, \vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n, d \in \mathbb{R})$, it is also possible to formulate an algorithm, where the input is a primal feasible affine program, and the output is either the primal-dual optimal solution, or a statement that the primal program solution values are unbounded.

Consider a Tucker tableau:

$$
\begin{array}{r|c|c|l}
 & \vec{x}^{\top} & \text{-1} & \\
\hline
\vec{y} & A & \vec{b} & = -\vec{t} \\
\hline
-1 & \vec{c}^{\top} & d & = f \\
\hline
\multicolumn{1}{r}{} & \multicolumn{1}{c}{\|} & \multicolumn{1}{c}{\|} & \\
\multicolumn{1}{r}{} & \multicolumn{1}{c}{\vec{s}^{\top}} & \multicolumn{1}{c}{g} &
\end{array}
$$

The top and right variables are called *independent* variables (primal and dual respectively), and the bottom and right variables are the *dependent* variables (primal and dual respectively). This tableau records a primal "solution" where each of the primal independent variables x_i are 0. The values of the primal dependent variables $(-t_j)$ are then exactly the values of $-b_j$, or $t_j = b_j$. We can think of this as been the solution that corresponds to the intersection of all the hyperplanes $x_i = 0$ in $\mathbb{R}^n$. Similarly, this tableau also records a dual solution where each one of the dual independent variables y_j are 0, and the dual dependent variables s_i are exactly the $-c_i$. Finally, the tableau records the solution values of $f(\vec{x}), g(\vec{y})$, and since all independent variables are set to 0, the solution is simply $-d$.

What then determines if such a tableau describes an optimal solution? An equivalent condition to a decision variable being feasible, is each independent and dependent coordinate is non-negative. Thus $\vec{x} = \vec{0}$ is feasible if each $t_j \geq 0$, since each $x_i = 0$ and thus $x_i \geq 0$. Similarly, a dual variable is feasible if each $s_i \geq 0$. If both the primal and dual solutions are feasible, by construction they yield the same solution value: $-d$. Thus by Theorem 1.2.6, the Weak Duality Theorem, both solutions must be optimal. However, if there is a x_i such that $c_i > 0$ but for each $j, 1 \leq j \leq m$ such that $a_{ji} \leq 0$, then for each j, we have that:

$$a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{ji}x_i + \ldots + a_{jn}x_n \leq b_j.$$

Since each $a_{ji} \leq 0$, for any value $r \in R_+$ $x_1 = 0, x_2 = 0, \ldots, x_i = 0, \ldots, x + n = 0$ is feasible. Then the objection solution value is $c_i \cdot x_i - d > -d$. We see that by making r arbitrarily large, that the $c_i \cdot x_i - d$ is arbitrarily large. Thus the primal solution values are unbounded.

Definition 1.2.32. A Tucker tableau is *primal feasible* if $\vec{b} \geq \vec{0}$. A Tucker tableau is *dual feasible* $\vec{c} \leq \vec{0}$. If the tableau is both primal and dual feasible, then it is *optimal*.

Remark 1.2.33. How then, does one obtain an optimal Tucker tableau from a primal feasible one? The intuition behind this process follows from the underlying geometry. Both the decision variables x_i, y_j and the slack variables s_i, t_j, represent a "slack" or distance from a bounding hyperplane: either a hyperplane of the form $-x_i = 0, y_j = 0$ or of the form $A_j\vec{x} = b_j$, $A_i^\top = c_i$. Thus a primal solution can be obtained by selecting n of the primal decision or slack variables to play the role of the original $\vec{x}$, that is to be set equal to 0. This selection corresponds to an intersection of n of the bounding hyperplanes.

We then discuss the process of obtaining a new feasible Tucker tableau from a feasible Tucker tableau. When we consider the primal program, we note that a feasible Tucker tableau corresponds to the following system of affine equalities:

$$
\begin{aligned}
a_{11} \cdot x_1 + a_{12} \cdot x_2 + \ldots + a_{1n} \cdot x_n - b_1 &= -t_1 \\
a_{21} \cdot x_1 + a_{22} \cdot x_2 + \ldots + a_{2n} \cdot x_n - b_2 &= -t_2 \\
\vdots &= \vdots \\
a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + \ldots + a_{mn} \cdot x_n - b_m &= -t_m \\
c_1 \cdot x_1 - c_2 \cdot x_2 + \ldots + c_n \cdot x_n - d &= w.
\end{aligned}
$$

Where w is a placeholder variable representing the value of $f(\vec{x})$ associated with this tableau. As discussed above, if we wish to replace one of the x_i with one of the t_j to represent a new solution, we could do so by simply solving for x_i in the affine equality containing t_j, and replacing all instances of x_i with the appropriate expressions.

For example, without loss of generality, we may replace x_1 with t_1. We note that:

$$\begin{aligned}
a_{11} \cdot x_1 + a_{12} \cdot x_2 + \ldots + a_{1n} \cdot x_n - b_1 &= -t_1 \\
-a_{12} \cdot x_2 - \ldots - a_{1n} \cdot x_n - t_1 + b_1 &= a_{11} \cdot x_1 \\
\frac{1}{a_{11}} \cdot t_1 + \frac{a_{12}}{a_{11}} \cdot x_2 + \ldots + \frac{a_{1n}}{a_{11}} \cdot x_n - \frac{b_1}{a_{11}} &= -x_1.
\end{aligned}$$

We then consider the affine equality:

$$\begin{aligned}
a_{j1} \cdot x_1 + a_{j2} \cdot x_2 + \ldots + a_{jn} \cdot x_n - b_j &= -t_j \\
a_{j1} \cdot \left(-\frac{1}{a_{11}} \cdot t_1 - \frac{a_{12}}{a_{11}} \cdot x_2 - \ldots - \frac{a_{1n}}{a_{11}} \cdot x_n + \frac{b_1}{a_{11}} \right) + a_{j2} \cdot x_2 + \ldots + a_{jn} \cdot x_n - b_j &= -t_j \\
-\frac{a_{j1}}{a_{11}} \cdot t_1 + \left(a_{j2} - \frac{a_{j1} a_{12}}{a_{11}} \right) \cdot x_2 + \ldots + \left(a_{jn} - \frac{a_{j1} a_{1n}}{a_{11}} \right) \cdot x_n + \left(\frac{a_{j1} b_1}{a_{11}} - b_j \right) &= -t_j
\end{aligned}$$

Finally, we may also rewrite the objective value:

$$\begin{aligned}
c_1 \cdot x_1 + c_2 \cdot x_2 \ldots x_n \cdot x_n - d &= w \\
c_1 \cdot \left(-\frac{1}{a_{11}} \cdot t_1 - \frac{a_{12}}{a_{11}} \cdot x_2 - \ldots - \frac{a_{1n}}{a_{11}} \cdot x_n + \frac{b_1}{a_{11}} \right) + c_2 \cdot x_2 + \ldots + c_n \cdot x_n - d &= w \\
-\frac{c_1}{a_{11}} \cdot t_1 + \left(c_2 - \frac{c_1 a_{12}}{a_1 1} \right) \cdot x_2 + \ldots + \left(c_2 - \frac{c_n a_{1n}}{a_1 1} \right) - \left(d - \frac{c_1 b_1}{a_{11}} \right) &= w
\end{aligned}$$

What happens to the dual variables in this case? We may consider the dual decision variables y_j to be weights placed on the hyperplanes defined by

$$a_{j1} \cdot x_1 + a_{j2} \cdot x_2 + \ldots + a_{jn} \cdot x_n = b_j$$

and the dual slack variables s_i to be weights placed on the hyperplanes

$$-x_i = 0.$$

Notice that if a feasible primal solution is optimal, then it should still be optimal if one were to remove the other hyperplanes that do not define this point (the slack hyperplanes) Thus a dual solution should exist that is an linear combination of the affine functionals which define the non-slack hyperplanes. (This was made explicit by Theorem 1.2.30)

Thus, for the initial Tucker tableau, the associated primal feasible solution is one where each $x_i = 0$, and the hydroplanes $a_{j1} \cdot x_1 + a_{j2} \cdot x_2 + \ldots + a_{jn} \cdot x_n = b_j$ are slack. Thus each y_j was also set to 0. When one exchanges x_i for t_j, one then exchanges s_i for y_j.

It can be illustrated in Figure 1.5

Figure 1.5: A Tucker pivot

$$\begin{array}{cc} & x \\ y & \boxed{\begin{array}{cc} p^* & k \\ q & w \end{array}} = -t \\ & \parallel \\ & s \end{array} \quad \mapsto \quad \begin{array}{cc} & t \\ s & \boxed{\begin{array}{cc} \frac{1}{p} & \frac{k}{p} \\ -\frac{q}{p} & w - \frac{qk}{p} \end{array}} = -x \\ & \parallel \\ & y \end{array}$$

Where other labels are unchanged.

Such an exchange of variables (as described above) is called a *Tucker pivot.*

It then suffices to describe a set of rules on how Tucker pivots are selected. This leads to what is referred to as the **Two Phase Simplex Algorithm.** Here we discuss each phase of the simplex algorithm and the steps involved. We also make the assumption of non-degeneracy.

Definition 1.2.34. The associated solution of a Tucker tableau is considered *degenerate* if any entries in the $\vec{b}$ column are 0. Equivalently, More than n of the defining hyperplanes intersect at that solution.

A program with no degenerate solutions is *non-degenerate*.

We will assume for now that our given programs are non-degenerate. Later we will address the case of degeneracy. **Phase I:** Phase I of the Simplex algorithm takes a tableau describing a primal problem with non-empty feasible region and returns a tableau that is feasible. Typically if the origin is a feasible primal solution, then this step is not necessary and in fact Phase I will return the same tableau as an output. However it may be the case that the solution associated with a given initial tableau is not primal feasible.

Example 1.2.35. Consider the following primal maximization problem:

$$\begin{aligned}
\text{Find } x_1, x_2 &\geq 0 \text{ such that} \\
f(x_1, x_2) &= x_1 + 5x_2 \text{ is maximized, subject to:} \\
x_1 + x_2 &\geq 1 \\
2x_1 + x_2 &\leq 5
\end{aligned}$$

We note that the inequality $x_1 + x_2 \geq 1$ may be re-written as $-x_1 - x_2 \leq -1$. Thus the associated Tucker tableau for this problem is:

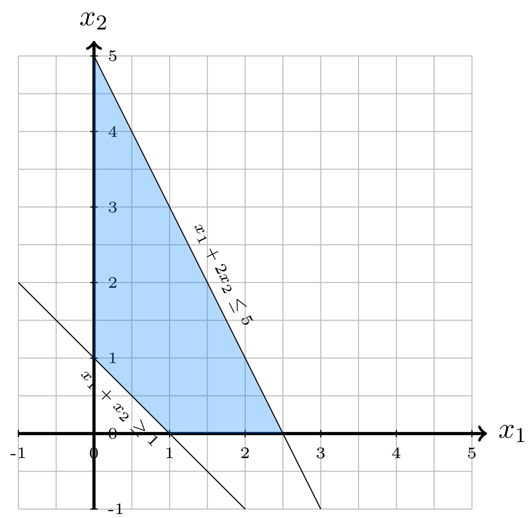
Figure 1.6: Origin infeasible tableau

$$\begin{array}{c|cc|c|l}
 & x_1 & x_2 & \text{-1} & \\
\hline
y_1 & -1 & -1 & -1 & = -t_1 \\
y_2 & 2 & 1 & 5 & = -t_2 \\
\hline
-1 & 1 & 5 & 0 & = f \\
\hline
 & \| & \| & \| & \\
 & s_1 & s_2 & g &
\end{array}$$

We see that the associated solution $x_1, x_2 = 0$ is not feasible, since $-0 - 0 \not\leq -1$. Moreover, we may observe this fact geometrically.

We see that the origin is *not* within the feasible region.

Figure 1.7: Origin infeasible, feasible region



Thus we need to select pivots such that the result is a tableau whose associated solution is feasible, or is unbounded:

1. Select b_ℓ such that ℓ is the largest index where b_ℓ is negative. If not such b_ℓ exists then the associated solution is feasible. STOP

2. Select $a_{\ell i}$ be any negative entry in this row. If each $a_{\ell i}$ is non-negative, then the original primal problem is infeasible. STOP.

3. For $a_{\ell i}$ and for each positive entry $a_{ki}, k \geq \ell$, form the quotient $\frac{b_k}{a_{ki}}$. Select $\frac{b_k}{a_{ki}}$ to be the smallest of these quotients.

4. Pivot on a_{ki}.

5. Goto 1.

We now sketch the proof that this process returns a feasible tableau:

Proof. When the pivot is made, for each $z \geq \ell$ there are several cases:

1. If $k = \ell, z = k, \ell$, then b_ℓ is replaced with $b'_\ell := \frac{b_\ell}{a_{\ell i}}$, since $b_\ell, a_{\ell i} < 0$, $b'_\ell > 0$.

2. If $k = \ell$, $z \neq k, \ell$, then b_z is replaced by $b'_z := b_z - \frac{b_\ell a_{zi}}{a_{\ell i}}$. If $a_{zi} < 0$, then $b'_z > 0$ since $b_z > 0$ and $b_\ell, a_{\ell i} < 0$. Otherwise:

$$\begin{aligned} \frac{b_\ell}{a_{\ell i}} &\leq \frac{b_z}{a_{zi}} \\ b_\ell a_{zi} &\geq b_z a_{\ell i}, \text{ since both } a_{zi} > 0, a_{\ell i} < 0, \\ b_z a_{\ell i} - b_\ell a_{zi} &\leq 0 \\ \frac{b_z a_{\ell i} - b_\ell a_{zi}}{a_{\ell i}} &\geq 0 \text{ since } a_{\ell i} < 0. \end{aligned}$$

3. If $k > \ell, z = k$, then b_z is replaced with $b'_z := \frac{b_z}{a_{zi}}$, and since $a_{ki} > 0$, the sign of this entry is preserved, and since $b_z > 0, b'_z > 0$.

4. If $k > \ell, z = \ell$, then b_z is replaced with $b'_z := b_z - \frac{b_k a_{zi}}{a_{ki}}$. Since $z = \ell$, it follows that $a_{zi} < 0$, similarly since $k > \ell, a_{ki}$, it follows that $b_k > 0$. Thus $-\frac{b_k a_{zi}}{a_{ki}} > 0$, and $b'_z > b_z$.

5. If $k > \ell, z \neq k, \ell$, then

$$\begin{aligned} \frac{b_k}{a_{ki}} &\leq \frac{b_z}{a_{zi}} \\ b_k a_{zi} &\leq b_z a_{ki}, \text{ since both } a_{zi}, a_{ki} > 0, \\ b_z a_{ki} - b_k a_{zi} &\geq 0 \\ \frac{b_z a_{ki} - b_k a_{zi}}{a_{ki}} &\geq 0. \end{aligned}$$

We see by cases $2, 3, 5$ where $z \neq \ell$, that b'_z is a non-negative (thus positive by non-degeneracy) value. Thus this process always preserves the positivity of entries $b_z, z > \ell$. In case 1, b'_ℓ is positive and so now the largest index j such that $b_j < 0$ is strictly bounded above by ℓ. In case 4, we move from an infeasible intersection of hyperplanes to a (also potentially infeasible) intersection of hyperplanes. But in doing so we strictly increase the value of the entry b_ℓ. Since there are at most $n+m$ hyperplanes, the number of possible vertices are also finite, and only a finite number of infeasible vertices. Thus this process cannot be infinite and in some iteration, $b'_\ell > 0$.

Thus, in a finite number of steps, we can strictly lower the largest index with a negative entry. Then since this index is itself finite, we can iterate this process until no negative entries are left.

$\square$

Example 1.2.36. Recall Example 1.2.35:

$$\begin{array}{c|cc|c|l} & x_1 & x_2 & \text{-1} & \\ \hline y_1 & -1 & -1 & -1 & = -t_1 \\ y_2 & 2 & 1 & 5 & = -t_2 \\ \hline -1 & 1 & 5 & 0 & = f \\ \hline & \| & \| & \| & \\ & s_1 & s_2 & g & \end{array}$$

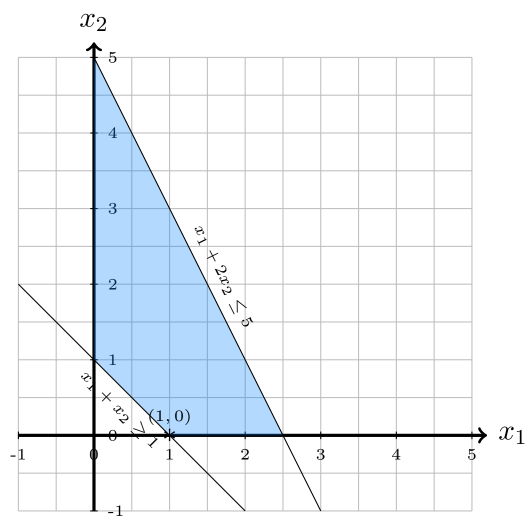
By our rules, we note that the first row is the only row where the entry in the $\vec{b}$ column is negative. Let us select the second column to be the picot column. Then, by our rules, the first row is our pivot row: Since both entries of the $\vec{b}$ column are non-negative, this is a feasible

Figure 1.8: Origin infeasible pivot

$$\begin{array}{c|cc|c|l} & x_1 & x_2 & \text{-1} & \\ \hline y_1 & -1 & -1^* & -1 & = -t_1 \\ y_2 & 2 & 1 & 5 & = -t_2 \\ \hline -1 & 1 & 5 & 0 & = f \\ \hline & \| & \| & \| & \\ & s_1 & s_2 & g & \end{array} \quad \mapsto \quad \begin{array}{c|cc|c|l} & x_1 & t_1 & \text{-1} & \\ \hline s_2 & 1 & -1 & 1 & = -x_2 \\ y_2 & 1 & 1 & 4 & = -t_2 \\ \hline -1 & -4 & 5 & -5 & = f \\ \hline & \| & \| & \| & \\ & s_1 & y_1 & g & \end{array}$$

tableau. We note that the solution associated to this tableau $(1, 0)$ is in fact feasible:

Figure 1.9: Origin infeasible, End of Phase I



Phase II: Here, we assume that Phase I is complete and that the given tableau is feasible. It then suffices to devise rules to select a pivot that will lead to a tableau where the associated solution is optimal, or the described program is unbounded. That is, no entry b_j is negative, and by the non-degeneracy assumption, each b_j is in fact positive.

1. Select a c_i, where $c_i > 0$. If each $c_i \leq 0$, then the current tableau is optimal. STOP.
2. For each positive entry a_{ji}, compute $\frac{b_j}{a_{ji}}$. If there are no positive entries a_{ji}, then the primal program is unbounded. STOP.
3. Let a_{ki} be the entry where the ratio $\frac{b_k}{a_{ki}}$ is the smallest amongst all valid entries. Pivot on a_{ki}.
4. Go to 1.

Here we again include a sketch of the proof that this algorithm terminates.

Proof. We first note that for each $z \neq k$, $b'_z := b_z - \frac{b_k a_{zi}}{a_{ki}}$. But as before, either $a_{zi} \leq 0$, in which case, $-\frac{b_k a_{zi}}{a_{ki}} \geq 0$, since $b_k, a_{ki} > 0$. Thus $b'_z \geq b_z > 0$. Otherwise, if $a_{z_i} > 0$:

$$\begin{aligned}
\frac{b_k}{a_{ki}} &\leq \frac{b_z}{a_{zi}} \\
b_k a_{zi} &\leq b_z a_{ki}, \text{ since both } a_{zi}, a_{ki} > 0, \\
b_z a_{ki} - b_k a_{zi} &\geq 0 \\
\frac{b_z a_{ki} - b_k a_{zi}}{a_{ki}} &\geq 0.
\end{aligned}$$

Thus $b'_z > 0$. We also observe that $b'_k := \frac{b_k}{a_{ki}} > 0$, since $b_k, a_{ki} > 0$. Thus the resulting tableau is feasible.

We also notice that $d' := d - \frac{c_i b_k}{a_{ki}}$, since $c_i, b_k, a_{ik} > 0, d' < d$. Thus each such pivot strictly improves the solution. Since the result of each such pivot is an improved solution, no vertex can be visited more than once by this method. Since again there are only finitely many vertices, this process must terminate.

□

Example 1.2.37. Recall the Lumbermill Problem 1.2.1. The primal and dual programs were encapsulated in the following tableau:

Figure 1.10: Lumbermill tableau

$$\begin{array}{c|ccc|c|l}
 & x_1 & x_2 & x_3 & -1 & \\
\hline
y_1 & 1 & 3 & 2 & 10 & = -t_1 \\
y_2 & 2 & 1 & 1 & 8 & = -t_m \\
\hline
-1 & 3 & 2 & 4 & 5 & = f \\
\hline
 & \| & \| & \| & \| & \\
 & s_1 & s_2 & s_3 & g &
\end{array}$$

Let us pick c_1 to be the pivot column, then the second row must be our pivot row:

$$
\begin{array}{c|ccc|c|l}
 & x_1 & x_2 & x_3 & \text{-1} & \\
\hline
y_1 & 1 & 3 & 2 & 10 & = -t_1 \\
y_2 & 2^* & 1 & 1 & 8 & = -t_2 \\
\hline
-1 & 3 & 2 & 4 & 5 & = f \\
\hline
 & \| & \| & \| & \| & \\
 & s_1 & s_2 & s_3 & g &
\end{array}
\quad \mapsto \quad
\begin{array}{c|ccc|c|l}
 & t_2 & x_2 & x_3 & \text{-1} & \\
\hline
y_1 & -\frac{1}{2} & \frac{5}{2} & \frac{3}{2} & 6 & = -t_1 \\
s_1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 4 & = -x_1 \\
\hline
-1 & -\frac{3}{2} & \frac{1}{2} & \frac{5}{2} & -7 & = f \\
\hline
 & \| & \| & \| & \| & \\
 & y_2 & s_2 & s_3 & g &
\end{array}
$$

We then select c_2 to be the pivot column which forces row one to be the pivot row:

$$
\begin{array}{c|ccc|c|l}
 & t_2 & x_2 & x_3 & \text{-1} & \\
\hline
y_1 & -\frac{1}{2} & \frac{5}{2}^* & \frac{3}{2} & 6 & = -t_1 \\
s_1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 4 & = -x_1 \\
\hline
-1 & -\frac{3}{2} & \frac{1}{2} & \frac{5}{2} & -7 & = f \\
\hline
 & \| & \| & \| & \| & \\
 & y_2 & s_2 & s_3 & g &
\end{array}
\quad \mapsto \quad
\begin{array}{c|ccc|c|l}
 & t_2 & t_1 & x_3 & \text{-1} & \\
\hline
s_2 & -\frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{12}{5} & = -x_2 \\
s_1 & \frac{3}{5} & -\frac{1}{5} & \frac{1}{5} & \frac{14}{5} & = -x_1 \\
\hline
-1 & -\frac{7}{5} & -\frac{1}{5} & \frac{11}{5} & -\frac{31}{5} & = f \\
\hline
 & \| & \| & \| & \| & \\
 & y_2 & y_1 & s_3 & g &
\end{array}
$$

Now c_3 is the only choice for pivot column, with row one been the only choice for pivot row.

$$
\begin{array}{c|ccc|c|l}
 & t_2 & t_1 & x_3 & \text{-1} & \\
\hline
s_2 & -\frac{1}{5} & \frac{2}{5} & \frac{3}{5}^{*} & \frac{12}{5} & = -x_2 \\
s_1 & \frac{3}{5} & -\frac{1}{5} & \frac{1}{5} & \frac{14}{5} & = -x_1 \\
\hline
-1 & -\frac{7}{5} & -\frac{1}{5} & \frac{11}{5} & -\frac{31}{5} & = f \\
\hline
 & \shortparallel & \shortparallel & \shortparallel & \shortparallel & \\
 & y_2 & y_1 & s_3 & g &
\end{array}
\quad \mapsto \quad
\begin{array}{c|ccc|c|l}
 & t_2 & t_1 & x_3 & \text{-1} & \\
\hline
s_2 & -\frac{1}{3} & \frac{2}{3} & \frac{3}{5} & 4 & = -x_2 \\
s_1 & \frac{8}{15} & -\frac{1}{3} & -\frac{1}{3} & 2 & = -x_1 \\
\hline
-1 & -\frac{2}{3} & -\frac{5}{3} & -\frac{11}{3} & -17 & = f \\
\hline
 & \shortparallel & \shortparallel & \shortparallel & \shortparallel & \\
 & y_2 & y_1 & s_3 & g &
\end{array}
$$

Since each entry in the $\vec{c}$ row is non-positive, we have achieved optimality. Thus the primal solution is $(2, 0, 4)$ and the dual solution is $(\frac{5}{3}, \frac{2}{3})$. We can illustrate these solutions:

Figure 1.11: Lumbermill Primal-Optimal Solution

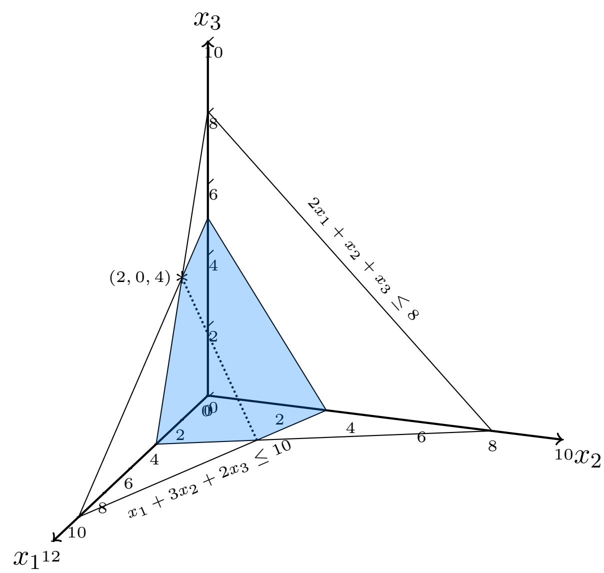
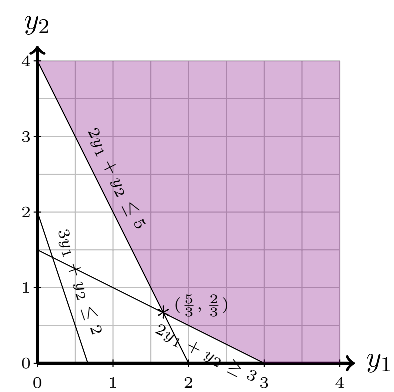
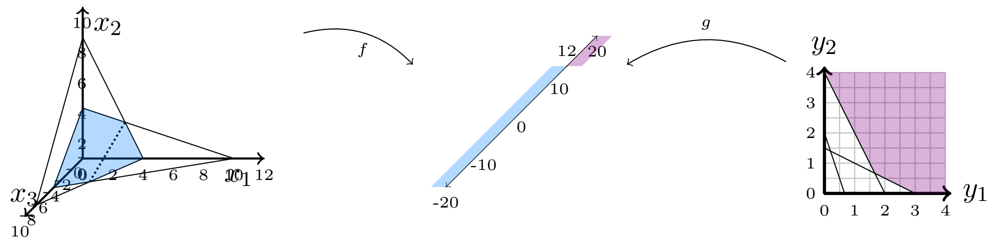


Figure 1.12: Lumbermill Dual-Optimal Solution



In fact, we may represent the feasible region of each of these spaces, the primal and dual variables, and the solutions, with the following diagram:

Figure 1.13: Lumbermill 3-space diagram



Example 1.2.38. Consider the following tableau (due to E.M.L. Beale [Bea55]):

$$
\begin{array}{c|cccc|c|l}
 & x_1 & x_2 & x_3 & x_4 & -1 & \\ \hline
y_1 & \frac{1}{4} & -8 & -1 & 9 & 0 & = -t_1 \\
y_2 & \frac{1}{2} & -12 & -\frac{1}{2} & 3 & 0 & = -t_2 \\
y_3 & 0 & 0 & 1 & 0 & 1 & = -t_3 \\ \hline
-1 & \frac{3}{4} & -20 & \frac{1}{2} & -6 & 0 & = f \\ \hline
 & \| & \| & \| & \| & \| & \\
 & s_1 & s_2 & s_3 & s_4 & g &
\end{array}
$$

Then consider the following sequence of pivots:

$$
\begin{array}{c|cccc|c|l}
 & x_1 & x_2 & x_3 & x_4 & -1 & \\ \hline
y_1 & \frac{1}{4}^* & -8 & -1 & 9 & 0 & = -t_1 \\
y_2 & \frac{1}{2} & -12 & -\frac{1}{2} & 3 & 0 & = -t_2 \\
y_3 & 0 & 0 & 1 & 0 & 1 & = -t_3 \\ \hline
-1 & \frac{3}{4} & -20 & \frac{1}{2} & -6 & 0 & = f \\ \hline
 & \| & \| & \| & \| & \| & \\
 & s_1 & s_2 & s_3 & s_4 & g &
\end{array}
\mapsto
\begin{array}{c|cccc|c|l}
 & t_1 & x_2 & x_3 & x_4 & -1 & \\ \hline
s_1 & 4 & -32 & -4 & 36 & 0 & = -x_1 \\
y_2 & -2 & 4^* & \frac{3}{2} & -15 & 0 & = -t_2 \\
y_3 & 0 & 0 & 1 & 0 & 1 & = -t_3 \\ \hline
-1 & -3 & 4 & \frac{7}{2} & -33 & 0 & = f \\ \hline
 & \| & \| & \| & \| & \| & \\
 & y_1 & s_2 & s_3 & s_4 & g &
\end{array}
\mapsto
$$

$$
\begin{array}{c|cccc|c|l}
 & t_1 & t_2 & x_3 & x_4 & -1 & \\
\hline
y_1 & -12 & 8 & 8^* & -84 & 0 & = -x_1 \\
s_2 & -\frac{1}{2} & -\frac{1}{4} & \frac{3}{8} & -\frac{15}{4} & 0 & = -x_2 \\
y_3 & 0 & 0 & 1 & 0 & 1 & = -t_3 \\
\hline
-1 & -1 & -1 & 2 & -18 & 0 & = f \\
\hline
 & \parallel & \parallel & \parallel & \parallel & \parallel & \\
 & s_1 & y_2 & s_3 & s_4 & g &
\end{array}
\quad \mapsto \quad
\begin{array}{c|cccc|c|l}
 & t_1 & t_2 & x_1 & x_4 & -1 & \\
\hline
s_3 & -\frac{3}{2} & 1 & \frac{1}{8} & -\frac{21}{2} & 0 & = -x_3 \\
s_2 & \frac{1}{16} & -\frac{1}{8} & -\frac{3}{64} & \frac{3}{16}^* & 0 & = -x_2 \\
y_3 & \frac{3}{2} & -1 & -\frac{1}{8} & -\frac{21}{2} & 1 & = -t_3 \\
\hline
-1 & 2 & -3 & -\frac{1}{4} & 3 & 0 & = f \\
\hline
 & \parallel & \parallel & \parallel & \parallel & \parallel & \\
 & s_1 & y_2 & y_1 & s_4 & g &
\end{array}
\quad \mapsto
$$

$$
\begin{array}{c|cccc|c|l}
 & t_1 & t_2 & x_1 & x_2 & -1 & \\
\hline
s_3 & 2^* & -6 & -\frac{5}{2} & 56 & 0 & = -x_3 \\
s_4 & \frac{1}{3} & -\frac{2}{3} & -\frac{1}{4} & \frac{16}{3} & 0 & = -x_4 \\
y_3 & -2 & 6 & \frac{5}{2} & -56 & 1 & = -t_3 \\
\hline
-1 & 1 & -1 & \frac{1}{2} & -16 & 0 & = f \\
\hline
 & \parallel & \parallel & \parallel & \parallel & \parallel & \\
 & s_1 & y_2 & y_1 & s_2 & g &
\end{array}
\quad \mapsto \quad
\begin{array}{c|cccc|c|l}
 & x_3 & t_2 & x_1 & x_2 & -1 & \\
\hline
s_1 & \frac{1}{2} & -3 & -\frac{5}{4} & 28 & 0 & = -t_1 \\
s_4 & -\frac{1}{6} & \frac{1}{3}^* & \frac{1}{6} & -4 & 0 & = -x_4 \\
y_3 & 1 & 0 & 0 & 0 & 1 & = -t_3 \\
\hline
-1 & -\frac{1}{2} & 2 & \frac{7}{4} & -44 & 0 & = f \\
\hline
 & \parallel & \parallel & \parallel & \parallel & \parallel & \\
 & s_3 & y_2 & y_1 & s_2 & g &
\end{array}
\quad \mapsto
$$

$$
\begin{array}{c|cccc|c|l}
 & x_3 & x_4 & x_1 & x_2 & -1 & \\
\hline
s_1 & -1 & 9 & \frac{1}{4} & -8 & 0 & = -t_1 \\
y_2 & -\frac{1}{2} & 3 & \frac{1}{2} & -12 & 0 & = -t_2 \\
y_3 & 1 & 0 & 0 & 0 & 1 & = -t_3 \\
\hline
-1 & \frac{1}{2} & -6 & \frac{3}{4} & -20 & 0 & = f \\
\hline
 & \parallel & \parallel & \parallel & \parallel & \parallel & \\
 & s_3 & s_4 & y_1 & s_2 & g &
\end{array}
$$

But we see that the entries of this final tableau is identical to the first tableau, up to a rearrangement of columns. Such behavior is referred to as *cycling*.

However, a result by Robert Bland [Bla77] insures that we may choose pivots in such a way such that cycling does not occur.

Theorem 1.2.39 (Bland's Anticycling Rule)**.** *When choosing Tucker pivot's, if one selects the column with smallest index of all valid column choices, and one selects the row with smallest index of all valid row choices, then the Simplex algorithm will terminate.*

Theorem 1.2.40 (Classical Fact XI: Simplex Algorithm)**.** *The two phase Simplex Algorithm, along with Bland's Anticycling rule terminates.*

Knowing that the simplex algorithm is guaranteed to terminate gives us alternative proofs to one of our previous results.

Alternative proof to Strong Duality 1.2.31. Suppose that there exists optimal solutions to both the primal and dual problem. Run the simplex algorithm. Since an optimal primal solution exists, the algorithm terminates, where the output is a tableau whose associated primal and dual solutions are optimal.

Thus, a dual solution is found, and by construction it has the same objective value as the primal solution. Thus all dual optimal solutions must have the same objective value. □

We can also now finally prove a powerful classification result about primal-dual affine programs.

Theorem 1.2.41 (Classical Fact VIII: Existence-Duality Theorem)**.** *Given a primal-dual affine program, exactly one of the following hold:*

1. *Both the primal and dual programs are infeasible.*
2. *The primal program is infeasible and the dual program is unbounded.*
3. *The dual program is infeasible and the primal program is unbounded.*

4. Both the primal and dual program achieve an optimal solution.

Proof. Consider the following cases:

- If the primal program is infeasible, and the dual program is also infeasible, then we satisfy (1) .
- Otherwise, suppose the primal program is infeasible and there is a feasible $\vec{y}' \in \mathbb{R}^m$. Note that the primal program being infeasible is equivalent to Corollary 1.2.26(1) (recall the discussion in the proof of Theorem 1.2.30). Thus by Corollary 1.2.26, there is a $\vec{y} \in \mathbb{R}^m$ such that $A^\top \vec{y} \geq_{\mathbb{R}^n} 0$ but $\vec{y}^\top \vec{b} < 0$. Thus, given $r \in \mathbb{R}^+$, $\vec{y}' + r\vec{y}$ is feasible, but by choosing r to be arbitrarily large, $g\left(\vec{y}' + r\vec{y}\right) = \left(\vec{y}' + r\vec{y}\right)^\top \vec{b}$ is arbitrarily small, thus the dual program is unbounded and (2) holds.
- If the primal program is feasible, then then by Theorem 1.2.23, the dual program cannot be unbounded. If the dual program is infeasible, then Theorem 1.5.2(1) does not hold (recall the discussion in the proof of Theorem 1.2.30). Thus, Theorem 1.5.2(2) holds, and there is a $\vec{x}'$ such that $A\vec{x}' \geq_{\mathbb{R}^m} 0$, but $\vec{c}^\top \vec{x} < 0$. Let $\vec{x} = -\vec{x}'$. Then given any feasible vector $\vec{x}''$, we note that $\vec{x}'' + r\vec{x}$ is feasible for any choice of $r \in \mathbb{R}^+$. But by allowing r to be arbitrarily large, $f\left(\vec{x}'' + r\vec{x}\right) = \vec{c}^\top \left(\vec{x}'' + r\vec{x}\right) - d$ is arbitrarily large. Thus the primal program is unbounded (3).
- Otherwise both the primal and dual program are feasible, and by Theorem 1.2.6, both programs are bounded. Since the primal program is feasible and unbounded, when we run the Simplex algorithm on it, it will return an optimal primal and optimal dual solution (4).

$\square$

1.3 Generalizations

This basic framework for affine programming models a particular type of affine optimization problems. However, there are certain types of problems which do not fall under this framework, and as such, we would like to be able to generalize this set-up.

1.3.1 Generalizing the Ring of Scalars

We notice that the notions of feasibility and optimality described above do not depend on the Archimedean or least upper bound properties of real numbers, only it's order. Thus, it is reasonable to believe that these notions may be generalized to an ordered ring R, rather than the real numbers $\mathbb{R}$.

Definition 1.3.1 (Ordered Ring [Lam01])**.** A ring R is *ordered* (sometimes called *trichotomy ordered*), if there is a non-empty subset $P \subset R$ called the "positives" with the following properties:

1. R can be partitioned into the disjoint union: $P \sqcup \{0\} \sqcup -P$ (this is called the trichotomy property).
2. Given $a, b \in P$, then $a + b \in P$.
3. Given $a, b \in P$, then $ab \in P$.

We say that $a \geq b$ if $a - b \in P \cup \{0\}$. We also say that $a > b$ if $a - b \in P$. We will also use $R_{\oplus}$ to denote $P \cup \{0\}$, ("the positives and zero").

A well known modification of affine programming is to change the underlying ring of scalars from $\mathbb{R}$ to $\mathbb{Z}$ [Sch86]. A famous example is the "Knapsack Problem" [Num55].

Problem 1.3.2. Suppose there were a knapsack capable of holding W weight. In it, we wish to place some of n items, each of which has weight w_i (> 0) and value v_i. We wish to find the quantities of each of the n items, quantities $x_1, \ldots, x_n, x_i \in \mathbb{Z}$ such that:

$$\begin{aligned} \textbf{Maximize: } f(x_1, \ldots, x_n) &:= \sum_{i=1}^{n} v_i \cdot x_i \\ \text{subject to: } \sum_{i=1}^{n} w_i \cdot x_i &\leq W \\ x_i &\geq 0 \end{aligned}$$

We see that this is exactly an affine programming problem as described before, except $\mathbb{R}$ is replaced with $\mathbb{Z}$, $A \in \mathbb{Z}^{1\times n}, A_{1i} = w_i, \vec{b} = [W]$ (a 1×1 matrix), $\vec{c} = [v_1, \ldots, v_n]^\top$ and $d = 0$.

Integer programming problems play a great role in modeling and solving real-world problems, and so are a well studied field of optimization. However, integers are not the only ring over which one may wish to do affine programming. Another ring is the hyperreals:

Remark 1.3.3 (A discussion of the Hyperreals ${}^*\mathbb{R}$)**.** A classical example of a non-standard ordered ring is an extension of the Real numbers called the *Hyperreals* (denoted by ${}^*\mathbb{R}$). The Hyperreals were originally a field constructed by Abraham Robinson in the early 1960's [Rob79] to do non-standard analysis, with a focus on the infinitesimal approach to calculus (and analysis) that mimicked the original approach of Leibniz.

The Hyperreals are constructed by placing equivalence classes on sequences of real numbers. This involves an object called an *ultrafilter*.

Definition 1.3.4 ([Rob79])**.** An *ultrafilter* of $\mathbb{N}$ is a subset $U \subseteq \mathcal{P}(\mathbb{N})$ such that:

- If $S \in U, S \subseteq S'$, $S' \in U$. ("Closed under super sets" property)
- If $S_1, S_2 \in U$, $S_1 \cap S_2 \in U$. ("Closed under finite intersections" property)

- If $S \subset \mathbb{N}$, $|S| < \infty$, then $S \notin U$. ("Cauchy" property)

- If $S \subset \mathbb{N}$, then $S \in U$ if and only if $S^c \notin U$. ("Ultrafilter" property)

A collection of subsets satisfying the first two properties is called a filter. The subsets in a filter are "huge" subsets.

Then we consider the ring of real-valued sequences, $\mathbb{R}^{\mathbb{N}}$ with pointwise sums and products.

Claim 1.3.5. *The set $M := \{\varphi \in \mathbb{R}^{\mathbb{N}} : \{n \in \mathbb{N} : \varphi(n) = 0\} \in U\}$ is a maximal ideal of $\mathbb{R}^{\mathbb{N}}$, where U is an ultrafilter.*

Proof. We first show that M is an ideal. Let $\varphi, \psi \in M$ and consider $\varphi + \psi$. We note that if $\varphi(n) = \psi(n) = 0$ then $(\varphi + \psi)(n) = 0$. Thus $\{n : (\varphi + \psi)(n) = 0\} \supseteq \{n : \varphi(n) = 0\} \cap \{n : \psi(n) = 0\} \in \mathcal{U}$. Thus $\{n : (\varphi + \psi)(n) = 0\} \in \mathcal{U}$ and M is closed under sums.

Let $\rho \in \mathbb{R}^{\mathbb{N}}$, and consider $\rho \cdot \psi$. We notice that $\{n : \rho(n)\psi(n) = 0\} \supseteq \{n : \rho(n) = 0\} \in \mathcal{U}$. Thus $\{n : \rho(n)\psi(n) = 0\} \in \mathcal{U}$ and $\rho \cdot \psi \in M$. We conclude that M is an ideal.

Now, we show that M is maximal. Suppose $M < M' \trianglelefteq \mathbb{R}^{\mathbb{N}}$. Let $r \in M' \backslash M$. Since $r \notin M$, $\{n \in \mathbb{N} : r(n) \neq 0\} \in U$. Then define $s : \mathbb{N} :\to \mathbb{R}$ such that

$$s(n) = \begin{cases} \frac{1}{r(n)}, & r(n) \neq 0 \\ 0 & \text{otherwise} \end{cases}.$$

Notice that

$$r \cdot s = \begin{cases} 1, & r(n) \neq 0 \\ 0 & \text{otherwise} \end{cases},$$

and $r \cdot s \in M'$ since $r \in M'$.

Then define $t : \mathbb{N} \to \mathbb{R}$ by

$$t(n) = \begin{cases} 1, & r(n) = 0 \\ 0 & \text{otherwise} \end{cases}.$$

Notice that $\{n \in \mathbb{N} : t(n) = 0\} \in U$, thus $t \in M \subset M'$. Thus, $r \cdot s + t = 1 \in M'$, contradicting M' been a proper ideal. Thus M is maximal. $\square$

Thus, we define ${}^*\mathbb{R} := \mathbb{R}^{\mathbb{N}}/M$, and by the discussion above, ${}^*\mathbb{R}$ is a field. Although the ultrafilter U is not a σ-algebra, we can also think of the equivalence class induced by M to be all sequences which agree "almost everywhere", meaning over a set that is contained in U. The collection of sequences which is zero "almost everywhere" then is exactly the set M.

We define the positives of ${}^*\mathbb{R}$ to be the classes of sequences which are positive "almost everywhere". Formally, $P := \{[a] \in {}^*\mathbb{R} : \{n : a(n) > 0\} \in U\}$. We then notice that by the complement properties of an ultrafilter, each sequence must be equal to, greater than, or less than zero, almost everywhere.

A 2008 economics paper [Piv08], uses hyperreals to model values over an infinite time span. For example in the lumber mill problem, the number of available large logs in year n could be thought of as a function $\ell_L \in \mathbb{R}^{\mathbb{N}}$, which may then be identified with a hyperreal number. The example problems 1.2.1, 1.2.2 can be then reformulated to find a sequence (or equivalence class of sequences) of either quantities of products (primal) or valuation of resources (dual), given a potentially ever changing resources required for each product (A), quantity of available logs ($\vec{b}$), sale price of our products ($\vec{c}$) and our sunk (or fixed) costs (d). Each of these could be reformulated as vectors or matrices with hyperreal entries.

Another natural ring of scalars for affine optimization problems is the ordered field of rational functions (i.e. quotients of polynomials) in several indeterminants.

Thus, it is natural to wish to extend affine programming to ordered rings.

1.3.2 Generalizing the Dimension or Rank

It is also natural to extend affine programming to situations of arbitrary dimension. Consider the scheduling problem over an infinite time horizon and non-stationary demand found in [DeN82]:

Problem 1.3.6. Given $i \in \mathbb{Z}_+$, where i is the ith period of scheduling, we may define I_i to be the net inventory of period i, P_i to be the net production in period i, and let d_i be the demand during period i, $\bar{I}_i, \bar{P}_i$ to be the upper bound of inventory and production in time i. Additionally, for each i we define k_i, s_i, the costs of producing goods and storing goods in period i respectively. We also define α be a discounting factor, representing the devaluation of money over time. Our goal here is to minimize the cost over an infinite period of time.

Then we have:

$$\begin{aligned}
\textbf{Minimize: } f(P_i, I_i)_{i=1}^{\infty} &:= \sum_{i \in \mathbb{Z}^+} (k_i P_i + s_i I_i)\alpha^{i-1} \text{ subject to:} \\
I_{i-1} + P_i - I_i &\geq d_i \\
P_i &\leq \bar{P}_i \\
I_i &\leq \bar{I}_i \\
P_i, I_i &\geq 0.
\end{aligned}$$

We see that the primal solution space is contained in $\mathbb{R}^{\mathbb{N}}$, and similarly, so is the space of constraints. Thus we may model this program by allowing the vector spaces in question to have arbitrary dimension.

1.3.3 Generalizing the Cones

Given an ordered ring R and X, Y^* left R-modules. We would like to define the analogue to the *non-negative cone* of our classical case.

Definition 1.3.7. A *positive cone* of a module X over an ordered ring R is a subset $C \subset X$ such that given $r, s \in R, r, s \geq 0$, and $a, b \in C$, then $r \cdot a + s \cdot b \in C$.

We also define the following specialized types of cones (C):

- If C contains no non-trivial sub-module of X, then C is said to be a *pointed* cone.
- If there is a $x \in X \backslash C$ and $r \in R, r \geq 0$, such that $r \cdot x \in C$, then we say that C is *perforated*. Otherwise C is unperforated.

Notice that our classical examples of non-negative cones were the non-negative span of some basis. However, it may be the case we wish to work more generally.

Given a topological set S, it is natural to consider the collection of functions $C^n(S)$ (the collection of nth differentiable continuous functions from $S \to \mathbb{R}$) as a vector space over $\mathbb{R}$. Then a good example for a non-negative cone is the collection of non-negative functions in $C^n(S)$. It is easy to check that this collection forms a cone. However, it is not necessarily the case that these functions may be expressed as the non-negative span of some basis of $C^n(S)$. Thus it may be that we will wish to generalize the types of cones over which we do affine programming.

1.4 General Framework

Thus, we establish the general setting over which affine programming takes place. Rather than deal only in finite dimensional real-vector spaces, we can generalize this idea to modules over a ring R. In order to do this, we must properly describe the appropriate features of these problems and their data.

If we allow R to be an arbitrary ring, rather than $\mathbb{R}$, then the natural analogues to $\mathbb{R}^n, \mathbb{R}^m$ would be a pair of (left) R-modules, X, Y^*. We will focus on the left structure of these objects, although there is no reason one cannot define the same concepts for right modules instead. In the traditional case, A is a $n \times m$ real-matrix, or equivalently $A \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m)$, and so the appropriate generalization would be to let $A \in \mathrm{Hom}_R(X, Y^*)$. It is clear that $\vec{b} \in Y^*$ is an element of the co-domain of A, and $d \in R$ as before.

Recall that in the classical case, $\vec{c} \in \mathbb{R}^n$. However, when we consider the role that $\vec{c}$ plays in the original problem, $\vec{c}$ acted on $\vec{x}$ via inner product to obtain a value in $\mathbb{R}$. Hence, the appropriate generalization of $\vec{c}$ is $\mathbf{c} \in X^* := \mathrm{Hom}_R(X, R)$. Similarly, while $\vec{x}$ should remain an element of the domain of A, $\vec{x} \in X$, $\vec{y}$ is a vector that acts on $\vec{b}$ via an inner product, and so $\vec{y}$ should be generalized to $\mathbf{y} \in \mathrm{Hom}_R(Y^*, R) =: Y^{**}$. (Recall that this also makes sense even in the classical case. When $\vec{b}$ represented some quantity of raw materials, $\vec{y}$ represented an evaluation of those materials, a functional from the space of materials to $\mathbb{R}$, the space of revenue).

We may then, define as before:

$$\begin{aligned}
f &: X \to R, \vec{x} \mapsto \mathbf{c}(\vec{x}) - d \\
g &: Y^* \to R, \mathbf{y} \mapsto \mathbf{y}(\vec{b}) - d. \\
\vec{t} &:= -A(\vec{x}) + \vec{b} \\
\mathbf{s} &:= \mathbf{y} \circ A - \mathbf{c}.
\end{aligned}$$

Which gives rise to a Tucker Tableau:

Figure 1.14: General Tucker tableau

$$
\begin{array}{r|c|c|l}
\multicolumn{1}{r}{} & \multicolumn{1}{c}{\vec{x}^\top} & \multicolumn{1}{c}{-1} & \\ \cline{2-3}
\mathbf{y} & A & \vec{b} & = -\vec{t} \\ \cline{2-3}
-1 & \mathbf{c} & d & = f \\ \cline{2-3}
\multicolumn{1}{r}{} & \multicolumn{1}{c}{\|} & \multicolumn{1}{c}{\|} & \\
\multicolumn{1}{r}{} & \multicolumn{1}{c}{\mathbf{s}} & \multicolumn{1}{c}{g} &
\end{array}
$$

1.5 Results

There are some facts about the classical affine primal-dual programming case which we would like to describe in this generalized setting. The goal will be to describe the minimal hypothesis necessary to prove these facts and to further specialize either the ring R, the dimension (or rank) of the modules, and the cones of the modules, until we may establish the appropriate generalization of these results.

1.5.1 Results about affine maps

Some of our classical results describe the property of the underlying affine maps. The classical results Proposition 1.2.4 (the Tucker key equation) and it's Corollary 1.2.5 are results which rely only one the ring structure of the scalars $\mathbb{R}$, and the module structure of the vector spaces $\mathbb{R}^n, \mathbb{R}^m$. We intend to generalize these to arbitrary rings and modules.

1.5.2 Results about Duality

Some of our classical results describe the relationship between primal-dual solutions. Theorem 1.2.6 (the Weak Duality Theorem) is a result about primal-dual solutions that requires the order of the real numbers and the positive orthants of $\mathbb{R}^n, \mathbb{R}^m$. We intend to generalize this to ordered rings and cones in modules.

Theorem (the Farkas Lemma), has a proof that classically requires the least upper bound property of the real numbers. Then Theorems 1.2.30 (the Complementary Slackness Theorem) and 1.2.31 (the Strong Duality Theorem) have proofs that require the conclusion of the Farkas' Lemma. We intend to generalize these to ordered division rings and their vector spaces.

1.5.3 Results Classifying Solutions

There are also a number of classical affine programming results which describe the nature and properties of feasible or optimal solutions to the primal and dual program. Corollaries 1.2.16, 1.2.18 (the Convexity of primal and dual, feasible and optimal solutions) and Theorem 1.2.21 (the existence of extreme point optimizers) are results about the properties of feasible or optimal solutions that require the order of the real numbers and the positive orthants of $\mathbb{R}^n, \mathbb{R}^m$. We intend to generalize these to ordered rings and cones in modules.

Theorem 1.2.41 (the Existence-Duality Theorem) is a result classifying the possible ways that a primal-dual program can have solutions. The proof classically requires the field properties of $\mathbb{R}$ as well as the finite-dimensionality of $\mathbb{R}^n, \mathbb{R}^m$. We intend to generalize this to ordered division rings and their (not necessarily finite-dimensional) vector spaces.

1.5.4 Results about Structure: Tucker Tableaux and Oriented Matroids

In the late 1960's [Roc69] R. T. Rockafellar observed that much of the results about affine programming duality, such as the Tucker Duality Equation, the Complementary Slackness Theorem, and the Strong Duality Theorem, could be encapsulated by statements about sign patterns in complementary subspaces of $\mathbb{R}^n$. In his paper *"The Elementary Vectors of a Subspace of* $\mathbb{R}^n$*"*, he described how the Tucker tableau representation of an affine programming problem gives rise to a natural representation of an affine programming problem and its equivalent tableau as an *oriented matroid*.

Definition 1.5.1. Given a set E (usually but not necessarily finite), a *signed set* or *sign vector* is an ordered pair $X := (X^+, X^-)$, where $X^+, X^- \subseteq E$ and $X^+ \cap X^- = \emptyset$. An *oriented matroid* is an ordered pair $\mathcal{M} := (E, \mathcal{C})$, where $\mathcal{C}$ is a collection of sign vectors such that:

1. The sign vector $(\emptyset, \emptyset)$ is not in $\mathcal{C}$.
2. Given $X, X' \in \mathcal{C}$ such that $X^+ \subseteq X'^+, X^- \subseteq X'^-$, then $X = X'$.
3. If $X \in \mathcal{C}$, then $-X := (X^-, X^+) \in \mathcal{C}$.
4. If $X, Y \in \mathcal{C}$, $X \neq \pm Y$, and there is an $e \in X^+ \cap Y^-$, then there is a $Z \in \mathcal{C}$ such that $Z^+ \subseteq (X^+ \cup Y^+) \backslash \{e\}$, and $Z^- \subseteq (X^- \cup Y^-) \backslash \{e\}$.

We call the elements of $\mathcal{C}$ oriented or signed *circuits*. We call supersets of a signed circuit a *sign vector*.

What is crucial to capturing the information of a primal-dual programming problem is the notion of duality in a matroid. Any oriented matroid $\mathcal{M} := (E, \mathcal{C})$, gives another oriented matroid of cocircuits $mathcalM(E, \mathcal{C}^*)$, the dual matroid of $\mathcal{M}$.

Definition 1.5.2. We say that two sign vectors X, Y are *orthogonal* if one of the following hold:

- $(X^+ \cap Y^+) \cup (X^- \cap Y^-) = \emptyset$ and $(X^+ \cap Y^-) \cup (X^- \cap Y^+) = \emptyset$.

- $(X^+ \cap Y^+) \cup (X^- \cap Y^-) \neq \emptyset$ and $(X^+ \cap Y^-) \cup (X^- \cap Y^+) \neq \emptyset$

Definition 1.5.3. Given a (non-oriented) matroid $\mathcal{M} := (E, \mathcal{C})$, the basis $\mathcal{B}$ is a set $\mathcal{B} \subseteq E$ such that given $e \in E \backslash \mathcal{B}$, $\mathcal{B} \cup \{e\}$ contains a unique circuit $e \in C \in \mathcal{C}$. We say that $\mathcal{B}$ is a maximal independent set.

Proposition 1.5.4. *Given an oriented matroid $\mathcal{M} := (E, \mathcal{C})$, we have the following:*

1. *There is a collection of cocircuits $\mathcal{C}^*$, such that the pair $\mathcal{M}^* := (E, \mathcal{C}^*)$ forms an oriented matroid.*

2. *Given $C^* \in \mathcal{C}^*$, $\mathcal{C}^*$ is orthogonal to each circuit $\mathbb{C}^{\in}\mathcal{C}$.*

3. $(\mathcal{M}^*)^* = \mathcal{M}$.

Theorem 1.5.5 (Primal-Dual Affine Program admits an oriented matroid structure)**.** *Given the classical Tucker tableau $\hat{E}data$, the collection of feasible primal solutions form a collection of covectors, and the collection of dual solutions form a collection of vectors, which in turn form a pair of orthogonal oriented matroids.*

Example 1.5.6 (The Lumbermill Production and Insurance Matroid)**.** Using the variable g, we homogenize the affine functionals which define the Lumbermill Production and Insurance Problems 1.2.1 1.2.2. This gives us the following linear functionals:

$$
\begin{aligned}
e_1 &:= x_1 \\
e_2 &:= x_2 \\
e_3 &:= x_3 \\
e_4 &:= -x_1 - 3x_2 - 3x_3 + 10g \\
e_5 &:= -2x_1 - x_3 - x_3 + 8g \\
f &:= -3x_1 - 2x_3 - 4x_3 \\
g &:= g.
\end{aligned}
$$

These 7 linear functionals live in $(\mathbb{R}^4)^* \cong \mathbb{R}^4$, and so we have the following structure: The circuits defined on $E := \{e_1, e_2, e_3, e_4, e_5, f, g\}$ are collections of (vectorspace) vectors minimal with respect to linear dependence. The non-negative vectors which contain f represent a non-negative linear combination of the remaining functionals which sum to 0. These represent the feasible solutions to the dual problem. Conversely, the cocircuits represent regions of $\mathbb{R}^4$ which are maximal with respect to lying on the kernels of the defining functionals. A non-negative cocircuit which contains g represents a feasible primal solution.

The goal then is to find a optimal pair of feasible primal and dual solutions.

A more detailed discussion follows in Chapter 5.

1.5.5 Results about Optimal Solutions: The Simplex Algorithm

Finally, one of the most important results of linear programming is the development of Theorem **??** (the *Simplex Algorithm*) as a means of obtaining the primal and dual optimal solutions. It also allows constructive proofs of the Strong Duality, and the Fundamental Theorem of Linear Programming. We intend to generalize this result to ordered division rings, where the co-domain is finite dimensional.

1.6 Potential Difficulty in an Abstract Situation

In the classical setting of affine programming, many of the standard results were proved using the properties of the real numbers, including the existence of an ordering that is both Archimedean and least-upper bound closed, and a commutative product. For example, in a finite dimensional real vector space, a closed and bounded region of $\mathbb{R}^n$ is compact, and all linear functions are continuous. Thus in such a case we are guaranteed an optimal solution.

The proof of the Farkas Lemma originally used the least upper bound properties of the reals to construct a separating hyperplane, which can be extended to an infinite dimensional case by the *Hahn Banach separation theorem* [Con90]. However, a general ordered ring may not be Archimedean ordered, and thus may not admit any of these features. One then is required to find alternative methods of obtaining these classical results.

Similarly, the non-commutative multiplication in the ring of scalars provides potential barriers in finding the existence of duality gaps and application of the Simplex Algorithm. These are all issues that must be addressed in a study of Abstract Affine Programming.

1.7 Summary of Results

Our project is the generalization from the classical setting of primal-dual affine programming using the Tucker tableau format, to the general case still using the Tucker tableau format, but now with a emphasis on functions and a basis free viewpoint. We also wish to generalize the ring of scalars, to be an ordered ring. Moreover, we generalize to a possibly infinite dimensional case. Using a generalization of the poset structure in the classical case, we show that Weak Duality still holds in these cases (Proposition 2.4.8). Moreover, when the ring of scalars is an ordered ring, one can show that a generalized version of the Fundamental Theorem of Affine Programing holds (Proposition 2.4.23). We can also define a notion of convexity in this setting, and show that our feasible and optimal solutions satisfy this notion of convexity (Proposition 2.4.18, Corollaries 2.4.19, 2.4.20).

When the ring of scalars is not a division ring, we show that the duality gap can also exist (Proposition 2.5.1). When we restrict to the case where the ring of sclalars is a division ring, we exhibit a counter example to the remaining results (Example 4.3.5). We then describe a number of hypothesis under which the Farkas' Lemma does hold (Theorems 4.3.10, 4.3.17, 4.3.25), and under these hypothesis, the following results (Propositions 5.3.1, 5.3.2, Theorem

5.3.4, Theorem 5.3.7) also hold. We then show that if certain “finite-type” hypothesis hold, then a generalization of the Exsistence-Duality Theorem also holds (Theorem 6.4.4).

Finally, we note that under these “finite-type” conditions, we may place an oriented matroid structure on the program (Proposition 6.2.3). Moreover, if the image of the module homomorphism A is in fact finite dimensional, we may describe a generalization of the Simplex algorithm that successfully terminates (Theorem 7.5.2).

Chapter 2

Ordered Rings and Modules

2.1 Introduction

In this chapter we begin by describing the features of affine programming which depend only on a ring and module structure, namely the Tucker Key Equation and the Tucker Duality Equation. Afterwards, we introduce ordered rings and some properties of ordered rings, and describe some properties of modules over ordered rings. We then show that when we define an affine program in this ordered ring setting, we obtain some new facts. We will prove the Weak Duality and the convexity of feasible and optimal solution sets. We will also prove some partial results leading up to some of our other key facts. We finally exhibit counter-examples to generalizations of some other well-known facts. We will assume throughout that our rings will be unital.

2.2 General Rings

We may define the variables and equations of affine programming in a more general context. Rather than deal in finite dimensional real-vector spaces, we can generalize this idea to modules over a ring R. In order to do this, we must properly describe the appropriate features of these problems and their data.

If we allow R to be an arbitrary ring, rather than $\mathbb{R}$, then the natural analogues to $\mathbb{R}^n, \mathbb{R}^m$ would be a pair of (left) R-modules, X, Y^*, where $Y^* = \mathrm{Hom}_R(Y, R)$ for a (left) R-module Y. We will focus on the left structure of these objects, although there is no reason one cannot define the same concepts for right modules instead. In the traditional case, A is a $n \times m$ real-matrix, or equivalently $A \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m)$, and so the appropriate generalization would be to let A be an $\mathbb{R}$-module homomorphism, i.e. $A \in \mathrm{Hom}_R(X, Y^*)$. It is clear that $\vec{b} \in Y^*$ is an element of the co-domain of A, and $d \in R$ as before.

Recall that in the classical case, $\vec{c} \in \mathbb{R}^n$. However, when we consider the role that $\vec{c}$ plays in the original problem, $\vec{c}$ acted on $\vec{x}$ via inner product to obtain a value in $\mathbb{R}$. Hence, the appropriate generalization of $\vec{c}$ is $\mathbf{c} \in X^* := \mathrm{Hom}_R(X, R)$. Similarly, while $\vec{x}$ should remain an element of the domain of A, $\vec{x} \in X$, $\vec{y}$ is a vector that acts on $\vec{b}$ via an inner product, and so $\vec{y}$ should be generalized to $\mathbf{y} \in \mathrm{Hom}_R(Y^*, R) =: Y^{**}$.

We may then, define as before:

$$\begin{aligned}
f &: X \to R, \vec{x} \mapsto \mathbf{c}(\vec{x}) - d \\
g &: Y^{**} \to R, \mathbf{y} \mapsto \mathbf{y}(\vec{b}) - d. \\
\vec{t} &:= -A(\vec{x}) + \vec{b} \\
\mathbf{s} &:= \mathbf{y} \circ A - \mathbf{c}.
\end{aligned}$$

All of which can be captured in the generalized Tucker tableau:

$$
\begin{array}{c|c|c|l}
 & \vec{x} & \text{-1} & \\
\hline
\mathbf{y} & A & \vec{b} & = -\vec{t} \\
\hline
-1 & \mathbf{c} & \mathrm{d} & = f \\
\hline
 & \| & \| & \\
 & \mathbf{s} & g &
\end{array}
$$

$$
\begin{tikzcd}
X \arrow[rr, "A"'] \arrow[rr, squiggly, shift left, "\vec{x}\to-\vec{t}"] \arrow[drrr, dashed, "\mathbf{s}"] \arrow[drrr, "\mathbf{c}"'] \arrow[drrr, squiggly, bend right, "f"'] & & Y^{*} \arrow[rr, "R^{\square}"] \arrow[dr, Rightarrow, "\mathbf{y}"] & & X^{*} \arrow[dl, Rightarrow, "\vec{x}"'] & & Y^{**} \arrow[ll, "A^{\vdash}"] \arrow[ll, squiggly, shift right, "\mathbf{y}\to\mathbf{s}"'] \arrow[dlll, dashed, "-\vec{t}"'] \arrow[dlll, "\vec{b}"] \arrow[dlll, squiggly, bend left, "g"] \\
 & & & R & & &
\end{tikzcd}
$$

Figure 2.1: Affine Maps of a General Affine Program

With all of this data, we may now restate the original classical result:

Proposition 2.2.1 (Generalized Fact I: Tucker Key Equation)**.** *Given generalized Tucker tableau data, then* $\mathbf{s}(\vec{x}) - g(\mathbf{y}) = \mathbf{y}(-\vec{t}) - f(\vec{x})$.

Proof. Consider the computations:

$$\mathbf{s}(\vec{x}) - g(\mathbf{y}) = \mathbf{y} \circ A(\vec{x}) - \mathbf{c}(\vec{x}) - \mathbf{y}(\vec{b}) + d.$$

$\square$

Corollary 2.2.2 (Generalized Tucker Duality Equation)**.** *Given the generalized Tucker tableau* *data, then* $g(\mathbf{y}) - f(\vec{x}) = \mathbf{s}(\vec{x}) + \mathbf{y}(\vec{t})$.

2.3 Some Properties of Ordered Rings and Modules

2.3.1 Properties of Ordered Rings

From Corollary 2.2.2 the duality equation holds over variables and maps defined over any module over any ring. In particular, it holds in the classical case of finite dimensional real vector spaces. However, we cannot define the primal maximization and dual minimization problem in this situation. In order to define $\leq$, an inequality or poset structure, for elements in a ring R, we require that R must be a trichotomy-ordered ring. This is necessary for the notions of "maximization" and "minimization" to be meaningful.

Remark 2.3.1. We notice some facts about ordered rings.

- Since R is partitioned into positives, the additive inverses of the positives (the negatives) and 0, and moreover the positives are closed under sums, no repeated sum of any positive element is 0. In other words, the additive order of any positive element, and consequently any non-zero element, is ∞.
- Recall that given $a, b \in R\backslash\{0\}$, ab is positive iff a, b are the same sign. Else ab is negative. Consequently R is a domain, since the product of non-zero elements will be non-zero. In particular, if the ordered ring R is unital, then $\mathbb{Z} \hookrightarrow R$ in the natural universal way.
- Given $a, b, c \in R$, $c, a-b \in P$ (i. e. $a > b$), we have that $c(a-b), (a-b)c \in P$, so $ca > cb$ and $ac > bc$. Similarly $-cb > -ca, b(-c) > a(-c)$. We can also check $a+c > b+c$, and $a-c > b-c$.

Most of the ordered rings one generally encounters are the standard $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and the number rings $\mathbb{Q}[\sqrt[n_1]{a_1}, \ldots, \sqrt[n_m]{a_m}]$, where $a_i, n_i \in \mathbb{Z}_+$. However, we can construct some examples of more general ordered rings.

Example 2.3.2. Let $\mathbb{R}[x]$ be the ring of polynomials over $\mathbb{R}$, and let

$$P := \{p(x) : p(x) \text{has a positive leading coefficient}\}.$$

We first notice that polynomials with positive leading coefficients are closed under sums and products. Moreover, each polynomial either has a positive leading coefficient, or a negative leading coefficient, or is 0. Thus, P satisfies the conditions of being a set of positives.

The type of ordered rings one is used to working with are called *Archimedean* rings. These are rings which exhibit the Archimedean Principle.

Definition 2.3.3. Given an ordered ring R, $a \in R$, we define the *absolute value function* $|\cdot| : R \to R$ via:

$$|a| := \begin{cases} a, & a \in P \\ -a, & a \in -P \\ 0, & a = 0 \end{cases} \cdot$$

Definition 2.3.4 ([Lam01])**.** An ordered ring R is *Archimedean* if given any $r \in P \subset R$ and ring elements $a, b \in P \subset R$, such that $a < r < b$, then there are natural numbers n, m such that $n \cdot a > r$ and $m \cdot r > b$. Any ring that does not exhibit this behavior is called *non-Archimedean*. The set of elements where $\{r \in R \backslash \{0\} : n \cdot |r| < 1_R, \ \forall \ n \in \mathbb{N}\}$ are the *infinitesimals* of R. We will denote these with R_ϵ

Similarly, the set of elements $\{r \in R : n \cdot 1_r < |r|, \ \forall n \in \mathbb{N}\}$ are the *infinities* if R, and 1_R is the unit of R. We will denote these with R_∞. (Non-Infinite elements of an ordered ring are called *finite*.)

Note that only non-Archimedean rings contain infinities or infinitesimals.

Theorem 2.3.5 ([Lam01])**.** *Let R be an Archimedean ring. Then:*

- *R is commutative*
- *R is order isomorphic to a unique subring of $\mathbb{R}$.*
- *The only order preserving automorphism of R is the identity map.*

This shows that Archimedean ordered rings are in fact very familiar rings, both in terms of the order property, and in terms of the algebraic structure. The difficulty arrives when one deals with non-Archimedean rings. We next exhibit some examples of non-Archimedean ordered rings.

Example 2.3.6. Our prior example of an ordered ring, $\mathbb{R}[x]$ with the leading coefficient order is non-Archimedean. Consider $x^n, n > 0$. Given any $m \in \mathbb{N}$, $x^n - m$, has a positive leading coefficient. Thus $x^n > m$, for each choice of m and x^n is an infinite of R.

Example 2.3.7 (The Hyperreals)**.** Recall the Hyperreals $^*\mathbb{R}$, Example 1.3.3.

We define the positives of $^*\mathbb{R}$ to be the classes of sequences which are positive "almost everywhere". Formally, $P := \{[a] \in {}^*\mathbb{R} : \{n : a(n) > 0\} \in U\}$. We then notice that by the compliment properties of an ultrafilter, each sequence must be equal to, greater than, or less than zero, almost everywhere.

The real numbers naturally inject into $^*\mathbb{R}$ by $r \in \mathbb{R} \mapsto [(r, r, r, \ldots)]$, which we can write as $[r]$ for convenience. It is clear that such an injection is an order preserving isomorphism. One can then consider the class of a sequence

$$[a] := [(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots)].$$

Given any positive real number r, the entries of $[r]$ will be greater than the entries of $[n][a]$, where n is a positive integer, almost everywhere. Thus $[a]$ is an infinitesimal in $^*\mathbb{R}$. Similarly the class

$$[b] = [(1, 2, 3, \ldots, n, \ldots)]$$

will be greater than any real number almost everywhere and $[b]$ is an infinite of ${}^*\mathbb{R}$.

Example 2.3.8. Consider the ring of polynomials in several indeterminants over $\mathbb{R}$, $\mathbb{R}[x_1, \ldots, x_n]$, with a lexicographical ordering, $x_1 > x_2 > \ldots > x_n$. Given two monomials, $q_1 = \prod_{i=1}^n x_i^{s_i}$, $q_2 = \prod_{i=1}^n x_i^{t_i}$, we say $q_1 > q_2$ if there is an index j such that $s_j > t_j$ and $s_i = t_i$ for all $i < j$.

We define the set of positives to be the collection of polynomials whose leading monomials have positive coefficients (leading with respect to the above ordering). This is a well defined set of positives: Since each non-zero polynomial has either a positive or negative leading coefficient, this partitions the ring into positives, negatives and 0. It is also clear that polynomials with positive leading coefficient are closed under sums and products.

We will see that non-Archimedean rings have somewhat strange properties which are potentially inconvenient for linear programming.

Example 2.3.9. Consider ${}^*\mathbb{R}$. Note that the set $[0, 1]$ is compact in $\mathbb{R}$ under its usual topology. However, if we equip ${}^*\mathbb{R}$ with the order topology, i.e. with a basis $\mathcal{B} := \{(a, b) : a, b \in {}^*\mathbb{R}\}$, then $[0, 1] \subseteq {}^*\mathbb{R}$ is no longer compact.

To see this, we note that every finite hyperreal number h has a *standard part*, a unique real number r such that $h = r + \iota$, where ι is either zero or an infinitesimal. Let ϵ be a positive infinitesimal, and consider the covering of $[0, 1]$, $C := \{(a - \epsilon, a + \epsilon) : a \in [0, 1]\}$. Notice then that each $(a - \epsilon, a + \epsilon)$ contains at most 1 real number. Thus no finite sub-covering of C will cover $[0, 1]$

In particular, the *shadow function*, $r + \iota \mapsto \iota$ is continuous on $[0, 1]$, but does not attain a maximum value.

Moreover, a non-Archimedean ring need not be commutative.

Example 2.3.10. Consider the ring $R := \mathbb{R}\langle x, y\rangle / \langle yx - 2xy\rangle$, i. e. the ring of polynomials over x, y with the skew-product $yx = 2xy$. We can establish a "lexicographical" ordering of the monomials of R, with $y^{n_1}x^{m_1} > y^{n_2}x^{m_2}$ if $n_1 > n_2$ or $n_1 = n_2, m_1 > m_2$. Then, we can define the positives P of R to be the polynomials with positive leading coefficient, where leading means greatest with respect to the above lexicographical ordering.

One can easily see that P satisfies the properties of a collection of positives, and R is non-commutative. This is contrary to our intuition when it comes to a ring with order, as we expect an ordered ring to be commutative. However, the order conditions force some regularity conditions between products ab and ba

In particular, David Hilbert exhibited the first example of a non-commutative division ring in 1903 [Hil03].

Example 2.3.11 (A Non-Commutative Ordered Division Ring)**.** Consider $R' := \mathbb{R}((x))$, be the Laurent series of x over $\mathbb{R}$. We can define a set of positives P' to be the collection of all series with positive leading coefficient. Then, define $R := R((y))$ to be the Laurent series of y over R', with the skew product $xy = 2yx$. We then give the monomials the lexicographical ordering where $y > x$ as before, and define the positives of R to be P, the collection of series with positive leading coefficient. As before we see that R is an ordered ring.

Thus, it suffices to show that this R is a division ring. We first show that R' is a division ring. Let $a \in R'$. We can write $a = a'x^m$, where $m \in \mathbb{Z}$ and $a' = \sum_{i=0}^{\infty} a_i x^i, a_i \in \mathbb{R}, a_0 \neq 0$. So without loss of generality, let $a = a'$. We can then define b such that $b = \sum_{j=0}^{\infty} b_j x^j$ where:

$$\begin{aligned} b_0 &= a_0^{-1} \\ b_{j+1} &= -a_j^{-1} \sum_{k=1}^{k} a_k b_{j-k}. \end{aligned}$$

Thus, ab may be written as:

$$\begin{aligned} ab &= \sum_{\ell=0}^{\infty} \sum_{i+j=\ell} a_i b_j x^\ell \\ &= 1 + \sum_{\ell=1}^{\infty} \left(-\sum_{i=1}^{\ell} = a_i b_{\ell-i} + \sum_{j=1}^{\ell} a_j b_{\ell-j} \right) x^\ell = 1. \end{aligned}$$

Similarly, given $a \in R$, we can define $b \in R$, such that $ab = 1$. Thus R is a division ring.

Definition 2.3.12. Let R be a ring. Then the *center* of a ring, denoted $Z(R)$, is the collection of elements $\{z \in R : az = za, \ \forall\ a \in R\}$.

Proposition 2.3.13. *Let R be an ordered unital ring with set of positives P. Let $a, b \in P$, such that, without loss of generality, $ab \leq ba$. Then there is no element $z \in Z(R)$ such that $ab < z < ba$.*

Proof. Suppose that this were not true, then $aba < za = az < aba$, a contradiction. $\square$

Corollary 2.3.14. *If R is a unital ring, and there are $a, b \in P$ such that ab is finite and $ba \geq ab$, then $ba = ab + \epsilon$, where ϵ is zero or infinitesimal.*

Proof. If ba were infinite, then since ab is finite, there is an $m \in \mathbb{Z}_+$ such $ab < m < ba$, which contradicts Proposition 2.3.13, (recall that $Z \hookrightarrow Z(R)$). Thus $ba - ab$ is finite. If this difference is not infinitesimal or zero, then there is a $n \in \mathbb{Z}_+$ such that $nba - nab > 1$. Since nab is finite, we may find an integer $m_1 \leq nab$ and moreover, we may find a maximum such integer (else nba would be infinite.)

There is an integer z in between $nab < z < nba$, since, $m_1 + 1 > nab$ but $m_1 + 1 < nba$. Since z is an integer, it is central, and by Proposition 2.3.13, this is a contradiction. $\square$

2.3.2 Modules and Cones over Ordered Rings

Now that we have established a notion of order for the ground ring R, we want to establish a notion of partial order for the modules X, Y^*. In the classical case, $A(\vec{x}) \leq \vec{b}$ meant each entry of $\vec{b} - A(\vec{x})$ was non-negative. Thus the collection of vectors in $\mathbb{R}^n$ with non-negative entries plays a similar role to the non-negatives of a positive ring. However this set of "positives" , along with its additive inverses and $\{0\}$ do not partition $\mathbb{R}^n$ the way the positives, its additive inverses and $\{0\}$ partition an ordered ring. The analogous structure in a module is a positive (or non-negative) cone. By convention, we also assume that R will be an ordered unital ring. In particular, this means that $\mathbb{Z} \hookrightarrow Z(R)$ for each choice of R.

Recall the definition of a *cone* of a module over an ordered ring:

Definition 2.3.15. A *positive cone* of a module X over an ordered ring R is a subset $C \subset X$ such that given $r, s \in R, r, s \geq 0$, and $a, b \in C$, $r \cdot a + s \cdot b \in C$.

We also define the following specialized types of cones (C):

- If C contains no non-trivial sub-module of X, then C is said to be a *pointed* cone [Zie13].
- If there is a $x \in X \backslash C$ and $r \in R, r \geq 0$, such that $r \cdot x \in C$, then we say that C is *perforated* [Fuc11]. Otherwise C is unperforated.

We can verify easily that in $\mathbb{R}^n$, the collection of vectors with non-negative entries satisfy the definition of a (unperforated) positive cone.

Definition 2.3.16. Let R be a ring, and M a (left) R-module. We say the *torsion sub-module* of M is the set $\{m \in M : ra = 0, r \in R\backslash\{0\}\}$, we then denote this set $\operatorname{tor}(M)$.

Proposition 2.3.17. *If R is an ordered ring, and X is a R-module with a positive cone C_X, then* $\operatorname{tor}(X) \cap C_X = \{0\}$.

Proof. Let $a \in \operatorname{tor}(X) \cap C_X$. If $a \neq 0$, then there is a minimal natural number n such that $n \cdot a = 0$. Consider any natural numbers n_1, n_2 such that $n_1 + n_2 = n$. We see that

$$\sum_{i=1}^{n_1} a \in C_X, \sum_{i=1}^{n_1} a = -\sum_{j=1}^{n_2} a = \sum_{j=1}^{n_2} (-a).$$

This contradicts C_X been a cone, and thus $a = 0$. □

In the classical case of linear programming, we wanted all of our vectors, and the differences between vectors, to be in their respective non-negative cones. By selecting a cone for both X, Y^*, we then selected appropriate cones for our programs. However, in the classical case, X, Y^* are finite dimensional real spaces, so the space of their duals X^*, Y^{**} are isomorphic (though not naturally) to the original spaces, and thus we may use the same cones for all 4 spaces. This may not be true generally. Thus, we want to choose cones for X, Y^*, that naturally give rise to cones in X^*, Y^{**}.

Definition 2.3.18. Let R be an ordered ring, and X be a left R-module. We call a cone of X, C_X, a *full cone* if C_X is a generating set for X. (That is, given $x \in X$, we can write $x = \sum_{i=1}^{n} r_i \cdot c_i$, where $r_i \in R, c_i \in C$, or in generating set notation, $\langle C_X \rangle = X$).

Proposition 2.3.19. *Given an ordered ring R, and X a module over R with a pointed cone C_X. Then C_X is a full cone if and only if the set $C_{X^*} := \{\varphi \in X^* : \varphi(v) \geq 0, \ \forall\ v \in C_X\}$ is a unperforated pointed positive cone of X^**

Proof. We first assume that C_X is an unperforated full cone. We verify the 3 properties of unperforated pointed cones.

- Given $\varphi, \psi \in C_{X^*}, \lambda, \mu \in R_+ := P$, and $v \in C_X$, $(\lambda\varphi + \mu\psi)(v) = \lambda\varphi(v) + \mu\psi(v)$. Both $\varphi(v), \psi(v) \geq 0$, by construction of C_{X^*}. Thus $\lambda\varphi(v), \mu\psi(v) \geq 0$, and so is their sum. We conclude that $(\lambda\varphi + \mu\psi) \in C_{X^*}$.

- Let $\varphi \in C_{X^*}$. Let $S \subset C_X$ be a minimal spanning, or generating, set of X and let $s \in S$. If $\varphi(v) = -\varphi(v)$ for $\varphi \in C_{X^*}, v \in C_X$, then $\varphi(s) = -\varphi(s) = 0$ for each $s \in S$, and thus φ must be the zero map.

- Given $\psi \notin C_{X^*}$, there is a $v \in C_X$ such that $\psi(v) < 0$. Then given $\lambda > 0$, $\lambda\psi(v) < 0$, and so $\lambda\psi \notin C_{X^*}$.

Conversely, if C_{X^*} is an unperforated cone, then consider $s \in C_X$. If $\{n \cdot s : n \in \mathbb{Z}\} \subsetneq C_X$, then we may find $s' \in C_X$ such that the generating sets: $\langle s \rangle \subsetneq \langle s, s' \rangle$. Thus, we may use Zorn's Lemma to extend $\{s\}$ to a maximal set S, $S \subseteq C_X$, and $\{n \cdot s : n \in \mathbb{Z}_{\oplus}, s \in S\} = C_Y$. If $\langle S \rangle \neq X$, then there is a $\varphi \in X^*$ such that $\varphi \neq 0$, but $\langle S \rangle \subseteq \mathrm{Ker}(\varphi)$. Thus $\langle S \rangle = Y$ and C_X is a full cone. $\square$

Example 2.3.20. Notice that the traditional positive cones in classical programming are full cones. Let $X = \mathbb{R}^n$, and consider $C_X := \{\vec{x} \in X : x_i \geq 0\}$, i. e. C_X is the set of vectors with non-negative entries.

The vectors $(1, 0, 0 \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, 0, \ldots, 0, 1)$ form a basis (i.e. a generating set) for $\mathbb{R}^n$, and each one is in C_X.

Proposition 2.3.21. *Let R be an ordered ring, X an R-module, and C_X be a cone in X. Then let ι be the inclusion map $\iota : C_X \to X$. Then if the following diagram commutes if and only if $\varphi = 0$, then C_X is a full cone.*

$$\begin{tikzcd} C_X \arrow[r, hook, "\iota"] \arrow[dr, dotted, "0"'] & X \arrow[d, "\varphi"] \\ & R \end{tikzcd}$$

Similarly, if C_X is a full cone, then this diagram commutes if and only if $\varphi = 0$.

Proof. If there is a φ such that the above diagram commutes but $\varphi \neq 0$, then since $\varphi(C_X) = 0$

but there is a $x' \in X$ such that $\varphi(x') \neq 0$, it follows that x' cannot be written as a linear combination of elements of C_X, and thus C_X is not full.

If C_X is a full cone, then for any $\varphi \neq 0$, there is a $x' \in X$ such that $\varphi(x') \neq 0$. Since x' is a linear combination of elements of C_X, it follows that $\varphi(C_X) \neq 0$. Thus the above diagram does not commute. □

2.4 Results with Cones and Programs

2.4.1 Feasibility

With cones for $X, Y^*, X^* = \mathrm{Hom}_R(X, R), Y^{**} := \mathrm{Hom}_R(Y^*, R)$ defined, we may reintroduce the notion of feasibility.

Definition 2.4.1. Let C_X be a full cone with C_{X^*} the associated dual cone. Define $\leq_X, \leq_{X^*}$

- If $a, b \in X$ such that $b - a \in C_X$, then $a \leq_X b$.
- If $a^*, b^* \in C_{X^*}$ such that $b^* - a^* \in C_{X^*}$, then $a^* \leq_{X^*} b^*$.

Similarly define $C_{Y^*}, C_{Y^{**}}$for Y^*, Y^{**}

Definition 2.4.2. Given R an ordered ring, X, Y left R-modules. Let $C_X \subset X, C_{Y^*} \subset Y^*$, be full cones, with associated dual cones $C_{X^*}, C_{Y^{**}}$. Then given $A \in \mathrm{Hom}_R(X, Y^*)$, $\vec{b} \in Y^*, \mathbf{c} \in X^*, d \in R$ we may define the following:

- A primal decision variable $\vec{x} \in X$ is *primal canonically feasible* if $-A(\vec{x}) + \vec{b} \in C_{Y^*}$ and if $\vec{x} \in C_X$.

- Similarly, a dual decision variable $\mathbf{y} \in Y^{**}$ is *dual canonically feasible* if $\mathbf{y} \circ A - \mathbf{c} \in C_{X^*}$, and if $\mathbf{y} \in C_{Y^{**}}$.

This definition of feasible coincides with the classical notion of feasible. However, there is a way to generalize the classical notion to a potentially more powerful version of feasibility:

Definition 2.4.3. Given R an ordered ring, X, Y^* left R-modules, $A \in \mathrm{Hom}_R(X, Y^*)$, $\vec{b} \in Y^*, \mathbf{c} \in X^*, d \in R$, with the conditions that there are non-perforated full cones $C_X \subset X, C_Y \subset Y$, and appropriate dual cones C_{X^*}, C_{Y^*} we define the following:

- A variable $\vec{x} \in X$ is *primal feasible* if $-A(\vec{x}) + \vec{b} \in C_Y$.
- Consequently, $\mathbf{y} \in Y^{**}$ is *dual feasible* if $\mathbf{y}' \circ A = \mathbf{c}$.

Notation 2.4.4. Given a ring (or module) $\Box$ with a set of non-negatives (positive cone), we denote $\Box_{\oplus}$ to be the collection of non-negatives in that ring (module).

This notation is due to Albert Tucker [NT93].

Remark 2.4.5. Let R be an ordered ring, let X, Y^* be left R-modules with non-negative trichotomy cones C_X, C_{Y^*}, $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$.

We then define $Y' := Y^* \oplus X$, let $C_{Y'} := C_{Y^*} \oplus C_X$, $A' := A \oplus -I_X \in \mathrm{Hom}_R(X \oplus X, Y')$, $\vec{b}' := (\vec{b}, 0)$. This information defines a second program, namely:

Any $\vec{x}$ is canonically feasible for the first program if and only if it is feasible for the second. Similarly a $\mathbf{y}$ is canonically feasible for the first program if and only if there is a $\mathbf{y}'$ feasible for the second program such that $\mathbf{y}'|_{Y^*} = \mathbf{y}$.

To see this, we let $\vec{x}$ be canonically feasible for the first program, then $A(\vec{x}) \leq_{Y^*} \vec{b}$ and $\vec{x} \geq_X 0_X$, which happens if and only if $A'(\vec{x}) \leq_{Y'} \vec{b}'$.

Similarly, $\mathbf{y} \circ A \geq_{X^*} \mathbf{c}$, if and only if $\mathbf{s} := \mathbf{y} \circ A - \mathbf{c} \in C_{X^*}$, which occurs if and only if $\mathbf{y}' := \mathbf{y} \oplus \mathbf{s}$ is feasible.

This generalization of feasibility has several advantages, first of which is that we only compare vectors with an inequality on Y. Thus, we only need to define a full cone for Y^*, and not for X. Another advantage is that the feasible solutions to the dual problem have an equality constraint, $\mathbf{y} \circ A = \mathbf{c}$. This makes computing the dual solutions, and determining feasibility, much easier.

With these ideas in mind, we will now formally define a primal-dual program.

Definition 2.4.6. A program $\mathscr{P}$ is a ordered septuple $\mathscr{P} := (R, X, Y^*, C_{Y^*}, A, \vec{b}, \mathbf{c}, d)$, where R is an (ordered) ring, X, Y^* are R modules, C_{Y^*} is a full cone for Y, $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R), d \in R$.

Then $\mathscr{P}$ induces a *primal problem*: Find $\vec{x} \in X$ such that $A(\vec{x}) \leq \vec{b}$ and $f(\vec{x}) = \mathbf{c}(\vec{x}) - d$ is minimized.

The program $\mathscr{P}$ also induces a *dual problem*: Find $\mathbf{y} \in Y^{**}_{\oplus}$ (which is what we previously called $C_{Y^{**}}$) such that $\mathbf{y} \circ A = \mathbf{c}$ and $g(\mathbf{y}) = \mathbf{y}(\vec{b}) - d$ is minimized.

We remark here that if R, X, Y^*, C_{Y^*} are understood, these entries may be suppressed for brevities sake. Additionally, for this paper, we will generally also suppress the use of d, where it is understood without loss of generality that $d = 0$.

Definition 2.4.7. Given a program $\mathscr{P}$, a problem is considered *feasible* if there are feasible solutions. Otherwise, the problem is *infeasible.*

2.4.2 Weak Duality

Since we have established the appropriate hypothesis to discuss ordered ring programming, we would like to know which of the classical duality results hold. The Tucker Duality Equation holds for any choice of rings and modules (not necessarily ordered), so it will also hold in this setting. Using the concepts and definitions delineated above, we may introduce a consequence of the Tucker Duality Equation, namely Weak Duality. We show that the weak duality holds under both of our notions of feasibility.

Proposition 2.4.8. *Given $R, X, Y^*, A, \vec{b}, \mathbf{c}, d, C_X, C_{Y^*}$ as above, where $\vec{x}, \mathbf{y}$ are both canonical and feasible, then $g(\mathbf{y}) \geq f(\vec{x})$.*

Proof. By the duality equation, $g(\mathbf{y}) - f(\vec{x}) = \mathbf{s}(\vec{x}) + \mathbf{y}(\vec{t})$. Since $\vec{x}$ is feasible, $\vec{t} \in C_{Y^*}$, and since $\mathbf{y}$ is feasible, $\mathbf{s} \in C_{X^*}$. Then $\mathbf{s}(\vec{x}) \geq_R 0$ since $\mathbf{s} \in C_{X^*}$. Similarly, $\mathbf{y}(\vec{t}) \geq_R 0$.

Thus $g(\mathbf{y}) - f(\vec{x}) \geq 0$, and therefore $g(\mathbf{y}) \geq f(\vec{x})$. □

Theorem 2.4.9. *Given $R, X, Y^*, A, \vec{b}, \mathbf{c}, d, C_X, C_{Y^*}$ as above. We require $A(\vec{x}) \leq_{Y^*} \vec{b}$, but $\vec{x}$ may or may not be in C_X. Additionally, we require that $\mathbf{y} \in C_{Y^{**}}$ and $\mathbf{y} \circ A = \mathbf{c}$. Then $g(\mathbf{y}) \geq f(\vec{x})$.*

Proof. Since $\mathbf{y} \circ A = \mathbf{c}$, we conclude that $\mathbf{s} = 0$. However, since $A(\vec{x}) \leq_{Y^*} \vec{b}$, we see that $\vec{t} \in C_{Y^*}$. Thus

$$f(\vec{x}) - g(\mathbf{y}) = \mathbf{s}(\vec{x}) + \mathbf{y}(\vec{t}) = 0 + \mathbf{y}(\vec{t}) \geq 0.$$

□

Corollary 2.4.10 (Generalized Fact II: Weak Duality)**.** *If $\vec{x}, \mathbf{y}$ are such that $\mathbf{y}(\vec{b}) = \mathbf{c}(\vec{x})$, then $\mathbf{y}, \vec{x}$ are optimal solutions for their respective problems.*

Naturally, we would like to know whether or not the converse of this statement is true, that is, if $\vec{x}^*, \mathbf{y}^*$ are optimal solutions to their respective problems, that $g(\mathbf{y}^*) = f(\vec{x}^*)$. However, we see that this is not generally the case.

Example 2.4.11. Let $R = X = Y^* = \mathbb{Z}$. Let $A = 2, \vec{b} = \vec{c} = 1, d = 0$. We see that the optimal solution for the primal program is $\vec{x} = 0$, where $f(\vec{x}) = 0$, but the optimal solution to the dual is $\vec{y} = 1$, with $g(\vec{y}) = 1$.

2.4.3 Convexity

In this setting, one may also describe a notion of convexity, and describe our feasible and optimal solutions in terms of convexity.

Definition 2.4.12. Let R be an ordered ring. Let X be a left R module, a *left skeletal line-segment* of X with *end points* $\vec{x}, \vec{y} \in X$ is the set $\{\lambda \cdot \vec{x} + (1-\lambda) \cdot \vec{y} : \lambda \in [0,1] \subseteq R\} \subset X$,where $[0,1] := \{r \in R : 0 \leq r \leq 1\}$.

The *open left skeletal line-segment* with end points $\vec{x}, \vec{y}$ then is the set $\{\lambda \cdot \vec{x} + (1-\lambda) \cdot \vec{y} : \lambda \in]0,1[\subseteq R\} \subset X$, where $]0,1[:= \{r \in R : 0 < r < 1\}$.

Definition 2.4.13. Let R be an ordered ring. Let X be a left R module. A *left full line segment* with endpoints $\vec{x}, \vec{y}$ denoted $[[\vec{x}, \vec{y}]]$ is the set

$$\bigcup_{\substack{r_i \in R_+, a_i \in X \\ \vec{y} = \vec{x} + r_i \cdot a_i}} \{\vec{x} + (r_i - \lambda) \cdot a_i : \lambda \in [0, r_i]\}.$$

The *open left full line segment* with endpoints $\vec{x}, \vec{y}$ denoted $]]\vec{x}, \vec{y}[[$ is the set

$$\bigcup_{\substack{r_i \in R_+, a_i \in X \\ \vec{y} = \vec{x} + r_i \cdot a_i}} \{\vec{x} + (r_i - \lambda) \cdot a_i : \lambda \in]0, r_i[\}.$$

Proposition 2.4.14. *If R is a division ring, and X is a R left vector space, then the left* *skeletal line segment with end points $\vec{x}, \vec{y}$ and the full line segment with the same endpoints, are the same set.*

Proof. We begin by noting that the skeletal line segment is always a subset of the full line segment, as $\vec{y} - \vec{x}, 1$ are valid choices for a_i, r_i.

Conversely, given any choice a_j, r_j such that $\vec{y} = \vec{x} + r_j \cdot a_j$, we note that given $\lambda \in [0, r_j]$, $r_j^{-1}\lambda \in [0,1]$. Thus the full line segment is contained in the skeletal line segment. $\square$

Example 2.4.15. Let $R = \mathbb{Z}$ and $X = \mathbb{Z}^2$. Then let $\vec{x} = (1,0)$ and $\vec{y} = (5,0)$. Since $[0,1] = \{0,1\}$ the skeletal line segment with end points $(1,0),(5,0)$ would just be $\{(1,0),(5,0)\}$. However, this does not coincide with what our notion of a "line-segment" should be.

Instead, when we consider the full line segment, we see that one may write $(5,0)$ as $(1,0) + 2 \cdot (2,0)$ or $(1,0) + 4 \cdot (1,0)$. Thus the full line segment is:

$$\begin{aligned} & \{(1,0),(5,0)\} \cup \{(1,0),(3,0),(5,0)\} \cup \{(1,0),(2,0),(3,0),(4,0),(5,0)\} \\ = \;& \{(1,0),(2,0),(3,0),(4,0),(5,0)\}, \end{aligned}$$

and the open full line segment is:

$$\{(2,0),(3,0),(4,0)\}$$

Definition 2.4.16. Consider the ring $R := \mathbb{R}[x]$ as an ordered ring, where the positives are the polynomials with positive leading coefficient. Consider the line segment $[0,x]$. Notice that a valid choice for r_i, a_i is $x, 1$. Thus, for any $\lambda \in [0,x]$, we have that $0 + \lambda \cdot 1 \in [[0,x]]$.

In particular, $[[0,x]] = \mathbb{R}_{\oplus} \cup \{ax - b : a, b \in \mathbb{R}, a \in (0,1], b \geq 0\}$. The elements $ax - b$ are greater than any real number in the ordering of R, and thus the line segment $[[0,x]]$ cannot be contained in $\mathbb{R}$.

Definition 2.4.17. Let R be an ordered ring. Given a left R-module X, a set $C \subseteq Y$ is *left convex* if given $\vec{x}, \vec{y} \in C$, then C contains the full line segment with endpoints $\vec{x}, \vec{y}$.

We now describe the relevance these results have to our study.

Proposition 2.4.18. *Let X, Y^* be left R modules, with positive cones C_X, C_{Y^*}, and $A \in \mathrm{Hom}_R(X, Y^*)$, then*

1. *The cones C_X, C_{Y^*} are left convex.*
2. *The pre-image of a left convex set under A is left convex.*
3. *The intersection of left convex sets is left convex.*

Proof. Let a $r \in R_\oplus$, $\lambda \in [0, r]$.

1. Let $\vec{x}, \vec{y} \in X$ with $\vec{y} = \vec{x} + r \cdot \vec{a}$. Since $r - \lambda$ is non-negative, $\vec{x} + (r - \lambda) \cdot \vec{y} \in C_X$ for each $\vec{x}, \vec{y} \in C_X$.
2. Let $\vec{x}_1, \vec{y}_1 \in A^{-1}(C)$, where $C \subseteq Y^*$ is a left convex set. Suppose that $\vec{y}_1 = \vec{x}_1 + r \cdot \vec{a}$. Then there are $\vec{x}_2, \vec{y}_2 \in C$ such that $A(\vec{x}_1) = \vec{x}_2, A(\vec{y}_1) = \vec{y}_2$. Then $A(\vec{x}_1 + (r - \lambda) \cdot \vec{a}) = \vec{x}_2 + (r - \lambda) \cdot A(\vec{a})$. Since $\vec{y}_2 = \vec{x}_2 + r \cdot A(\vec{a})$, this completes this argument.
3. Let C, D be convex sets $\vec{x}_1, \vec{x}_2 \in C \cap D$, then $\lambda \cdot \vec{x}_1 + (r - \lambda) \cdot \vec{x}_2 \in C$ and D, and so is in $C \cap D$.

$\square$

Corollary 2.4.19 (Generalized Fact III Part 1: Convexity of feasible regions)**.** *The feasible regions (primal and dual) for a canonical, or non-canonical linear programming problem is left convex.*

Proof. The primal feasible region for a canonical program is the intersection of the positive cone C_X and the pre-image of $\vec{b} - C_{Y^*}$. Since each cone is convex, the pre-images of convex sets are convex and the intersection of convex sets is convex, it suffices to show that $\vec{b} - C_{Y^*}$ is convex.

To see this, consider that given $\vec{b} - \vec{w}, \vec{b} - \vec{v} \in \vec{b} - C_{Y^*}$, any choice $r \in R_+$ $\vec{a} \in Y^*$ such that $\vec{b} - \vec{w} + r \cdot \vec{a} = \vec{b} - \vec{v}$, satisfies $\vec{v} = \vec{w} + r \cdot \vec{a}$. Since C_{Y^*} contains the line segment with end points $\vec{v}, \vec{w}$, $\vec{b} - C_{Y^*}$ contains the line segment with end points $\vec{b} - \vec{w}, \vec{b} - \vec{v}$.

We also notice that $A^{-1}(\vec{b} - C_{Y^*})$ is the non-canoncial primal feasible region and so is also left convex.

For the dual feasible of the canonical program, the region is the intersection of C_{Y^*} and the pre image under $A^{\vdash}$ of $\mathbf{c} - C_{X^*}$. As before, this intersection is left convex. For the non-canonical program, the feasible region is the intersection of C_{Y^*} and the pre-image of $\mathbf{c}$ under $A^{\top}$, which is again left convex. □

Corollary 2.4.20 (Generalized Fact III Part 2: Convexity of optimal solutions)**.** *The optimal solutions of the primal and dual program form a left convex set.*

Proof. Given a primal-dual affine program, let $\vec{x} \in X$ be an optimal solution, with objective value $o := \mathbf{c}(\vec{x})$. Then the collection of all optimizers are all feasible points that obtain the value o. That is, the intersection of the feasible region and $\mathbf{c}^{-1}(o)$. Thus the set of all optimizers for the primal problem is left convex.

Similarly, the set of optimizers for the dual problem is also left convex. □

Definition 2.4.21 (Extreme Point)**.** Given a left R-module Y, An *extreme point* p of a convex set $C \subset Y$ is a point that is not contained in any open full line segment contained in C.

Theorem 2.4.22 (Generalized Fact IV: Fundamental Theorem of Affine Programming)**.** *Let*

R be an ordered ring, X a left R-module and $C \subseteq X$ a convex set. Let $\mathbf{c} \in \mathrm{Hom}(X, R)$ and $\vec{x} \in C$ be a maximizer (minimizer) for $\mathbf{c}$. If $\vec{x}$ is contained in any open generalized line segment contained in C, then each element of those line segments is a maximizer (minimizer) of $\mathbf{c}$ as well.

Proof. Let $\vec{v}_1, \vec{v}_2 \in C$, and let $r \in R_{\oplus}, a \in Y$ such that $\vec{v}_2 = \vec{v}_1 + r \cdot \vec{a}$. Then let $\vec{x} = \vec{y}_1 + (r - \lambda) \cdot \vec{y}_2, \lambda \in (0, r)$. Suppose there were an $\vec{x}' = \vec{y}_1 + (r - \lambda') \cdot \vec{a}, \lambda' \in [0, r]$ such that $\mathbf{c}(\vec{x}') < \mathbf{c}(\vec{x})$. Without loss of generality, let $\lambda' < \lambda$. It then follows that:

$$\begin{aligned}
\mathbf{c}(\vec{x}) - \mathbf{c}(\vec{x}') &> 0 \\
\mathbf{c}(\vec{y}_1 + (r - \lambda) \cdot \vec{a}) - \mathbf{c}(\vec{y}_1 + (r - \lambda') \cdot \vec{a}) &> 0 \\
(\lambda' - \lambda)\mathbf{c}(\vec{a}) &> 0 \\
\mathbf{c}(\vec{a}) &< 0.
\end{aligned}$$

Since $\lambda < r$, then $\mathbf{c}(\vec{y}_1 + (r - \lambda) \cdot \vec{a}) = \mathbf{c}(\vec{y}_1) + (r - \lambda)\mathbf{c}(\vec{a}) < \mathbf{c}(\vec{y}_1)$, contradicting the maximality of $\vec{x}$. Since no interior point of a skeletal line segment may be maximal, and each point in the interior of a full line segment is contained in a skeletal line segment, no interior point of a full line segment may be considered maximal either. $\square$

Proposition 2.4.23. *Let R be an ordered division ring. Let Y be an R vector space, and let $\hat{f}_1, \ldots \hat{f}_n$ be a collection of affine functionals (we may write $\hat{f}_i := f_i + b_i$, where $f_i \in \mathrm{Hom}_R(Y, R)$ and $b_i \in R$). Let $g \in \mathrm{Hom}_R(Y, R)$. Given a point $\vec{v} \in Y$ such that $\hat{f}_i(\vec{v}) > 0$ for each $1 \leq i \leq n$, $\vec{v}$ is an optimizer of g if and only if $g = 0$.*

Proof. Suppose $g \neq 0$, then there is a $\vec{y}$ such that $g(\vec{y}) = 1$. Let $c_i := \hat{f}_i(\vec{y})$. Then define $c \in R$

such that $c > 0$ and $c < \min_{1 \leq i \leq n} c_i |\hat{f}_i(\vec{y})|^{-1}$. We then notice that:

$$\begin{aligned} |\hat{f}_i(c \cdot \vec{y})| &= c|\hat{f}_i(\vec{y})| \\ &< c_i |\hat{f}_i(\vec{y})|^{-1} |\hat{f}_i(\vec{y})| \\ &= c_i. \end{aligned}$$

Thus

$$\begin{aligned} \hat{f}_i(\vec{v} + c \cdot \vec{y}) &\leq \hat{f}_i(\vec{v}) - |\hat{f}_i(c \cdot \vec{y})| \\ &= c_i - |\hat{f}_i(c \cdot \vec{y})| > 0. \end{aligned}$$

This shows that $\vec{v} + c \cdot \vec{y}$ is also non-negative when evaluated by each $\hat{f}_i$. But $g(\vec{v} + c \cdot \vec{y}) = g(\vec{v}) + c > g(\vec{v})$. Thus $\vec{v}$ cannot maximize g, and we can use a symmetric argument to show that $\vec{v}$ does not minimize g.

□

Theorem 2.4.24. *Let R be an ordered division ring. Let X be an R vector space, and let $\hat{f}_1, \ldots \hat{f}_n$ be a collection of affine functionals (we may write $\hat{f}_i := f_i + b_i$, where $f_i \in \mathrm{Hom}_R(Y, R)$ and $b_i \in R$). Let $g \in \mathrm{Hom}(X, R)$. If there is a $\vec{v} \in X$ such that $f_i(\vec{v}) \geq 0$ such that $\vec{v}$ optimizes (without loss of generality maximizes) $g \in \mathrm{Hom}_R(X, R)$, then there is a $\vec{v}'$ that optimizes g such that $\vec{v}'$ is contained in a maximal intersection of $\mathrm{Ker}(\hat{f}_i)$, and $\hat{f}_i(\vec{v}') \geq 0$.*

Proof. We proceed via induction on n. Suppose $n = 1$, then by Proposition 2.4.23, no point $\vec{v}, \hat{f}_1(\vec{v}) \geq 0$ may be an maximizer of g if $\hat{f}_i(\vec{v}) > 0$. Thus if $\vec{v}$ is an maximizer of g, then $\vec{v} \in \mathrm{Ker}(\hat{f}_1)$ and is contained in a maximal intercession of $\hat{f}_i$.

Otherwise suppose that this is true for a collection of $n-1$ affine functionals. By Proposition 2.4.23, we have that $\vec{v}$ must be contained in the kernel of some functional, without loss of generality, $\hat{f}_1$. Let $\vec{b} \in Y$ such that $\hat{f}_1(\vec{b}) = -c_1$. Then $\vec{v}$ maximizes g subject to $\hat{f}_i(\vec{v}) \geq 0$ if and only if $\vec{v} - \vec{b}$ maximizes $\tilde{g} := g + g(\vec{b})$ subject to $\tilde{f}_i(\vec{v}) \geq 0$, where $\tilde{f}_i(\vec{v}) := \hat{f}_i(\vec{v} + \hat{f}_i(\vec{b})$. Thus we may assume without loss of generality, that $\hat{f}_1$ is linear.

Thus $\vec{v}$ lies in the kernel of a linear space, thus we may restrict each $\hat{f}_i, 2 \leq i \leq n$ to $\mathrm{Ker}(\hat{f}_i)$. By the induction hypothesis, there is a $\hat{v}'$ that lies on a maximal intersection of $\mathrm{Ker}(\hat{f}_i), 2 \leq i \leq n$, and since $\vec{v}' \in \mathrm{Ker}(\hat{f}_1)$ as well, it lies on a maximal intersection of kernels.

□

2.4.4 Partial Results

Several of our classical results only hold under more sophisticated hypothesis than modules over ordered rings. However, some of these classical results are stated as equivalences, and one of the implications holds in this setting.

As a consequence of the Weak Duality, we may prove a weak version of one of our key facts. We first require some definitions.

Definition 2.4.25. Given $R, X, Y^*, A, \vec{b}, \mathbf{c}, d, C_X, C_{Y^*}$ as above, defining an affine program, the solution values of the primal program are *unbounded* if given any $r \in R$, we may find a feasible $\vec{x} \in X$ such that $f(\vec{x}) \geq r$.

Similarly the solution values of the dual program are *unbounded* if given any $r \in R$, we may find a feasible $\mathbf{y} \in Y^*$ such that $g(\mathbf{y}^*) \leq r$.

Corollary 2.4.26 (Weak Existence-Duality Theorem)**.** *Given $R, X, Y^*, A, \vec{b}, \mathbf{c}, d, C_X, C_{Y^*}$ as above, defining an affine program, then we have the following implications:*

1. *If the solutions to the primal program are unbounded, then the dual program is infeasible.*
2. *If the solutions to the dual program are unbounded, then the primal program is infeasible.*

Proof. We proceed by contrapositive. If the dual program is feasible, then there is a feasible solution $\mathbf{y}' \in Y^*$. Then each feasible primal solution $\vec{x} \in X$ satisfies

$$f(\vec{x}) \leq_R g(\mathbf{y}')$$

and thus the primal solutions cannot be unbounded. Similarly, if the primal program admits a feasible solution, than the dual program cannot be unbounded either.

$\square$

Here we state and prove one direction of a generalization of the classical Complementary Slackness Theorem (Theorem 1.2.27).

Proposition 2.4.27. *Let R be an ordered ring, X, Y^* left R modules, $A \in \mathrm{Hom}_R X, Y^*$, $\mathbf{c} \in \mathrm{Hom}_R(X, R)$, $\vec{b} \in Y^*$. Let $\vec{t}, \mathbf{s}$ be the appropriate slack variables (2.2.1). Then if $\mathbf{s}(\vec{x}) = \mathbf{f}(\vec{t}) = 0$ for feasible solutions $\vec{x} \in X, \mathbf{y} \in Y^{**}$, then $\vec{x}, \mathbf{y}$ are optimal solutions.*

Proof. By the Tucker Duality Equation, Theorem 2.2.2, $f(\vec{x}) = g(\mathbf{y})$ and then by the Weak Duality, Proposition 2.4.8, both $\vec{x}$ and $\mathbf{y}$ are optimal. $\square$

Another statement which holds partially in this setting is a generalization of the Farkas' Lemma 1.5.2. The original Farkas' Lemma is presented as an exclusive or statement (which may be stated as an equality of statements). Here, we state and prove one generalized direction.

Proposition 2.4.28. *Let R be an ordered ring, X, Y^* be left R modules, $A \in \mathrm{Hom}_R X, Y^*$, $\mathbf{c} \in \mathrm{Hom}_R(X, R)$. If there is a $\mathbf{y} \in Y^{**}_{\oplus}$ such that $\mathbf{y} \circ A = \mathbf{c}$, then for each $\vec{x} \in X$ such that $A(\vec{x}) \leq_{Y^*} 0$, it follows that $\mathbf{c}(\vec{x}) \leq_R 0$.*

Proof. Suppose that this does not hold. That is, there is a $\mathbf{y} \in Y^*_{\oplus}$ such that $\mathbf{y} \circ A = \mathbf{c}$, but there is a $\vec{z} \in X$ such that $A(\vec{z}) \leq_{Y^*} 0$ but $\mathbf{c}(\vec{z}) >_R 0$. By our hypothesis, $\mathbf{c} = \mathbf{y} \circ A$, and we also note that $A(-\vec{z}) \geq_{Y^*} 0$, so by definition of $Y_{\oplus}^{**}$, $(\mathbf{y} \circ A)(-\vec{z}) \geq_R 0$, but $\mathbf{c}(-\vec{z}) <_R 0$, which is a contradiction. □

2.5 Counterexamples

However, not all of the nine classical facts generalize properly in this setting. Here, we exhibit some counter-examples to our classical results.

Proposition 2.5.1. *Let R be an ordered ring, but not a division ring. Then there is a primal-dual program such that both problems are feasible, and $f(\vec{x}) < g(\mathbf{y})$ for all feasible choices of $\vec{x}, \mathbf{y}$.*

Proof. Let $X, Y^* := R$ as modules. Then let A be right multiplication by a, a positive non-invertible element of R. Let $\vec{b} = 1$, $\mathbf{c} = id_R$, $d = 0$. The primal program then, is to maximize $x \in R$ subject to $xa < 1$. Conversely, the dual program is to minimize $y \in R$ subject to $ay > 1$. For any pair of feasible solutions $x, y \in R$, we have that $xa < 1 < ay$, and so by Proposition 2.3.13 $x \neq y$, and thus $x < y$.

□

Example 2.5.2. Let R, X, Y^* be $\mathbb{Z}$. Let A be multiplication by 2, and $\mathbf{c} = id_R$ or multiplication by 1, $\vec{b} = 1$. For any $x \in \mathbb{Z}^+$, $A(x) > 1$, thus $x = 0$ maximizes $\mathbf{c}(x)$ with objective value 0. However, $yA > 1$ for any feasible choice of y. Thus $y1$ is minimized with objective value 1 when $y = 1$. Thus both primal and dual problems obtain optimal solutions that are not equal.

Example 2.5.3. Let R, X, Y^* be defined to be $R := S^{-1}\mathbb{Z}$ where $S := \{3^n : n \in \mathbb{N}\}$. Let $A, \vec{b}, \mathbf{c}$ be as above. Then any feasible x, $Ax < 1$, since 2 is still not invertible. However, there is an n such that $\frac{1}{3^n} < \frac{1-2x}{2}$. Thus $x + \frac{1}{3n}$ is non-negative, but $2(x + \frac{1}{3n}) = 2x + \frac{2}{3^n} < 1$. Thus the primal program does not obtain an optimal solution, but all solutions are strictly bounded above by $\frac{1}{2}$.

Conversely, if y is feasible, $yA > 1$. But, there is a n such that $\frac{1}{3^n} < \frac{2y-1}{2}$. We see that $y - \frac{1}{3^n}$ is an improved solution for the dual problem. So the dual also obtains no optimal solution, but are bounded strictly below by $\frac{1}{2}$.

These examples lead us to the following characterizations

Corollary 2.5.4. *Let $R, X, Y^*, A, \vec{b}, \mathbf{c}$ be the same as in the proof of Proposition 2.5.1 except that R has the property that 1_R is the smallest positive. Then the primal and dual program obtains optimal solutions, which do not give the same objective value.*

Proof. Since a is positive and non-invertible, $a > 1_R$. Thus $xa > 1_R$ for any positive choice of x and $\mathbf{c}(x)$ is maximized when $x = 0$.

Similarly, any positive y allows $ya > 1_R$, thus y is minimized when $y = 1$. □

This difference $g(\mathbf{y}) - f(\vec{x})$ for optimal $\vec{x}, \mathbf{y}$ is called a *duality gap*.

This example can be used as counterexamples to two classical results. If a program admits a positive duality gap, then this is a contradiction to the Strong Duality Theorem (Theorem 1.2.31). Otherwise, then both programs are feasible but do not admit optimal solutions, which is a violation of the Fundamental Theorem of Linear Programming (Theorem 1.2.21).

In the classical case, it follows that the optimal solutions to a primal program falls on the intersection of bounding hyperplanes. Consider the following example:

Example 2.5.5. Let $R = \mathbb{Z}$, $X = \mathbb{Z}^2, Y^* = \mathbb{Z}^2$. Then let $A = \begin{pmatrix} 4 & 6 \\ 6 & 4 \end{pmatrix}, \vec{b} = \begin{pmatrix} 9 \\ 9 \end{pmatrix}, \mathbf{c} = x_1 + x_2$.

Suppose we let this be a canonical program (that is to be feasible, $x_1, x_2 \geq 0$). The primal program is maximized when $x_1 = 1, x_2 = 1$. However, this solution does not lie on the kernels of any of the defining affine functions, $\alpha_1 = x_1, \alpha_2 = x_2, \alpha_3 = -6x_1 - 4x_2 + 9, \alpha_4 = -4x_1 - 6x_2 + 9$.

Additionally, in this setting, we can provide a counterexample for a generalized version of the Existence-Duality Theorem 1.2.41:

Example 2.5.6. Let $R = \mathbb{Z}, X = \mathbb{Z}, Y^* = \mathbb{Z}^2$ with the standard non-negative cones. Let $A = \begin{pmatrix} 2 \\ -2 \end{pmatrix} \vec{b} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \mathbf{c} = 0, d = 0$.

We note that for a $x \in X = \mathbb{Z}$ to be feasible, $2x \leq 1$ and $-2x \leq -1$, which can only occur if $2x = 1$. Since 2 is not a unit in $\mathbb{Z}$, this cannot happen and thus the primal problem is infeasible. However, we note for in the dual problem, we require that:

$$\begin{aligned} 2y_1 - 2y_2 &\geq 0 \\ y_i &\geq 0 \end{aligned}$$

while minimizing $g(\vec{y}) = y_1 - y_2$. Clearly the feasible solution values are bounded below by 0. In fact $\vec{y} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a feasible and optimal solution. Thus the primal problem may be infeasible while the dual problem is not unbounded. This contradicts the statement of the Existence-Duality Theorem.

Finally, we provide a counterexample to a well-used fact in classical programming, that closed bounded intervals are compact. Although not one of our general facts, it is nevertheless used often in the discourse of classical programming to exhibit the existence of a maximizer in closed bounded sets. We demonstrate here that such sets may not be compact in the general case, and this we must use alternative methods to exhibit maximizers (or minimizers)

Example 2.5.7. Let $R =^* \mathbb{R}$ and consider $[0,1] \subseteq^* \mathbb{R}$. Let $\epsilon \in^* \mathbb{R}$ be an infinitesimal in ${}^*\mathbb{R}$. It follows that the collection of sets $\mathcal{O} := \{(a-\epsilon, a+\epsilon), a \in [0,1]\}$ forms a covering of $[0,1]$.

But, note that given distinct $a, b \in [0,1] \cap \mathbb{R}$, $(a-\epsilon, a+\epsilon) \cap (b-\epsilon, b+\epsilon) = \emptyset$, else, there is a $c \in (a-\epsilon, a+\epsilon) \cap (b-\epsilon b+\epsilon)$, and $|b-a| < 2\epsilon$, which is a contradiction since both a, b are real. Thus each subcover of $\mathcal{O}$ must contain $(a-\epsilon, a+\epsilon)$ for each $a \in \mathbb{R} \cap [0,1]$, and since there are infinitely many such a, no finite subcover exists.

2.6 Conclusions

We show that our first generalized fact, the Tucker Key Equation (Proposition 2.2.1), can be proved in the most general of settings, and any ring and modules over said ring. By requiring that our rings be ordered, we are able to prove two of our generalized facts: the Weak Duality (Proposition 2.4.8), and the Convexity of Feasible and Optimal solutions (Corollaries 2.4.19, 2.4.20). We also proved some partial versions of our other key facts such as the weak Existence-Duality Theorem (Corollary 2.4.26), weak Complementary Slackness (Proposition 2.4.27) and weak Farkas Lemma (Proposition 2.4.28). We also provided counter-examples to some of our later key facts, in particular Proposition 2.5.1. Thus additional hypothesis is required to obtain the remainder of our generalized key facts.

Chapter 3

Cones

3.1 Introduction

We have previously introduced the concept of a cone in a module over an ordered ring. These cones play the role of the non-negative orthant in classical linear programming. In this section, we exhibit some examples of cones and describe their properties. We then demonstrate some counterexamples to classical results with respect to these cones. Finally, we introduce a specific type of cone, the *orthant cone* which plays the role of the classical non-negative orthant.

3.2 Counterexamples

We first give an example of a natural choice of a positive cone.

Example 3.2.1. Let S be a set and let X be a collection of $\mathbb{R}$-valued functions closed under linear combinations. Let $C := \{\varphi \in \mathbb{R}^S : \varphi(S) \subseteq \mathbb{R}_\oplus\}$, the collection of non-negative functions (recall that $\mathbb{R}_\oplus$ is the set of all non-negative real numbers). Given non-negative real numbers

a, b and $\varphi, \psi \in C$, $a \cdot \varphi(x) + b \cdot \psi(x) \in \mathbb{R}_{\oplus}$ for each $x \in S$. Thus C is a cone.

This is a choice of positive cone which arises in many generalizations of affine programming.

Example 3.2.2. Recall the production and inventory example, Example 1.3.6. The variables P_i, I_i could be thought of as a single variable $x : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$. Similarly, the bounds $\bar{P}_i, \bar{I}_i$ can be thought of as a function $b : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$. Then the inequality $P_i \geq \bar{P}_i, I_i \geq \bar{I}_i$ can be rewritten $x - b$ is non-negative, or $x - b \in C$ where $C := \{\varphi \in \mathbb{R}^{\mathbb{Z}\times\mathbb{Z}} : \varphi(\mathbb{Z} \times \mathbb{Z}) \subseteq \mathbb{R}_{\oplus}\}$.

In some cases, one can even show that such a cone is a full cone.

Proposition 3.2.3. *Let S be a set (with a topology), and let X be either $C^0(S)$ or $\mathbb{R}^S$. Then C, the collection of all non-negative functions forms a full cone.*

Proof. Let B be a basis of X as a $\mathbb{R}$ vector space. Given any $b \in B$, we note that $B \backslash b \cup \{b^+, b^-\}$ forms a generating set of X. Thus, we may find a generating set for X that consists of only non-negative functions, and the collection of all non-negative functions C is a generating set of X. $\square$

However, using these cones, we may find counter-examples to classical results.

Example 3.2.4. Let $X := C^0([0, 1])$, the collection of all continuous real valued functions defined on $[0, 1]$, and $Y^* := \mathbb{R}^{[0,1]}$. Let A be the inclusion map, $\vec{b}$ be χ_0, where $\chi_0(s) := \begin{cases} 1 & s = 0 \\ 0 & s \neq 0 \end{cases}$, and $\mathbf{c} : X \to \mathbb{R}$ by $\mathbf{c}(x) := x(0)$ for each $x \in X$. Let the cones in both X, Y^* be the collection of non-negative functions. Thus we may define a canonical primal dual affine programming problem.

Notice that $x = 0$ is the only feasible primal solution. For $s \in [0, 1], s \neq 0$, $x(s) = 0$ else $b(s) - x(s) < 0$. For $s = 0$, if $x(s) > 0$, then x is not continuous ($x^{-1}(0, 2x(s)) = x(s)$, that is, the pre image of an open set is closed and not open). Thus $\mathbf{c}(x)$ is maximized with value 0.

However, consider $\mathbf{y} \in Y^{**}$. We require that $\mathbf{y} \circ A - \mathbf{c} \in X^*_\oplus$. Since $X \subseteq Y^*$, $Y^{**} \subseteq X^*$. Since A is inclusion, we have that $\mathbf{y} - \mathbf{c} \in X^*$, that is $\mathbf{y} = \mathbf{c} + \mathbf{s}$, where $\mathbf{s} \in X^*_\oplus$. Thus $\mathbf{y}(\vec{b}) = \mathbf{c}(\vec{b}) + \mathbf{s}(\vec{b}) \geq \mathbf{c}(\vec{b}) = 1$. Thus, $\mathbf{y}(\vec{b})$ is minimized when $\mathbf{y} = \mathbf{c}$ with value 1.

This exhibits a positive duality gap between the primal and dual solutions.

Another counterexample to classical results is when the choice of cone is not closed.

Example 3.2.5. Let $R = \mathbb{R}$, and let $X, Y^* := \mathbb{R}^2$. Let $A \in \mathrm{Hom}_R(X, Y^*)$ be define by the matrix $\begin{pmatrix} \frac{1}{2} & 1 \\ 2 & 1 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \in Y^*$, $\mathbf{c} \in X^*$ be defined to be $\mathbf{c}\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) := x_1$. Then, define C_X to be the traditional positive orthant, but define C_{Y^*} to be $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in Y^*$ such that $y_1 = y_2 = 0$ or $y_1, y_2 > 0$.

Then consider the primal maximization problem. The feasible region is define by the non-negative cone C_X, and $A^{-1}(C_{Y^*})$, the union of the singleton $\begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}$ with the set $\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in X : \frac{1}{2}x_1 + x_1 < 1, 2x_1 + x_2 < 2\}$. This feasible region contains the set $\{\vec{x} \in X : x_2 = 0, 0 \leq x_1 < 1\}$. Note that $x_1 \not\geq 1$, else $x_2 + 2x_1 \not< 2$. Thus, the objective function takes on values arbitrarily close to 1, but is not maximized at any point.

So, although the primal program is feasible and bounded, it does not admit an optimal solution. This is a violation of Theorem **??**.

Thus, in order to obtain generalizations of the classical results, one must specify the type of cone used.

3.3 Orthant Cones

3.3.1 Definition of Orthant Cone

In classical affine programming, the cone one uses is the positive orthant, the non-negative span of a basis. This then should be the analogous cone in the general case.

Definition 3.3.1 ([ML98]). Given a category and a collection of objects $O_i, i \in I$, their *co-product*, is an object $\coprod_{i \in I} O_i$ along with morphisms $\iota_j : O_j \to \coprod_{i \in I} O_i$ satisfying the universal mapping propertyL i.e. given any other object Z, and morphisms $f_j : O_j \to Z$, there exists a unique $f : \coprod_{i \in I} O_i \to Z$ such that $f_j = f \circ \iota_j$.

Definition 3.3.2. Let R be a ring, a *free module* is the module $\coprod_{i \in I} R_i$, the co-product of a collection of R, indexed by the set I.

Definition 3.3.3. Let R be a ring, and let X be a free R module. Given a fixed index $j \in I$, we define $I_i : R_i \to R$ to be the identification isomorphism. Then by the co-product structure of X, we have:

$$\begin{tikzcd} X \arrow[dr, dotted, "\exists! \hat{x}_j"] & \\ R_i \arrow[u, hook, "\iota_i"] \arrow[r, "\delta_{ij} I_i"'] & R \end{tikzcd}$$

where δ_{ij} is the kronecker delta. We call the collection of of $\hat{x}_j$, $\hat{\mathcal{X}}$, a *collection of spanning maps*.

Definition 3.3.4. Let R be an ordered ring, and let X be a free module over R. An *orthant cone* is a full cone C_X such that $\hat{x}_i(C_X) \subseteq R_\oplus$ for each $\hat{x}_i \in \hat{\mathcal{X}}$.

In other words,

$$\begin{tikzcd} C_X \arrow[r, "\iota"] \arrow[d, "\hat{x}_i"] & X \arrow[d, "\hat{x}_i"'] \\ R_\oplus \arrow[r, "\iota"'] & R \end{tikzcd}$$

commutes for each $\hat{x}_i$.

This is also the non-negative span of $\mathcal{X}$ (a basis of X).

We note that the inclusion $\iota_i : R_i \to X$ maps $\iota_i(R_i) \subseteq C_X$, for an orthant cone C_X. To see this, suppose there were a $r \in (R_i)_\oplus$ such that $\iota_i(r) \notin C_X$. Then there is a $\hat{x}_j$ such that $\hat{x}_j(\iota_i(r)) \notin R_\oplus$. But $\hat{x}_j \circ \iota_i = \delta_{ij} I_i$, and this is a contradiction. This leads us to our next proposition.

Proposition 3.3.5. *Let R be an ordered ring and $X \cong \coprod_{i \in I} R_i$ a free module over R. Then an orthant cone of X, C_X is the co-product $\coprod_{i \in I} (R_\oplus)_i$.*

Proof. Let M be a monoid and for each index $i \in I$, let $f_i : R_i \to M$ be a monoid homomorphism. Then consider they $\hat{f}_i : C_X \to M$ defined by $\hat{f}_i := f_i \circ I_i^{-1} \circ \hat{x}_i$, where $\hat{x}_i$ is a spanning map and $I_i : R_i \to R$ is the identification isomorphism. Then, $\hat{f}_i \circ \iota_i = f_i \circ I_i^{-1} \circ \hat{x}_i \iota_i = f_i$. Thus C_X satisfies the definition of a co-product, and by the universal mapping property, it is unique (up to isomorphism). $\square$

3.3.2 Satisfying Feasibility Conditions

With this notion of a positive cone, we state some new ways to view feasibility.

Definition 3.3.6. Given R an ordered ring, X, Y^* left R free-modules, and a fixed $\hat{\mathcal{Y}}^*$, a collection of spanning maps (Definition 3.3.3). Then each $A \in \mathrm{Hom}_R(X, Y)$ induces a collection of *row-like projections*,

$$\mathcal{A} := \{\hat{y}_i \circ A : \hat{y} \in \mathcal{Y}^*\}.$$

Remark 3.3.7. Since each free module gives rise to such a collection of spanning maps, one can always define a full cone as the collection of all elements that are non-negative when evaluated by these spanning maps. In this thesis, we will focus primarily on such cones. These cones are the closest analogue to the positive orthant of finite dimensional spaces.

Then, with respect to an orthant cone of Y^*, ($Y^*_\oplus := C_Y$) we have:

Proposition 3.3.8. *Let R be an ordered ring, and let X, Y^* be left R modules, where the dual module Y^* is a free module. Let C_{Y^*} be an orthant cone. Then given $w \in X, v \in Y^*$, and $A \in \mathrm{Hom}_R(X, Y^*)$, the following are equivalent:*

1. $A(w) \leq_{Y^*} v$.
2. $\alpha_i(w) \leq_R \hat{y}_i(v)$.

Proof. Notice that $A(w) \leq_{Y^*} v$ if and only if $v - A(w) \in Y^*_\oplus$, then:

$$\begin{aligned} v - A(w) &\in Y^*_\oplus, \text{ if and only if} \\ \hat{y}_i(v) - \alpha_i(Y^*) &\in R_\oplus, \text{ if and only if} \\ \alpha_i(w) &\leq \hat{y}(v). \end{aligned}$$

∎

Proposition 3.3.9. *Let R be an ordered ring, Y^* a left R free module. Each linear functional $\mathbf{u} : Y^* \to R$ can be defined as $\mathbf{u} = \sum_{i \in I} \hat{y}_i \cdot u_i$, where $u_i \in R$.*

Proof. Notice that given such a $\mathbf{u} : Y^* \to R$, there is an induced map $\mathbf{u}_i : R_i \to R$ via $\mu_i := \mathbf{u}\iota_i$. Any such map is a scalar of the identity map, i. e. $\mu_i = I_i \cdot u_i$.

Then notice that each $v \in Y^*$ is written as $\sum_{j\in I} \iota_i(r_i)$, where $r_i \in R_i$ and only finitely many of the $r_i \neq 0$. It follows that:

$$\begin{aligned}
\mathbf{u}(v) &= \mathbf{u}(\sum_{j\in J} \iota_j(r_j)) \\
&= \sum_{i\in I} \mathbf{u}(\iota_i(r_i)) \\
&= \sum_{i\in I} \mu_i(r_i) \\
&= \sum_{i\in I} I_i(r_i) \cdot u_i \\
&= \sum_{i\in I} \hat{y}_i(v) \cdot u_i.
\end{aligned}$$

□

Proposition 3.3.10. *Let R be an ordered ring, Y^* a left R free module. A function $\mathbf{u} : Y^* \to R$ is in $Y^{**}_\oplus$ (the non-negatives of $\mathrm{Hom}_R(Y^*, R)$) if and only if $\mathbf{u} = \sum_{i\in I} \hat{y}_i \cdot u_i$ and $u_i \geq 0$ for each $i \in I$.*

Proof. We recall that each $\hat{y}_i$ is an order preserving map. Thus $\mathbf{u} \in Y^{**}_\oplus$ (the non-negatives of $\mathrm{Hom}_R(Y^*, R)$) if and only if $\sum_{i\in I} \hat{y}_i(v) \cdot u_i \in R_\oplus$ for each $v \in Y^*_\oplus$. So given any $\hat{y}_i, \hat{y}_i(v) \in R_\oplus$, since $\hat{y}_i$ is order preserving, thus $\hat{y}_i(v) \cdot u_i \in R_\oplus$ if and only if each $u_i \in R_\oplus$. For such a collection of u_i,

$$\mathbf{u}(v) = \sum_{i\in I} \hat{y}_i(v) \cdot u_i \in R_\oplus$$

and $\mathbf{u} \in Y^{**}_\oplus$.

Conversely, for each index i, define $v_i := \iota_i(1)$. That is, the element in Y such that $\hat{y}_j(v) = \delta_{ij}$ (δ_{ij} being the Kronecker delta). It is clear that $v_i \in Y^*_\oplus$, and $\mathbf{u}(v_i) = \hat{y}_i(v_i) \cdot u_i$, so $\hat{y}_i(v_i) \cdot u_i \in R_\oplus$

if and only if $u_i \in R_{\oplus}$. Since this is true for index i, each u_i is non-negative. $\square$

Remark 3.3.11. We may formulate the feasibility conditions described in Definition 2.4.3. We let R be an ordered ring, X a left BR module and Y^* a left R free module. We also let $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$, and let C_{Y^*} be an orthant cone of Y^*.

Then the following statements:

- A variable $\vec{x} \in X$ is *primal feasible* if $-A(\vec{x}) + \vec{b} \in C_Y$;
- Consequently, $\mathbf{y}$ is *dual feasible* if $\mathbf{y} \circ A = \mathbf{c}$;

may be written as:

- A primal variable $\vec{x} \in X$ is *primal feasible* if $-\alpha_i(\vec{x}) + \hat{y}_i(\vec{b}) \geq_R 0$ for each $i \in I$.
- A dual variable $\mathbf{y} \in Y^{**}$ is *dual feasible* if $\mathbf{c} = \sum_{i \in I} \alpha_i \cdot y_i$, $y_i := \mathbf{y}(\iota_i(1_R))$ where ι_i is the canonical inclusion, and $y_i \geq_R 0$.

3.3.3 Non-Orthant Cones

We show here that the cones in Examples 3.2.4, 3.2.5 are not orthant cones.

Claim 3.3.12. *The cone of nonnegative continuous functions in Example 3.2.4 is not an orthant cone, when S is a compact Hausdorff set.*

Proof. Let S be a compact Hausdorff space, and consider $X := C^0(S)$. Let C_X be the collection of non-negative continuous functions $S \to \mathbb{R}$. We will let $\mathcal{B} := \{b_i\}$ be a (Hamel) basis of X. Consider the constant function 1. We may write $1 = \sum_{i=1}^{n} r_i \cdot b_i$ where $r_i \in \mathbb{R}$ and b_i are some of the basis elements. Then let b_j be a basis element, not one of these b_i. Since

b_j is continuous, we have that b_j reaches a maximum on S. Let $m := \max_{s \in S} b_j(s)$. Then the function $\left(\sum_{i=1}^{n} mr_i \cdot b_i -\right) b_j$ is a non-negative function, but this function is *not* a non-negative linear combination of basis elements. Since the choice of basis was arbitrary, this shows that any basis will admit non-negative functions that are not non-negative linear combination of basis elements, and thus C_X is not an orthant cone. $\square$

Claim 3.3.13. *The cone C_Y in Example 3.2.5 is not an orthant cone.*

Proof. We first note that in $\mathbb{R}^n$, orthant cones are closed. To see this, let $e_1, \ldots, e_n$ be a basis for $\mathbb{R}^n$, and let C be their non-negative span. We note that the maps f_i that send $\vec{x}$ to the coefficient of e_i for each basis element e_i is continuous (since $\mathbb{R}^n$ is finite dimensional). Thus, the half space C_i, the pre image of $[0, \infty)$ under f_i are closed sets. Each C_i is precisely the vectors where the coefficients of e_i are non-negative, and so C is the intersection of all the C_i and is closed.

We then note that the cone C_Y is not closed. To see this, consider the sequence $\{s_n\}, s_n := (1, \frac{1}{n})$. Each $s_n \in C_Y$, but the s_n converge to $(1, 0)$, which is not in C_Y, and so C_Y is not closed.

$\square$

3.4 Conclusion

In this chapter, we begun by examining several possible examples of cones. But we then demonstrate that not every possible cone allows us to prove generalized versions of our classical facts. Thus, we introduce a notion of an *orthant cone*, the analogue of a positive orthant in the classical setting in Definition 3.3.4. With this cone, we also introduce the notions of row-like projections in Definition 3.3.6. We then demonstrate that several of examples of cones that were introduced here were not orthant cones, in particular, we demonstrated this in Examples 3.2.4, 3.2.5.

Chapter 4

Farkas' Lemma and Generalizations

4.1 Introduction

We saw that the existence of a positive duality gap for a generalized primal-dual affine program (Proposition 2.5.1) is largely a result of the ring and cone structures. That is, the non-invertibility of ring elements, or a non-orthant cone, gives rise to the existence of such a gap. In this chapter, we make the assumption that the ring of scalars R is an ordered division ring and that all the cones are orthant cones. We then investigate the circumstances under which the duality gap may be closed. When R is a division ring, both modules X, Y^* will be vector-spaces over R. Thus, one must investigate both the dimension of X, Y^*, and the structure of the linear transformation A. We then investigate the circumstances under which the **Farkas' Lemma** holds.

The Farkas' Lemma is then key to proving the remaining duality results. The original statement was proved in 1902 by Gyula Farkas' [Far02]:

Theorem 4.1.1 (Classical Farkas' Lemma). *Given $A \in \mathbb{R}^{m \times n}, \vec{c} \in \mathbb{R}^n$, then exactly one of the following hold:*

(a) There is an $\vec{x} \in \mathbb{R}^n$, such that $A\vec{x} \leq_{\mathbb{R}^m} \vec{0}$ and $\vec{c}^\top \vec{x} >_{\mathbb{R}} 0$.

(b) There is a $\vec{y} \in \mathbb{R}^m$, where $\vec{y} \geq_{\mathbb{R}^m} \vec{0}$ and $A^\top \vec{y} = \vec{c}$.

We see that the Farkas' Lemma is a statement about primal solutions ($\vec{x}$) and dual solutions ($\vec{y}$), and relates them to each other. Here, we will state general versions of the Farkas' Lemma, and prove them under various hypotheses. The original proof of the Farkas' Lemma required the use of the least upper bound axiom of the real numbers in order to create a separating hyperplane. Since we do not have the same topological properties of the real numbers to rely on here, our techniques will be purely algebraic. These proofs were inspired by David Bartl's paper, [Bar07]. In a 2012 paper [Bar12a], Bartl conjectured that for the Farkas' Lemma to hold in an infinite dimensional case, the row-like projection maps would have to be linearly independent. We prove this, and strengthen the result.

Throughout this chapter, we will use the definition of feasibility given in Definition 2.4.3 unless otherwise stated.

4.2 Tools

In this section, we develop some of the tools necessary to discuss the Farkas' Lemma in a general setting. We follow loosely the structure of [Bar07], but provide original proofs unless otherwise stated. The results in [Bar07] depend on the hypothesis that the dimension of the co-domain $\dim_R(Y^*)$ is finite. In particular, this allows proof techniques such as induction, and assume that $Y^* \cong Y$. Here we attempt to avoid these conditions.

We begin with a lemma that relates the functionals $\mathbf{c} \in \mathrm{Hom}_R(X, R)$ to the functionals

$\mathrm{Hom}_R(Y^*, R)$.

Lemma 4.2.1 (Fundamental Lemma)**.** *Let R be a division ring. Let X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \mathbf{c} \in \mathrm{Hom}_R(X, R)$. Then the following statements are equivalent:*

(a) $\mathrm{Ker}(A) \subseteq \mathrm{Ker}(\mathbf{c})$.

(b) There exists a $\mathbf{u} \in \mathrm{Hom}_R(Y^*, R)$ such that $\mathbf{c} = \mathbf{u} \circ A$.

Proof. Consider the following:

(b) $\Longrightarrow$ (a) This is clear, since if (b) holds, for each $\vec{x} \in \mathrm{Ker}(A)$, $\alpha_i(\vec{x}) = \pi_i \circ A(\vec{x}) = 0$. Thus $\sum_{i=1} u_i \alpha(\vec{x}) = 0$.

(a) $\Longrightarrow$ (b) We wish to construct a map $\mathbf{u} : Y^* \to R$ such that

$$\begin{tikzcd} X \arrow[r, "A"] \arrow[dr, "c"'] & Y^* \arrow[d, "\exists \mathbf{u}"] \\ & R \end{tikzcd}$$

commutes. Recall by Proposition 3.3.9 that we may write $\mathbf{u}$ as a linear combination $\mathbf{u} = \sum_{i \in I} \hat{y}_i \cdot u_i, u_i \in Y^*$.

In order for this to hold, we need that for each $\vec{y}$ an element of Y^*, that $c(A^{-1}(\vec{y})) = \{\mathbf{u}(\vec{y})\}$, where A^{-1} is the inverse image under the function A. So suppose there was an element of Y^*, $\vec{y}$ where $\vec{x}, \vec{x}' \in A^{-1}(\vec{y})$ such that $c(\vec{x}) \neq c(\vec{x}')$. Then $c(\vec{x} - \vec{x}') \neq 0$, but $A(\vec{x} - \vec{x}') = A(\vec{x}) - A(\vec{x}') = \vec{y} - \vec{y} = 0$. This contradicts (a). Thus (a) $\Longrightarrow$ (b).

$\square$

In particular, by the above proof, we see that $\mathbf{c}$ is a (potentially infinite) linear combination of the $\alpha_i \in \mathcal{A}$, the collection of row-like projection maps (Definition 3.3.6). Although somewhat abusive of the term linear combination, this gives us a way to think of the action that $\mathbf{u}$ has on the α_i and on A.

Definition 4.2.2. Let R be an ordered division ring, X a left R vector space and $f_i : X \to R_i, i \in I$ be an indexed collection of linear functionals, such that given $w \in X$, only finitely many $f_i(w) \neq 0$, and let $g : X \to R$ be a linear functional. We say that g is a *(potentially infinite)* linear combination of the $\{f_i\}$ if there is a $\lambda : \prod_{i \in I} R_i \to R$ such that

$$\lambda\left(\sum_{f_i(w)\neq 0} f_i(w)\right) = g(w).$$

Definition 4.2.3. Let R be an ordered division ring, X a left R vector space and $f_i : X \to R_i, i \in I$ be an indexed collection of linear functionals, such that given $w \in X$, only finitely many $f_i(w) \neq 0$. This collection is *linear independent* if no functional may be written as a (potentially infinite) linear combination of the others.

We can see the usefulness of this result, as it allows us to express $\mathbf{c}$ in terms of a composition of maps. Since feasible solutions to the dual problem are exactly the non-negative functionals $\mathbf{u} \in \mathrm{Hom}_R(Y^*, R)$ such that $\mathbf{u} \circ A = \mathbf{c}$, the utility of this lemma is clear.

Next, we prove a generalization of Fredholm's Theorem [Ion27]

Theorem 4.2.4 (Generalization of Fredholm's Theorem)**.** *Let R be an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y$. Then the system*

$$A(\vec{x}) = \vec{b}$$

has no solutions if and only if:

$$\textit{There exists a } \lambda \in \mathrm{Hom}_R(Y^*, R) : \lambda \circ A = 0, \textit{and } \lambda(\vec{b}) \neq 0.$$

Proof. For the forward direction, notice that $A(\vec{x}) \neq \vec{b}$ for any $\vec{x}$ means $\vec{b} \in Y \backslash A(X)$. Thus, we may define a basis $\mathcal{B}_X$ for $A(X)$ and notice that $\{\vec{b} \cup \mathcal{B}_X$ is linear independent in Y^*. Thus, this set can be extended to $\mathcal{B}$ a basis for Y^*. We then define $\lambda \in \mathrm{Hom}_R(Y^*, R)$ on this basis via $\lambda(\vec{b}) := 1$, and 0 on each of the other basis elements. This is a well-defined linear map, where $A(X) \subseteq \mathrm{Ker}(\lambda)$, and $\lambda(\vec{b}) \neq 0$.

Conversely, if $\lambda \circ A = 0$ and $\lambda(\vec{b}) \neq 0$ then $\vec{b}$ is not in the image of A, thus there can be no solution to the equation $A(\vec{x}) = \vec{b}$.

□

Proposition 4.2.5. *Let R be a division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*)$, and $\mathcal{A}$ be the collection of row-like projection maps induced by A (Definition 3.3.6). Then the statements:*

(a) A is surjective.

(b) The collection $\alpha_i \in \mathcal{A}$ are linearly independent.

are equivalent.

Proof. Consider the following:

(a) $\Longrightarrow$ (b) If A is surjective, a linear combination of the α_i, can be written as

$$\sum_{i \in I} \alpha_i \cdot r_i = \sum_{i \in I} (\hat{y}_i \circ A) \cdot r_i, r_i \in R$$

Then we note that for each $v \in Y$, there is a $w \in X$ such that $A(w) = v$. In particular, we may define for $v_i := \iota_i(1_R)$, for each index i, where ι_i is the inclusion map $\iota_i : R_i \to Y$, such that $\hat{y}_i \iota_i = id_{R_i}$. For such a v_i,

$$\begin{aligned} \hat{y}_i(v_i) &= \hat{y}_i(\iota_i(1_{R_i})) \\ &= \delta_{ii} I_i(1_{R_i}) = 1. \end{aligned}$$

However:

$$\begin{aligned} \hat{y}_j(v_i) &= \hat{y}_j(\iota_i(1_{R_i})) \\ &= \delta_{ji} I_i(1_{R_i}) = 0, \end{aligned}$$

where δ_{ij} is the Kronecker delta. Thus, $\left(\sum_{i \in I} \alpha_i r_i\right)(v_i) = r_i + 0$. It follows that this sum $\sum_{i \in I} \alpha_i r_i = 0$ if and only if each $r_i = 0$, and thus the α_i are linearly independent.

(b) $\implies$ (a) Conversely, suppose that there was an $\vec{b} \in Y$ such that $\vec{b} \notin A(X)$. Then, by Fredholm's Theorem (Theorem 4.2.4), there is a $\lambda \in \mathrm{Hom}_R(Y, R)$ such that $\lambda \circ A = 0$ but $\lambda(\vec{b}) \neq 0$. Notice that $\lambda \neq 0$, yet:

$$\begin{aligned} 0 = \lambda \circ A &= \sum_{i \in I} \hat{y}_i \circ A \cdot u_i, \text{ where } r_i \in R, \text{ by Proposition 3.3.9} \\ &= \sum_{i \in I} \alpha_i \cdot u_i. \end{aligned}$$

Since $\lambda \neq 0$, the u_i are not all zero, but $\displaystyle\sum_{i \in I} \alpha_i \cdot u_i = 0$, and the α_i are not linearly independent.

□

4.3 The Farkas' Lemma

We now have the tools necessary to prove our generalized Farkas' Lemma. In general the Farkas' Lemma has a number of different equivalent statements. The one we will prove is stated as follows:

Theorem 4.3.1 (Generalized Fact V: Farkas' Lemma)**.** *Given R an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \mathbf{c} \in \mathrm{Hom}_R(X, R)$. Then (under some hypothesis) the following are equivalent:*

(a) For each $\vec{x}, \in X$, if $A(\vec{x}) \leq_{Y^*} 0$, then $\mathbf{c}(\vec{x}) \leq_R 0$.

(b) The map $\mathbf{c} = \sum_i \alpha_i \cdot u_i, u_i \geq_R 0$. Equivalently, there is a $\mathbf{y} \in Y_\oplus^{**}$ such that $\mathbf{y} \circ A = \mathbf{c}$.

Recall that by Proposition 2.4.28, we have already shown (b) $\implies$ (a), and thus it suffices to find the conditions where (a) $\implies$ (b).

Remark 4.3.2. We first notice that $(b) \implies (a)$ under any circumstance. If (b) holds, then a vector $\vec{x} \in X$ satisfies $A(\vec{x}) \leq_Y 0$ if and only if each $\alpha_i(\vec{x}) \leq_R 0$. Then

$$\mathbf{c}(\vec{x}) = \sum_{i \in I} \alpha_i(\vec{x}) \cdot u_i$$

where each $u_i \geq_R 0$ and each $\alpha_i(\vec{x}) \leq_R 0$, and so $\mathbf{c}(\vec{x}) \leq_R 0$.

Moreover, we notice that under any circumstances, if 4.3.1(a) is satisfied, then given $\vec{x} \in \mathrm{Ker}(A)$, Notice $A(\vec{x}) = A(-\vec{x}) = 0$, so by (a), $\mathbf{c}(\vec{x}), \mathbf{c}(-\vec{x}) \geq 0$. Thus $\mathbf{c}(\vec{x}) = 0, \mathrm{Ker}(A) \subseteq \mathrm{Ker}(c)$. So by the Fundamental Lemma (Lemma 4.2.1), $\mathbf{c} = \sum_i \alpha_i \cdot u_i$.

Thus, when we present the proof of the Farkas' Lemma under various hypotheses, we will only prove $(a) \implies (b)$. Moreover, we will always assume that $\mathbf{c} = \lambda \circ A$ for some $\lambda \in \mathrm{Hom}_R(Y^*, R)$. Equivalently, we assume that $\mathbf{c}$ is a linear combination of the α_i.

With this in mind, we also note that since $\mathbf{c}$ is a linear combination of α_i, that if any basis element of X were in the kernel of each α_i, it would also be in the kernel of $\mathbf{c}$. Such a basis element would not contribute to the problem in any way, and so without loss of generality, we may assume that each basis element, for each basis we pick, is not in the kernel of some α_i. In other words, we may write

$$X = X' \oplus \mathrm{Ker}(A)$$

and since $\mathrm{Ker}(A) \subseteq \mathrm{Ker}(\mathbf{c})$, we can think of $\mathbf{c} = \lambda \circ (A|_{X'})$. So without loss of generality, we may assume $\mathrm{Ker}(A) = \{0\}$.

Remark 4.3.3. This notion of the generalized Farkas' Lemma is done with respect to the orthant cones in X, Y^*, and Y^{**}. However, one could state the generalized Farkas' Lemma in terms of general cones in a module over an ordered ring:

Given R an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \mathbf{c} \in \mathrm{Hom}_R(X, R)$. Then the following are equivalent:

(a) If, $A(\vec{x}) \leq_{Y^*} 0$ for each $\vec{x} \in X$, then $\mathbf{c}(\vec{x}) \leq_R 0$.

(b) The map $\mathbf{c} = \mathbf{y} \circ A$, where $\mathbf{y} \in Y^{**}_{\oplus}$

These inequalities can be made sensible with respect to any ordered ring and cones, but as we have seen, may not be a true statement in a more general setting.

This statement is not true for all possible choices of $Y^*, X, A, \mathbf{c}$. We begin by presenting the proof provided by [D. Bartl], where $\dim_R(Y^*) < \infty$. We then show circumstances under which the Farkas' Lemma fails, and various hypothesis under which it holds.

Theorem 4.3.4 (Finite Farkas' Lemma ([Bar12b]))**.** *Let R be an ordered division ring, X, Y^* be left R vector spaces such that $\dim_R(Y^*) < \infty$, $A \in \mathrm{Hom}_R(X, Y^*)$, and $\mathbf{c} \in \mathrm{Hom}_R(X, R)$. The following are equivalent* *(a) For each $\vec{x}, \in X$, if $A(\vec{x}) \leq_{Y^*} 0$, then $\mathbf{c}(\vec{x}) \leq_R 0$.*

(b) The map $\mathbf{c} = \sum_i \alpha_i \cdot u_i, u_i \geq_R 0$. Equivalently, there is a $\mathbf{y} \in Y^{**}_{\oplus}$ such that $\mathbf{y} \circ A = \mathbf{c}$.

Proof. Notice that $(b) \implies (a)$ trivially under any given hypothesis.

We proceed via induction on $\dim_R(Y)$. If $\dim_R(Y) = 1$, then $A : X \to Y$ is a linear functional. By Remark 4.3.2, if (a) holds, then we may write $\mathbf{c} = A \cdot u$ for some $u \in R$. Then since $A(\vec{x}) \leq_{Y^*} 0 \implies \mathbf{c}(\vec{x}) \leq 0$, this u cannot be negative. Thus $u \geq_R 0$.

Thus, we assume that $(a) \Leftrightarrow (b)$ for $Y', \dim_R(Y') = m$ and suppose $\dim_R(Y^*) = m+1$. Notice that if (a) does not hold for Y^*, $(a) \implies (b)$ holds vacuously. Moreover, if $\alpha_i(\vec{x}) \leq_R 0$ for each $i \in [m]$ implies $\mathbf{c}(\vec{x}) \leq_R 0$, then

$$\mathbf{c} = \sum_{i=1}^{m} \alpha_i \cdot u_i, u_i \geq_R 0$$

and thus

$$\mathbf{c} = \sum_{i=1}^{m+1} \alpha_i \cdot u_i, u_i \geq_R 0, u_{m+1} = 0,$$

and we are done. Thus, we may assume this does not happen, i. e. that there is a $\vec{z} \in X$ such that

$$\alpha_i(\vec{z}) \leq_R 0, i \in [m], \text{ but } \mathbf{c}(\vec{z}) >_R 0.$$

It follows that $\alpha_{m+1}(\vec{z}) >_R 0$, else (a) does not hold. Without loss of generality, we may assume $\alpha_{m+1}(\vec{z}) = 1$.

We then define $T \in \mathrm{Hom}_R(X, X)$ to be

$$T(\vec{x}) := \vec{x} - \alpha_{m+1}(\vec{x})\vec{z}.$$

To verify that T is a linear map, given $a, b \in R, \vec{x}, \vec{y} \in X$, we get:

$$\begin{aligned}
T(a\vec{x} + b\vec{y}) &= (a\vec{x} + b\vec{y}) - \alpha_{m+1}(a\vec{x} + b\vec{y})\vec{z} \\
&= a\vec{x} + b\vec{y} - (a\alpha_{m+1}(\vec{x}) - b\alpha_{m+1}(\vec{y}))\vec{z} \\
&= a\vec{x} + b\vec{y} - a\alpha_{m+1}(\vec{x})\vec{z} - b\alpha_{m+1}(\vec{y})\vec{z} \\
&= a\vec{x} - a\alpha_{m+1}(\vec{x})\vec{z} + b\vec{y} - b\alpha_{m+1}(\vec{y})\vec{z} \\
&= aT(\vec{x}) + bT(\vec{y}).
\end{aligned}$$

Consider then that given any $\vec{x} \in X$:

$$\begin{aligned}
\alpha_{m+1} \circ T(\vec{x}) &= \alpha_{m+1}(\vec{x} - \alpha_{m+1}(\vec{x})\vec{z}) \\
&= \alpha_{m+1}(\vec{x}) - \alpha_{m+1}(\vec{x})\alpha_{m+1}(\vec{z}) \\
&= 0.
\end{aligned}$$

Let $\beta_i := \alpha_i|_{T(X)}$ and $\gamma := \mathbf{c}|_{T(X)}$. Notice that since $\beta_{m+1} = 0$, then by (a), if for a given $\vec{x} \in T(X)$, $\beta_i(\vec{x}) \leq_R 0$, then $\gamma(\vec{x}) \leq 0$. In fact, we may delete β_{m+1} and restrict to the m dimensional subspace induced by $\hat{y}_1, \ldots \hat{y}_m$.

So by the induction hypothesis:

$$\begin{aligned}
\gamma &= \sum_{i=1}^{m} \beta_i \cdot u_i, \ u_i \geq_R 0, \text{ so given } \vec{x} \in X, \\
\mathbf{c}(\vec{x} - \alpha_{m+1}(\vec{x})\vec{z}) &= \sum_{i=1}^{m} \alpha_i(\vec{x} - \alpha_{m+1}(\vec{x})\vec{z}) \cdot u_i \\
\mathbf{c}(\vec{x}) - \alpha_{m+1}\mathbf{c}(\vec{z}) &= \sum_{i=1}^{m} \alpha_i(\vec{x}) \cdot u_i - \alpha_{m+1}(\vec{x}) \sum_{i=1}^{m} \alpha_i(\vec{z}) \cdot u_i \\
\mathbf{c}(\vec{x}) &= \sum_{i=1}^{m} \alpha_i(\vec{x}) \cdot u_i + \alpha_{m+1} \left(\mathbf{c}(\vec{z}) - \sum_{i=1}^{m} \alpha_i(\vec{z}) \cdot u_i \right)
\end{aligned}$$

But by construction, each $\alpha_i(\vec{z}) \leq_R 0$ for $i \in [m]$, and $\mathbf{c}(\vec{z}) >_R 0$, thus

$$u_{m+1} := \mathbf{c}(\vec{z}) - \sum_{i=1}^{m} \alpha_i(\vec{z}) \cdot u_i >_R 0$$

and $\mathbf{u}$ is a non-negative linear combination of α_i. $\square$

We would clearly like to be able to extend this result to an infinite-dimensional case. However, here we run into some difficulty:

Example 4.3.5 (A Counterexample to "unrestricted" Farkas' Lemma)**.** We, we let $R = \mathbb{R}$, and let $Y^* = X = \mathbb{R}^{\mathbb{N}}$. Consider the following illustration of $A, \mathbf{c}$, where we think of the nth row as the action on the nth entry of X:

$$A : \begin{array}{c} \alpha_0 = \\ \alpha_1 = \\ \alpha_2 = \\ \alpha_3 = \\ \vdots \end{array} \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & \cdots \\ -1 & 0 & 0 & 0 & \cdots & 0 & \cdots \\ 0 & -1 & 0 & 0 & \cdots & 0 & \cdots \\ 0 & 0 & -1 & 0 & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \end{pmatrix}$$

$$\mathbf{c} : \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n & \cdots \end{pmatrix}.$$

So, we first verify that this choice of $A, \mathbf{c}$ satisfy the Farkas' Lemma. First, note that given a $\vec{x}$, in order for each $\alpha_i(\vec{x}), i > 0$ to be non-positive, each entry x_n must be non-negative $x_n \geq 0$. Under these circumstances, for $\alpha_0(\vec{x})$ to be non-negative, each $x_n = 0$. Thus $A(\vec{x}) \leq_Y 0$ if and only if $\vec{x} = 0$. $\mathbf{c}(\vec{x}) = 0$, and so the hypothesis to (a) is satisfied.

However, when we consider any non-negative linear combination of the α_i

$$\gamma := \sum_{i=0}^{n} u_i \alpha_i, u_i \geq_{\mathbb{R}} 0 : \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_n & \cdots \end{pmatrix}$$

We see that γ_i is bounded above by u_0. Yet the entries of $\mathbf{c}$ are unbounded. Thus we cannot express $\mathbf{c}$ as a non-negative linear combination of the α_i and (b) fails.

Notice though that by defining $c_i = -i$, we can express $\mathbf{c}$ as a linear combination of the α_i. We also note that this is not unique.

Thus, we need to show under what circumstances the Farkas' Lemma does hold. We begin categorizing some necessary conditions for the Farkas' Lemma to be true, and proving it with some additional hypotheses.

Remark 4.3.6. Recall that given any $\mathcal{A}' \subseteq \mathcal{A}$ induces a linear transformation $A' \in \mathrm{Hom}_R(X, Y')$, where $Y' \leq Y^*$ is the subspace of Y^* defined by $\coprod_{\alpha_i \in \mathcal{A}} R_i$. Equivalently, we may think of the subspace of Y^* defined by the basis elements that the maps $\alpha_i \in \mathcal{A}'$ project onto.

Proposition 4.3.7. *Given R an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*)$ and $\mathbf{c} \in \mathrm{Hom}_R(X, R)$, then the following are equivalent:*

(a) $A(\vec{x}) \leq_{Y^*} 0 \implies \mathbf{c}(\vec{x}) \leq_R 0$ *for each* $\vec{x} \in X$.

(b) There is a linear transformation A' induced by $\mathcal{A}' \subseteq \mathcal{A}$ such that $A'(\vec{x}) \leq_{Y'} 0 \implies \mathbf{c}(\vec{x}) \leq_R 0$.

Recall that $\mathcal{A}$ is the collection of row-like projection maps (Definition 3.3.6).

Proof. If (a) holds, then allowing $A' = A$, (b) holds.

If (b) holds, then given any $\vec{x} \in X$, we have that:

$$\begin{aligned}
A(\vec{x}) &\leq_{Y'} 0 \text{ only if} \\
\alpha_i(\vec{x}) &\leq_R 0, \text{ for each } \alpha_i \in \mathcal{A}, \text{ which implies} \\
\alpha_j(\vec{x}) &\leq_R 0, \text{ for each } \alpha_j \in \mathcal{A}', \text{ which occurs only if} \\
A'(\vec{x}) &\leq_{Y^*} 0.
\end{aligned}$$

Thus $A(\vec{x}) \leq_{Y^*} 0 \implies A'(\vec{x}) \leq_{Y'} 0 \implies \mathbf{c} \leq_R 0$. □

Proposition 4.3.8. *Given R an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*)$ and $\mathbf{c} \in \mathrm{Hom}_R(X, R)$, the following are equivalent:*

(a) $A(\vec{x}) \leq_{Y^*} 0 \implies \mathbf{c}(\vec{x}) \leq_R 0$ *for each* $\vec{x} \in X$.

(b) $A'(\vec{x}) \leq_{Y'} 0 \implies \mathbf{c}(\vec{x}) \leq_R 0$, *where A' is induced by $\mathcal{A}' := \{\alpha_i \in \mathcal{A} : \alpha_i \neq 0\}$.*

Proof. If $A(\vec{x}) \leq_{Y^*} 0$, then each $\alpha_i(\vec{x}) \leq_R 0$, in particular, for the non-zero α_i as well. Conversely, since each $\alpha_j = 0$ is automatically less than or equal to 0 for each $\vec{x}$, $A(\vec{x}) \leq_Y 0$ only if $\alpha_i(\vec{x}) \leq 0$ for each non-zero α_i.

Thus $A(\vec{x}) \leq_{Y^*} 0$ if and only if $A'(\vec{x}) \leq_{Y'} 0$, and if one implies $\mathbf{c}(\vec{x}) \leq_R 0$, so must the other.

□

Corollary 4.3.9. *If given R an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y)$ and $\mathbf{c} \in \mathrm{Hom}_R(X, R)$, the Farkas' Lemma holds if and only if the Farkas' Lemma holds for A', Y', where A' is induced by $\mathcal{A}' := \{\alpha_i \in \mathcal{A} : \alpha_i \neq 0\}$, the row-like projection maps (Definition 3.3.6).*

So without loss of generality, we may assume each $\alpha_i \neq 0$.

We will now prove some different versions of the Farkas' Lemma. The first is the linearly independent Farkas' lemma, conjectured by Bartl in [Bar12a].

Theorem 4.3.10 (Linearly Independent Farkas' Lemma). *Given R an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*)$, where the collection of induced row-like projections $\mathcal{A}$(Definition 3.3.6) is a linearly independent set, and $\mathbf{c} \in \mathrm{Hom}_R(X, R)$, the following are equivalent:*

(a) For each $\vec{x}, \in X$, if $A(\vec{x}) \leq_{Y^*} 0$, then $\mathbf{c}(\vec{x}) \leq_R 0$.

(b) The map $\mathbf{c} = \sum_i \alpha_i \cdot u_i, u_i \geq_R 0$. Equivalently, there is a $\mathbf{y} \in Y^{**}_{\oplus}$ such that $\mathbf{y} \circ A = \mathbf{c}$.

Proof. Recall that we may assume that there is a map $\lambda \in \mathrm{Hom}_R(Y^*, R)$ such that $\mathbf{c} = \lambda \circ A$. Since the α_i are linearly independent, by Proposition 4.2.5 the map A is surjective. Thus given each index i, there is a $x_i \in X$ such that $A(x_i) = v_i$, where $v_i = \iota_i(1_R)$. Recall that $\hat{y}_j(v_i) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. Thus:

$$\begin{aligned}\lambda(v_i) &= \sum_{j \in I} \hat{y}_j(v_i) \cdot u_i \\ &= u_i.\end{aligned}$$

So $\alpha_i(-x_i) = \hat{y}_i(A(-x_i)) = -\hat{y}_i(v_i) = -1 \leq_R 0$, and $\alpha_j(-x_i) = \hat{y}_j(A(-x_i)) = -\hat{y}_j(v_i) = 0 \leq_R 0$. Thus by our hypothesis $\mathbf{c}(-x_i) = \lambda \circ A(-x_i) = -u_i \leq 0$, and so $u_i \geq 0$. Since this is true for each index i, we have that $\mathbf{c}$ is a non-negative linear combination of the α_i.

$\square$

We then present a version of the Farkas' Lemma under hypothesis which allow us to prove the Farkas' Lemma with respect to a subspace of X, then extend it with Zorn's Lemma.

For the following definitions, we will abuse notation and define the following for both singletons and sets.

Definition 4.3.11. Given $\alpha \in \mathcal{A}$, the collection of row-like projection maps (Definition 3.3.6) we define the *support*, $\operatorname{supp}(\alpha)$ to be $\operatorname{supp}(\alpha) := \{x \in \mathcal{X} : \alpha(x) \neq 0\}$. Given $\mathcal{A}' \subseteq \mathcal{A}$, we define $\operatorname{supp}(\mathcal{A}') := \bigcup_{\alpha \in \mathcal{A}'} \operatorname{supp}(\alpha)$.

Definition 4.3.12. Given an $x \in \mathcal{X}$, we define the *co-support*, $\operatorname{cosupp}(x)$, to be $\operatorname{cosupp}(x) := \{\alpha \in \mathcal{A} : \alpha(x) \neq 0\}$. Given $X \subseteq \mathcal{X}$, we define $\operatorname{cosupp}(X) = \bigcup_{x \in X} \operatorname{cosupp}(x)$. Notice that this will be finite for any $x \in \mathcal{X}$.

Definition 4.3.13. We define the *foundation*, $\operatorname{found}(\alpha)$, to be $\operatorname{found}(\alpha) := \{\alpha' \in \mathcal{A} : \operatorname{supp}(\alpha) \cap \operatorname{supp}(\alpha') \neq \emptyset\}$. Similarly, define $\operatorname{found}(A') = \bigcup_{\alpha \in \mathcal{A}'} \operatorname{found}(\alpha)$.

Definition 4.3.14. Finally, we define the *roof*, $\operatorname{roof}(x)$, to be $\operatorname{roof}(x) = \operatorname{found}(\operatorname{cosupp}(x))$ and $\operatorname{roof}(X) = \bigcup_{x \in X} \operatorname{roof}(x)$.

Proposition 4.3.15. *Let R be an ordered division ring, X, Y^* be left R vector spaces, $A \in \operatorname{Hom}_R(X, Y^*)$ and $\mathcal{A}$ be the induced collection of row-like projection maps (Definition 3.3.6). Then, given $\mathcal{A}' \subseteq \mathcal{A}$, where $0 \neq \alpha \in \mathcal{A}'$, then $\mathcal{A}' \subseteq \operatorname{found}(\mathcal{A}')$.*

Proof. Given $\alpha \in \mathcal{A}'$ $\operatorname{supp}(\alpha) \neq \emptyset$ so $\operatorname{supp}(\alpha) \cap \operatorname{supp}(\alpha) \neq \emptyset$ and $\alpha \in \operatorname{found}(\mathcal{A}')$. □

Corollary 4.3.16. *Let $x \in \mathcal{X}$, if $\operatorname{roof}(x) \neq \emptyset$, then $\operatorname{cosupp}(x) \subseteq \operatorname{roof}(x)$.*

Theorem 4.3.17 (Extendable Farkas' Lemma)**.** *Let R be an ordered division ring, X, Y^* be left R vector spaces. Also let $\mathbf{c} \in \operatorname{Hom}_R(X, R), A \in \operatorname{Hom}_R(X, Y^*)$ such that:*

- *$|\operatorname{roof}(x)| < \infty$ for each $x \in \mathcal{X}$.*
- *Let $X_1, X_2 \subseteq \mathcal{X}$, and $\mathbf{c}_i = \mathbf{c}|_{\operatorname{span}(X_i)}$ such that we may express $\mathbf{c}_i$ to be a non-negative linear combination of α_i. Then for a fixed μ_1, we may choose coefficient function μ_2*

such that $\mu_1|_{X_1 \cap X_2} = \mu_2|_{X_1 \cap X_2}$*, and*

$$\mathbf{c}_i = \sum_{\alpha_k \in \text{cosupp}(X_i)} \alpha_k \mu_i(\alpha_k).$$

Then the following are equivalent:

(a) The implication $A(\vec{x}) \leq 0 \implies \mathbf{c}(\vec{x}) \leq 0$ *holds for each* $\vec{x} \in X$*.*

(b) We may write $\mathbf{c} = \sum_{i \in Y^*} \alpha_i \cdot u_i$ *where* $u_i \in R_\oplus$*.*

Proof. We proceed via Zorn's Lemma. We define a family of pairs $\mathcal{F}$ where the (X', μ'), where $X' \subseteq \mathcal{X}$ and $\mu' : \text{cosupp}(X') \to R_\oplus$ such that

$$\sum_{\alpha_i \in \text{cosupp}(X')} \alpha_i(x)\mu'(\alpha_i) = \mathbf{c}(x)$$

for each $x \in X' \cup \{x'' : \text{cosupp}(x'') \subseteq \text{roof}(X')\}$.

This family has a natural poset structure. We say the pairs $(X', \mu') \leq (X'', \mu'')$ if $X' \subseteq X''$ and $\mu''|_{\text{cosupp}(X')} = \mu'$. We then verify that this is a partial ordering.

- Clearly $X' \subseteq X', \mu'|_{\text{cosupp}(X')} = \mu'$. Thus $(X', \mu') \leq (X', \mu')$.
- If $(X', \mu') \leq (X'', \mu''), (X'', \mu'') \leq (X', \mu')$, then $X' \subseteq X'' \subseteq X'$ and $X' = X''$. Then $\text{cosupp}(X') = \text{cosupp}(X'')$ and $\mu' = \mu''$.
- If $(X', \mu') \leq (X'', \mu''), (X'', \mu'') \leq (X''', \mu''')$, then $X' \subseteq X'' \subseteq X'''$, and

$$\begin{aligned} \mu'''|_{\text{cosupp}(X')} &= (\mu'''|_{\text{cosupp} X'})|_{\text{cosupp}(X')} \\ &= \mu''|_{\text{cosupp}(X')} \\ &= \mu'. \end{aligned}$$

Thus $(X', \mu') \leq (X''', \mu''')$.

We also need to show that every chain in $\mathcal{F}$ has a maximal element. Let $\{(X_i, \mu_i)\}_{i \in I}$ be a chain. We claim that (X', μ') is a maximal element, where $X' := \bigcup_{i \in I} X_i$ and $\mu'(\alpha) := \mu_j(\alpha)$, where $\alpha \in \operatorname{cosupp}(X_j)$.

To verify that μ' is a well defined function, given an $\alpha \in \operatorname{cosupp}(X')$, and any μ_k, μ_ℓ such that $\alpha \in \operatorname{cosupp}(X_k) \cap \operatorname{cosupp}(X_\ell)$, it is the case that $k \leq \ell$ or $\ell \leq k$. Without loss of generality, say $k \leq \ell$, then

$$\mu_\ell(\alpha) = \mu_\ell|_{\operatorname{cosupp}(X_k)}(\alpha) = \mu_k(\alpha).$$

So given any (X_i, μ_i) in our chain, we see that $X_i \subseteq X'$, and $\mu'|_{\operatorname{cosupp}(X_i)} = \mu_i$. Thus (X', μ') is a maximal element for the chain.

Finally, to show that $\mathcal{F}$ is non-empty, we let $x \in \mathcal{X}$. Consider the roof of x, $\operatorname{roof}(x)$. We may assume without loss of generality, that $\operatorname{cosupp}(x) \neq 0$ for each $x \in X$, else there is an $x' \in X$ such that $\alpha(x') = 0$ for each $\alpha \in \mathcal{A}$, and since $\mathbf{c}$ is a linear combination of the α's, $\mathbf{c}(x') = 0$, and we may disregard it, as x' contributes to neither $\mathbf{c}$, nor the α's.

So consider $\operatorname{roof}(x)$ and $\operatorname{supp}(\operatorname{roof}(x))$. This induces a new program. We let X' be the span of $\operatorname{supp}(\operatorname{roof}(x))$, and let Y' be induced by the projection maps associated with $\operatorname{roof}(x)$, we let $\mathbf{c}' := \mathbf{c}|_{X'}$, and A' be induced by $\operatorname{roof}(x)$. Since $\operatorname{roof}(x)$ is finite, Y' is finite dimensional. Moreover, since the implication of (a) holds for X, it also holds for each subspace of X. Thus, we may use the finite version of the Farkas' Lemma to state that $\mathbf{c}'$ is a non-negative linear combination of $\alpha \in \operatorname{roof}(x)$. That is, if $\operatorname{roof}(x) = \{\alpha_1, \ldots \alpha_m\}$, $\mathbf{c}' = \sum_{i=1}^m \alpha_i \cdot u_i, u_i \in R_\oplus$.

Notice that $\operatorname{cosupp}(x) \subseteq \operatorname{roof}(x)$. Then given any $\beta \in \mathcal{A} \backslash \operatorname{roof}(x)$, $\beta(x) = 0$. Thus given any linear combination $\gamma := \sum_{\alpha_j \in \mathcal{A}} v_j \alpha_j$, if $v_j = u_j$ for each $\alpha_j \in \operatorname{cosupp}(x)$, hence see that:

$$\begin{aligned}
\gamma(x) &= \sum_{\alpha_i \in \operatorname{cosupp}(x)} \alpha_i(x) \cdot v_i + \sum_{\alpha_j \notin \operatorname{cosupp}(x)} \alpha_j(x) \cdot v_j \\
&= \sum_{\alpha_i \in \operatorname{cosupp}(x)} x_i \cdot u_i + 0 \\
&= \mathbf{c}'(x).
\end{aligned}$$

If there was a $y \in X$ such that $\operatorname{cosupp}(y) \subseteq \operatorname{roof}(x)$, then for the same linear combinations, $\mathbf{c}'(y) = \gamma(y)$ as well. We notice that all such y would be in the support of the roof of x.

So, define X_0 to be $y \in X$ such that $\operatorname{cosupp}(y) \subseteq \operatorname{cosupp}(x)$. Since $x \in X_0, X_0$ is not empty. For each $\alpha_i \in \operatorname{cosupp}(x)$, we define $\mu_0(\alpha_i) := u_i$. Finally, we notice that given $y \in \mathcal{X}$ such that $\operatorname{cosupp}(y) \subseteq \operatorname{cosupp}(x)$,

$$\mathbf{c}(y) = \mathbf{c}'(y) = \sum_{\alpha_i \in \operatorname{roof}(x)} \alpha_i(y) \cdot u_i = \sum_{\alpha_i \in \operatorname{cosupp}(x)} \alpha_i(y) \cdot \mu_0(\alpha_i).$$

Thus $\mathcal{F}$ is not empty.

So, we may use Zorn's Lemma and suppose that (X^*, μ^*) is a maximal element of $\mathcal{F}$. Also suppose that $X^* \neq X$. Let $x \in X^* \backslash X$. If $\operatorname{cosupp}(x) \subseteq \operatorname{roof}(X^*)$, then, μ^* is already defined for each $\alpha \in \operatorname{cosupp}(x)$, and we may simply replace X^* with $X^* \cup \{x\}$, contradicting the maximality of X^*.

So consider $x \in X \backslash X^*$ such that $\operatorname{cosupp}(x) \backslash \operatorname{roof}(X^*) \neq \emptyset$. We once again, we to use the finite Farkas' Lemma to create a new maximal element, contradicting the maximality of (X^*, μ^*).

We define X' to be the span of $\operatorname{supp}(\operatorname{roof}(x))$. We also define Y' to be the subspace of Y induced by the projection maps associated with $\mathcal{A}' := \operatorname{roof}(x)$. We define A' to be the map induced by $\mathcal{A}'$. Once again, we notice that Y' is finite dimensional.

Again, hypothesis (a) of the Farkas' Lemma is satisfied, as X' is a subspace of X, $\mathcal{A}'$ is composed of all projections that are non-zero on X'. Thus, given $\mathcal{A}' = \{\alpha_1, \ldots, \alpha_n\}$, we have that

$$\mathbf{c}' = \sum_{\alpha_i \in \mathcal{A}'} \alpha_i \cdot u_i, u_i \in R_{\oplus},$$

and $\mathbf{c}' = \mathbf{c}|_{X'}$.

In fact, by the hypothesis, we have that we may select such a u_i such that $u_i = \mu^*(\alpha_i)$ when α_i is in cosupp(X^*). Thus, consider

$$\mu_1 : \text{cosupp}(X^*) \cup \text{cosupp}(x) \to R$$

where $\mu_1|_{\text{cosupp}(X^*)} = \mu^*$ and $\mu_1(\alpha_i) = u_i$ for $\alpha_i \in \text{cosupp}(x)$. This is well defined since we have chosen the u_i to agree on cosupp(X_i). Define $\mathbf{c}_1$ to be

$$\mathbf{c}_1(\vec{x}) := \sum_{\alpha_i \in \text{cosupp}(\mathcal{X} \cup \{x\})} \alpha_i(\vec{x}) \mu_1(\alpha_i).$$

This is a well defined linear map. Notice also that for any $z \in \mathcal{X}$ such that cosupp$(z) \subseteq$ cosupp(x), $z \in X'$ and

$$\mathbf{c}_1(z) = \sum_i \alpha_i(z)\mu_1(\alpha_i) = \sum_{\alpha_i \in \text{cosupp}(x)} \alpha_i \mu_1(\alpha_i) = \mathbf{c}'(z) = \mathbf{c}(z).$$

Thus $(X^* \cup \{x\}, \mu_1) > (X^*, \mu^*)$ which violates the maximality of (X^*, μ^*). Thus $X^* = \mathcal{X}$, and $\mathbf{c}^* = \mathbf{c}$, which is a non-negative linear combination of the α_i. This completes the proof

$\square$

This provides an alternative proof of Theorem 4.3.10.

Proof. Since each α_i is linearly independent of the other α_j, we can choose a basis $\mathcal{X}$ for X such that $x_i \in \mathcal{X}$ is supported by at most one of the $\alpha_i \in \mathcal{A}$, and by the Remark 4.3.2, we may assume that it is supported by exactly one $\alpha_i \in \mathcal{A}$. Thus $\operatorname{roof}(x_i) = \{\alpha_i\}$. So for each $\mathbf{c}'$ defined on $X' < X$, given a $x_j \in \mathcal{X}, x_j \notin X'$, we can extend $\mathbf{c}'$ to x_j by defining $\mathbf{c}'' := \mathbf{c}' + \alpha_j \cdot (\alpha_j(x_j)^{-1}\mathbf{c})$.

□

Next we describe ways to "partition" an affine programming problems into sub-problems, and show that the Farkas' Lemma holds if and only if each of the sub-problems hold.

Proposition 4.3.18. *Let R be an ordered division ring, let X, Y be left R vector spaces. Then, let $A \in \mathrm{Hom}_R(X, Y)$, where $\mathcal{A}$ is the collection of row-like projection maps (Definition 3.3.6). We then let $L := \{\alpha_k \in \mathcal{A}\}$ such that L is maximal with respect to linear independence. Then for each $\alpha_\ell \in L$, there is a $x_\ell \in X$ such that $\alpha_k(\ell) = \delta_{\ell k}$ (where $\delta_{\ell k}$ is the Kronecker delta), and the collection of x_ℓ forms a basis for X.*

Proof. We first construct the collection $\{x_\ell\}$. Notice that for each α_ℓ, α_ℓ is linearly independent of each $\alpha_k \in L, k \neq \ell$. Thus, by the Fundamental Lemma (Lemma 4.2.1) there is a $x_\ell \in \bigcap_{k \neq \ell} \mathrm{Ker}(\alpha_k) \backslash \mathrm{Ker}(\alpha_\ell)$. Without loss of generality we may choose x_i such that $\alpha_\ell(x_i) = 1$.

To see that these x_i are linearly independent, suppose we may write $0 = \sum_{i \in L' \subseteq L} x_i \cdot r_i, r_i \in R$. Then for each index k,

$$0 = \alpha_k(0) = \alpha_k(\sum_{i \in L' \subseteq L} r_i \cdot x_i) = r_k.$$

Thus each r_k must be 0.

To show that these x_i span X, we then let $0 \neq \vec{x} \in X$, and consider that $\alpha_i(\vec{x})$ will be non-zero for only finitely many elements of $\alpha_i \in \mathcal{A}$. Without loss of generality, we label $\alpha_1, \dots \alpha_m$ to be the elements of L where $\alpha_i(\vec{x})$ is non-zero.

We first note that such a collection of α_i exist, else, if $\alpha_k(\vec{x}) = 0$ for each $\alpha_k \in L$, then since each α_j is a linear combination of the α_k, $\alpha_j(\vec{x}) = 0$ as well, and by our convention, $\vec{x} = 0$.

We then claim

$$\vec{x} = x' := \sum_{i=1}^{m} \alpha_i(\vec{x}) \cdot x_i,$$

where i indexes the α_i, x_i such that $\alpha_i(\vec{x}) \neq 0$. Certainly for each such $\alpha_i, \alpha_i(x') = \alpha_i(\alpha_i(\vec{x}) \cdot x_i) = \alpha_i(\vec{x})$. Then since each $\alpha_j \in \mathcal{A}$ is a linear combination

$$A(\vec{x}) = A\left(\sum_{\alpha_\ell \in L} \alpha_\ell(\vec{x}) \cdot x_\ell\right) = \sum_{i=1}^{m} \alpha_i(\vec{x}) \cdot v_i = \sum_{i=1}^{m} \alpha_i(x') \cdot v_i = A(x').$$

Thus $A(\vec{x}) = A(x')$, and by our convention, A is injective. Thus $\vec{x} = x'$, and the collection x_ℓ spans X.

$\square$

Definition 4.3.19. Given any collection of vectors over R, $C := \{a_i\}$, we may define an equivalence relationship between $a_1 \sim a_2$ if there is a subset $C' \subseteq C \backslash \{a_1\}$ such that $a_2 \in C'$ and

$$a_1 = \sum_{a_i \in C'} a_i r_i \cdot, r_i \neq 0.$$

To see that this is an equivalence relationship, notice that:

- $a_1 = a_1 \cdot 1$.
- If $a_1 = \sum_{a_i \in C'} a_i r_i$, then $a_2 = a_1 r_2^{-1} - \sum_{a_i \in C' \backslash \{a_2\}} a_i r_i r_2^{-1}$.

- If $a_1 = \sum_{a_i \in C'} a_i r_i$ and $a_2 = \sum_{a_j \in C''} a_j r_j$, where $a_3 \in C''$. Then either $a_3 \in C''$, or

$$a_1 = \sum_{a_i \in C' \setminus \{a_2\}} a_i r_i + \left(\sum_{\alpha_j \in C''} a_j r_j \} \right) r_2.$$

We call $[a_1]$ to be the *linear dependence class* of a_1, and recall that these classes partition C.

Remark 4.3.20. We notice then that given $A \in \mathrm{Hom}_R(X, Y^*)$, $\mathcal{A}$ can be partitioned into linear dependence classes. (Recall $\mathcal{A}$ is the collection of row-like projection Definition 3.3.6).

Proposition 4.3.21. *Given R an ordered division ring, X, Y^* R-vector spaces with $\dim_R(Y^*) = \infty$, $A \in \mathrm{Hom}_R(X, Y^*)$, one can find a basis for X such that $\mathrm{supp}([\alpha_1]) \cap \mathrm{supp}([\alpha_2]) = \emptyset$ when $[\alpha_1] \neq [\alpha_2]$.*

Proof. Given a partition of $\mathcal{A}$ (Recall $\mathcal{A}$ is the collection of row-like projection maps Definition 3.3.6) via linear dependence classes, consider a collection $L \subseteq \mathcal{A}$ where L is maximal with respect to linear independence. Such define the classes $[\tilde{\alpha_i}] = [\alpha_i] \cap L$, and define $B_i := \{x_k : \alpha_k \in [\tilde{\alpha_k}]\}$, where the x_k are constructed as in Proposition 4.3.18. Since the classes $[\alpha_i]$ partition $\mathcal{A}$, the $[\tilde{\alpha_i}]$ partition L, and thus the disjoint union of the B_i form a basis for X by Proposition 4.3.18.

Thus consider any pair $[\alpha_i]$, B_i, we claim that $\mathrm{supp}([\alpha_i]) = B_i$, with respect to the basis B. Notice that given $x_k \in B_k$, there is a $\alpha_k \in [\alpha_i]$ such that $\alpha_k(x_k) = 1$. Thus $B_i \subseteq \mathrm{supp}([\alpha_i])$.

Conversely, for any $x_j \notin B_i$, $\alpha_k(x_j) = 0$ for each $\alpha_j \in [\tilde{\alpha_i}]$. Since the remaining elements of $[\alpha_i]$ are linear combinations of these α_k, these will be zero as well, and so $\mathrm{supp}([\alpha_i]) \subseteq B_i$.

We then note that distinct classes $[\alpha_1], [\alpha_2]$ induce distinct subsets of the basis B_1, B_2, that are disjoint. This completes the proof. $\square$

Remark 4.3.22. By our various conventions, we have partitioned the row-like projection maps $\mathcal{A}$, based on linear dependence, and in doing so induced a partition of a basis of X. We also see that partitioning $\mathcal{A}$ is equivalent to partitioning the spanning projection maps of Y^*, or equivalently, partitioning the basis of Y.

In other words, given a map $\alpha_1 \in A$, we get a collection of linearly dependent maps $[\alpha_1]$, which induce both a partition of a basis of X, B_1, and of spanning projection maps of the codomains $\{\hat{y}_i : \alpha_i \in [\alpha_1]\}$. This allows us to deconstruct the program into a collection of subprograms, with $X_1 := \text{span}(B_1), Y_1 := \coprod_{\alpha_i \in [\alpha_1]} R_i$, A_1 the linear transformation induced by $\mathcal{A}_1 := [\alpha_1]$ and $\mathbf{c}_1 := \mathbf{c}|_{X_1}$.

Definition 4.3.23. Let R be an ordered division ring, let X, Y^* be left R vector spaces, and let $A \in \text{Hom}_R(X, Y)$ and $\mathbf{c} \in \text{Hom}(X, R)$. Label the linear dependence classes of A, (Definition 4.3.19) $\mathcal{A}_j, j \in J$. For each index j, we define the *subprogram* $(X, Y^*, A, \mathbf{c})_j$ to be $(X_j, Y_j^*, A_j, \mathbf{c}_j)$ described above. The collection $\{(X, Y^*, A, \mathbf{c})_j\}_{j \in J}$ is the *linear dependence decomposition* of $(X, Y^*, A, \mathbf{c}_j)$.

Proposition 4.3.24. *Let R be an ordered division ring, let X, Y be left R vector spaces, and let $A \in \text{Hom}_R(X, Y)$ and $\mathbf{c} \in \text{Hom}(X, R)$. If $\mathbf{c} = \lambda \circ A$ for some $\lambda \in \text{Hom}_R(Y^*, R)$, then $\mathbf{c}_j = \lambda \circ A_j$. Equivalently, $\mathbf{c}_j$ is a linear combination of $\alpha_k \in \mathcal{A}_j$.*

Proof. If $\mathbf{c} = \lambda \circ A$, then given $\vec{x}_j \in X_j$, each $\alpha_i(\vec{x}) = 0$ for $\alpha_i \notin \mathcal{A}_j$. Thus $\mathbf{c}$ is a linear combination of the $\alpha_i \in A_j$, and so $\mathbf{c}_j = \mathbf{c}|_{X_j} = \lambda \circ A_j$.

$\square$

Recall the Generalized Farkas' Lemma:

Given R an ordered ring, X, Y^* be left R vector spaces, $A \in \text{Hom}_R(X, Y^*)$, $\mathbf{c} \in \text{Hom}_R(X, R)$. Then the following are equivalent:

(a) For each $\vec{x}, \in X$, if $A(\vec{x}) \leq_{Y^*} 0$, then $\mathbf{c}(\vec{x}) \leq_R 0$.

(b) The map $\mathbf{c} = \sum_i \alpha_i \cdot u_i, u_i \geq_R 0$. Equivalently, there is a $\mathbf{y} \in Y^{**}_{\oplus}$ such that $\mathbf{y} \circ A = \mathbf{c}$.

Proposition 4.3.25. *Given X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*)$, $\mathbf{c} \in \mathrm{Hom}_R(X, R)$., we have that condition (a) of the Generalized Farkas' Lemma holds for $(X, Y^*, A, \mathbf{c})$, if and only of it holds for each subprogram (Definition 4.3.23): $(X, Y^*, A, \mathbf{c})_j$. Similarly with condition (b).*

Proof. Suppose (a) holds for each $(X, Y^*, A, \mathbf{c})_j$. Then given any $\vec{x} \in X$ such that $A(\vec{x}) \leq_{Y^*} 0$, we may write $\vec{x} = \sum_{j \in J} x_j, x_j \in X_j$ (with only finitely many $x_j \neq 0$.) Then since the $\mathcal{A}_j$ partition the α_j, $A(\vec{x}) \leq_Y 0$ if and only if $A_j(x_j) \leq_{Y_j} 0$. Thus each $\mathbf{c}_j(x_j) \leq_R 0$. Then it follows that

$$\mathbf{c}(\vec{x}) = \sum_{j \in J} \mathbf{c}_j(x_j) \leq_R 0.$$

If (b) holds for each index j, $\mathbf{c}_j$ can be written as $\lambda_j \circ A_j$, then we define $\mathbf{c} := \lambda \circ A$, where $\lambda = \oplus_{j \in J} \lambda_j$. Since each λ_i maps the basis elements $\iota_i(1)$ to non-negative elements of R, so to does λ, and λ is non-negative.

□

Theorem 4.3.26 (Decomposable Farkas' Lemma)**.** *Given X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*)$, $\mathbf{c} \in \mathrm{Hom}_R(X, R)$, such that the conclusion of the Farkas' Lemma holds for each induced subprogram $(X, Y^*, A, \mathbf{c})_j$, (that is: $(a)_j \iff (b)_j$) then the following are equivalent:*

(a) If $A(\vec{x}) \leq_{Y^*} 0$ for any $\vec{x} \in X$ i.e. $\alpha_i(\vec{x}) \leq_R 0$ for a choice $\vec{x}$, then $\mathbf{c}(\vec{x}) \leq_R 0$

(b) The map $\mathbf{c} = \sum_i \alpha_i \cdot u_i, u_i \geq_R 0$. In other words, there is a $\mathbf{u} \in \mathrm{Hom}_R(Y^*, R)_{\oplus}$ such that $\mathbf{u} \circ A = \mathbf{c}$.

Proof. If, (a) does not hold for any of the subprograms of $(X, Y^*, A, \mathbf{c})$, then there is an index k and a $\vec{x} \in X_k$ such that $A_k(\vec{x}) \leq_{Y_k} 0$ but $\mathbf{c}_k(\vec{x}) > 0$. Since $A_k(\vec{x}) \leq_{Y_k} 0$ only if $A(\vec{x}) \leq_{Y^*} 0$

and $\mathbf{c}(\vec{x}) = \mathbf{c}_k(\vec{x})$, (a) fails for $(X, Y^*, A, \mathbf{c})$ and the implication $(a) \implies (b)$ holds vacuously.

Otherwise (a) holds for each $(X, Y^*, A, \mathbf{c})_j$, and so (b) holds for each subprogram as well. Then by Proposition 4.3.25, (b) holds for $(X, Y^*, A, \mathbf{c})$ as well.

□

Corollary 4.3.27. *If given $(X, Y^*, A, \mathbf{c})$, each $|\mathcal{A}_j| < \infty$, then the Farkas' Lemma holds.*

Proof. By Theorem 4.3.4, the Farkas' Lemma holds for each subprogram. Then by Corollary 4.3.27, it hold for the main program as well. □

Corollary 4.3.28. *This gives an alternative proof to Theorem 4.3.10*

Notice that the converse to Corollary 4.3.27 is **not** true.

Example 4.3.29. We, we let $R = \mathbb{R}$, and let $X := \mathbb{R}^{\mathbb{N}} \oplus \mathbb{R}^2, Y^* := \mathbb{R}^{\mathbb{N}} \oplus \mathbb{R}$. We then define $A_1 : R^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ and $\mathbf{c}$ to be the the counterexample to the Farkas' Lemma (Example 4.3.5):

$$A_1 : \begin{array}{r} \alpha_0 = \\ \alpha_1 = \\ \alpha_2 = \\ \alpha_3 = \\ \vdots \end{array} \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & \cdots \\ -1 & 0 & 0 & 0 & \cdots & 0 & \cdots \\ 0 & -1 & 0 & 0 & \cdots & 0 & \cdots \\ 0 & 0 & -1 & 0 & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \end{pmatrix}$$

$$\mathbf{c}_1 : \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n & \cdots \end{pmatrix}.$$

We also define $A_2 : \mathbb{R}^2 \to \mathbb{R}$ via $(w_1, w_2) \mapsto w_1 - w_2$, and $\mathbf{c} : \mathbb{R}^2 \to \mathbb{R}$ via $(w_1, w_2) \mapsto w_1 + w_2$.

We note that given $A((0_{X_1}, (1, 1))) = (0_{Y_1}, 0)$, but $\mathbf{c}((0_{X_1}, (1, 1))) = 2$ Since (a) does not hold, $(a) \implies (b)$ is vacuously true. However, the subprogram $(X, Y, A, \mathbf{c})_1$ does not satisfy the Farkas' Lemma.

Proposition 4.3.30. *Given R an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y)$, $\mathbf{c} \in \mathrm{Hom}_R(X, R)$, we may choose a collection of $\hat{y}_i : Y \to R_i, \iota_i : R_i \to Y$ such that the Farkas' Lemma holds.*

Proof. We notice that $A(X)$ is a linear subspace of Y^*, and thus, we may choose $\hat{y}_i, \iota_i$ (or equivalently, a basis for Y) such that a sub collection of these maps (basis elements) span $A(X)$. We note that by our convention, we may ignore any non-zero α_i, and so with respect to these maps (this basis), we may assume that A is onto. By Proposition 4.2.5, the α_i are linearly independent and by Theorem 4.3.10, the Generalized Farkas' Lemma holds. □

Theorem 4.3.31. *Let R be an ordered division ring, let X, Y^* be left R vector spaces, let $A \in \mathrm{Hom}_R(X, Y)$, and $\mathbf{c} \in \mathrm{Hom}_R(X, R)$, such that $\dim_R(X_j), |\mathcal{A}_j| < \infty$ for the each of the induced subprograms (Definition 4.3.23). Then the Farkas' Lemma holds for $(X, Y \oplus X, A \oplus -I, \mathbf{c})$. Equivalently, the Farkas' Lemma holds when we impose a non-negative constraint on $\vec{x}$.*

Proof. If each $\dim_R(X_j), |\mathcal{A}_j| < 0$, then when we consider the program $(X, Y^* \oplus X, A \oplus (-I), \mathbf{c})$, we note that given any subprogram $(X, Y \oplus X, A \oplus (-I), \mathbf{c})_j$, $\dim_R(Y_j \oplus X_j) < \infty$, and so by Theorem 4.3.4 the Generalized Farkas' Lemma holds for each subprogram. Thus by Corollary 4.3.27, it holds for the given program as well.

□

4.4 Conclusion

In this chapter, we followed the results of [Bar07], and proved some generalizations of technical results he proved in his paper. We then used these technical results to show that the generalized Farkas' Lemma 4.3.1 holds under several finite-type circumstances (Proposition 4.3.10, Corollary 4.3.27) proving a conjecture of Bartl's as well. We also show how the Farkas' Lemma may be extended with Zorn's Lemma (Theorem 4.3.17). In this chapter, we assumed nothing more about our ring structure other than that they were ordered division rings. In particular, we did not assume the commutativity of a product.

Chapter 5

Theorems of the Alternative and Duality Results

5.1 Introduction

Here, we state some consequences of the Farkas' Lemma, which are dubbed **Theorems of the Alternative**. They are named this way since historically, they deal with the circumstances where there either exists a vector in the primal solution space X with certain properties, or a functional in the dual solution space Y^* with another property, but not both. However, some of the implications in the infinite case do not hold, or only hold with additional hypothesis. In this section, we assume that R is an ordered division ring, X, Y^* be R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \mathbf{c} \in \mathrm{Hom}_R(X, R)$. We shall also use the convention that $X^* := \mathrm{Hom}_R(X, R)$ and $Y^{**} := \mathrm{Hom}_R(Y^*, R)$. Moreover, we shall assume that the conclusion of the Generalized Farkas' Lemma holds. This section also follows loosely the structure of [Bar07].

Additionally, we will use these theorems to prove some of the duality results of affine optimization. In particular, we prove the generalized facts Complementary Slackness 5.3.4 and the Strong Duality Theorem 5.3.7, as well as prove partial results that lead to the Existence-Duality Theorem.

5.2 Theorems of the Alternative.

Lemma 5.2.1. *Let R be a division ring, X, Y^* be left R vector spaces, $\vec{b} \in Y^*$ and $\mathbf{c} \in$ $\mathrm{Hom}_R(X, R)$. Then the system of inequalities*

$$A(\vec{x}) \leq_{Y^*} \vec{b}$$

has no solution if and only if there is a $\lambda \in Y^{**}$, such that $\lambda \geq_{Y^{**}} 0, \lambda \circ A = 0$ and $\lambda(\vec{b}) < 0$.

Proof. Let $X' := X \oplus R$, $Y' = Y^*$, and $A' : X' \to Y'$, $(\vec{x}, r) \mapsto A(\vec{x}) - r\vec{b}$. Then $A(\vec{x}) \leq_{Y^*} \vec{b}$ has a solution if and only if there is a solution to the system

$$A'((\vec{x}, r)) \leq_{Y'} 0$$

where $r > 0$.

We then define $\gamma : X' \to R$ via $\gamma((\vec{x}, r)) = r$. This will play the role of $\mathbf{c}$ in the Farkas' Lemma.

Then, we suppose that $A(\vec{x}) \leq_{Y^{**}} \vec{b}$ has no solution. By our previous observation, this occurs if and only if any solution to $A'((\vec{x}, r)) \leq_{Y}' 0$ implies $\gamma((\vec{x}, r)) \leq 0$.

Thus hypothesis (a) of the Farkas' Lemma is satisfied, where A' plays the role of A, and γ the role of $\mathbf{c}$. This implies that there is a $\lambda \in Y_{\oplus}^{**}$ such that $\lambda \circ A' = r$. So $\lambda(A(\vec{x})) = 0$ for each $\vec{x} \in X$, and $\lambda(\vec{b}) = \lambda(0 - (-1)\vec{b}) = -1 < 0$.

$\square$

Corollary 5.2.2. *The system of equations*

$$A(\vec{x}) = \vec{b}$$

has no solutions if and only if there is a $\lambda \in Y^{**}$*, such that :* $\lambda \circ A = 0$ *and* $\lambda(\vec{b}) < 0$.

Proof. Let $Y_i := Y^*$, and $Y' := Y_1 \oplus Y_2$, and $A' := A \oplus -A$. Then it follows that:

$$\begin{aligned} A(\vec{x}) &= \vec{b}, \text{ has no solution if and only if,} \\ A(\vec{x}) &\leq_{Y^*} \vec{b}, \text{ has no solution, and} \\ -A(\vec{x}) &\leq_{Y^*} -\vec{b}, \text{ has no solution, in other words} \\ A'(\vec{x}) &\leq_{Y'} (\vec{b}, -\vec{b}), \text{ has no solution.} \end{aligned}$$

So by Lemma 5.2.1, there is a $\lambda' \in \mathrm{Hom}_R(Y', R), \lambda'$ in the induced non-negative cone of $\mathrm{Hom}_R(Y', R)$ such that $\lambda'(A') = 0, \lambda'((\vec{b}, -\vec{b})) < 0$. Let $\lambda_i := \lambda'|_{Y_i}$, and let $\lambda := \lambda_1 - \lambda_2$. We may verify that

$$\lambda \circ A = \lambda_1 \circ A - \lambda_2 \circ A = \lambda' \circ (A \oplus -A) = 0,$$

and

$$\lambda(\vec{b}) = \lambda_1(\vec{b}) - \lambda_2(\vec{b}) = \lambda'((\vec{b}, -\vec{b})) < 0.$$

Conversely, If $\lambda \circ A = 0, \lambda(\vec{b}) \neq 0$, then $\vec{b} \notin A(X)$.

$\square$

This gives us the following useful technical result:

Proposition 5.2.3. *Let R be an ordered division ring, X, Y_1, Y_2 be left R vector spaces. Define*

$A_i \in \mathrm{Hom}_R(X, Y_i)$, $\vec{b}_i \in Y_i$. *Then the system of linear inequalities*

$$A_1(\vec{x}) = \vec{b}_1, A_2(\vec{x}) \leq_{Y_2} \vec{b}_2$$

has no solution then there is a $\lambda_i : Y_i \to R, \lambda_2 \in (Y_2^*)_\oplus$ *(the non-negatives of* $\mathrm{Hom}_R(Y_2, R)$*), such that*

$$\lambda_1 \circ A_1 + \lambda_2 \circ A_2 = 0$$

and $\lambda_1(\vec{b}_1) + \lambda_2(\vec{b}_2) <_R 0$.

Proof. Similarly, we consider $A' : X \to Y_1 \oplus Y_1 \oplus Y_2$ via $A' := A_1 \oplus -A_1 \oplus A_2$ bounded by $(\vec{b}_1, -\vec{b}_1, \vec{b}_2)$. The result then follows similarly as above.

□

The next result is a semi-generalization of Motzkin's Theorem [Mot36].

Theorem 5.2.4 (Motzkin's Theorem)**.** *Let* $A_1 \in \mathbb{R}^{n_1 \times m}$ *and* $A_2 \in \mathbb{R}^{n_2 \times m}$. *Then the system of linear inequalities*

$$A_1\vec{x} \leq_{\mathbb{R}^{n_1}} 0, A_2\vec{x} \leq_{\mathbb{R}^{n_2}} 0$$

has no solution for any $\vec{x} \in \mathbb{R}^m$ *if and only if there are vectors* $\lambda_i \in \mathbb{R}^{n_1}_\oplus, \lambda_2 \in \mathbb{R}^{n_2}_\oplus$ *such that*

$$\lambda_1^\top A_1 + \lambda_2^\top A_2 = 0$$

and $\lambda_1 \neq 0$.

Notice however, that this result has being partitioned into two separate implications. The original statement of Motzkin's Theorem presumed that the co-domain of A, Y^*, was a finite dimensional space, and as part of the hypothesis, that each $\alpha_i(\vec{x})$ were strictly less than the co-ordinate b_i. However, as Y^* is potentially infinite dimensional, finitely many of the $\alpha_i(\vec{x})$

and b_i will be non-zero for any choice of $\vec{x}, \vec{b}$, and this hypothesis can only hold if Y^* is finite dimensional as well. Thus we prove a more general version of this theorem where the two statements are are not generally equivalent.

Theorem 5.2.5 (Generalized Motzkin's Theorem Part 1)**.** *Let R be an ordered division ring, X, Y_1, Y_2 be left R vector spaces. Let $A_i \in \mathrm{Hom}_R(X, Y_i)$ for $i = 1, 2$. Then if the system of linear inequalities*

$$A_1(\vec{x}) <_{Y_1} 0, A_2(\vec{x}) \leq_{Y_2} 0$$

has no solution then there is a $\lambda_i : Y_i \to R, \lambda_i \in (Y_i^*)_\oplus, \lambda_1 \neq 0$, such that

$$\lambda_1 \circ A_1 + \lambda_2 \circ A_2 = 0.$$

Proof. We define $Y := Y_1 \oplus Y_2$, $A : X \to Y$ to be $A_1 \oplus A_2$. Then given any $\vec{b} \in (Y_1)_\oplus$, we define $A' : X \oplus R \to Y, (\vec{x}, r) \mapsto (A_1(\vec{x}) - r\vec{b}, A_2(\vec{x}))$. Then the system

$$A_1(\vec{x}) <_{Y_1} 0, A_2(\vec{x}) \leq_{Y_2} 0$$

has no solution, only if $A'((\vec{x}, r) \leq_{Y^*} 0$ implies $r \leq_R 0$.

We then define $\gamma : X \oplus R \to R$ via $(\vec{x}, r) \mapsto r$. Then there is a $\lambda \in Y_\oplus^{**}$ such that $\gamma = \lambda \circ A'$. As before, $\lambda((A_1(\vec{x}) - r\vec{b}), A_2(\vec{x})) = \lambda((A(\vec{x}), 0) - r\lambda(\vec{b}, 0) + \lambda(0, A_2(\vec{x})) = r$. Thus by defining $\lambda_i := \lambda|_{Y_i}$, we have that $\lambda_1 \circ A_1 + \lambda_2 \circ A_2 = 0$, and since $\lambda((\vec{b}, 0)) = 1$, $\lambda_1 \neq 0$.

□

Theorem 5.2.6 (Generalized Motzkins Theorem Part 2)**.** *Let R be an ordered division ring, X, Y_1, Y_2 be left R vector spaces. Define $A_i \in \mathrm{Hom}_R(X, Y_i)$. Then the system of linear inequalities*

$$A_1(\vec{x}) <_{Y_1} 0, A_2(\vec{x}) \leq_{Y_2} 0$$

has no solution if then there is a $\lambda_i : Y_i \to R, \lambda_i \in (Y_i^*)_\oplus, \lambda_1 \neq 0$, *such that*

$$\lambda_1 \circ A_1 + \lambda_2 \circ A_2 = 0$$

and $\mathrm{Ker}(\lambda_1) = 0$. *(Where* $(Y_i)^*_\oplus$ *denotes the non-negatives of* $\mathrm{Hom}_R(Y_i, R)$*).*

Proof. Suppose under these hypothesis there was a $x' \in X$ such that:

$$A_1(x') <_{Y_1} 0, A_2(x') \leq_{Y_2} 0$$

Then consider that $A_i(-x') \in (Y_i)_\oplus$. Thus $\lambda_i(A_i(x')) = -\lambda_i(A(-x')) \leq_R 0$, and since λ_1 has Kernel $\{0\}$, and $A_1(x') \neq 0$, $\lambda_1(A_1(x')) < 0$ and $\lambda_1 \circ A_1 + \lambda_2 \circ A_2 <_R 0$ which is a contradiction.

□

5.3 Duality Results

5.3.1 Partial results for the Existence-Duality Theorem

Here we can prove some facts that relate to the Existence-Duality Theorem **??**.

Proposition 5.3.1 (Primal infeasible implies dual unbounded)**.** *Let* R *be an ordered division ring, let* X, Y^*, *be left* R *vector spaces,* $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$.

Then, if (P) (maximize $\mathbf{c}(\vec{x})$ *subject to:* $A(\vec{x}) \leq_{Y^*} \vec{b}$*) is infeasible, then (D) (minimize* $\mathbf{y}(\vec{b})$ *subject to:* $\mathbf{y} \circ A = \mathbf{c}$*) is infeasible or is unbounded.*

Proof. Suppose (D) is feasible. By Lemma 5.2.1 (P) infeasible, or

$$A(\vec{x}) \leq_{Y^*} 0$$

has no solution only if there is a $\lambda \in Y_{\oplus}^{**}$ such that $\lambda(\vec{b}) = -a <_R 0$.

Then, given a feasible solution $\mathbf{y} \in Y_{\oplus}^{**}$, and any $r \in R_{\oplus}$, $\mathbf{y} + \lambda \cdot r \in Y_{\oplus}^{*}$, but $\mathbf{y} + \lambda \cdot r(\vec{b}) = \mathbf{y}(\vec{b}) - a \cdot r$, and since r is arbitrary, (D) is unbounded. □

Proposition 5.3.2 (Dual infeasible implies primal unbounded)**.** *Let R be an ordered division ring, let X, Y^*, be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$.*

Then, if (D) (minimize $\mathbf{y}(\vec{b})$ subject to: $\mathbf{y} \circ A = \mathbf{c}$) is infeasible, then (P) (maximize $\mathbf{c}(\vec{x})$ subject to: $A(\vec{x}) \leq_{Y^*} \vec{b}$) is infeasible or is unbounded.

Proof. Suppose (P) is feasible. By the Farkas' Lemma (Lemma 4.3.1), $\mathbf{c}$ is not a non-negative linear combination of α_i and so (b) of the Farkas' Lemma fails. Thus (a) fails as well, and there is a $x' \in X$ such that $A(x') \leq_Y 0$ but $\mathbf{c}(x') = a >_R 0$. Thus given a feasible solution $\vec{x} \in X$, and $r \in R_{\oplus}$:

$$\begin{aligned} A(\vec{x} + r \cdot x') = A(\vec{x}) + rA(x') \quad &\leq_{Y^*} \quad \vec{b}. \\ \mathbf{c}(\vec{x} + r \cdot x') \quad &= \quad \mathbf{c}(\vec{x}) + r \cdot a. \end{aligned}$$

Since r is arbitrary, (P) is unbounded □

Remark 5.3.3. With the above results, and the weak Existence-Duality 2.4.26, we have established that if either the primal or dual program is infeasible, then the remaining program must also be infeasible or unbounded, and that ether program being unbounded forces the other program to be infeasible.

Thus with the current level of generalization, only the case where both programs are feasible and bounded has yet to be addressed.

5.3.2 Complementary Slackness and Strong Duality

In order to prove a generalization of the Strong Duality theorem, we introduce a generalization of one of our key duality results, the Generalized Complementary Slackness Theorem (Recall the classical Complementary Slackness Theorem, Theorem **??**).

Theorem 5.3.4 (Generalized Fact VI: Complementary Slackness Theorem)**.** *Let R be an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$. Then let $\vec{x}^* \in X$ satisfy $A(\vec{x}^*) \leq \vec{b}$. Recall that $\vec{b} = \sum_{i \in I} \iota_i(b_i), b_i \in R$, with finitely many $b_i \neq 0$. Define then the sets $I^= := \{i : \alpha_i(\vec{x}^*) = b_i\}, I^< := \{j : \alpha_j(\vec{x}^*) < \vec{b}_j\}$.*

Then $\vec{x}^*$ is an optimal solution to the primal maximization problem if and only if $\mathbf{c} \in \{\sum_{i \in I^=} \alpha_i \cdot u_i$ where $u_i \geq_R 0\}$.

Proof. We begin with the if statement. Suppose $\mathbf{c} = \sum_{i \in I^=} \alpha_i \cdot u_i$ as above. Define $\mathbf{y} := \sum_{i \in I^=} \hat{y}_i \cdot u_i$, and notice $\mathbf{y} \in Y^{**}_{\oplus}$. Recall that the slack variable $\vec{t} := \vec{b} - A(\vec{x}^*) \in Y^*$ (as in 2.2.2). It is clear that for $i \in I^=$, $\hat{y}_i(\vec{t}) = 0$. Thus $\mathbf{y}(\vec{t}) = 0$. Thus by Corollary 2.4.10, both $\vec{x}^*$, $\mathbf{y}$ are optimal solutions to their respective problems.

Conversely, we let A_J denote the transformation induced by $\{\alpha_j : j \in I^<\}$, and again let $\vec{x}$ be feasible, $\vec{x} = \vec{x}^* + \vec{x}'$ where $\vec{x}^*$ is now assumed to be optimal by hypothesis, and $\vec{x}'$ is the appropriate vector. In other words, for any $\vec{x}' \in X$, $A(\vec{x}^* + \vec{x}') \leq \vec{b} \implies \mathbf{c}(\vec{x}^* + \vec{x}') \leq \mathbf{c}(\vec{x}^*)$. We can rephrase this to mean if $A_I(\vec{x}') \leq_{Im(A_I)} 0$ and $A_J(\vec{x}') \leq_{Im(A_J)} \vec{b}_J - A_J(\vec{x}^*)$ then $\mathbf{c}(\vec{x}') \leq 0$, where $b_J := \displaystyle\sum_{j \in I^<} \rho_j(b_j)$.

We then want to show

$$A_I(\vec{x}') \leq_{A_I} 0 \implies \mathbf{c}(\vec{x}') \leq_R 0.$$

If $A_J(\vec{x}') \leq b_J - A_J(\vec{x}^*)$ holds, then by the previous paragraph, whenever $A_I(\vec{x}') \leq 0$, $\mathbf{c} \leq 0$.

Thus, we assume that for some $\vec{x}'$, this does not hold. Then consider that $b_j - \alpha_j(\vec{x}^*) > 0$ for each $j \in I^<$ and moreover, $I^<$ is a finite set. Thus for some $\epsilon > 0$ $A_J(\epsilon\vec{x}') \leq \vec{b}_J - A_J(\vec{x}^*)$. Also for the same $\epsilon\vec{x}'$, $A_I(\epsilon\vec{x}') \leq 0$ if and only if $A_I(\vec{x}') \leq 0$. Thus $\mathbf{c}(\epsilon\vec{x}') \leq 0$ which only happens iff $\mathbf{c}(\vec{x}') \leq 0$.

So by the Farkas Lemma, $\mathbf{c} = \sum_{i \in I^=} \alpha_i \cdot u_i$ where $u_i \geq 0$ and we are done. □

Remark 5.3.5. To see how this generalizes the classical Complementary Slackness Theorem **??**. We notice that by our reformed notion of feasible, there is no slack variable $\mathbf{s}$, since we require that $\mathbf{y} \circ A = \mathbf{c}$. Thus for two variables to be complementary slack, it suffices to show that $\mathbf{y}(\vec{t}) = 0$, where $\vec{t} := \vec{b} - A(\vec{x}$.

But $\mathbf{y}(\vec{t}) = \sum_{i \in I} t_i \cdot y_i$, where $y_i \geq 0$. For this to be 0, then y_i can only be non-zero when $t_i = 0$. In other words, $\mathbf{c}$ must be a linear combination of the $\alpha_i \in \mathcal{A}$, where $\alpha_i(\vec{x}) = b_i$, as described above.

We then prove one of the most important duality results, the **Strong Duality Theorem**. We will prove a result using The Complementary Slackness Theorem 5.3.4, and the Strong Duality was a corollary.

Proposition 5.3.6. *Let R be an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$. Then let $\vec{x}^* \in X$ be feasible (i.e. $A(\vec{x}^*) \leq_{Y^*} \vec{b}$), such that $\vec{x}^*$ is a maximizer for $\mathbf{c}$, then there is a $\mathbf{y}^* \geq_{Y^{**}} 0$ such that $\mathbf{y}^* \circ A = \mathbf{c}$ and $\mathbf{y}^*(\vec{b}) = \mathbf{c}(\vec{x}^*)$.*

Proof. By Lemma 5.3.4, we have that $\mathbf{c} = \sum_{i \in I^=} \alpha_i y_i$, i. e. $\mathbf{c}(e_j) = 0, j \in I^<$. Let $\mathbf{y}^* : Y \to G$

be defined by $\mathbf{y}^*(b_ie_i) = y_ib_i, i \in I^=$, and $\mathbf{y}^*(b_je_j) = 0, j \in I^<$. Then:

$$\mathbf{y}^*\left(\vec{b}\right) = \mathbf{y}^*\left(\sum_{i\in I^=} b_ie_i\right) + \mathbf{y}^*\left(\sum_{j\in I^<} b_je_j\right) = \sum_{i\in I^=} y_ib_i + 0 = \sum_{i\in I^=} y_i\alpha_i(\vec{x}^*) = \mathbf{c}(\vec{x}^*)$$

since $\alpha_i(\vec{x}^*) = b_i$ for $i \in I^=$. □

Theorem 5.3.7 (Generalized Fact VII: Strong Duality)**.** *Let R be an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$. Then let $\vec{x}^* \in X$, $\mathbf{y}^* \in Y^*$ such that they are feasible optimizers of their respective programs. Then $\mathbf{c}(\vec{x}^*) = \mathbf{y}^*(\vec{b})$.*

Proof. Since (P) has an optimizer, Proposition 5.3.6 allows us to construct a $\mathbf{y}'$ that is an optimizer for (D). Thus $\mathbf{c}(\vec{x}^*) = \mathbf{y}'(\vec{b}) = \mathbf{y}^*(\vec{b})$ for any optimizer $\mathbf{y}^*$. □

Remark 5.3.8. Since we use the Farkas' Lemma, and several consequences of the Farkas' Lemma, to prove the Strong Duality Theorem, it is natural to ask, whether or not the Farkas' Lemma is a necessary condition for the Strong Duality to hold. The answer is actually hidden in the proofs of Propositions 5.3.1, 5.3.2. If (a) for the Farkas' Lemma fails, then there is a $\vec{x} \in X$ such that $A(\vec{x}) \leq_{Y^*} 0$ and $\mathbf{c}(\vec{x}) >_R 0$, and in this case (P) is unbounded. Similarly, if (b) fails, then $\mathbf{c}$ cannot be written as a $\mathbf{y} \circ A$, where $\mathbf{y} \in Y^{**}_{\oplus}$. Thus, both (a) and (b) need to hold for either program to have optimizing solutions. In any case where one holds but the other does not, at least one of these programs will not have any optimizing solutions.

5.3.3 Counterexamples to Generalizations of Classical Results

We see that the existence of an optimal solution to the primal problem results in an optimal solution to the dual. However, we note that the converse is not necessarily true:

Example 5.3.9. Let $R = \mathbb{R}, X = Y^* = \mathbb{R}^{\mathbb{N}}$

$$A_1 : \begin{array}{r} \alpha_0 = \\ \alpha_1 = \\ \alpha_2 = \\ \alpha_3 = \\ \vdots \end{array} \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & \cdots \\ -1 & 0 & 0 & 0 & \cdots & 0 & \cdots \\ 0 & -1 & 0 & 0 & \cdots & 0 & \cdots \\ 0 & 0 & -1 & 0 & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \end{pmatrix}, \vec{b} := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}$$

$$\mathbf{c} := \begin{pmatrix} \frac{1}{2} & \frac{2}{3} & \cdots \frac{n}{n+1} & \cdots \end{pmatrix}$$

We see that we can write $\mathbf{c}$ as a sum $\mathbf{c} = \sum \alpha_i y_i$, where

$$\mathbf{y} := \begin{pmatrix} 1 \\ \frac{1}{2} \\ \vdots \\ \frac{1}{n} \\ \vdots \end{pmatrix}.$$

Moreover, that this is an optimal solution, since if $y_0 < 1$, then for some index N, $c_N > y_0$ and we cannot write c_N as a non-negative linear combination of $\alpha_{i,N}$.

However, given any feasible $\vec{x} \in X$, we can think of $\vec{x} = \begin{pmatrix} x_1 & x_2 & \cdots \end{pmatrix}$, with only finitely many $x_i \neq 0$, and $\sum x_i \leq 1$. Suppose x_N is the largest such index. Then

$$\mathbf{c}(\vec{x}) \leq \frac{N}{N+1},$$

but

$$\mathbf{c}\left(\vec{x}^* := \begin{pmatrix} 0 & 0 & \cdots & x_{N+1} = 1 & \cdots \end{pmatrix}\right) = \frac{N+1}{N+2}.$$

Thus the primal program has no optimal solution.

Notice that this is also a counterexample to a generalized Existence-Duality Theorem. Both programs are feasible, and thus bounded. However, only the dual program achieves optimality, while the primal program does not.

We apply some more hypothesis to achieve the desired result.

Theorem 5.3.10. *Let R be an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$. If $\mathbf{y}^* \in \mathrm{Hom}_R(Y^*, R)$ is a linear map such that , $A \circ \mathbf{y}^* = \mathbf{c}, \mathbf{y}^* \geq_{Y^*} 0$ and $\mathbf{y}^*(\vec{b})$ is a minimizer, then there is a $\vec{x^*} \in X$ such that $\mathbf{c}(\vec{x^*}) = \mathbf{y}^*(\vec{b})$ and $A(\vec{x^*}) \leq_{Y^*} \vec{b}$, so long as $|\mathrm{supp}(\mathbf{y}^*)| < \infty$.*

Proof. Let $\mathbf{y}^*$ be as described in the hypothesis. For each ι_i, define $y_i := \mathbf{y}^*(\iota_i(1))$ then let $J^+ := \{i : y_i > 0\}$ and $J^0 =:= \{j : y_j = 0\}$. Define $b_i := \rho_i(\vec{b})$. Note that by our hypothesis, $|J^+| < \infty$. We let Π^+, Π^0 denote the projections from Y onto the subspaces of Y, Y^+, Y^0 induced by J^+, J^0 respectively.

We want to find a $\vec{x^*}$ such that $\alpha_i(\vec{x^*}) = b_i, i \in J^+$ and $\alpha_j(\vec{x^*}) \leq b_j, j \in J^0$. Then such a $\vec{x^*}$ would have the property

$$\mathbf{c}(\vec{x^*}) = \sum \alpha_i \cdot y_i = \sum_{i \in J^+} \alpha_i(\vec{x^*}) \cdot y_i = \sum_{i \in J^+} b_i \cdot b_i = \mathbf{y}^*(\vec{b}).$$

Let $\vec{d_+} := \Pi(\vec{b}), \vec{d_0} := \Pi(\vec{b})$, and let A^+ be the map $A \circ \Pi^+$, and similarly define $A^0 = \Pi^0 \circ A$.

Then by Proposition 5.2.3, $A^+(\vec{x}) = \vec{d_+}, A^0(\vec{x}) \leq \vec{d_0}$ has no solution if and only if f there is a $\lambda_1 \in Y_1^*$, and $\lambda_2 \in Y_2^*, \lambda_2 \geq 0$ (which induces a $\lambda \in Y^*, \lambda = \lambda_1 \oplus \lambda_2$) such that

$$\lambda \circ A = \lambda_1 \circ A^+ + \lambda_2 \circ A^0 = 0$$

and $\lambda(\vec{b}) = \lambda_1(\vec{d_+}) + \lambda_2(\vec{d_0}) < 0$.

Then consider $\mathbf{y}' := \mathbf{y}^* + \lambda \cdot \epsilon$, where $\epsilon < \min(\{-\lambda(\iota_i)^{-1} y_i : i \in J^+\})$. Notice that for such an ϵ, $\mathbf{y}^* + \lambda \cdot \epsilon > 0$ (as $|J^+| \leq \infty$). We note that $\mathbf{y}'(\vec{b}) < \mathbf{y}(\vec{b})$, and $\mathbf{y}'(\vec{x}) = \mathbf{y}(\vec{x}) + 0 \cdot \epsilon = \mathbf{y}(\vec{b}) = \mathbf{c}(\vec{x})$. This contradicts the minimality of $\mathbf{y}$.

Thus $A^+(\vec{x}) = \vec{d}_+, A^0(\vec{x}) \leq \vec{d}_0$ has a solution, and such a $\vec{x}$ exists.

□

However, as in the case of the Farkas' Lemma, we can extend this result to certain infinite cases.

Proposition 5.3.11. *Let R be an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$. Consider the partition of $\mathcal{A}$, the collection of row-like projection maps induced by A, (Definition 3.3.6), into linear dependence classes $\{\mathcal{A}_j\}$. Let $\vec{b}_j$ be $\Pi_j(\vec{b})$, where $\Pi_j : Y^* \to Y_i$ be the projection from Y to the induced subspace. Let $\mathbf{y}_j = \mathbf{y}|_{Y_j}$.*

Then $\mathbf{y}$ is an optimal solution to (D), if and only if $\mathbf{y}_j$ is an optimal solution to $(D)_j$: Minimize $\mathbf{y}_j(\vec{b}_j)$ subject to $\mathbf{y}_j \in (Y_j)^*_\oplus$, $\mathbf{c}_j = \mathbf{y}_j \circ A_j$.

Proof. If $\mathbf{y}$ is optimal, suppose for some index k, $\mathbf{y}_k$ is not optimal for it's program, that is $\exists\ \mathbf{y}'_k$ such that $\mathbf{y}'_k(\vec{b}_k) < \mathbf{y}_k(\vec{b}_k)$, but $\mathbf{y}'_k \circ A_k = \mathbf{c}_k$.

Then define $\mathbf{y}' : Y^* \to R$ via $\mathbf{y}' := \mathbf{y}_k + \sum_{j \neq k} \mathbf{y}_j$. This is well defined since each vector in Y will only be nonzero for finitely many $\mathbf{y}_j$. Moreover,

$$\mathbf{y}'(\vec{b}) = \mathbf{y}_k(\vec{b}_k) + \sum_{j \neq k} \mathbf{y}_j(\vec{b}_j) < \sum_j \mathbf{y}_j(\vec{b}_j) = \mathbf{y}(\vec{b}).$$

This contradicts the minimality of $\mathbf{y}$. Hence each $\mathbf{y}_k$ is a minimizer of their respective program.

Similarly, if each $\mathbf{y}_j$ is a minimizer of their respective programs, and $\mathbf{y}$ is not, then we may take the improved solution $\mathbf{y}'$, and consider their restriction to each subprogram and derive a contradiction. $\square$

Proposition 5.3.12. *Let R be an ordered division ring, X, Y^* be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$. Consider the partition of $\mathcal{A}$, the collection of row-like projection maps induced by A, (Definition 3.3.6), into linear dependence classes $\{\mathcal{A}_j\}$. Let $\vec{b}_j$ be $\Pi_j(\vec{b})$, where $\Pi_j : Y^* \to Y_i$ be the projection from Y^* to the induced subspace. Let $\mathbf{y}_j = \mathbf{y}|_{Y_j}$.*

If $\vec{x}$ is an optimal solution to (P), if and only if $\vec{x}_j$ is an optimal solution to (P)$_j$: Maximize $\mathbf{c}_j(\vec{x}_j)$ subject to $A_j(\vec{x}_j) \leq \vec{b}_j$.

Proof. If $\vec{x}$ is a maximizer, then by Theorem 5.3.4, there is a minimizer $\mathbf{y} \in Y_{\oplus}^{**}$, which by Proposition 5.3.11 implies each $\mathbf{y}_j$ is a minimizer. But if any $\mathbf{c}_j(\vec{x}_j) < \mathbf{y}_j(\vec{b}_j)$, then $\mathbf{c}(\vec{x}) = \sum_j \mathbf{c}_j(\vec{x}_j) < \mathbf{y}(\vec{b})$. Thus each $\vec{x}_j$ is a maximizer.

Similarly, if each $\vec{x}_j$ is a maximizer, then there is a collection of minimizers $\mathbf{y}_j$ such that $\mathbf{y}_j(\vec{b}_j) = \mathbf{c}_j(\vec{x}_j)$, and so $\mathbf{c}(\vec{x}) = \displaystyle\sum_j \mathbf{c}_j(\vec{x}_j) = \sum_j \mathbf{y}_j(\vec{b}_j) = \mathbf{y}(\vec{b})$. Moreover since only finitely many of these $\vec{b}_j$ are non-zero, this sum is well defined. $\square$

Corollary 5.3.13. *Consider the partition of $\mathcal{A}$, the collection of row-like projection maps induced by A, (Definition 3.3.6), into linear dependence classes $\{\mathcal{A}_j\}$. If each $|\mathcal{A}_j| < \infty$, then (P) has an optimal solution if and only if (D) has an optimal solution.*

Proof. If (P) has an optimal solution, then by Proposition 5.3.6, there is an optimal solution to (D)

If (D) has an optimal solution, then by Proposition 5.3.11, each subprogram has an optimal solution. Since each $|\mathcal{A}_j| < \infty$, each $|J_j^+| < \infty$. Thus by Theorem 5.3.10, each subprogram has an optimal solution to $(\mathrm{P})_j$, and by Proposition 5.3.12, (P) has an optimal solution. $\square$

5.4 Conclusion

In this chapter, we assume that one of the hypothesis of the generalized Farkas' Lemma 4.3.1 holds. Using the generalized Farkas' Lemma, we prove some theorems of the alternative, which are largely technical results. We then use these to prove partial results to the Existence Duality Theorem: Proposition 5.3.1, and Proposition 5.3.2. We then prove the generalized Complementary Slackness Theorem 5.3.4. We use the generalized Complimentary Slackness Theorem to prove Proposition 5.3.6, a result that states that the existence of a primal optimal solution gives rise to a dual solution with equal objective value. The generalized Strong Duality Theorem 5.3.7 follows as a consequence.

Chapter 6

An Oriented Matroid Solution

6.1 Introduction

Recall that out of the major linear programming duality theorems, the only one we have not proved is that a bounded feasible program is guaranteed to have an optimal solution. In fact, we have shown that under the hypothesis where the Farkas' Lemma holds, this still may not be true. In Example 5.3.9, we saw that the primal solution is feasible, and the dual program is not only feasible but has an optimal solution, bounding the primal problem. Yet the primal program has no optimal solution. As Example 5.3.9 represents an affine program where $\mathcal{A}$ consists of a single infinite linear dependence class, a reasonable conjecture is that if each linear dependence class is finite, a feasible, bounded program obtains an optimal primal solution, which in turn induces an optimal dual solution by Proposition 5.3.6.

In order to proceed, we need a view of our linear programs which describes the underlying problem independent of the algebraic and topological features R may have. Many of our standard methods of finding optimal solutions for an affine program depend on the properties of the real numbers such as the compactness of bounded spaces, or the commutativity of multiplication. For a general division ring R, both the algebraic operations and the underlying topology are potentially unwieldy and may lack the properties that we desire or are convenient.

However, as R is ordered, there is a natural, simple, and near-binary description of our feasible and optimal solutions. The value of a functional is either positive or zero, or negative. A linear combination of affine maps have either positive or zero, or negative coefficients. A vector lies on one side of a hyperplane, or another, or on it. These facts hold independently of any other properties of R. Thus a structure that captures these relations is the ideal lens through which to look at these affine programs.

We propose here that **oriented matroids** are exactly the structure that captures the pertinent information and relations. By reducing everything to sets and set containment, we rid ourselves from the distracting features that our ring and our spaces may possess. We then refer to the established theory of Oriented Matroid Programming to show that a finite Oriented Matroid Program that is bounded and feasible has an optimal solution. Finally, we use the decomposition established by Theorem 4.3.25 to show that any bounded feasible potentially infinite affine program that can be decomposed into finite sub-programs has a pair of optimal solutions, for the primal and dual problem (Theorem 6.4.2). We then prove that under these hypothesis, we can prove the generalized Existence-Duality Theorem 6.4.4.

6.2 Oriented Matroids

Recall the definition of an oriented matroid Definition 1.5.1. Here, we describe some general examples of oriented matroids that will be relevant to our discussion:

Example 6.2.1. Given a vector-space X over an ordered division ring R, and a collection of vectors $E := \{e_i\} \subseteq X$, we can define $\mathcal{M} := (E, \mathcal{C})$, where $\mathcal{C}$ is the collection of sign vectors

$C = (C^+, C^-)$ where the following hold:

$$0 = \left(\sum_{e_i \in C^+} e_i \cdot a_i\right) - \left(\sum_{e_j \in C^-} e_j \cdot a_j\right), a_i, a_j \in R, A_i, a_j >_R 0$$

and moreover, $\mathcal{M}$ is the collection of all non-empty, minimal (with respect to containment) such sets.

We may verify that such a collection forms an oriented matroid:

1. By construction, the empty set is not in $\mathcal{M}$.
2. Similarly, the condition of minimality shows that no sign vector is contained in another.
3. Since a sign vector C corresponds to a linear combination $\left(\sum_{e_i \in C^+} e_i \cdot a_i\right) - \left(\sum_{e_j \in C^-} e_j \cdot a_j\right)$, $a_i, a_j > 0$ that sums to 0, the sum $\left(\sum_{e_i \in C^-} e_i \cdot a_i\right) - \left(\sum_{e_j \in C^+} e_j \cdot a_j\right)$, $a_i, a_j > 0$ also sums to 0, and corresponds to $-C_1$.
4. Given C_1, C_2, if $e \in C_1^+, C_2^-$, these correspond to two separate linear combinations

$$\left(a \cdot e + \sum_{e_i \in C_1^+ \setminus \{e\}} e_i \cdot a_i\right) - \left(\sum_{e_j \in C_1^-} e_j \cdot a_j\right), a_i, a_j, a > 0,$$

and

$$\left(\sum_{e_i \in C_2^+} e_i \cdot b_i\right) - \left(b \cdot e + \sum_{e_j \in C_2^- \setminus e} e_j \cdot b_j\right), b_i, b_j, b > 0,$$

both sums summing to 0. By multiplying via the appropriate scalar, we may assume that $a = b$, and thus we have that

$$0 = \left(\sum_{e_i \in (C_1^+ \cup C_2^+) \setminus \{e\}} e_i \cdot c_i\right) - \left(\sum_{e_j \in (C_1^- \cup C_2^-) \setminus \{e\}} e_j \cdot c_j\right), c_i, c_j > 0,$$

where the c_i, c_j are the scalar multiples of the a_i, a_j, b_i, b_j. Since this is a linear combination that sums to 0, there is some minimal collection of such vectors that do the same, and thus there is a circuit which satisfies axiom 4.

This example is relevant since our linear dependence classes $\mathcal{A}_j$ are a collection of vectors, and thus by letting $E := \mathcal{A}_j$, we are able to place an oriented matroid structure on this class. We should then also be able to place a dual oriented matroid structure on $\mathcal{A}_j$ as well, via Proposition 1.5.4.

Example 6.2.2. In the previous example, we showed that given a collection of vectors $E \subset X$ a vector-space over a division ring R, that the collection of vectors minimal with respect to linear dependence forms a collection of circuits, or an oriented matroid over E. Here, we investigate the cocircuits of an affine program, as described in Example 6.2.1.

Consider a subset $S \subset E := \mathcal{A}_j$ that are linearly independent, such that their span has co-dimension 1. Notice then that given any vector $v \in E$ not in this span, $S \cup \{v\}$ is a linearly independent set that spans X, and then given any additional vector $\vec{x} \in E$, this collection of vectors is now not linearly independent, and more importantly, it is minimal with respect to this fact.

We then define the collection of cocircuits as follows: Given each subset S as above, S defines a hyperplane, the kernel to f_S, where f_S is an appropriate linear functional. We then define C' to be $((C')^+, (C')^-)$, where $(C')^+ := \{e \in E : F_S(e) > 0\}$ and $(C')^- := \{e \in E : f_S(e) < 0\}$. We then define the collection of all such C' to be $\mathcal{C}^*$, and show that these are cocircuits over E.

Then take any cocircuit C', and circuit C. Either the support of C and C' do not intersect (equivalently, the support of C is contained in the span of S). Thus $C \perp C'$. Otherwise, suppose that there was a $v \in C^+ \cap (C')^+$. Hence, $f(v) > 0$. Moreover, since there is a linear combination of elements $0 = \sum_{v_i \in C^+} a_i \cdot v_i + \sum_{v_j \in C^-} a_j \cdot v_j$, $a_i > 0, a_j < 0$ associated with C, we have that the coefficient of v in this sum, a, is positive as well. Since

$$f(0) = f\left(\sum_{v_i \in C^+} a_i \cdot v_i + \sum_{v_j \in C^-} a_j \cdot v_j\right) = 0,$$

there must be a $a_k \in R\backslash\{0\}, v_k \in C^+ \cup C^-$, such that $f(a_k \cdot v_k) < 0$, i.e. $a_k < 0, v_k \in C^+$ or $a_k > 0, v_k < 0$. Either way, $f(v_k) \neq 0$, so v_k is in the span of S, and either $v_k \in C^- \cap (C')^+$ or $v_k \in C^+ \cap (C')^-$. Thus $C \perp C'$.

Now that we have a notion of a oriented matroid, we would like to relate our affine programming problem.

Proposition 6.2.3. *Given an ordered division ring R and a R-left vector space X, then any finite collection of linear maps $\alpha_i : X \to R$, which are in the same linear dependence class, gives rise to a primal and dual oriented matroid.*

Proof. Let E be the collection of linear maps α_i, and impose on it the circuits and cocircuit structure we described above, since $\mathrm{Hom}_R(X, R)$, is a R-vector space. □

Remark 6.2.4. Here, we note that a covector can be interpreted to be encoding "geometric" information about a region of X. Given a $x \in X$, we can associate a sign vector S to x such that $\alpha_i \in S^+$, if $\alpha_i(x) > 0$ and $\alpha_i \in S^-$ if $\alpha_i(x) \in S^-$.

We verify that our notions of co-vector correspond to this. Let g be a linear combination of α_i corresponding to the signs of a circuit C, i.e. when g is 0. Then consider $g(w) = 0$. If sign vectors C^+, C^- do not intersect with S^+, S^-, then g is a linear combination of functionals which vanish at w. So naturally their sum would be 0, and $(S^+, S^-) \perp C$.

Conversely, if say, $\alpha_1 \in S^+ \cap C^+$, then $\alpha_1(x) > 0$ and the coefficient of α_1 is the sum g is positive. Thus the summand $c_1 \cdot \alpha_1(x)$ is positive, and there must be a $c_k \cdot \alpha_k$ in the sum such that $c_k < 0, \alpha_k(x) > 0$ or $c_k > 0$ and $\alpha_k < 0$. Thus $C^- \cap S^+$ or $C^+ \cap S^-$ are non-empty, and $(S^+, S^-) \perp C$.

What then is a cocircuit? We previously defined it in terms of a hyperplane, the zero set of some linear functional that was maximal with respect to not spanning X. Here then, it would similarly be the zero set of a functional, but where that functional is determined by evaluating at x. In other words, the kernel of $F_x : \alpha_i \mapsto \alpha_i(x)$. Thus, these must coincide with $x \in X$ such that α_i that vanish on these x are maximal with respect to not spanning $\mathrm{Hom}_R(X, R)$.

We then notice that the intersection of the kernels of α_i that span $\mathrm{Hom}_R(X, R)$ would just be 0, then the x which fit the description above form a 1-dimensional subspace of X, (a line).

This notion may be extended to collection of affine functionals. We do this by linearizing the functionals.

Remark 6.2.5. Given a system of affine functionals $\{\alpha_i' : X \to R\}$, where each $\alpha_i' := \alpha_i + b_i$, where only finitely many of the b_i are non-zero, we can define a linear functional $\alpha_i^* : X \oplus R \to R, (x, f) \mapsto \alpha_i(x) + f \cdot b_i$. We see that this captures all of the essential information of the system above. By fixing $r := 1$ We have the same system as we originally had, and by letting f be any positive number, we obtain the same solutions of any systems of equality or inequality up to a scalar multiple.

We can then encode the intersection of affine hyperplanes as a sign vector. We define $E^* := \{\alpha_i^*\} \cup \{id_R\}$. We then can describe each region of X with respect to these hyperplanes by considering co-vectors S where $f \in S^+$.

We can then consider the following problem:

Problem 6.2.6. Let $E := \{-\alpha_i^*\} \cup \{\mathbf{c} \oplus 0\} \cup \{-id_R\}$. Notice that a (matroid) vector S measures the signs of a linear combination of these functionals that sum to 0. We see that

$g \in S^+, S^- = \emptyset$ and $f := id_R \notin S^+ \cup S^-$ would be a linear combination where $\mathbf{c}$ is a non-negative linear combination of the α_i, a feasible solution to the dual problem.

Conversely consider S to be a co-vector. By requiring that $id_R \in S^+$, $S^- = \emptyset$ and $-\mathbf{c} \notin S^+ \cup S^-$, we find a region of X such that each α_i is equal to or less than its bound b_i, that is, a feasible solution to the primal problem. Moreover, by requiring $id_R \in S^+$, we are essentially fixing a positive value for f, which reduces the line represented by this cocircuit to a point.

In either case, we are looking for a solution to this problem that "maximizes" one of our given functionals. Either a region of X that maximizes $\mathbf{c}$, or linear combination of functionals that maximizes $-f \cdot \vec{b}$ or minimizes $\vec{b}$.

What we have shown is that our affine program gives rise to an oriented matroid program. The goal is now to identify the appropriate circuit and cocircuit that are the optimal solutions to their respective problems. We noted before that finding a solution to both problems is potentially impossible if $\mathcal{A}$ (the collection of row-like projection maps from Definition 3.3.6) is infinite. In Example 5.3.9, we see that the existence of a dual solution did not give rise to the existence of a primal solution. However, in this example, $\mathcal{A}$ consisted of a single linear dependence class. We have shown that a solution to an affine program can be described as the sum of the solutions to the programs induced by the linear dependence classes. Thus, we may show that if $|\mathcal{A}| < \infty$ forces the existence of feasible primal and dual solutions to give rise to optimal primal and dual solutions, then we will have shown that for such affine programs, both programs feasible imply both programs are bounded, and that both programs have optimal solutions.

These results about oriented matroid programs have been shown by Robert Bland and James Lawrence in their respective dissertations [BLV78] [Law75]. We will outline the relevant part of their work here, and generalize when appropriate, as their work was done with finite dimensional real spaces in mind. We will also focus on the primal problem, as the existence of the primal optimizer gives rise to dual optimizer, as seen in Proposition 5.3.6.

6.3 Oriented Matroid Programming

We begin by defining an oriented matroid program, and the geometry of the program in terms of the oriented matroid.

Definition 6.3.1 ([BLVS+99])**.** Let E be a non-empty set (usually a collection of functionals) as in Example 6.2.1).

1. An *oriented matroid program* is a triple $(P) := (\mathcal{M}, f, g)$, where $\mathcal{M}$ is an oriented matroid defined on $E := E_n \cup \{f, g\}$ where g is not a *loop* (a circuit that is a singleton), f is not a coloop (a cocircuit that is a singleton) and $f \neq g$.
2. The dual of a matroid program is the triple $(D) := (\mathcal{M}^*, f, g)$.

Definition 6.3.2 ([BLVS+99])**.** Let E_n be a set of order n, and let $E := E_n \cup \{f\} \cup \{g\}$. Let $\mathcal{M}$ be an oriented matroid defined on E, such that g is not a loop, f is not a co-loop, and $f \neq g$.

- The *feasible region*, F_r of the primal problem is the collection of all co-vectors C' such that $e_i \notin (C')^-$ for any $e_i \in E_n$, and $g \in (C')^+$.
- The *face at infinity* F^∞ is the collection of all co-vectors C' such that $e_i \notin (C')^-, g \notin (C')^+ \cup (C')^-$.

Here, g represents $\mathbf{c}$ and f represents $-id_R$ from Remark 6.2.5.

We examine the motivation of these definitions. The feasible region is the collection of points in x which are not on the "negative" side of each hyperplane, equivalently the points that satisfy each of the inequalities of the original program, and $g \in (C')^+$ forces the bound vector ($\vec{b}$) to be positive, thus reflects the regional inequalities of the program. The "face at infinity" is the equivalence classes formed by parallel planes. This is exactly the co-vectors formed when $g = 0$, as each "parallel" functional, $\alpha + g \cdot b_1, \alpha + g \cdot b_2$ are the same, and any region of $X \oplus R$ will fall on one side or the other, or on the plane formed by the kernel of $\alpha + g \cdot b_1$ if and only if it does the same for $\alpha + g \cdot b_2$.

The face of infinity then represents the space of possible "slopes" or "directions" that each of our given affine functionals can take. A co-vector in the face at infinity would be one of minimal support, or maximal in terms of being zero (which is equated with lying on a hyperplane) would represent points at infinity, which are the directions (or slopes) one can take along these hyperplanes.

We then need to describe traveling in a direction in this space, in terms of our matroid:

Definition 6.3.3 ([BLVS$^+$99]). The *composition* of two sign vectors $C_1 \circ C_2$ is defined as $(C_1 \circ C_2)^+ = C_1^+ \cup (C_2^+ \backslash C_1^-)$, and $(C_1 \circ C_2)^- = C_1^- \cup (C_2^- \backslash C_1^+)$. We observe that this is a associative, but not commutative operation.

Remark 6.3.4. In the general theory of oriented matroids, the composition operator has many uses, here we focus on the aspects that are pertinent to affine programming.

Consider a feasible co-vector C (represented by point $x \in X$) and a co-vector on the face at infinity, D (represented by a point at infinity a). Then consider $C \circ D$. This represents a region of X (represented by a point $x' \in X$) such that x' lies on the same side of all hyperplanes as x did. In addition it lies on the same side of all hyperplanes as a does, unless x was on the opposite side of the same hyperplane. Then x' may be viewed as a point on the line from x to a, that does not cross any of the other hyperplanes represented by our ground set.

Here lies the key to verifying an optimal solution. If D is a co-vector on the face at infinity, such that $f \in D^+$, then evaluating f at the points represented by a (the appropriate intersection of the remaining hyperplanes still yields a point $\vec{a} \in R^n$) is positive. Thus, increasing in the direction of a (or equivalently, adding a positive scalar multiple of $\vec{a}$ to x) increases the value of the objective function. This insight is how we will define our certificate of optimality.

Definition 6.3.5 ([BLVS+99]). Consider an oriented matroid program:

- The *directions* are the co-vectors D where $g \notin D^+ \cup D^-$. A direction is *increasing, decreasing* or *constant* if $f \in D^+, f \in D^-, f \notin D^+ \cup D^-$ respectively.
- For a given feasible co-vector C, a feasible direction is a direction D such that $C \circ D$ is feasible.
- A feasible co-vector C^o is optimal, if there are no feasible increasing directions for C^o.

Again, the definition of feasible is entirely intuitive and consistent with our definitions. We are looking for the corner point of the feasible polytope such that each direction either decreases the objective function, keeps it constant, or takes you out of the feasible polytope.

Lemma 6.3.6 (Three Painting Lemma [BLV78] [Law75]). *Given an Oriented Matroid $\mathcal{M}$, with ground set E, consider a partition of*

$$E = B \sqcup G \sqcup R, e \in B$$

and let $e \in B$. Then exactly one of the following hold:

- *There is a circuit C_1 such that $e \in C_1^+ \cup C_1^- \subseteq B \cup G$ and $C_1^- \cap B = \emptyset$.*
- *There is a cocircuit C_2 such that $e \in C_2^+ \cup C_2^- \subseteq B \cup R$ and $C_2^- \cap B = \emptyset$.*

The proof of the Three Painting Lemma was done by both Bland and Lawrence. It is omitted, as it requires many definitions and concepts of matroid theory that are not enlightening here.

Corollary 6.3.7. *Exactly one of the following hold:*

(a) There is a cocircuit C_1 such that, $f, g \in C_1^+$ and $C_1^- = \emptyset$, or there is a circuit C_2 such that $f, g \in C_2^+, C_2^- = \emptyset$.

(b) There is a cocircuit C_{1_o}, and a circuit C_{2_o}, such that $f \in C_{1_o}^+, C_{1_o}^- \subseteq \{g\}$, $g \in C_{2_o}^+, C_{2_o}^- \subseteq \{f\}$, and $(C_{1_o}^+ \cup C_{1_o}^-) \cap (C_{2_o}^+ \cup C_{2_o}^-) \subseteq \{f, g\}$.

Proof. Suppose there exists C_1 such that $f, g \in C_1^+$ and $C_1^- = \emptyset$, and a circuit C_2, where $g \in C_2^+$ and $C_2^- \subseteq \{f\}$. Then since C_1, C_2 are orthogonal, and $g \in C_1^+ \cap C_2^+$, then $A^+ \cap C_2^-$ must be non empty and $f \in C_2^-$. Then we let R be empty, let $B = E_n \cup \{g\}$ and $G = \{f\}$. Consider $e := g$. The cocircuit C_1 satisfies the first condition of the Three Painting Lemma, and the circuit C_2 satisfies the second, which is a contradiction. Similarly if there were a circuit C_2 satisfying (a), we use a symmetric argument.

Then, assuming neither event in (a) holds, then by the above argument, we have C_{1_o}, C_{2_o} such that $f \in C_{1_o}^+, C_{1_o}^- \subseteq \{g\}$, $g \in C_{2_o}^+, C_{2_o}^- \subseteq \{f\}$, and $(C_{1_o}^+ \cup C_{1_o}^-) \cap (C_{2_o}^+ \cup C_{2_o}^-)$, as both events in (a) fail. Then by the Three Painting Lemma, if we color E_n blue, then no element of E_n can be shared by C_{1_o}, C_{2_o}, and only f, g may be shared.

□

Remark 6.3.8. If C_1 is a cocircuit such that $f, g \in C_1^+$ and $C_1^- = \emptyset$, then C_1' defined by $(C_1')^+ := C_1^+ \backslash \{g\}, (C_1')^- = \emptyset$ is an increasing direction for C_1, and $C_1 \circ C_1' = C_1$. Thus, we can increase in the direction of C_1' indefinitely and increase f so the primal problem is unbounded.

Similarly, if there exists a circuit C_2 such that $f, g \in C_2^+, C_2^- = \emptyset$, then the dual problem is unbounded.

Theorem 6.3.9 (Certificate of Optimality[BLVS+99])**.** *Let C_1 be a given feasible cocircuit, and let C_2 be a dual-feasible circuit ($f \in C_2^+, C_2^- \cap E_n = \emptyset$) such that $(C_1^+ \cup C_1^-) \cap (C_2^+ \cup C_2^-) \subseteq$*

$\{f, g\}$. *Then C_1 is optimal.*

Proof. Let C_1, C_2 be given as stated above. Then let Z be an increasing direction for C_1. Since Z is an increasing direction, $g \notin Z^+ \cap Z^-$, $f \in Z^+$. Since $f \in Z^+ \cap C_2^+$, there must be a $h \in E_n$ such that $h \in Z^+ \cap C_2^+$ (no $h \in E_n$ will be in C_2^-, since C_2 is feasible). So $h \notin C_1^+ \cup C_1^-$, else $h \in (C_1^+ \cup C_1^-) \cap (C_2^+ \cup C_2^-)$. Thus, $h \in (C_1 \circ Z)^-$, and Z is not a feasible direction. Thus no increasing direction for C_1 is feasible and C_1 is optimal.

□

Putting together the last few results gives us the following final result:

Theorem 6.3.10 ([BLVS+99]). *Given an oriented matroid program where both the primal and dual program are bounded (i. e. both programs are feasible), there is a co-vector C_1 which is optimal.*

To translate this back into the language of affine programing, we have found, of all corner points formed by intersections of bounding hyperplanes, a point which is optimal .

6.4 The Existence Duality Theorem

We will now show that if a polytope is bounded by a finite number of hyperplanes, then the feasible polytope is itself closed, the interior is exactly the points that do not lie on any of the hyperplanes, and the boundary are exactly the points that do lie on at least one of the hyperplanes.

We note that by the Fundamental Theorem of Linear Programming (Theorem 2.4.22), a point can only be an optimizer if it is not contained in an open full line segment of the feasible region. Thus, it suffices to consider the points maximal with respect to lying on the bounding hyperplanes. However, these are exactly the cocircuits of the underlying oriented matroid. Thus, we have the following result:

Proposition 6.4.1. *The solution to a finite oriented matroid program is a solution to the affine program.*

Finally, we noted that by Proposition 5.3.12, given an affine program, and the resulting decomposition into sub-programs via linear dependence classes, that the primal program has an optimal solution if and only if each subprogram has an optimal solution. Our work with these subprograms also made it clear that the overall problems, both primal and dual, are feasible if and only if the sub programs are all feasible. Thus we have the following final result.

Theorem 6.4.2. *Let R be an ordered division ring, let X, Y^* be left R vector-spaces, and let $A \in \mathrm{Hom}_R(X, Y^*)$, $\vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(C_1, R)$, such each linear dependence class of $\mathcal{A}$ (the collection of row-like projection maps Definition 3.3.6) is finite.*

Then if primal and dual programs are both feasible, there exists a $\vec{x}^* \in X$ that is an optimal solution to the primal problem.

Proof. Since each subprogram is finite and feasible, by Theorem 6.3.10, each subprogram has an optimal $\vec{x}^* \in X_i$ and thus by Proposition 5.3.12, there is an optimal $\vec{x}^* \in X$ that maximizes $\mathbf{c}$. □

Corollary 6.4.3. *Let R be an ordered division ring, let X, Y^* be left R vector-spaces, and let $A \in \mathrm{Hom}_R(X, Y^*)$, $\vec{b} \in Y^*, \mathbf{c} \in \mathrm{Hom}_R(X, R)$, such each linear dependence class of $\mathcal{A}$ is finite.*

Then if there are feasible solutions to both the primal and dual problem, then there exists a $\mathbf{y}^* \in \mathrm{Hom}_R(Y^*, R)$ that is an optimal solution to the dual problem.

Proof. The existence of a primal solution gives rise to a dual solution by Proposition 5.3.6.

□

We may finally state a full generalization of one of our classical results **??**, the Existence-Duality Theorem:

Theorem 6.4.4 (Generalized Fact VIII: Existence-Duality Theorem)**.** *Given generalized Tucker tableau data such each linear dependence class of $\mathcal{A}$ is finite, then exactly one of the following hold:*

1. *Both the primal program and the dual program are infeasible.*
2. *The primal program is infeasible and the dual program is unbounded.*
3. *The dual program is infeasible and the dual program is unbounded.*
4. *Both the primal and dual programs admit optimal solutions.*

Proof. Here, we use several of our previously established partial results:

1. If both programs are infeasible, we satisfy (1).
2. If the dual program is unbounded, then the primal program is infeasible by Corollary 2.4.26. Conversely, if the primal program is infeasible, but the dual program is feasible, then by Proposition 5.3.1, the dual program is unbounded.
3. If the primal is unbounded, then the dual program is infeasible by Corollary 2.4.26. Conversely, if the dual program is infeasible, but the primal program is feasible, then by Proposition 5.3.2, the primal program is unbounded.
4. Finally, if both programs are feasible, then the primal program admits an optimal solution by Theorem 6.4.2, the primal program admits an optimal solution, and by Corollary 6.4.3, the dual program admits an optimal solution as well.

□

6.5 Conclusion

In this chapter we show that if each linear dependence class of $A \in \mathrm{Hom}_R(X, Y^*)$ (as defined in Definition 4.3.19) is finite then each of the sub-programs the original problem decomposes into can be molded by a finite oriented matroid (Proposition 6.2.3). Then, using the results of Bland and Lawrence (Thorem 6.3.10), we show that each of these oriented matroid programs which are bounded and feasible give rise to a primal optimal solution. Then by Proposition 5.3.6, they give rise to a dual solution as well. We then combine these results, and several results from previous chapters to prove the generalized Existence-Duality theorem.

Chapter 7

Oriented Matroid Programs as Tucker Tableaux and the Simplex Algorithm

7.1 Introduction

In his 1969 paper "The elementary vectors of a subspace of $\mathbb{R}^n$", R. T. Rockafellar [Roc69] conjectured that Tucker Tableaux were the proper way to interpret an oriented matroid program. In Chapter 4 of this work, we have shown how to interpret a general affine program as an oriented matroid program. In this chapter, we will show how to encode this information in a Tucker tableau, and how to describe a Tucker pivot and the Simplex Algorithm.

7.2 The Tucker Tableau

Recall that by Theorem 5.3.4, a vector $\vec{x} \in X$ is an optimal solution if and only if we may write $\mathbf{c}$ as a linear combination of the α_i projection functionals, where $\alpha_i(\vec{x}) = \hat{v}_i(\vec{b})$. In other words, the linear functionals where equality is achieved. Additionally, the discussion in Chapter 4 shows that we can describe a region of X in terms of the affine functionals $-\alpha_i + b_i$, by whether or not a point in this region evaluates to a positive number, a negative number or 0. This information is encoded in a sign vector, where the sign vectors of minimal support are called "co-vectors" and are the candidates for optimal solutions.

Now, a traditional Tucker tableau [NT93] is a 4 compartment array $A \in \mathbb{R}^{m\times n}, \vec{b} \in \mathbb{R}^m, \vec{c} \in \mathbb{R}^n, d \in \mathbb{R}$. Each row is meant to encode a projection map $A_i \cdot \vec{x}$, bounded above by $\vec{b}_i$, and the vector $\vec{c}$ encodes the linear functional $\mathbf{c}$, and d is the affine component of $f(\vec{x}) = \mathbf{c}(\vec{x}) - d$. A given tableau then records the value of $f(\vec{0})$, and the associated dual solution. If the primal program is feasible (no $\vec{b}_i < 0$), and each $\vec{c}_i \leq 0$, then $\vec{0} \in \mathbb{R}^n$ is an optimal solution for this tableau and this tableau is optimal.

The situation may then be illustrated as follows:

$$
\begin{array}{c|ccc|c|l}
 & x_1 & \cdots & x_n & \text{-1} & \\ \cline{2-5}
y_1 & a_{11} & \cdots & a_{1n} & b_1 & = -t_1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
y_m & a_{m1} & \cdots & a_{mn} & b_m & = -t_m \\ \cline{2-5}
-1 & c_1 & \cdots & c_n & d & = f \\ \cline{2-5}
\multicolumn{1}{c}{} & \| & \cdots & \| & \multicolumn{1}{c}{\|} & \\
\multicolumn{1}{c}{} & s_1 & \cdots & s_n & \multicolumn{1}{c}{g} &
\end{array}
$$

We can interpret $\vec{x} = 0$ as lying on the intersection of the hyperplanes $x_i = 0$. In this way, we can generalize the information captured by a traditional Tucker tableau in order to describe an oriented matroid Tucker tableau.

Given the information for an affine program, (let R be an ordered division ring, X, Y^*, be left R vector spaces, $A \in \mathrm{Hom}_R(X, Y^*), \mathbf{c} \in \mathrm{Hom}(X, R), \vec{b} \in Y^*$), it gives rise to an oriented matroid $\mathcal{M}$ on $E = \{f \cdot b_i - \alpha_i\} \cup \{f, g\}$, ($b_i := \hat{v}_i(\vec{b})$) where the circuits represent collections of linear (affine) functionals, minimal with respect to linear dependence. Recall that circuits C where $g \in C^+$ represent linear combinations of the α_i that sum to $\mathbf{c}$, and co-circuits C where $f \in C^+$ represent regions of the domain maximal with respect to lying on hyperplanes.

Note that we may think of E as a collection of functionals defined on $X \oplus R$, where f is the projection $f : X \oplus R \to R$, a dummy variable. Then each b_i may be replaced with $f \cdot b_i$. For shorthand, we define $\hat{\alpha}_i := f \cdot b_i - \alpha_i$.

Thus, we can define an oriented matroid Tucker tableau as follows:

Definition 7.2.1 (Oriented matroid Tucker tableau)**.** Let $\bar{X} \subseteq E \backslash \{f, g\}, \bar{Y} = E \backslash (\bar{X} \cup \{f, g\})$, where $\bar{X}$ is a collection of the $\hat{\alpha}_i$ such that the hyperplanes associated with $\alpha_i \hat{\in} \bar{X}$ is maximal. Equivalently, there is a circuit, whose support is contained in $\bar{X} \cup \{f, g\}$, which in turn is the support of a vector. Since we view g as $\mathbf{c}$ which does not share support with f, this linear dependence holds if and only if $\mathbf{c}$ is a linear combination of the associated α_i. We then define $C : \bar{X} \to \{+, -, 0\}$ such that $C(\hat{\alpha}_i)$ is the sign of the coefficient of α_i in the equation $\displaystyle\sum_{\hat{\alpha}_i \in \bar{X}} \alpha_i \cdot c_i + \mathbf{c} = 0$ (i.e. the opposite sign of c_i in the equation $\displaystyle\sum_{\hat{\alpha}_i \in \bar{X}} \hat{\alpha}_i \cdot c_i + f \cdot c_f + \mathbf{c} = 0$).

Then, since $\bar{X} \cup \{f, g\}$ is minimal with respect to linear dependence, when we fix a value for f (without loss of generality, $f = 1$), this describes a region in X, (the intersection of the kernels of affine functionals $b_i - \alpha_i$). Let us call this intersection K. The region K can be described by which side of each bounding hyperplane it lies on (for each $\hat{\alpha}_i$, the sign of $\hat{\alpha}_i(k, 1), k \in K$). Recall that this describes a co-vector of $\mathcal{M}$, where f is positive. Since $\hat{\alpha}_i(k, 1)$ is defined to be 0 for $\hat{\alpha}_i \in \bar{X}$, it suffices to record this information for $\bar{Y} := E \backslash (\bar{X} \cup \{f, g\})$. We then define $B : \bar{Y} \to \{+, -, 0\}$ to record the sign of $\hat{\alpha}_j(k, 1), k \in K, \hat{\alpha}_j \in \bar{Y}$.

We then define $\mathbf{A} : \bar{X} \times \bar{Y} \to \{+, -, 0\}$ as follows: Given $\hat{\alpha}_i \in \bar{X}, \hat{\alpha}_j \in \bar{Y}$, consider

$((\bar{X} \cup \{f\}) \cup \{\hat{\alpha_j}\}$. The $\alpha_\ell, \hat{\alpha_\ell} \in \bar{X}$ are linearly independent, so we may either write α_j uniquely as a linear combination of the α_ℓ, or α_j is independent of the α_ℓ. Let $\mathbf{A}(\alpha_i, \alpha_j)$ be the coefficient of α_i in the equation $\displaystyle\sum_{\hat{\alpha_\ell} \in \bar{X}} \alpha_\ell \cdot c_\ell = \alpha_j$.

The tableau then is the ordered tuple $(\bar{X}, \bar{Y}, \mathbf{A}, B, C)$.

We now make some observations. The functions $\mathbf{A}, B, C$ are meant to mirror the matrix A and vectors $\vec{b}, \vec{c}$. In this way we capture much of the essential information of the Tucker tableau. The d is suppressed in this case, as its only purpose is to record the actual value of the optimal solution(s) which is not captured in an Oriented Matroid. In the same way a standard Tucker tableau captures the situation where we attempt to evaluate the linear program at the origin, this version of the Tucker tableau captures the situation where we lie on a maximal number of bounding hyperplanes. It turns out that by Theorem 5.3.4, these determine exactly the α_i whose linear combination form $\mathbf{c}$.

Then, if B is a non-negative function, the region K which is encoded by the co-vector with support contained in $\bar{Y}$ is a feasible region, since each $\alpha_j(k) \leq b_k, k \in K$, and so a non-negative B records a feasible primal solution(s). If C is a non-positive function, then we may write

$$\begin{aligned} \mathbf{c} + \sum_{\hat{\alpha_\ell}} \alpha_\ell \cdot c_i &= 0 \\ \sum_{\hat{\alpha_\ell}} \alpha_\ell \cdot (-c_i) &= \mathbf{c} \end{aligned}$$

as a non-negative linear combination of the $\alpha_i \in \mathcal{A}$. Thus a non-positive C function records a feasible dual solution. If B is non-negative and C is non-positive, then we capture a region K, which is feasible, and has an associated feasible dual solution made up of the projections of the constraints that K meets. Thus by Theorem 5.3.4, K is a region of optimal solutions, and since $\mathbf{c}(k) = y \circ A(k) = y(\vec{b})$, the associated dual solution is optimal as well, by the Weak Duality Theorem.

Example 7.2.2 (The Lumber Mill Problem)**.** Recall the Lumbermill Problem 1.2.1. The primal-dual problem was be encapsulated in the following Tucker tableau:

$$
\begin{array}{c|ccc|c|l}
 & x_1 & x_2 & x_3 & \text{-1} & \\
\hline
y_1 & 1 & 3 & 2 & 10 & = -t_1 \\
y_2 & 2 & 1 & 1 & 8 & = -t_2 \\
\hline
-1 & 3 & 2 & 4 & 5 & = f \\
\hline
 & \shortparallel & \shortparallel & \shortparallel & \shortparallel & \\
 & s_1 & s_2 & s_3 & g &
\end{array}
$$

Then, as we see, $\mathbf{c}(0,0,0) = 0$. Since the right hand column is non-negative, this solution is feasible, but since $\vec{c}$ is a not a non-positive vector, this solution is not optimal.

However, we can consider $X = \mathbb{R}^3, Y^* = \mathbb{R}^5$ and consider the following bounds:

$$
\begin{aligned}
(-1)x_1 + 0x_2 + 0x_3 &\leq 0 \\
0x_1 + (-1)x_2 + 0x_3 &\leq 0 \\
0x_1 + 0x_2 + (-1)x_3 &\leq 0 \\
1x_1 + 3x_2 + 2x_3 &\leq 10 \\
2x_1 + 1x_2 + 1x_3 &\leq 8.
\end{aligned}
$$

Thus A is a linear transformation from $X \to Y^*$ such that:

$$\begin{aligned}
\alpha_1(x_1, x_2, x_3) &= -x_1 \quad , \quad b_1 = 0 \\
\alpha_2(x_1, x_2, x_3) &= -x_2 \quad , \quad b_2 = 0 \\
\alpha_3(x_1, x_2, x_3) &= -x_3 \quad , \quad b_3 = 0 \\
\alpha_4(x_1, x_2, x_3) &= x_1 + 3x_2 + 2x_3 \quad , \quad b_4 = 10 \\
\alpha_5(x_1, x_2, x_3) &= 2x_1 + x_2 + x_3 \quad , \quad b_5 = 8.
\end{aligned}$$

With $\hat{\alpha_i} := f \cdot b_i - \alpha_i$.

The primal problem is then maximize $\mathbf{c}, \mathbf{c}(x_1, x_2, x_3) = 3x_1 + 2x_2 + 4x_3$, and the dual problem is to find $\mathbf{y} : \mathbb{R}^5 \to \mathbb{R}$ such that $\mathbf{y} \circ A = \mathbf{c}, \mathbf{y} \geq_{Y^{**}} 0$ and $\mathbf{y}(\vec{b})$ is minimized. In other words, we wish to find $y_1, \ldots y_5 \in \mathbb{R}_{\oplus}$ such that $\mathbf{c} = \sum_{i=1}^{n} y_i \alpha_i$ and $\sum_{i=1}^{5} y_i b_i$ is minimized.

We note that $\mathbf{c} = 3\hat{\alpha_1} + 2\hat{\alpha_2} + 4\hat{\alpha_3}$, so $C = +$. Similarly, when we fix $f > 0$, the intersection of the kernels of $\hat{\alpha_1}, \hat{\alpha_1}, \hat{\alpha_1}$ in $\mathbb{R}^3$ is the origin, and $\hat{\alpha_4}(0, 0, 0)$ and $\hat{\alpha_4}(0, 0, 0) > 0$. Thus $B = +$.

We have now recorded both a primal and dual solution. The primal solution is the origin of $\mathbb{R}^3$ and the dual solution is $y = -3\alpha_1 - 2\alpha_2 - 4\alpha_3$. Since for these three $\alpha_j, \alpha_j(0, 0, 0) = b_j = 0$, we have written $\mathbf{c}$ as a linear combination of the α_i.

However, although this region is a feasible region for the primal problem, it is infeasible for the dual problem, as it not a non-negative linear combination of the α_i.

Finally, we note that we can write α_4, α_5 as a non-negative linear combination of the $\alpha_1, \alpha_2, \alpha_3$. This gives us the following tableau:

$$\begin{array}{c|ccc|c|}
 & \hat{\alpha}_1 & \hat{\alpha}_2 & \hat{\alpha}_3 & \text{-1} \\ \hline
\hat{\alpha}_4 & + & + & + & + \\
\hat{\alpha}_5 & + & + & + & + \\ \hline
-1 & + & + & + & \\ \hline
\end{array}$$

Figure 7.1: Oriented matroid Tucker tableau

Where $\bar{X} = \{\hat{\alpha_1}, \hat{\alpha_2}, \hat{\alpha_3}\}, \bar{Y} = \{\hat{\alpha_4}, \hat{\alpha_5}\}$, and $\mathbf{A}, B, C$ are captured as above.

7.3 Tucker Pivot

It is natural to ask then, what plays the role of a Tucker pivot? Previously we would have swapped one of the primal slack variables, t_j with one of the primal decision variables x_i. This would have also resulted in a simultaneous swapping of a dual slack variable s_i, with a dual decision variable y_j.

However, with our newfound understanding, we see that this simply represents swapping one hyperplane for another, or swapping one functional of the linear combination that forms $\mathbf{c}$ for another. Recall that for any $\hat{\alpha_i} \in \bar{X}, \hat{\alpha_j} \in \bar{Y}$ where $\mathbf{A}(\hat{\alpha_i}, \hat{\alpha_j}) \neq 0$, we have shown that we may write α_j as a linear combination of the α_ℓ. Thus, we may replace α_i with α_j in the circuit and co-vector both. The tucker pivot then, is the induced change in $\mathbf{A}, B, C$ by the change in $\bar{X}, \bar{Y}$.

Example 7.3.1. Consider if we exchange $\hat{\alpha_1}$ for $\hat{\alpha_4}$, so $\bar{X} = \{\hat{\alpha_2}, \hat{\alpha_3}, \hat{\alpha_4}\}$, $\bar{Y} = \{\hat{\alpha_1}, \hat{\alpha_5}\}$.

We first note that $\mathbf{c} = 7\alpha_2 + 2\alpha_3 + 3\alpha_4$. So $C = -$. The region of $\mathbb{R}^3$ where $f > 0$ and $\hat{\alpha_2}, \hat{\alpha_3}, \hat{\alpha_4} = 0$, is the point $(10, 0, 0)$. Notice that $\hat{\alpha_1}(10, 0, 0) = 10 > 0, \hat{\alpha_5}(10, 0, 0) = -12 < 0$. So $B = (+, -)$.

The tucker pivot where one exchanges $\hat{\alpha_1}$ and $\hat{\alpha_4}$ results in the following tableau:

$$\begin{array}{r|ccc|c}
 & \hat{\alpha}_4 & \hat{\alpha}_2 & \hat{\alpha}_3 & \text{-1} \\
\hline
\hat{\alpha}_1 & - & + & + & + \\
\hat{\alpha}_5 & + & + & + & - \\
\hline
-1 & - & - & + & \\
\hline
\end{array}$$

Figure 7.2: Oriented matroid Tucker pivot

We observe that we obtain a new primal solution, the intersection of the hyperplanes defined by $\bar{X}$, and a dual solution, defined by the linear combination of the functionals, also defined by $\bar{X}$. Since we are given $\hat{\alpha_i} \in \bar{X}$, $C(\hat{\alpha_i}) \neq +$, we have written $\mathbf{c}$ as a non-negative linear combination of the α_i, and so we have a feasible dual solution. However, since $B(\hat{\alpha_5}) = -$, we have that this bound is exceeded, and this is not a primal feasible solution.

Remark 7.3.2. This gives us the essential tools for the Simplex Algorithm, defined on a finite Oriented Matroid Program. [BLVS+99]:

Input: A feasible (B-nonnegative) tableau, $T = (\bar{X}, \bar{Y}, \mathbf{A}, B, C)$.

I. If T is *optimal* (B-nonnegative, C-nonpositive) STOP. The current tableau is optimal.

II. If T is unbounded, (there is a $\hat{\alpha_i} \in \bar{X}$ such that $C(\hat{\alpha_i}) = +$, $\mathbf{A}(\hat{\alpha_i}, \hat{\alpha_j}) = -$ for some $\hat{\alpha_j} \in \bar{Y}$, and $\mathbf{A}(\hat{\alpha_i}, \hat{\alpha_k}) \neq +$ for any other $\hat{\alpha_k} \in \bar{Y}$. STOP. The current tableau is unbounded.

III. Choose $\hat{\alpha_i} \in \bar{X}$ such that $C(\hat{\alpha_i}) = +$.

IV. Choose $\hat{\alpha_j} \in \bar{Y}$ such that the co-vector induced by $\bar{Y} \cup \{\hat{\alpha_i}\} \backslash \{\hat{\alpha_j}\}$ is feasible.

V. Return to (I).

Proof. This is an adaptation of Blands Simplex Algorithm found in [BLVS+99] □

Example 7.3.3. Recall the Lumbermill Problem and its tableau:

$$
\begin{array}{c|ccc|c}
 & \hat{\alpha}_1 & \hat{\alpha}_2 & \hat{\alpha}_3 & -1 \\
\hline
\hat{\alpha}_4 & + & + & + & + \\
\hat{\alpha}_5 & + & + & + & + \\
\hline
-1 & + & + & + & \\
\hline
\end{array}
$$

If we swap $\hat{\alpha_1}$ with $\hat{\alpha_5}$, some simple computation shows that the new tableau will be:

$$
\begin{array}{c|ccc|c}
 & \hat{\alpha}_5 & \hat{\alpha}_2 & \hat{\alpha}_3 & -1 \\
\hline
\hat{\alpha}_4 & + & + & + & + \\
\hat{\alpha}_1 & + & + & + & + \\
\hline
-1 & - & + & + & \\
\hline
\end{array}
$$

This corresponds to a primal solution $\vec{x} = (4, 0, 0)$ and dual solution $\vec{y} = (0, -\frac{1}{2}, -\frac{7}{2}, 0, \frac{3}{2})$. However, as our dual solution is not yet feasible, we continue to pivot.

Suppose we switch $\hat{\alpha_3}$ and $\hat{\alpha_4}$. We would get:

$$
\begin{array}{c|ccc|c}
 & \hat{\alpha}_5 & \hat{\alpha}_2 & \hat{\alpha}_4 & -1 \\
\hline
\hat{\alpha}_3 & - & - & + & + \\
\hat{\alpha}_1 & + & - & - & + \\
\hline
-1 & - & - & - & \\
\hline
\end{array}
$$

The region defined by the intersecting hyperplanes (when $f = 1$) in $\mathbb{R}^3$ is $\vec{x} = (2, 0, 4)$. We see that $\hat{\alpha_1}(2, 0, 4) = 2$, and $\hat{\alpha_4}(2, 0, 4) = 4$, both positive. Moreover, we may write $\mathbf{c} = \frac{11}{3}\alpha_2 + \frac{5}{3}\alpha_4 + \frac{2}{3}\alpha_5$. Thus this point also corresponds to a feasible dual solution, $\vec{y} = (0, \frac{11}{3}, 0, \frac{5}{3}, \frac{2}{3})$.

We note that by Bland's Anti-Cycling rules, one can show that the Simplex Algorithm will terminate. [BLVS+99]

7.4 Tucker Tableau with entries in R

Here, we describe the Tucker tableau in terms of elements of the ground division ring R. Given $A : X \to Y^*, \vec{b}, \mathbf{c}$, as before, we define $\hat{\alpha_i}, \bar{X}, \bar{Y}$ as before. We then define $\mathbf{A}, B, C$ as not merely the sign of the appropriate coefficients, but the coefficients themselves. We also define a new value D, where D is the coefficient of f in the sum:

$$\sum_{\hat{\alpha_i} \in \bar{X}} \alpha_i \cdot c_i + g + f \cdot D = 0.$$

Then, the Tucker pivot is exactly as described before, exchanging $\hat{\alpha_i} \in \bar{X}$ with $\hat{\alpha_j} \in \bar{Y}$. We note that by definition

$$\hat{\alpha_j} = f \cdot B(\hat{\alpha_j}) - \sum_{\hat{\alpha_k} \in \bar{X}} \hat{\alpha_k} \cdot \mathbf{A}(\hat{\alpha_k}, \hat{\alpha_j}).$$

Suppose that $\mathbf{A}(\hat{\alpha_i}, \hat{\alpha_j}) \neq 0$, and for shorthand, we use $\mathbf{A}_{ij} := \mathbf{A}(\hat{\alpha_i}, \hat{\alpha_k}), B_j := B(\hat{\alpha_j}), C_i := C(\hat{\alpha_i})$. Then:

$$
\begin{aligned}
\hat{\alpha_j} &= f \cdot B_j - \sum_{\hat{\alpha_k} \in \bar{X}} \hat{\alpha_k} \mathbf{A}_{kj} \\
\hat{\alpha_i} \mathbf{A}_{ij} &= f \cdot B_j - \sum_{\hat{\alpha_k} \in \bar{X}, k \neq i} \hat{\alpha_k} \mathbf{A}_{kj} - \hat{\alpha_j} \\
\hat{\alpha_i} &= f \cdot B_j \mathbf{A}_{ij}^{-1} - \sum_{\hat{\alpha_k} \in \bar{X}, k \neq i} \hat{\alpha_k} \cdot \mathbf{A}_{kj} \mathbf{A}_{ij}^{-1} - \hat{\alpha_j} \mathbf{A}_{ij}^{-1}.
\end{aligned}
$$

With this, we consider the tableau defined on $\bar{X}' := \bar{X} \cup \{\hat{\alpha_j}\} \backslash \{\hat{\alpha_i}\}$, $\bar{Y}' := \bar{Y} \cup \{\hat{\alpha_i}\} \backslash \{\hat{\alpha_j}\}$ (the tucker pivot discussed above). We notice that for $\hat{\alpha_\ell} \in \bar{Y} \backslash \{\hat{\alpha_j}\}$,

$$
\begin{aligned}
\hat{\alpha_\ell} &= f \cdot B_\ell - \sum_{\hat{\alpha_k} \in \bar{X}} \hat{\alpha_k} \cdot \mathbf{A}_{k\ell} \\
\hat{\alpha_\ell} &= f \cdot B_\ell - \sum_{\hat{\alpha_k} \in \bar{X}, k \neq i} \hat{\alpha_k} \cdot \mathbf{A}_{k\ell} - \left(f \cdot B_j \mathbf{A}_{ij}^{-1} - \sum_{\hat{\alpha_k} \in \bar{X}, k \neq i} \hat{\alpha_k} \mathbf{A}_{kj} \mathbf{A}_{ij}^{-1} - \hat{\alpha_j} \mathbf{A}_{ij}^{-1} \right) \mathbf{A}_{i\ell} \\
\hat{\alpha_\ell} &= f \cdot \left(B_\ell - B_j \mathbf{A}_{ij}^{-1} \mathbf{A}_{i\ell} \right) - \sum_{\hat{\alpha_k} \in \bar{X}, k \neq i} \hat{\alpha_k} \cdot \left(\mathbf{A}_{k\ell} - \mathbf{A}_{kj} \mathbf{A}_{ij}^{-1} \mathbf{A}_{i\ell} \right) + \hat{\alpha_j} \cdot \left(\mathbf{A}_{ij}^{-1} \mathbf{A}_{i\ell} \right).
\end{aligned}
$$

Thus we have that:

$$
\begin{aligned}
\mathbf{A'}_{k\ell} &= \mathbf{A}_{k\ell} - \mathbf{A}_{kj} \mathbf{A}_{ij}^{-1} \mathbf{A}_{i\ell}, \quad \text{if } \ell \neq i, k \neq j. \\
&= -\mathbf{A}_{ij}^{-1} \mathbf{A}_{i\ell}, \quad \text{if } \ell \neq, k = j. \\
&= \mathbf{A}_{kj} \mathbf{A}_{ij}^{-1}, \quad \text{if } \ell = i, k \neq j. \\
&= \mathbf{A}_{ij}^{-1}, \quad \text{if } \ell = j, i = k.
\end{aligned}
$$

Similarly,

$$B'_\ell \;=\; \begin{cases} B_\ell - B_j \mathbf{A}_{ij}^{-1}\mathbf{A}_{i\ell}, & \text{if } \ell \neq i. \\ B_j \mathbf{A}_{ij}^{-1}, & \text{otherwise.} \end{cases}$$

We can also use this method to compute C':

$$\begin{aligned} g &= \sum_{\hat{\alpha_k} \in \bar{X}} \alpha_k(x) \cdot C_k - f \cdot D \\ &= \sum_{\hat{\alpha_k} \in \bar{X}, k \neq i} \alpha_k(x) \cdot C_k + \left(f \cdot B_j \mathbf{A}_{ij}^{-1} - \sum_{\hat{\alpha_k} \in \bar{X}, k \neq i} \hat{\alpha_k} \cdot \mathbf{A}_{kj}\mathbf{A}_{ij}^{-1} - \hat{\alpha_j}\mathbf{A}_{ij}^{-1}. \right) \cdot C_i - f \cdot D \\ &= \sum_{\hat{\alpha_k} \in \bar{X}, k \neq i} \alpha_k(x) \cdot \left(C_k - \mathbf{A}_{kj}\mathbf{A}_{ij}^{-1} C_i \right) - \hat{\alpha_j} \cdot \mathbf{A}_{ij}^{-1} C_i - f \cdot \left(D - B_j \mathbf{A}_{ij}^{-1} C_i \right). \end{aligned}$$

From this, we get that:

$$\begin{aligned} C'_k &= \begin{cases} C_k - \mathbf{A}_{kj}\mathbf{A}_{ij}^{-1} C_i, & \text{if } k \neq j, \\ -\mathbf{A}_{ij}^{-1} C_i, & \text{otherwise.} \end{cases} \\ D' &= D - B_i \mathbf{A}_{ij}^{-1} C_i. \end{aligned}$$

All of this can be encapsulated in the following figure:

$$\begin{array}{c|cc|c|} & \hat{\alpha_i} & \hat{\alpha_k} & -1 \\ \hline \hat{\alpha_j} & \mathbf{A}_{ij} & \mathbf{A}_{kj} & B_j \\ \hat{\alpha_\ell} & \mathbf{A}_{i\ell} & \mathbf{A}_{k\ell} & B_\ell \\ \hline -1 & C_i & C_k & D \\ \hline \end{array}$$

Figure 7.3: Generalized Tucker Tableau

$$
\begin{array}{c|cc|c|}
 & \hat{\alpha}_i & \hat{\alpha}_k & -1 \\ \hline
\hat{\alpha}_j & \mathbf{A}_{ij} & \mathbf{A}_{kj} & B_j \\
\hat{\alpha}_\ell & \mathbf{A}_{i\ell} & \mathbf{A}_{k\ell} & B_\ell \\ \hline
-1 & C_i & C_k & D \\ \hline
\end{array}
\longrightarrow
\begin{array}{c|cc|c|}
 & \hat{\alpha}_j & \hat{\alpha}_k & -1 \\ \hline
\hat{\alpha}_i & \mathbf{A}_{ij}^{-1} & \mathbf{A}_{kj}\mathbf{A}_{ij}^{-1} & B_j\mathbf{A}_{ij}^{-1} \\
\hat{\alpha}_\ell & -\mathbf{A}_{ij}^{-1}\mathbf{A}_{i\ell} & \mathbf{A}_{k\ell} - \mathbf{A}_{kj}\mathbf{A}_{ij}^{-1}\mathbf{A}_{i\ell} & B_\ell - B_j\mathbf{A}_{ij}^{-1}\mathbf{A}_{i\ell} \\ \hline
-1 & -\mathbf{A}_{ij}^{-1}C_i & C_k - \mathbf{A}_{kj}\mathbf{A}_{ij}^{-1}C_i & D - B_j\mathbf{A}_{ij}^{-1}C_i \\ \hline
\end{array}
$$

Figure 7.4: Generalized Tucker Pivot

7.5 The Simplex Algorithm

We make some observations, given a feasible Tucker tableau (each $B_j \geq 0$ for $\hat{\alpha}_j \in \bar{Y}$):

1. If each $C_i \leq 0$ for each $\hat{\alpha}_i \in \bar{X}$, then the tableau is optimal.
2. If there is a $C_i > 0$, but each $\mathbf{A}_{ij} \leq 0, \hat{\alpha}_j \in \bar{Y}$, then the primal problem is unbounded. (By Proposition 4.3.21, we can find a $w_i \in X$ for each $\hat{\alpha}_i \in \bar{X}$, such that $\alpha_k(w_i) = -\delta_{ki}$. Then $A(-w_i) \leq 0, \mathbf{c}(-w_i) > 0$, and the primal is unbounded).

For the remainder of this chapter, we will assume that $A \in \mathrm{Hom}_R(X, Y^*)$, the collection of row-like projections $\mathcal{A}$ (Definition 3.3.6) is finite ($|\mathcal{A}| < \infty$). That is, there are a finite number of α_i, and so we may number them $\alpha_1 \ldots, \alpha_{|\mathcal{A}|}$. We will also assume that the primal problem is feasible.

We can now discuss the Simplex Algorithm.

I) Select the minimum index i such that $\hat{\alpha}_i \in \bar{X}$, C_i is positive. If no such $\hat{\alpha}_i$ exists, then STOP. The tableau is optimal (as discussed before.)

II) Select an index j such that $\hat{\alpha}_j \in \bar{Y}$, $B_j\mathbf{A}_{ij}^{-1}$ is minimal with respect to $\mathbf{A}_{ij}$ positive. If there are multiple candidates for $\hat{\alpha}_j$, pick the one of lowest subscript. Note that if each

$\mathbf{A}_{ij} \le 0$ for the given $\hat{\alpha}_i$, then by our above observation the primal problem is infeasible, which contradicts our assumption that the tableau was feasible.

III) Perform a Tucker pivot with respect to $\hat{\alpha}_i, \hat{\alpha}_j$.

IV) Return to (I) as before.

Remark 7.5.1. We observe some facts about the Simplex Algorithm.

1. First, each Tucker pivot done this way results in a feasible tableau. Suppose that it did not. First, since $\mathbf{A}_{ij} > 0, B'_j = B_j \mathbf{A}_{ij}^{-1} \ge 0$. So suppose there is a $B'_\ell < 0$. We recall that:

$$\begin{aligned} B'_\ell = B_\ell - B_j \mathbf{A}_{ij}^{-1} \mathbf{A}_{i\ell} &< 0 \\ B_\ell &< B_j \mathbf{A}_{ij}^{-1} \mathbf{A}_{i\ell} \\ B_\ell \mathbf{A}_{i\ell}^{-1} &< B_j \mathbf{A}_{ij}^{-1}. \end{aligned}$$

   Noting that if $\mathbf{A}_{i\ell}^{-1} \le 0$, then $B'_\ell > 0$. Thus $B_\ell \mathbf{A}_{i\ell}^{-1} < B_j \mathbf{A}_{ij}^{-1}$, and this is a contradiction.

2. We also notice that given any choice $\bar{X}$ (equivalently $\bar{Y}$), the Tucker pivot is uniquely determined. Thus if the algorithm does not ever terminate, it is due to *cycling*. Let $\bar{X}_0$ denote an initial choice of $\bar{X}$, and given $\bar{X}_N$, let $\bar{X}_{N+1}$ denote the set $\bar{X}$ after a Tucker pivot, if the algorithm does not terminate. Then, consider a $\bar{X}_0$ such that the simplex algorithm never terminates. Since $\mathcal{A}$ is finite, there are only a finite number of choices for $\bar{X}_N$. Thus, by the pigeonhole principle, there is a $\bar{X}_N, M \in \mathbb{Z}^+$, such that $\bar{X}_N = \bar{X}_{N+M}$. But since $\bar{X}_{N+1}$ is uniquely determined, it follows that $\bar{X}_{N'+M} = \bar{X}_{N'}$ for any $N' > N$.

3. By our choices, $D' = D - B_j \mathbf{A}_{ij}^{-1} C_i$, and since $B_j, \mathbf{A}_{ij}^{-1}, C_i$ are each non-negative, D is non-increasing. Thus if cycling occurs, each D in the tableau induced by $\bar{X}_{N'}, T(\bar{X}_{N'})$ for $N' \ge N$ must be the same.

In order to show that the algorithm terminates (and returns an optimal or unbounded tableau) we require that cycling does not occur.

Theorem 7.5.2 (Generalized Fact IX: Simplex Algorithm)**.** *The Simplex Algorithm (as described above) does not cycle.*

This can be shown with an adaptation of the lexicographical perturbation rule, originally developed by Dantzig in 1951 [Dan].

Proof. Suppose that cycling does occur. Without loss of generality, we may assume that each $\hat{\alpha_i}$ both leaves and enters $\bar{X}$ at some pivot. That is, we may remove all the rows and columns that do not leave this pivot within this cycle. Since D is non-increasing , and in each pivot the value of C and $\mathbf{A}$ are positive, this shows that the value for B is negative. Thus, each $\hat{\alpha_i}$ leaves $\bar{Y}$ and enters $\bar{X}$, each $B(\hat{\alpha_i}) = 0$ for some choice of $\bar{X}$, and for that choice of $\bar{X}$, $\hat{\alpha_i}\left(\bigcap_{\hat{\alpha_k}\in\bar{X}} \mathrm{Ker}(\hat{\alpha_k})\right) = 0$. Since $\bar{X}$ is chosen to be maximal with respect to linear independence, and $\bar{X}$ is finite, this shows that $\mathrm{Ker}(\hat{\alpha_i})$ intersects the $\mathrm{Ker}(\hat{\alpha_k}), \hat{\alpha_k} \in \bar{X}$ at exactly the same point the $\hat{\alpha_k}$ intersect. (We may consider this intersection a single point by Remark 4.3.2). We consider the existence of such a point a *degenerate point*. We see by the above arguments that cycling occurs only if there exist degenerate points.

Then, by finite induction, we conclude that all the $\hat{\alpha_i}$ intersect at the same point. This is a necessary condition for cycling to occur. Thus for $\hat{\alpha_i}$, we may replace it with $\hat{\alpha_i}' := \hat{\alpha_i} - f \cdot \epsilon_i$, where $\epsilon_i \in R_+$. Since the original problem contained a finite number of functionals, we may pick these ϵ_i to be sufficiently small, and in such a way, such that no new degenerate points are created.

We do this for each degenerate point, each such that no new degenerate points are introduced. This adjusted program has no degenerate points, and so no cycling occurs. Thus this *perturbed program* terminates.

If the perturbed program is infeasible, then there is a $\vec{x} \in X$ such that $\alpha_i(\vec{x'}) \leq 0$, for each index i, $\mathbf{c}(\vec{x'}) > 0$, and a feasible point $\vec{x}$, where $\hat{\alpha}_i{}'(\vec{x}) \geq 0$, but $\hat{\alpha_i'}(\vec{x}) = \hat{\alpha}_i(\vec{x}) - f \cdot \epsilon_i$, and so such a $\vec{x}$ is feasible for the original program as well. The existence of a feasible $\vec{x}$, and a $\vec{x'}$ as described above, shows that the primal program is unbounded. (See proof of Proposition 5.3.2).

Otherwise the perturbed program is optimal, and returns a collection of functionals $\bar{X}$ such that we may write $\mathbf{c}$ as a non-negative linear combination of the $\alpha_i, \hat{\alpha}_i{}' \in \bar{X}$, and the intersection of the kernels of the functionals in $\bar{X}$ is feasible. But again $\hat{\alpha}_i = 0$ if and only if $\hat{\alpha}_i{}' = f \cdot \epsilon_i$. Thus this point is feasible in the original program, and the unperturbed functionals whose linear combination form $\mathbf{c}$ also correspond to hyperplanes whose intersections form the maximizer for the primal problem.

$\square$

Corollary 7.5.3. *We obtain an alternative proof of the Strong Duality Theorem (Theorem 5.3.7) and Theorem 6.4.2, where the collection of (non-zero) row-like projection maps $\mathcal{A}$ (Definition 3.3.6) is finite.*

Proof. Suppose that, by the hypothesis of Theorem 5.3.7, that the primal problem is feasible and bounded. Then we may encode the data from this program into the Tucker tableau, and by Theorem 7.5.2, the Simplex Algorithm terminates at an optimal tableau. Consider $\bar{X}$, we first note that we have encoded a function C such that:

$$\begin{aligned}
\mathbf{c} + \sum_{\hat{\alpha}_i \in \bar{X}} \hat{\alpha}_i \cdot (-C_i) + f \cdot D &= 0 \\
\mathbf{c} + \sum_{\hat{\alpha}_i \in \bar{X}} (b_i - \alpha_i) \cdot (-C_i) + D &= 0 \\
\sum_{\hat{\alpha}_i \in \bar{X}} (b_i) \cdot (-C_i) &= -D.
\end{aligned}$$

Since each $-C_i \geq_R 0$, this represents a feasible solution to the dual problem, whose objective value is $-D$.

Conversely, we record a region $K \subseteq X$ that is 0 for each $\hat{\alpha}_i \in \bar{X}$. Then for $k \in K$,

$$\begin{aligned} \mathbf{c}(k) + \sum_{\hat{\alpha}_i \in \bar{X}} \hat{\alpha}_i(k) \cdot (-C_i) + f \cdot D &= 0 \\ \mathbf{c}(k) + D &= 0 \\ \mathbf{c}(k) &= -D. \end{aligned}$$

Thus we have a feasible primal solution k whose objective value is $-D$. Then, by the Weak Duality Theorem, (Corollary 2.4.10), both of these solutions are optimal and give the same objective value.

□

Example 7.5.4. We frequently model linear or affine programming problems in terms of a constant number of resources or constraints, where the objective value relative to each constraint is fixed as well. This however does not reflect the actual reality of a production process. It is certainly possible that the amount of resources available fluctuates relative to other factors, in the market place or in terms of other resources. The amount of resources it takes to craft a product, and how much that product sells for may also fluctuate relative to other factors. It is natural then, to consider an affine programming problem where the coefficients are *functions* rather than real numbers.

So we consider (abstractly) a company who produces two products, product 1 and product 2. These are constrained by quantity of two inputs, input 1 and input 2. However, rather than being fixed we define the quantity of these inputs to be the (rational) function of two parameters, V, the availability of resources to the company on the market, and R the willingness of the company to expose the workforce to risk. Moreover, the amount of input it takes to produce these resources decreases as these factors increase.

So, let R be the ordered ring of rational functions $\mathbb{R}(V, R)$ as in Example 2.3.8 (note that any commutative ordered ring can be extended to an ordered field of fractions [Lam01]). We give this ring the lexicographical ordering $V > R$. Then, $X \cong R^2$, and since we have two non-negativity constraints, and two input constraints, $Y^* \cong R^4$. Let $\mathbf{c} : X \to R$ be $\mathbf{c}(x_1, x_2) := 5x_1 + 6x_2$, and let the input and non-negativity constraints be modeled by:

$$A := \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 3 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ R^2V \\ RV + 10 \end{pmatrix} =: \vec{b}.$$

We can then define the $\hat{\alpha_i}$:

$$\begin{aligned} \hat{\alpha_1} &= x_1 \\ \hat{\alpha_2} &= x_2 \\ \hat{\alpha_3} &= f \cdot R^2V - x_1 \cdot 1 - x_2 \cdot 3 \\ \hat{\alpha_4} &= f \cdot (RV + 10) - x_1 \cdot 2 - x_2 \cdot 2. \end{aligned}$$

If we initialize the tableau with $\bar{X} = \{\hat{\alpha_1}, \hat{\alpha_2}\}$, we have:

$$\begin{array}{c|cc|c} & \hat{\alpha}_1 & \hat{\alpha}_2 & -1 \\ \hline \hat{\alpha}_3 & 1 & 3 & R^2V \\ \hat{\alpha}_4 & 2 & 2 & RV+10 \\ \hline -1 & 5 & 6 & 0 \\ \hline \end{array}$$

This tableau is not optimal, or infeasible. Thus, we choose a $\hat{\alpha}_1$ to be the lowest index functional to leave $\bar{X}$ and we compare $\frac{RV+10}{2} < \frac{R^2V}{1} = R^2V$. Both of these are positive, so $\hat{\alpha}_4$ leaves $\bar{Y}$. This gives rise to the following tableau:

$$\begin{array}{c|cc|c} & \hat{\alpha}_4 & \hat{\alpha}_2 & -1 \\ \hline \hat{\alpha}_3 & -\frac{1}{2} & 2 & R^2V - \frac{RV}{2} - 5 \\ \hat{\alpha}_1 & \frac{1}{2} & 1 & \frac{RV}{2} + 5 \\ \hline -1 & -\frac{5}{2} & 1 & -\frac{5RV}{2} - 25 \\ \hline \end{array}$$

Since $1 > 0$, this tableau is not optimal. However, it is also not infeasible. Thus $\hat{\alpha}_2$ is our only candidate for leaving $\bar{X}$, and when we compare $\frac{RV}{2} + 5 < \frac{R^2V}{2} - \frac{RV}{4} - \frac{5}{2}$, we see that $\hat{\alpha}_1$ is the functional that leaves $\bar{Y}$. We then obtain:

$$\begin{array}{c|cc|c} & \hat{\alpha}_4 & \hat{\alpha}_1 & -1 \\ \hline \hat{\alpha}_3 & -\frac{3}{2} & -2 & R^2V - \frac{3RV}{2} - 15 \\ \hat{\alpha}_2 & \frac{1}{2} & 1 & \frac{RV}{2} + 5 \\ \hline -1 & -3 & -1 & -3RV - 30 \\ \hline \end{array}$$

Thus, where $\hat{\alpha_1}, \hat{\alpha_4}$ are both zero, the points $x_1 = 0$ units of product 1 and , $x_2 = \frac{RV}{2} + 5$ units of product 2, and the revenue is maximized at:

$$5 \cdot 0 + 6 \cdot \left(\frac{RV}{2} + 5\right) = 3RV + 30,$$

a function of the availability of resources, and willingness to accept risk.

There is a natural dual interpretation to this problem along the same lines as the Lumber Mill Problem 1.2.1. Recall that the quantity of input 1 is represented by the function R^2V and the quantity of input 2 is $RV + 10$. As before, we would like to find values for these resources, y_1, y_2 that minimize the value of our resources $^2V \cdot y_1 + (RV + 10) \cdot y_2$, subject to:

$$\begin{aligned} 1 \cdot y_1 + 2 \cdot y_2 &\geq 5 \\ 3 \cdot y_1 + 2 \cdot y_2 &\geq 6 \\ y_1, y_2 &\geq 0. \end{aligned}$$

From the above pivots, we see that this occurs when

$$y_1 = 0, y_2 = 3.$$

These are exactly the coefficients of α_3, α_4 respectively such that $\alpha_3 \cdot y_1 + \alpha_4 \cdot y_2 = \mathbf{c}$.

We verify that:

$$\begin{aligned}\alpha_3 \cdot y_1 + \alpha_4 \cdot y_2 &= 0 + 3) \cdot (RV + 10) \\ &= 3RV + 30.\end{aligned}$$

7.6 Conclusion

Here, we adapt the known algorithm for solving oriented matroid programs [BLVS+99] and place it within a Tucker tableau. We then expand on this and develop a Tucker tableau way of encoding an affine programming problem in a finite dimensional program over an ordered division ring. We then describe the generalized Simplex Algorithm 7.5.2 in this setting. Throughout this, we did not at any point require that our division rings have any property other than ordered, in particular this means that we did not require commutativity as one of our properties.

Chapter 8

Future Direction

One of the most important results of affine programming is the Farkas' Lemma, which allows us to prove the Complementary Slackness Theorem and the Strong Duality theorem. Moreover, as it provides a non-constructive proof of these results, it is crucial in the study of infinite dimensional programming, where algorithmic proofs such as the Simplex algorithm are not viable.

Although several generalizations of the Farkas' Lemma were proved, we have not established a necessary and sufficient condition where it holds, nor have we exhausted all the possible hypothesis under which it could hold. One possible direction of research would be to conclusively establish all the conditions under which the Farkas' Lemma, and in turn the Complementary Slackness and Strong Duality theorems hold.

Another potential direction is a careful examination of cones. In Chapter 3, we discuss some potential choices for positive cones, especially in terms of function spaces. However, we limited the discussion to orthant cones, after providing a counter example to Strong Duality in this setting. However, it is possible that our classical facts could in fact hold in this setting, with the proper hypothesis. One would then ask, using non-orthant cones, under which hypothesis our later basic facts may be generalized. Such a generalization would have natural applications to subjects such as infinite coloring and matching problems in graph theory.

We have also seen finite affine programming problems can be modeled by a finite oriented matroid. It can also be shown that over the cone of non-negative function $X \to R, |X| = \infty$, an affine programming problem can be modeled by an infinite oriented matroid. One possible avenue of research would be to study infinite oriented matroids and see what hypothesis must hold in order for our generalized facts to hold. Then to study other types of positive cones, and to see what the analogous combinatorial structure would be, and to ask when do the generalized facts hold under these conditions.

Bibliography

[Bar07] David Bartl. Farkas' Lemma, other theorems of the alternative, and linear programming in infinite-dimensional spaces: a purely linear-algebraic approach. *Linear Multilinear Algebra*, 55(4):327–353, 2007.

[Bar12a] David Bartl. Farkas' Lemma, Gale's theorem, and linear programming: the infinite case in an algebraic way. *Global Journal of Mathematical Sciences*, 1(1), 2012.

[Bar12b] David Bartl. A very short algebraic proof of the Farkas Lemma. *Math. Methods Oper. Res.*, 75(1):101–104, 2012.

[Bea55] E.M.L. Beale. Cycling in the dual simplex algorithm. *Naval Research Logistics Quaterly*, 2(0082187 (18,514e)):269–275 (1956), 1955.

[Bla77] Robert G. Bland. New finite pivoting rules for the simplex method. *Mathematics Operations Research*, 2(2):103–107, 1977.

[BLV78] Robert G. Bland and Michel Las Vergnas. Orientability of matroids. *J. Combinatorial Theory Ser. B.*, 24(1):94–123, 1978.

[BLVS$^+$99] Anders Björner, Michel Las Vergnas, Bernd Sturmfels, Neil White, and Günter M. Ziegler. *Oriented matroids*, volume 46 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 1999.

[Con90] John B. Conway. *A course in functional analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[Dan] G. B. Dantzig. Maximization of linear function of variables subject to linear inequalities.

[DeN82] E.V. DeNardo. *Dynamic programming*. Dover Books on Computer Science Series. Dover Publications, Incorporated, 1982.

[DF04] David S. Dummit and Richard M. Foote. *Abstract algebra*. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2004.

[Far02] Julius Farkas. Theorie der einfachen ungleichungen. *Journal fur die reine und angewandte Mathematik*, 124:1–27, 1902.

[Fuc11] László Fuchs. *Partially Ordered Algebraic Systems*. Dover Publications, 2011.

[Hil03] David Hilbert. *Grundlagen der geometrie*. Leipzig B.G. Teubner, 1903.

[Ion27] D.V. Ionescu. *Sur une classe d'équations fonctionnelles*. impr. et librairie E. Privat, 1927.

[JGKR63] Robin Robinson John G. Kennedy and Robert W. Ritchie. *New Directions in Mathematics*. Prentince-Hall Inc, 1963.

[Lam01] T. Y. Lam. *A first course in noncommutative rings*, volume 131 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001.

[Law75] James Franklin Lawrence. *ORIENTED MATROIDS*. ProQuest LLC, Ann Arbor, MI, 1975. Thesis (Ph.D.)–University of Washington.

[ML98] Saunders Mac Lane. *Categories for the working mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1998.

[Mot36] T.S. Motzkin. *Beiträge zur Theorie der linearen Ungleichungen*. Buchdr. Azriel, 1936.

[NT93] Evar D. Nering and Albert W. Tucker. *Linear programs and related problems*. Computer Science and Scientific Computing. Academic Press Inc., Boston, MA, 1993. With a foreword by George B. Dantzig, With 1 IBM-PC floppy disk (5.25 inch; HD).

[Num55] Dantzig, tobias. numbers: The language of science. new york: The macmillan company, 1954. 340 p. *Science Education*, 39(2):182–182, 1955.

[Piv08] Marcus Pivato. Sustainable preferences via nondiscounted, hyperreal intergenerational welfare functions. MPRA Paper 7461, University Library of Munich, Germany, March 2008.

[Rob79] Abraham Robinson. *Selected papers of Abraham Robinson. Vol. II*. Yale University Press, New Haven, Conn., 1979. Nonstandard analysis and philosophy, Edited and with introductions by W. A. J. Luxemburg and S. Körner.

[Roc69] R. Tyrell Rockafellar. *Combinatorial mathematics and its applications*, volume 1967 of *Proceedings of the Conference held at the University of North Carolina at Chapel Hill, April 10-14*. The University of North Carolina Press, Chapel Hill, N.C., 1969.

[Roy10] Halsey Royden. *Real Analysis*. Pearson, 4 edition, 2010.

[Sch86] Alexander Schrijver. *Theory of linear and integer programming*. John Wiley & Sons, Inc., New York, NY, USA, 1986.

[vN98] John von Neumann. *Continuous geometry*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1998. With a foreword by Israel Halperin, Reprint of the 1960 original, Princeton Paperbacks.

[Zie13] Günter M. Ziegler. *Lectures on Polytopes.* Springer, 2013.

Index