Caltech Division of the Humanities and Social Sciences

Alternative Linear Inequalities

KC Border
September 2013
v. 2020.10.15::09.50

Contents

1 Executive Summary

In these notes we present some basic results on the existence of solutions to systems of linear equalities and inequalities. We usually express these as matrix equations such as $Ax = b$ or inequalities such as $Ax \leqq b$. Occasionally we may write them out as a collection of expressions of the form $a_1x_1 + \cdots + a_nx_n \leqslant b$, or such.

The usual ordering on $\boldsymbol{R}$ is denoted $\geqslant$ or $\leqslant$. On $\boldsymbol{R}^{\mathrm{n}}$, the partial ordering $x \geqq y$ means $x_i \geqslant y_i$, $i = 1,\dots,n$, while $x \gg y$ means $x_i > y_i$, $i = 1,\dots,n$. We may occasionally write $x > y$ to mean $x \geqq y$ and $x \neq y$. A vector x is **nonnegative** if $x \geqq 0$, **strictly positive** if $x \gg 0$, and **semipositive** if $x > 0$. I shall try to avoid using the adjective "positive" by itself, since to most mathematicians it means "nonnegative," but to many nonmathematicians it means "strictly positive." Define $\boldsymbol{R}^{\mathrm{n}}_{+} = \{x \in \boldsymbol{R}^{\mathrm{n}} : x \geqq 0\}$ and $\boldsymbol{R}^{\mathrm{n}}_{++} = \{x \in \boldsymbol{R}^{\mathrm{n}} : x \gg 0\}$, the **nonnegative orthant** and **strictly positive orthant** of $\boldsymbol{R}^{\mathrm{n}}$ respectively.

$$
\begin{aligned}
x \geqq y &\iff x_i \geqslant y_i,\ i = 1, \dots, n \\
x > y &\iff x_i \geqslant y_i,\ i = 1, \dots, n \text{ and } x \neq y \\
x \gg y &\iff x_i > y_i,\ i = 1, \dots, n
\end{aligned}
$$

Figure 1. Partial orderings on $\boldsymbol{R}^{\mathrm{n}}$.

In these notes I shall adopt David Gale's [13] notation, which does not distinguish between row and column vectors.

> *This means that if A is an $m \times n$ matrix, and x is a vector, and I write Ax, you infer that x is an n-dimensional column vector, and if I write yA, you infer that y is an m-dimensional row vector. The notation yAx means that x is an n-dimensional column vector, y is an m-dimensional row vector, and yAx is the scalar $yA \cdot x = y \cdot Ax$. The expression xy will always denote the scalar product $x \cdot y$ of two vectors of the same dimension.*

The theorems on the existence of solutions are in the form of alternatives, that is, "an opportunity for choice between two things, courses, or propositions, either of which may be chosen, but not both" [23]. The archetypal example is Theorem 1 below, which says that either $Ax = b$ has a solution x, or else the system $pA = 0$, $p \cdot b > 0$ has a solution p, but not both.

There are two kinds of solutions we might look for. If all we know is that A is matrix of real numbers and that b is a vector of real numbers, then the best we can hope for is to find a real vector x satisfying $Ax = b$. But if A is a matrix of rational numbers and b is a vector of rational numbers, then we might hope to find a rational vector x satisfying $Ax = b$. (In these notes we are never interested in complex numbers.) Thus we present two kinds of theorems, one for the case where A and b are known only to be real, and the other for the case where A and b are known to be rational. The simplest and perhaps most intuitive proofs rely on the geometry of linear inequalities, and involve the use of separating hyperplane theorems. They yield real solutions only. Another class of proofs is algebraic, and is based on what happens when we actually try to solve the equations or inequalities by the method of elimination. These proofs will yield rational solutions when A and b are rational. The disadvantage of these algebraic proofs is that they are not very intuitive, and a little tedious.

I shall start by stating the main results and shall present the proofs later.

1.1 Solutions of systems of equalities

If A has an inverse (which implies $m = n$), then the system $Ax = b$ always has a unique solution, namely $\bar{x} = A^{-1}b$. But even if A does not have an inverse, the system may have a solution, possibly several—or it may have none. This brings up the question of how to characterize the existence of a solution. The answer is given by the following theorem. Following Riesz–Sz.-Nagy [26, p. 164] and wikipedia, I shall refer to it as the Fredholm Alternative, as Fredholm [10] proved it in 1903 in the context of integral equations. But I do note that Marlow [19, p. 86] refers to it as Gale's Theorem. It does appear in Gale [13, Theorem 2.5, p. 41].

1 Fredholm Alternative *Let A be an $m \times n$ real matrix and let $b \in \boldsymbol{R}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists an $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax = b \tag{1}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ *satisfying*

$$\begin{aligned} pA &= 0 \\ p \cdot b &> 0. \end{aligned} \tag{2}$$

Note that there are many trivial variations on this result. For instance, multiplying p by -1, I could have written (2) as $pA = 0$ & $p \cdot b < 0$. Or by replacing A by its transpose, we could rewrite (1) with $xA = b$ and (2) with $Ap = 0$. Keep this in mind as you look at the coming theorems.

The following corollary about linear functions is true in quite general linear spaces. Wim Luxemburg refers to this result as the **Fundamental Theorem of Duality**.

2 Corollary *Let* $p^0, \dots, p^m \in \boldsymbol{R}^{\mathrm{n}}$ *and suppose that* $p^0 \cdot v = 0$ *for all* v *such that* $p^i \cdot v = 0$, $i = 1, \dots, m$. *Then* p^0 *is a linear combination of* $p^1, \dots, p^m$. *That is, there exist scalars* $\lambda_1, \dots, \lambda_m$ *such that* $p^0 = \sum_{i=1}^m \lambda_i p_i$.

It is also true that the Fredholm alternative holds for the field $\mathbb{Q}$ of rational numbers. We shall say that a matrix is rational if each of its entries is a rational number. The following theorem is on its face neither weaker nor stronger than Theorem 1.

3 Theorem (Rational Fredholm Alternative) *Let* A *be an* $m \times n$ *rational matrix and let* $b \in \mathbb{Q}^{\mathrm{m}}$. *Exactly one of the following alternatives holds. Either there exists an* $x \in \mathbb{Q}^{\mathrm{n}}$ *satisfying*

$$Ax = b \tag{3}$$

or else there exists $p \in \mathbb{Q}^{\mathrm{m}}$ *satisfying*

$$\begin{aligned} pA &= 0 \\ p \cdot b &> 0. \end{aligned} \tag{4}$$

Note that (4) is positively homogeneous, that is if p satisfies (4), then so does λp for any $\lambda > 0$. Consequently, if each component of p is rational, we may without loss of generality assume that the denominator of each component is positive, so multiplying by a positive common denominator will give another solution of (4), but in *integers* $\mathbb{Z}$. Thus we have the following result.

4 Theorem (Rational-Integer Fredholm Alternative) *Let* A *be an* $m \times n$ *rational matrix and let* $b \in \mathbb{Q}^{\mathrm{m}}$. *Exactly one of the following alternatives holds. Either there exists an* $x \in \mathbb{Q}^{\mathrm{n}}$ *satisfying*

$$Ax = b \tag{5}$$

or else there exists an integral $p \in \mathbb{Z}^{\mathrm{m}}$ *satisfying*

$$\begin{aligned} pA &= 0 \\ p \cdot b &> 0. \end{aligned} \tag{6}$$

Finally, we can produce this result.

5 Theorem (Rational-Real-Integer Fredholm Alternative) *Let* A *be an* $m \times n$ *rational matrix and let* $b \in \mathbb{Q}^{\mathrm{m}}$. *Exactly one of the following alternatives holds. Either there exists an* $x \in \boldsymbol{R}^{\mathrm{n}}$ *satisfying*

$$Ax = b \tag{7}$$

or else there exists an integral $p \in \mathbb{Z}^{\mathrm{m}}$ *satisfying*

$$\begin{aligned} pA &= 0 \\ p \cdot b &> 0. \end{aligned} \tag{8}$$

That is, if there is no real solution of (7), then there is an integer solution of (8). The proof is trivial: Clearly we cannot have both (7) and (8) hold, for then $0 = 0x = pAx = pb > 0$, a contradiction. But if (8) fails, then Theorem 4 guarantees a rational, and *a fortiori* real solution to (7).

1.2 Nonnegative solutions of systems of equalities

The next theorem is one of many more or less equivalent results on the existence of solutions to linear inequalities. It is often known as Farkas's Lemma, and so is Corollary 11. Julius Farkas [8] proved them both in 1902.

6 Farkas's Alternative *Let A be an $m \times n$ real matrix and let $b \in \boldsymbol{R}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &= b \\ x &\geqq 0 \end{aligned} \tag{9}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &\geqq 0 \\ p \cdot b &< 0. \end{aligned} \tag{10}$$

Note that by replacing p by $-p$ we can replace (10) by

$$\begin{aligned} pA &\leqq 0 \\ p \cdot b &> 0. \end{aligned} \tag{10$'$}$$

We also have the following results.

7 Theorem (Rational Farkas's Alternative) *Let A be an $m \times n$ rational matrix and let $b \in \mathbb{Q}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists $x \in \mathbb{Q}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &= b \\ x &\geqq 0 \end{aligned} \tag{11}$$

or else there exists $p \in \mathbb{Q}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &\geqq 0 \\ p \cdot b &< 0. \end{aligned} \tag{12}$$

8 Theorem (Rational-Integer Farkas's Alternative) *Let A be an $m \times n$ rational matrix and let $b \in \mathbb{Q}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists $x \in \mathbb{Q}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &= b \\ x &\geqq 0 \end{aligned} \tag{13}$$

or else there exists integral $p \in \mathbb{Z}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &\geqq 0 \\ p \cdot b &< 0. \end{aligned} \tag{14}$$

9 Theorem (Real-Rational-Integer Farkas's Alternative) *Let A be an $m \times n$ rational matrix and let $b \in \mathbb{Q}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &= b \\ x &\geqq 0 \end{aligned} \tag{15}$$

or else there exists integral $p \in \mathbb{Z}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &\geqq 0 \\ p \cdot b &< 0. \end{aligned} \tag{16}$$

1.3 Solutions of Inequalities

The next result may be found in Gale [13, Theorem 2.7, p. 46] (in transposed form with A multiplied by -1) and in Bachem and Kern [2, Theorem 4.1, p. 46]. (Bachem–Kern refer to it as Farkas's Lemma, but most authors reserve that for results on nonnegative solutions to inequalities.)

10 Proposition *Let A be an $m \times n$ matrix and let $b \in \boldsymbol{R}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists an $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax \leqq b \tag{17}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &= 0 \\ p \cdot b &< 0 \\ p &> 0 \end{aligned} \tag{18}$$

1.4 Nonnegative Solutions of Inequalities

Once we have a result on nonnegative solutions of equalities, we get one on nonnegative solutions of inequalities almost free. This is because the system

$$\begin{aligned} Ax &\leqq b \\ x &\geqq 0 \end{aligned}$$

is equivalent to the system

$$\begin{aligned} Ax + z &= b \\ x &\geqq 0 \\ z &\geqq 0, \end{aligned}$$

or in partitioned matrix form

$$\begin{aligned} [A \ I] \begin{bmatrix} x \\ z \end{bmatrix} &= b \\ \begin{bmatrix} x \\ z \end{bmatrix} &\geqq 0. \end{aligned}$$

The next result follows from Farkas's Alternative 6 using this transformation.

11 Corollary (Farkas's Alternative) *Let A be an $m \times n$ matrix and let $b \in \boldsymbol{R}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists an $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &\leqq b \\ x &\geqq 0 \end{aligned} \tag{19}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &\geqq 0 \\ p \cdot b &< 0 \\ p &> 0 \end{aligned} \tag{20}$$

Both versions of Farkas's Lemma are subsumed by the next result, which is also buried in Farkas [8].

12 Farkas's Alternative *Let A be an $m \times n$ real matrix, let B be an $\ell \times n$ matrix, let $b \in \boldsymbol{R}^{\mathrm{m}}$, and let $c \in \boldsymbol{R}^{\ell}$. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &= b \\ Bx &\leqq c \\ x &\geqq 0 \end{aligned} \tag{21}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ and $q \in \boldsymbol{R}^{\ell}$ satisfying

$$\begin{aligned} pA + qB &\geqq 0 \\ q &\geqq 0 \\ p \cdot b + q \cdot c &< 0. \end{aligned} \tag{22}$$

13 Farkas's Alternative *Let A be an $m \times n$ matrix, let $b \in \boldsymbol{R}^{\mathrm{m}}$, and let I and E partition $\{1, \dots, m\}$. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} (Ax)_i &= b_i \quad i \in E \\ (Ax)_i &\leqslant b_i \quad i \in I \\ x &\geqq 0 \end{aligned} \tag{23}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &\geqq 0 \\ p \cdot b &< 0 \\ p_i &\geqq 0 \quad i \in I. \end{aligned} \tag{24}$$

We can also handle strict inequalities. The next theorem is due to Gordan [14] in 1873.

14 Gordan's Alternative *Let A be an $m \times n$ matrix. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax \gg 0. \tag{25}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &= 0 \\ p &> 0 \end{aligned} \tag{26}$$

We can rewrite Corollary 14 as follows.

15 Corollary *Let A be an $m \times n$ matrix. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax \gg 0 \tag{27}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &= 0 \\ p \cdot \mathbf{1} &= 1 \\ p &> 0. \end{aligned} \tag{28}$$

(The second alternative implies that p is a probability vector.)

Dantzig [4, p. 139] attributes the nest result to Jean Ville [30]. It may also be found in Gale [13, Theorem 2.10, p. 49].

16 Corollary (Ville's Alternative) *Let A be an $m \times n$ matrix. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &\gg 0 \\ x &\geqq 0. \end{aligned} \tag{29}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &\leqq 0 \\ p &> 0 \end{aligned} \tag{30}$$

The following result was proved by Stiemke [27] in 1915.

17 Stiemke's Alternative *Let A be an $m \times n$ matrix. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax > 0 \tag{31}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &= 0 \\ p &\gg 0. \end{aligned} \tag{32}$$

A variation on Stiemke's Alternative points to another whole class of theorems that I first encountered in Morris [20]. Here is Corollary A1 from his paper.

18 Theorem (Morris's Alternative) *Let A be an $m \times n$ matrix. Let $\mathcal{S}$ be a family of nonempty subsets of $\{1, \dots, m\}$. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ and a set $S \in \mathcal{S}$ satisfying*

$$\begin{aligned} Ax &\geqq 0 \\ A_i \cdot x &> 0 \quad \textit{for all } i \in S \end{aligned} \tag{33}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &= 0 \\ p &\geqq 0 \\ \sum_{i \in S} p_i &> 0 \quad \textit{for all } S \in \mathcal{S}. \end{aligned} \tag{34}$$

19 Remark Observe that Stiemke's Alternative corresponds to the case where $\mathcal{S}$ is the set of all singletons: (33) reduces to Ax being semipositive. And $\sum_{i \in S} p_i > 0$ for the singleton $S = \{i\}$ simply says $p_i > 0$. Requiring this for all singletons asserts that $p \gg 0$.

Finally we come to another alternative, Motzkin's Transposition Theorem [21], proven in his 1934 Ph.D. thesis. This statement is taken from his 1951 paper [22].[1]

20 Motzkin's Transposition Theorem *Let A be an $m \times n$ matrix, let B be an $\ell \times n$ matrix, and let C be an $r \times n$ matrix, where B or C may be omitted (but not A). Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &\gg 0 \\ Bx &\geqq 0 \\ Cx &= 0 \end{aligned} \tag{35}$$

or else there exist $p^1 \in \boldsymbol{R}^{\mathrm{m}}$, $p^2 \in \boldsymbol{R}^{\ell}$, and $p^3 \in R^{\mathrm{r}}$ satisfying

$$\begin{aligned} p^1 A + p^2 B + p^3 C &= 0 \\ p^1 &> 0 \\ p^2 &\geqq 0. \end{aligned} \tag{36}$$

Motzkin expressed (36) in terms of the transpositions of A, B, and C. The reason that A may not be omitted is that without some strict inequalities, $x = 0$, $p_2 = 0$, $p_3 = 0$ solves both systems (35) and (36).

Stoer and Witzgall [28] also provide a rational version of Motzkin's theorem, which can be recast as follows.

21 Motzkin's Rational Transposition Theorem *Let A be an $m \times n$ rational matrix, let B be an $\ell \times n$ rational matrix, and let C be an $r \times n$ rational matrix, where B or C may be omitted (but not A). Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &\gg 0 \\ Bx &\geqq 0 \\ Cx &= 0 \end{aligned} \tag{37}$$

or else there exist $p^1 \in \mathbb{Z}^{\mathrm{m}}$, $p^2 \in \mathbb{Z}^{\ell}$, and $p^3 \in \mathbb{Z}^{\mathrm{r}}$ satisfying

$$\begin{aligned} p^1 A + p^2 B + p^3 C &= 0 \\ p^1 &> 0 \\ p^2 &\geqq 0. \end{aligned} \tag{38}$$

1.5 Tucker's Theorem

Tucker [29, Lemma, p. 5] proves the following theorem that is related to Theorems of the Alternative, but not stated as an alternative. See Nikaidô [24, Theorem 3.7, pp. 36–37] for a proof of Tucker's Theorem using the Stiemke's Alternative, and vice-versa.

[1] Motzkin [22] contains an unfortunate typo. The condition $Ax \gg 0$ is erroneously given as $Ax \ll 0$.

22 Tucker's Theorem *Let A be an $m \times n$ matrix. Then there exist $x \in \boldsymbol{R}^{\mathrm{n}}$ and $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying*

$$\begin{aligned} Ax &= 0 \\ x &\geqq 0 \\ A'p &\geqq 0 \\ A'p + x &\gg 0, \end{aligned} \tag{39}$$

where A' is the transpose of A.

To get an idea of the connection between Stiemke's Alternative and Tucker's Theorem, consider the transposed version of Stiemke's Alternative 17. It has two "dual" systems of inequalities

$$A'p > 0 \tag{31$'$}$$

and

$$Ax = 0, \qquad x \gg 0 \tag{32$'$}$$

exactly one of which has a solution. Tucker's Theorem replaces these with the weaker systems

$$A'p \geqq 0, \tag{31$''$}$$

$$Ax = 0, \qquad x \geqq 0. \tag{32$''$}$$

These always have the trivial solution $p = 0$, $x = 0$. What Tucker's Theorem says is that there is a solution $(\bar{p}, \bar{x})$ of (31''–32'') such that if the i^{th} component $(A'\bar{p})_i = 0$, then the i^{th} component $\bar{x}_i > 0$; and if $\bar{x}_i = 0$, then the i^{th} component $(A'\bar{p})_i > 0$. Not only that, but since $A\bar{x} = 0$, we have $(A'\bar{p}) \cdot \bar{x} = \bar{p}A\bar{x} = 0$, so for each i we cannot have both $(A'\bar{p})_i > 0$ and $\bar{x}_i > 0$. Thus we conclude that $A'\bar{p}$ and $\bar{x}$ exhibit **complementary slackness**:

$$(A'\bar{p})_i > 0 \text{ if and only if } \bar{x}_i = 0, \text{ and } \bar{x}_i > 0 \text{ if and only if } (A'\bar{p})_i = 0.$$

Tucker's Theorem is also a statement about nonnegative vectors in complementary orthogonal linear subspaces. The requirement that $Ax = 0$ says that x belong to the null space (kernel) of A. The vector $A'p$ belongs to the range of A'. It is well-known (see, e.g., Theorem 53 in § 7.2 of my notes on linear algebra [3]) that the null space of A and the range of A' are complementary orthogonal linear subspaces. Moreover every pair of complementary orthogonal subspaces arises this way. (Let A be the orthogonal projection onto one of the subspaces. Thus we have the following equivalent version of Tucker's Theorem, which appears as Corollary 4.7 in Bachem and Kern [2], and which takes the form of an alternative.

23 Corollary *Let M be a linear subspace of $\boldsymbol{R}^{\mathrm{m}}$, and let $M^\perp$ be its orthogonal complement. For each $i = 1, \ldots, m$, either there exists $x \in \boldsymbol{R}^{\mathrm{m}}$ satisfying*

$$x \in M, \quad x \geqq 0, \quad x_i > 0 \tag{40}$$

or else there exists $y \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$y \in M^\perp, \quad y \geqq 0, \quad y_i > 0. \tag{41}$$

2 Basic nonnegative linear combinations

Let $A = \{x_1, \dots, x_n\}$ be a finite set of vectors in a vector space. Let us say that the linear combination $y = \sum_{i=1}^n \lambda_i x_i$ **depends on B** if $B = \{x_i \in A : \lambda_i \neq 0\}$. By convention, we also agree that the zero vector depends on the empty set. Let us say that y is a **basic linear combination** of the x_i's if it depends on a linearly independent set. You should know from linear algebra that every linear combination can be replaced by a basic linear combination. The trick we want is to do it with nonnegative coefficients. The next result is true for general (not necessarily finite dimensional) vector spaces.

24 Lemma *A nonnegative linear combination of a set of vectors can be replaced by a nonnegative linear combination that depends on a linearly independent subset.*

That is, if $x_1, \dots, x_n$ are vectors and $y = \sum_{i=1}^n \lambda_i x_i$ where each λ_i is nonnegative, then either $y = 0$, or there exist nonnegative scalars $\beta_1, \dots, \beta_n$ such that $y = \sum_{i=1}^n \beta_i x_i$ and $\{x_i : \beta_i > 0\}$ is linearly independent.

Proof: Since the empty set is vacuously independent, our convention covers the case of $y = 0$. We treat the remaining case by induction on the number of vectors x_i on which nonzero y depends.

So let $\mathbb{P}[n]$ be the proposition: A nonnegative linear combination of not more than n vectors can be replaced by a nonnegative linear combination that depends on a linearly independent subset.

The validity of $\mathbb{P}[1]$ is easy. If $y \neq 0$ and $y = \lambda_1 x_1$, where $\lambda_1 \geqslant 0$, then we must in fact have $\lambda_1 > 0$ and $x_1 \neq 0$. That is, y depends on the linearly independent subset $\{x_1\}$.

We now show that $\mathbb{P}[n-1] \implies \mathbb{P}[n]$. So assume $y = \sum_{i=1}^n \lambda_i x_i$ and that each $\lambda_i > 0$, $i = 1, \dots, n$. If $x_1, \dots, x_n$ itself constitutes an independent set, there is nothing to prove, just set $\beta_i = \lambda_i$ for each i. On the other hand, if $x_1, \dots, x_n$ are dependent, then there exist numbers $\alpha_1, \dots, \alpha_n$, not all zero, such that

$$\sum_{i=1}^n \alpha_i x_i = 0.$$

We may assume that at least one $\alpha_i > 0$, for if not we simply replace each α_i by $-\alpha_i$.

Now consider the following expression

$$\begin{aligned} y &= \sum_{i=1}^n \lambda_i x_i - \gamma \underbrace{\sum_{i=1}^n \alpha_i x_i}_{=0} \\ &= \sum_{i=1}^n \big(\lambda_i - \gamma\alpha_i\big) x_i. \end{aligned}$$

When $\gamma = 0$, this reduces to our original expression. Whenever $\gamma > 0$ and $\alpha_i \leqslant 0$, then $\lambda_i - \gamma\alpha_i > 0$, so the only coefficients that we need to worry about are those with $\alpha_i > 0$. We will choose $\gamma > 0$ just large enough so that at least one of the coefficients $\lambda_i - \gamma\alpha_i$ becomes zero and none become negative. Now for $\alpha_i > 0$,

$$\lambda_i - \gamma\alpha_i \geqslant 0 \iff \gamma \leqslant \frac{\lambda_i}{\alpha_i}.$$

Thus by setting

$$\bar\gamma = \min\left\{\frac{\lambda_i}{\alpha_i} : \alpha_i > 0\right\}$$

we are assured that

$$\lambda_i - \bar{\gamma}\alpha_i \geqslant 0 \text{ for all } i = 1, \ldots, n \quad \text{and} \quad \lambda_i - \bar{\gamma}\alpha_i = 0 \text{ for at least one } i.$$

Thus

$$y = \sum_{i=1}^{n} \big(\lambda_i - \bar{\gamma}\alpha_i\big) x_i$$

expresses y as a linear combination depending on no more than $n-1$ of the x_i's. Thus by the induction hypothesis $\mathbb{P}[n-1]$, we can express y as a linear combination that depends on a linearly independent subset. ∎

25 Remark The above proof is highly instructive and is typical of the method we shall use in the study of inequalities. We started with two equalities in n variables

$$y = \sum_{i=1}^{n} \lambda_i x_i$$
$$0 = \sum_{i=1}^{n} \alpha_i x_i.$$

We then took a linear combination of the two equalities, namely

$$1y + \gamma 0 = 1 \sum_{i=1}^{n} \lambda_i x_i + \gamma \sum_{i=1}^{n} \alpha_i x_i,$$

where the coefficients 1 and γ were chosen to eliminate one of the variables, thus reducing a system of equalities in n variables to a system in no more than $n-1$ variables. Keep your eyes open for further examples of this techniques! (If you want to be pedantic, you might remark as Kuhn [15] did, that we did not really "eliminate" a variable, we just set its coefficient to zero.)

The first application of Lemma 24 is Carathéodory's theorem on convex hulls in finite dimensional spaces.

26 Carathéodory's Convexity Theorem *In $\boldsymbol{R}^{\mathrm{m}}$, every vector in the convex hull of a set can be written as a convex combination of at most $m+1$ vectors from the set.*

Proof: Let A be a subset of $\boldsymbol{R}^{\mathrm{m}}$, and let x belong to the convex hull of A. Then we can write x as a convex combination $x = \sum_{i=1}^{n} \lambda_i x_i$ of points x_i belonging to A. For any vector y in $\boldsymbol{R}^{\mathrm{m}}$ consider the "augmented" vector $\hat{y}$ in $\boldsymbol{R}^{\mathrm{m}+1}$ defined by $\hat{y}_j = y_j$ for $j = 1, \ldots, m$ and $\hat{y}_0 = 1$. Then it follows that $\hat{x} = \sum_{i=1}^{n} \lambda_i \hat{x}_i$ since $\sum_{i=1}^{n} \lambda_i = 1$. Renumbering if necessary, by Lemma 24, we can write $\hat{x} = \sum_{i=1}^{k} \alpha_i \hat{x}_i$, where $\hat{x}_1, \ldots, \hat{x}_k$ are independent and $\alpha_i > 0$ for all i. Since an independent set in $\boldsymbol{R}^{\mathrm{m}+1}$ has at most $m+1$ members, $k \leqslant m+1$. But this reduces to the two equations $x = \sum_{i=1}^{k} \alpha_i x_i$ and for the 0^{th} component $1 = \sum_{i=1}^{k} \alpha_i$. In other words, x is a convex combination of $k \leqslant m+1$ vectors of A. ∎

27 Remark We shall find the mapping that takes a vector x in $\boldsymbol{R}^{\mathrm{m}}$ to the vector $\hat{x} = (1, x)$ in $\boldsymbol{R}^{\mathrm{m}+1}$ quite useful. I wish I had a good name for it.

28 Corollary *The convex hull of a compact subset of $\boldsymbol{R}^{\mathrm{m}}$ is compact.*

Proof: Let K be compact and define the mapping from $K^{m+1} \times \Delta_m$ (where as you may recall, Δ_m is the unit simplex $\{\alpha \in \boldsymbol{R}^{\mathrm{m}+1} : (\forall i = 0,\dots,m)\ [\,\alpha_i \geqslant 0\,]\ \&\ \sum_{i=0}^m \alpha_i = 1\}$) into $\boldsymbol{R}^{\mathrm{m}}$ by

$$\big(x_0,\dots,x_m,(\alpha_0,\dots,\alpha_m)\big) \mapsto \alpha_0 x_0 + \cdots + \alpha_m x_m.$$

By Carathéodory's Theorem its image is the convex hull of K. The mapping is continuous and its domain is compact, so its image is compact. ∎

The next application of Lemma 24 is often asserted to be obvious, but is not so easy to prove. It is true in general Hausdorff topological vector spaces, but I'll prove it for the Euclidean space case.[2]

Recall that a **cone** is a set closed under multiplication by nonnegative scalars. A **finitely generated cone** is the *convex* cone generated by a finite set.

29 Lemma *Every finitely generated cone is closed.*

Proof for the finite dimensional case: Let $A = \{x_1,\dots,x_k\}$ be a finite subset of $\boldsymbol{R}^{\mathrm{m}}$ and let $C = \{\sum_{i=1}^k \lambda_i x_i : \lambda_i \geqslant 0,\ i = 1,\dots,k\}$ be the finitely generated cone generated by A. Let y be the limit of some sequence y_n in C,

$$y_n \to y.$$

By Lemma 24 we can write each y_n as a nonnegative linear combination of an independent subset of the x_i's. Since there are only finitely many such subsets, by passing to a subsequence we may assume without loss of generality that each y_n depends on the same independent subset $\{x_1,\dots,x_p\}$. We can find vectors $z_1,\dots,z_{m-p}$ so that $\{x_1,\dots,x_p,z_1,\dots,z_{m-p}\}$ is a basis for $\boldsymbol{R}^{\mathrm{m}}$. We can now write

$$y_n = \sum_{i=1}^p \lambda_{n,i} x_i + \sum_{j=1}^{m-p} 0 z_j$$

for each n where each $\lambda_{n,i} \geqslant 0$, and

$$y = \sum_{i=1}^p \lambda_i x_i + \sum_{j=1}^{m-p} \alpha_j z_j.$$

Since $y_n \to y$ and the coordinate mapping is continuous, we must have $\lambda_{n,i} \to \lambda_i \geqslant 0$, for $i = 1,\dots,p$, and $0 \to \alpha_j = 0$ for $j = 1,\dots,m-p$, so that y belongs to C. (For a proof that the coordinate mapping is continuous, which is often taken for granted, see my on-line note at http://www.its.caltech.edu/~kcborder/Notes/Coordinates.pdf.) ∎

3 Solutions of systems of equalities

Consider the system of linear equations

$$Ax = b,$$

where $A = \big[a_{i,j}\big]$ is an $m \times n$ matrix, $x \in \boldsymbol{R}^{\mathrm{n}}$, and $b \in \boldsymbol{R}^{\mathrm{m}}$. There are two or three interpretations of this matrix equation, and, depending on the circumstances, one may be more useful than the

[2] The general proof relies on the fact that the span of any finite set in a Hausdorff tvs is a closed subset, and that every m-dimensional subspace of a tvs is linearly homeomorphic to $\boldsymbol{R}^{\mathrm{m}}$.

other. The first interpretation is as a system of m equations in n variables

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= b_1 \\ &\vdots \\ a_{i,1}x_1 + \cdots + a_{i,n}x_n &= b_i \\ &\vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= b_m. \end{aligned}$$

or equivalently as a condition on m inner products,

$$A_i \cdot x = b_i, \qquad i = 1, \ldots, m$$

where A_i is the i^{th} row of A.

The other interpretation is as a vector equation in $\boldsymbol{R}^{\mathrm{m}}$,

$$x_1 A^1 + \cdots x_n A^n = b,$$

where A^j is the j^{th} column of A.

Likewise, the system

$$pA = c$$

can be interpreted as a system of equalities in the variables $p_1, \ldots, p_m$, which by **transposition** can be put in the form $A'p = c$, or

$$\begin{aligned} a_{1,1}p_1 + \cdots + a_{m,1}p_m &= c_1 \\ &\vdots \\ a_{1,j}p_1 + \cdots + a_{m,j}p_m &= c_j \\ &\vdots \\ a_{1,n}p_1 + \cdots + a_{m,n}p_m &= c_n \end{aligned}$$

or equivalently as a condition on n inner products,

$$A^j \cdot p = c_j, \quad j = 1, \ldots, n,$$

where A^j is the j^{th} column of A. Or we can interpret it as a vector equation in $\boldsymbol{R}^{\mathrm{n}}$,

$$p_1 A_1 + \cdots + p_m A_m = c,$$

where A_i is the i^{th} row of A.

30 Definition *A vector $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)$ is a* ***solution*** *of the system* Do I need this definition?

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= b_1 \\ &\vdots \\ a_{i,1}x_1 + \cdots + a_{i,n}x_n &= b_i \\ &\vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= b_m. \end{aligned}$$

if the statements

$$\begin{aligned} a_{1,1}\bar{x}_1 + \cdots + a_{1,n}\bar{x}_n &= b_1 \\ &\vdots \\ a_{i,1}\bar{x}_1 + \cdots + a_{i,n}\bar{x}_n &= b_i \\ &\vdots \\ a_{m,1}\bar{x}_1 + \cdots + a_{m,n}\bar{x}_n &= b_m. \end{aligned}$$

are all true. The system is ***solvable*** *if a solution exists.*

If A has an inverse (which implies $m = n$), then the system $Ax = b$ always has a unique solution, namely $\bar{x} = A^{-1}b$. But even if A does not have an inverse, the system may have a solution, possibly several—or it may have none. This brings up the question of how to characterize the existence of a solution. The answer is given by the following theorem. Following Riesz–Sz.-Nagy [26, p. 164] I shall refer to it as the Fredholm Alternative, as Fredholm [10] proved it in 1903 in the context of integral equations. I'll bet there is an earlier proof.

31 Theorem (Fredholm Alternative) *Let A be an $m \times n$ matrix and let $b \in \boldsymbol{R}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists an $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax = b \tag{42}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &= 0 \\ p \cdot b &> 0. \end{aligned} \tag{43}$$

Proof: It is easy to see that both (42) and (43) cannot be true, for then we would have

$$0 = 0 \cdot x = pAx = p \cdot b > 0,$$

a contradiction. Let M be the subspace spanned by the columns of A. Alternative (42) is that b belongs to M. If this is not the case, then by the strong Separating Hyperplane Theorem there is a nonzero vector p strongly separating $\{b\}$ from the closed convex set M, that is $p \cdot b > p \cdot z$ for each $z \in M$. Since M is a subspace we have $p \cdot z = 0$ for every $z \in M$, and in particular for each column of A, so $pA = 0$ and $p \cdot b > 0$, which is just (43). ∎

Proof using orthogonal decomposition: Using the notation of the above proof, decompose b as $b = b_M + p$, where $b_M \in M$ and $p \in M_\perp$. (In particular, $pA = 0$.) Then $p \cdot b = p \cdot b_M + p \cdot p = p \cdot p$. If $b \in M$, then $p \cdot b = 0$, but if $b \notin M$, then $p \neq 0$, so $p \cdot b = p \cdot p > 0$. ∎

32 Remark There is another way to think about the Fredholm alternative, which was expounded by Kuhn [15]. Either the system $Ax = b$ has a solution, or we can find weights $p_1, \ldots, p_m$ such that if we weight the equations

$$\begin{aligned} p_1(a_{1,1}x_1 + \cdots + a_{1,n}x_n) &= p_1 b_1 \\ &\vdots \\ p_i(a_{i,1}x_1 + \cdots + a_{i,n}x_n) &= p_i b_i \\ &\vdots \\ p_m(a_{m,1}x_1 + \cdots + a_{m,n}x_n) &= p_m b_m. \end{aligned}$$

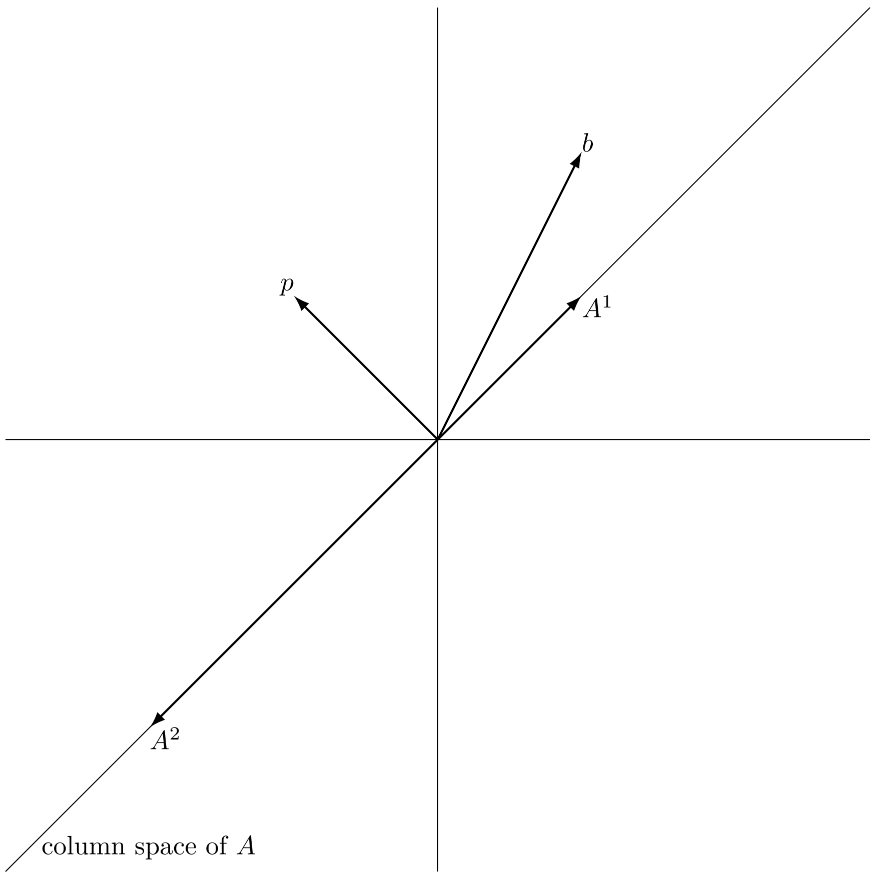

Figure 2. Geometry of the Fredholm Alternative

and add them up

$$(p_1 a_{1,1} + \cdots + p_m a_{m,1})x_1 + \cdots + (p_1 a_{1,n} + \cdots + p_m a_{m,n})x_n = p_1 b_1 + \cdots + p_m b_m$$

we get the **inconsistent** system

$$0x_1 + \cdots + 0x_n = p_1 b_1 + \cdots + p_m b_m > 0.$$

But this means that the original system is inconsistent too. Thus solvability is equivalent to consistency.

We can think of the weights p being chosen to "eliminate" the variables x from the left-hand side. Or as Kuhn points out, we do not eliminate the variables, we merely set their coefficients to zero.

The following corollary about linear functions is true in quite general linear spaces, but we shall first provide a proof using some of the special properties of $\boldsymbol{R}^{\mathrm{n}}$. Wim Luxemburg refers to this result as the **Fundamental Theorem of Duality**.

33 Corollary *Let $p^0, \ldots, p^m \in \boldsymbol{R}^{\mathrm{n}}$ and suppose that $p^0 \cdot v = 0$ for all v such that $p^i \cdot v = 0$, $i = 1, \ldots, m$. Then p^0 is a linear combination of $p^1, \ldots, p^m$. That is, there exist scalars $\lambda_1, \ldots, \lambda_m$ such that $p^0 = \sum_{i=1}^m \lambda_i p_i$.*

Proof: Consider the matrix A whose columns are $p^1, \ldots, p^m$, and set $b = p^0$. By hypothesis alternative (43) of Theorem 31 is false, so alternative (42) must hold. But that is precisely the conclusion of this theorem. ∎

Proof using orthogonal decomposition: Let $M = \operatorname{span}\{p^1, \ldots, p^m\}$ and orthogonally project p^0 on M to get $p^0 = p^0_M + p^0_\perp$, where $p^0_M \in M$ and $p^0_\perp \perp M$. That is, $p^0_\perp \cdot p = 0$ for all $p \in M$. In particular, $p^i \cdot p^0_\perp = 0$, $i = 1, \ldots, m$. Consequently, by hypothesis, $p^0 \cdot p^0_\perp = 0$ too. But

$$0 = p^0 \cdot p^0_\perp = p^0_M \cdot p^0_\perp + p^0_\perp \cdot p^0_\perp = 0 + \|p^0_\perp\|.$$

Thus $p^0_\perp = 0$, so $p^0 = p^0_M \in M$. That is, p^0 is a linear combination of $p^1, \ldots, p^m$. ∎

4 Nonnegative solutions of systems of equalities

The next theorem is one of many more or less equivalent results on the existence of solutions to linear inequalities. It is often known as Farkas's Lemma, and so is Corollary 36. Julius Farkas [8] proved them both in 1902.[3]

34 Farkas's Alternative *Let A be an $m \times n$ matrix and let $b \in \boldsymbol{R}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &= b \\ x &\geqq 0 \end{aligned} \tag{44}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &\geqq 0 \\ p \cdot b &< 0. \end{aligned} \tag{45}$$

[3] According to Wikipedia, Gyula Farkas (1847–1930) was a Jewish Hungarian mathematician and physicist (not to be confused with the linguist of the same name who was born about half a century later), but this paper of his, published in German, bears his Germanized name, Julius Farkas.

Proof: (44) $\implies \neg$(45) : Assume $x \geqq 0$ and $Ax = b$. Premultiplying by p, we get $pAx = p \cdot b$. Now if $pA \geqq 0$, then $pAx \geqslant 0$ as $x \geqq 0$, so $p \cdot b \geqslant 0$. That is, (45) fails.

$\neg$(44) $\implies$ (45) : Let $C = \{Ax : x \geqq 0\}$. If (44) fails, then b does not belong to C. By Lemma 29, the finitely generated cone C is closed, so by the Strong Separating Hyperplane Theorem there is some nonzero p such that $p \cdot z \geqslant 0$ for all $z \in C$ and $p \cdot b < 0$. Therefore (45). ∎

Note that there are many trivial variations on this result. For instance, multiplying p by -1, I could have written (45) as $pA \leqq 0$ & $p \cdot b > 0$. Or by replacing A by its transpose, we could rewrite (44) with $xA = b$ and (45) with $Ap \geqq 0$. Keep this in mind as you look at the coming theorems.

5 Solutions of systems of inequalities

Proposition 10 is restated here in slightly modified form. The condition $p \cdot b < 0$ is replaced by $p \cdot b = -1$, which is simply a normalization.

35 Proposition (Solution of Inequalities) *Let A be an $m \times n$ matrix and let $b \in \boldsymbol{R}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists an $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax \leqq b \tag{46}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &= 0 \\ p \cdot b &= -1 \\ p &\geqq 0. \end{aligned} \tag{47}$$

Note that in case (47), we actually have $p > 0$. (Why?)

Proof: Clearly (46) and (47) are inconsistent: If (47) holds, then $pAx = 0x = 0$. Also $p > 0$, so (46) implies $pAx \leqslant p \cdot b$. Thus $0 = pAx \leqslant p \cdot b = -1$, a contradiction.

So suppose (47) fails. That is, the system of equations

$$p \begin{bmatrix} A & b \end{bmatrix} = \begin{bmatrix} 0 & -1 \end{bmatrix}$$

has no nonnegative solution. Then by the transposed version of Farkas's Alternative 34, the system

$$\begin{bmatrix} A & b \end{bmatrix} \begin{bmatrix} \hat{x} \\ \xi \end{bmatrix} \geqq 0$$

$$\begin{bmatrix} 0 & -1 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \xi \end{bmatrix} < 0$$

has a solution $\begin{bmatrix} \hat{x} \\ \xi \end{bmatrix}$. In other words,

$$A\hat{x} + \xi b \geqq 0.$$

Divide both sides by $-\xi < 0$, set $x = -\hat{x}/\xi$, and rearrange to conclude

$$Ax \leqq b,$$

which is just (46). ∎

For the next result recall that $p > 0$ means that p is semipositive: $p \geqq 0$ and $p \neq 0$.

36 Corollary (Farkas's Alternative) *Let A be an $m \times n$ matrix and let $b \in \boldsymbol{R}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists an $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &\leqq b \\ x &\geqq 0 \end{aligned} \tag{48}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &\geqq 0 \\ p \cdot b &< 0 \\ p &> 0 \end{aligned} \tag{49}$$

37 Exercise Prove the corollary by converting the inequalities to equalities as discussed on page 5 and apply the Farkas Lemma 34. □

Both versions of Farkas's Lemma are subsumed by the next result, which is also buried in Farkas [8].

38 Farkas's Alternative *Let A be an $m \times n$ matrix, let B be an $\ell \times n$ matrix, let $b \in \boldsymbol{R}^{\mathrm{m}}$, and let $c \in \boldsymbol{R}^{\ell}$. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &= b \\ Bx &\leqq c \\ x &\geqq 0 \end{aligned} \tag{50}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ and $q \in \boldsymbol{R}^{\ell}$ satisfying

$$\begin{aligned} pA + qB &\geqq 0 \\ q &\geqq 0 \\ p \cdot b + q \cdot c &< 0. \end{aligned} \tag{51}$$

39 Exercise Prove this version of Farkas's Alternative. Hint: Consider the system

$$\begin{bmatrix} A & 0 \\ B & I \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} b \\ c \end{bmatrix}$$
$$x \geqq 0$$
$$z \geqq 0$$

and apply Farkas's Alternative 34. □

We can reformulate the above as follows.

40 Farkas's Alternative *Let A be an $m \times n$ matrix, let $b \in \boldsymbol{R}^{\mathrm{m}}$, and let I and E partition $\{1, \ldots, m\}$. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} (Ax)_i &= b_i \quad i \in E \\ (Ax)_i &\leqslant b_i \quad i \in I \\ x &\geqq 0 \end{aligned} \tag{52}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ *satisfying*

$$\begin{aligned} pA &\geqq 0 \\ p \cdot b &< 0 \\ p_i &\geqq 0 \quad i \in E. \end{aligned} \tag{53}$$

We can also handle strict inequalities. See Figure 3 for a geometric interpretation of the next theorem, which is due to Gordan [14] in 1873.

41 Gordan's Alternative *Let A be an $m \times n$ matrix. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax \gg 0. \tag{54}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ *satisfying*

$$\begin{aligned} pA &= 0 \\ p &> 0 \end{aligned} \tag{55}$$

There are two ways (55) can be satisfied. The first is that some row of A is zero, say row i—then $p = e^i$ satisfies (55). If no row of A is zero, then the finitely generated cone $\langle A_1 \rangle + \cdots + \langle A_m \rangle$ must contain a nonzero point and its negative. That is, it is not pointed. Gordan's Alternative says that if the cone is pointed, that is, (55) fails, then the generators (rows of A) lie in the same open half space $\{x > 0\}$.

There is another, algebraic, interpretation of Gordan's Alternative in terms of consistency and solvability. It says that if (54) is not solvable, then we may multiply each equality $p \cdot A^j = 0$ by a multiplier x_j and add them so that the resulting coefficients on p_i, namely $A_i \cdot x$ are all strictly positive, but the right-hand side remains zero, showing that (54) is inconsistent.

42 Exercise Prove Gordan's Alternative. Hint: If x satisfies (54), it may be scaled so that in fact $Ax \geqq \mathbf{1}$, where $\mathbf{1}$ is the vector of ones. Write $x = u - v$ where $u \geqq 0$ and $v \geqq 0$. Then (54) can be written as

$$-A(u - v) \leqq -\mathbf{1}. \tag{54$'$}$$

Now use Corollary 36. □

We can rewrite Corollary 41 as follows.

43 Corollary *Let A be an $m \times n$ matrix. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax \gg 0 \tag{56}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ *satisfying*

$$\begin{aligned} pA &= 0 \\ p \cdot \mathbf{1} &= 1 \\ p &> 0. \end{aligned} \tag{57}$$

(The second alternative implies that p is a probability vector.)

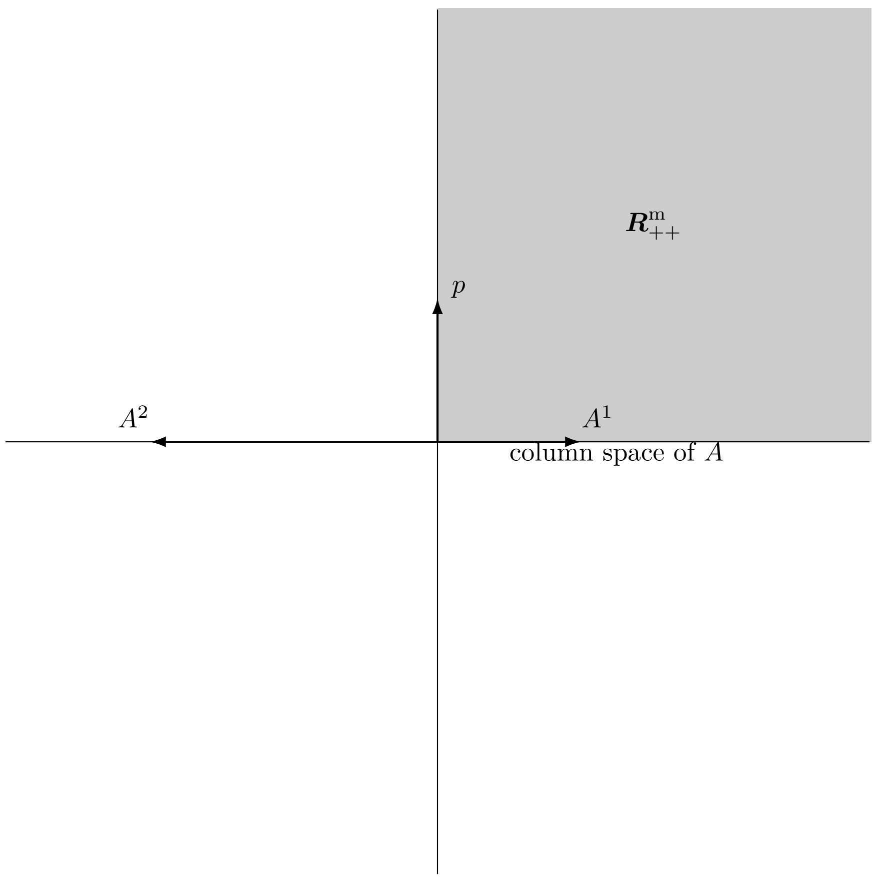

Figure 3. Geometry of Gordan's Alternative 41

44 Corollary (Yet Another Alternative) *Let A be an $m \times n$ matrix. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &\gg 0 \\ x &\geqq 0. \end{aligned} \tag{58}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &\leqq 0 \\ p &> 0 \end{aligned} \tag{59}$$

45 Exercise Prove the Corollary. Hint: If (58) holds we may normalize x so that $Ax \geqq \mathbf{1}$. Use Corollary 36. □

The following result was proved by Stiemke [27] in 1915.

46 Stiemke's Alternative *Let A be an $m \times n$ matrix. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax > 0 \tag{60}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &= 0 \\ p &\gg 0. \end{aligned} \tag{61}$$

Proof: (60) $\implies \neg$(61) : Clearly both cannot be true, for then we must have both $pAx = 0$ (as $pA = 0$) and $pAx > 0$ (as $p \gg 0$ and $Ax > 0$).

$\neg$(60) $\implies$ (61) : Let $\Delta = \{z \in \boldsymbol{R}^{\mathrm{m}} : z \geqq 0 \text{ and } \sum_{j=1}^{n} z_j = 1\}$ be the unit simplex in $\boldsymbol{R}^{\mathrm{m}}$. In geometric terms, (60) asserts that the span M of the columns $\{A^1, \dots, A^n\}$ intersects the nonnegative orthant $\boldsymbol{R}^{\mathrm{m}}_{+}$ at a nonzero point, namely Ax. Since M is a linear subspace, we may rescale x so that Ax belongs to $M \cap \Delta$. Thus the negation of (60) is equivalent to the disjointness of M and Δ.

So assume that (60) fails. Then since Δ is compact and convex and M is closed and convex, there is a hyperplane strongly separating Δ and M. That is, there is some nonzero $p \in \boldsymbol{R}^{\mathrm{n}}$ and some $\varepsilon > 0$ satisfying

$$p \cdot y + \varepsilon < p \cdot z \quad \text{for all } y \in M,\ z \in \Delta.$$

Since M is a linear subspace, we must have $p \cdot y = 0$ for all $y \in M$. Consequently $p \cdot z > \varepsilon > 0$ for all $z \in \Delta$. Since the j^{th} unit coordinate vector e^j belongs to Δ, we see that $p_j = p \cdot e^j > 0$, That is, $p \gg 0$.

Since each $A^i \in M$, we have that $p \cdot A^i = 0$, i.e.,

$$pA = 0.$$

This completes the proof. ∎

Note that in (61), we could rescale p so that it is a strictly positive probability vector. Also note that the previous proofs separated a single point from a closed convex set. This one separated the entire unit simplex from a closed linear subspace. There is another method of proof we could have used.

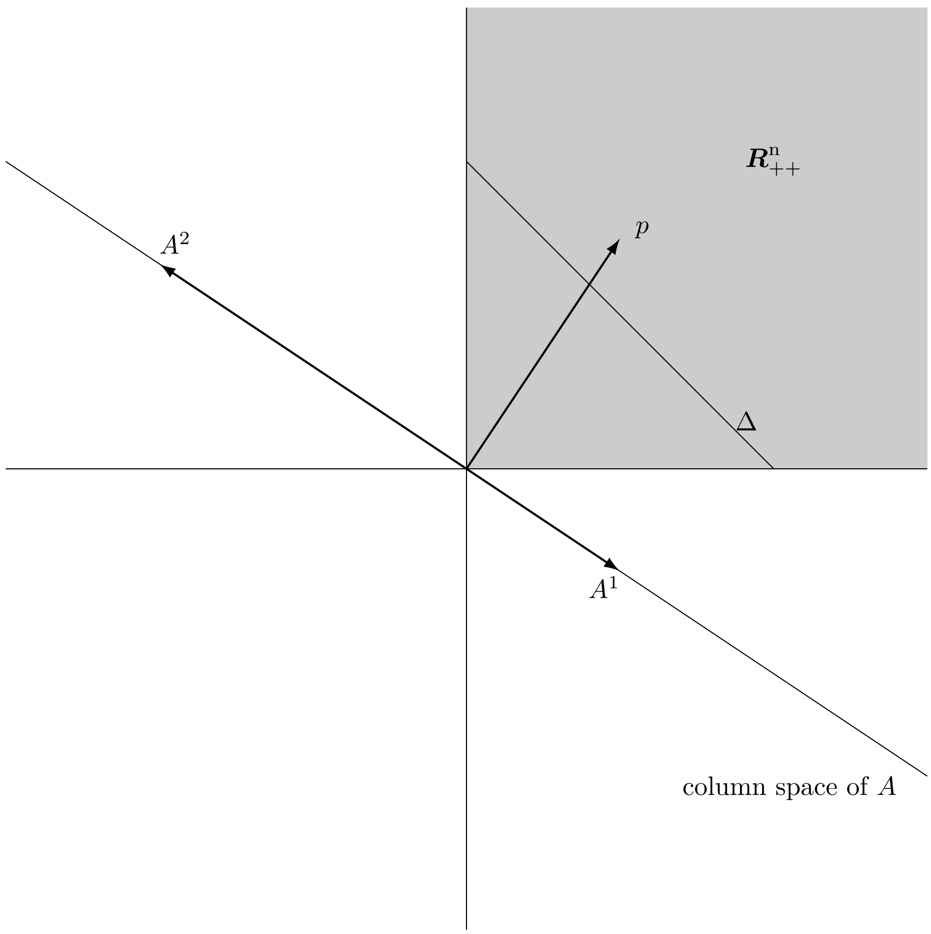

Figure 4. Geometry of the Stiemke Alternative

Alternate proof of Stiemke's Theorem: If (60) holds, then for some coordinate i, we may rescale x so that $Ax \geqq e^i$, or equivalently,

$$-A(u-v) \leqq -e^i, \qquad u \geqq 0,\ v \geqq 0.$$

Fixing i for the moment, if this fails, then we can use Corollary 36 to deduce the existence of p^i satisfying $pA = 0$, $p^i \cdot e^i > 0$, and $p^i > 0$.

Now observe that if (60) fails, then for each $i = 1, \dots, m$, there must exist p^i as described. Now set $p = p^1 + \cdots + p^m$ to get p satisfying (61). ∎

Finally we come to another alternative, Motzkin's Transposition Theorem [21], proven in his 1934 Ph.D. thesis. This statement is take from his 1951 paper [22].[4]

47 Motzkin's Transposition Theorem *Let A be an $m \times n$ matrix, let B be an $\ell \times n$ matrix, and let C be an $r \times n$ matrix, where B or C may be omitted (but not A). Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned} Ax &\gg 0 \\ Bx &\geqq 0 \\ Cx &= 0 \end{aligned} \tag{62}$$

or else there exist $p^1 \in \boldsymbol{R}^{\mathrm{m}}$, $p^2 \in \boldsymbol{R}^{\ell}$, and $p^3 \in R^{\mathrm{r}}$ satisfying

$$\begin{aligned} p^1A + p^2B + p^3C &= 0 \\ p^1 &> 0 \\ p^2 &\geqq 0. \end{aligned} \tag{63}$$

Motzkin expressed (63) in terms of the transpositions of A, B, and C.

48 Exercise Prove the Transposition Theorem. □

We are now in a position to prove Theorem 18, which is restated here for convenience:

49 Theorem (Morris's Alternative) *Let A be an $m \times n$ matrix. Let $\mathcal{S}$ be a family of nonempty subsets of $\{1, \dots, m\}$. Exactly one of the following alternatives holds. Either there exists $x \in \boldsymbol{R}^{\mathrm{n}}$ and a set $S \in \mathcal{S}$ satisfying*

$$\begin{aligned} Ax &\geqq 0 \\ A_i \cdot x &> 0 \quad \text{for all } i \in S \end{aligned} \tag{64}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &= 0 \\ p &\geqq 0 \\ \sum_{i \in S} p_i &> 0 \quad \text{for all } S \in \mathcal{S}. \end{aligned} \tag{65}$$

Proof: It is clear that (64) and (65) cannot both be true.

So assume that (64) fails. Then for each $S \in \mathcal{S}$, let A_S be the $|S| \times n$ matrix with rows A_i for $i \in S$, and let B_S be the matrix of the remaining rows. Then there is no x satisfying $A_S x \gg 0$ and $B_S x \geqq 0$. So by Motzkin's Transposition Theorem 47 there is $q^S \in \boldsymbol{R}^{|S|}$ and $q^{S^c} \in \boldsymbol{R}^{|S^c|}$ satisfying $q^S \gg 0$, $q^{S^c} \geqq 0$, and $q^S A_S + q^{S^c} B_S = 0$. Let $p^S \in \boldsymbol{R}^{\mathrm{m}}$ be defined $p_i^S = q_i^S$ for $i \in S$ and $p_i = q_i^{S^c}$ for $i \in S^c$. Then $p^S A = 0$, and $\sum_{i \in S} p_i^S > 0$. Now define $p = \sum_{S \in \mathcal{S}} p^S$ and note that it satisfies (65). ∎

[4] Motzkin [22] contains an unfortunate typo. The condition $Ax \gg 0$ is erroneously given as $Ax \ll 0$.

6 The Gauss–Jordan method

The **Gauss–Jordan method** is a straightforward way to find solutions to systems of linear equations using elementary row operations.

Give a cite. Apostol [1]?

50 Definition *The three **elementary row operations** on a matrix are:*

- *Interchange two rows.*
- *Multiply a row by a nonzero scalar.*
- *Add one row to another.*

It is often useful to combine these into a fourth operation.

- *Add a nonzero scalar multiple of one row to another row.*

We shall also refer to this last operation as an elementary row operation.[5]

You should convince yourself that each of these four operations is reversible using only these four operations, and that none of these operations changes the set of solutions.

Consider the following system of equations.

$$\begin{aligned} 3x_1 + 2x_2 &= 8 \\ 2x_1 + 3x_2 &= 7 \end{aligned}$$

The first step in using elementary row operations to solve a system of equations is to write down the so-called augmented coefficient matrix of the system, which is the 2×3 matrix of just the numbers above:

$$\left[\begin{array}{cc|c} 3 & 2 & 8 \\ 2 & 3 & 7 \end{array}\right]. \tag{66'}$$

We apply elementary row operations until we get a matrix of the form

$$\left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & b \end{array}\right]$$

which is the augmented matrix of the system

$$\begin{aligned} x_1 &= a \\ x_2 &= b \end{aligned}$$

and the system is solved. (If there is no solution, then the elementary row operations cannot produce an identity matrix. There is more to say about this in Section 10.) There is a simple algorithm for deciding which elementary row operations to apply, namely, the **Gauss–Jordan elimination algorithm**.

First we multiply the first row by $\frac{1}{3}$, to get a leading 1:

$$\left[\begin{array}{cc|c} 1 & \frac{2}{3} & \frac{8}{3} \\ 2 & 3 & 7 \end{array}\right]$$

[5] The operation 'add $\alpha \times$ row k to row i' is the following sequence of truly elementary row operations: multiply row k by α, add (new) row k to row i, multiply row k by $1/\alpha$.

We want to eliminate x_1 from the second equation, so we add an appropriate multiple of the first row to the second. In this case the multiple is -2, the result is:

$$\left[\begin{array}{cc|c} 1 & \frac{2}{3} & \frac{8}{3} \\ 2-2\cdot 1 & 3-2\cdot\frac{2}{3} & 7-2\cdot\frac{8}{3} \end{array}\right] = \left[\begin{array}{cc|c} 1 & \frac{2}{3} & \frac{8}{3} \\ 0 & \frac{5}{3} & \frac{5}{3} \end{array}\right]. \tag{67$'$}$$

Now multiply the second row by $\frac{3}{5}$ to get

$$\left[\begin{array}{cc|c} 1 & \frac{2}{3} & \frac{8}{3} \\ 0 & 1 & 1 \end{array}\right].$$

Finally to eliminate x_2 from the first row we add $-\frac{2}{3}$ times the second row to the first and get

$$\left[\begin{array}{cc|c} 1-\frac{2}{3}\cdot 0 & \frac{2}{3}-\frac{2}{3}\cdot 1 & \frac{8}{3}-\frac{2}{3}\cdot 1 \\ 0 & 1 & 1 \end{array}\right] = \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \end{array}\right], \tag{68$'$}$$

so the solution is $x_1 = 2$ and $x_2 = 1$.

7 A different look at the Gauss–Jordan method

David Gale [13] gives another way to look at what we just did. The problem of finding x to solve

$$\begin{aligned} 3x_1 + 2x_2 &= 8 \\ 2x_1 + 3x_2 &= 7 \end{aligned}$$

can also be thought of as finding a coefficients x_1 and x_2 to solve the vector equation

$$x_1 \begin{bmatrix} 3 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}.$$

That is, we want to write $b = \begin{bmatrix} 8 \\ 7 \end{bmatrix}$ as a linear combination of $a^1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $a^2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. One way to do this is to begin by writing b as a linear combination of unit coordinate vectors $e^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, which is easy:

$$8 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 7 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}.$$

We can do likewise for a^1 and a^2:

$$3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \qquad 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

We can summarize this information in the following ***tableau***.[6]

$$\begin{array}{c||cc|c} & a^1 & a^2 & b \\ \hline e^1 & 3 & 2 & 8 \\ e^2 & 2 & 3 & 7 \end{array} \tag{66}$$

There is a column for each of the vectors a^1, a^2, and b. There is a row for each element of the basis e^1, e^2. A *tableau* is actually a statement. It asserts that the vectors listed in the column titles can be written as linear combinations of the vectors listed in the row titles, and that the coefficients of the linear combinations are given in the matrix. Thus $a^1 = 3e^1 + 2e^2$, $b = 8e^1 + 7e^2$, etc. So far, with the exception of the margins, our *tableau* looks just like the augmented coefficient matrix (66), as it should.

But we don't really want to express b in terms of e^1 and e^2, we want to express it in terms of a^1 and a^2, so we do this in steps. Let us replace e^1 in our basis with either a^1 or a^2. Let's be unimaginative and use a^1. The new *tableau* will look something like this:

$$\begin{array}{c||cc|c} & a^1 & a^2 & b \\ \hline a^1 & ? & ? & ? \\ e^2 & ? & ? & ? \end{array}$$

Note that the left marginal column now has a^1 in place of e^1. We now need to fill in the *tableau* with the proper coefficients. It is clear that $a^1 = 1a^1 + 0e^2$, so we have

$$\begin{array}{c||cc|c} & a^1 & a^2 & b \\ \hline a^1 & 1 & ? & ? \\ e^2 & 0 & ? & ? \end{array}$$

I claim the rest of the coefficients are

$$\begin{array}{c||cc|c} & a^1 & a^2 & b \\ \hline a^1 & 1 & \frac{2}{3} & \frac{8}{3} \\ e^2 & 0 & \frac{5}{3} & \frac{5}{3} \end{array} \tag{67}$$

That is,

$$a^1 = 1a^1 + 0e^2, \qquad a^2 = \frac{2}{3}a^1 + \frac{5}{3}e^2, \qquad b = \frac{8}{3}a^1 + \frac{5}{3}e^2.$$

or

$$\begin{bmatrix} 3 \\ 2 \end{bmatrix} = 1 \begin{bmatrix} 3 \\ 2 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \frac{2}{3} \begin{bmatrix} 3 \\ 2 \end{bmatrix} + \frac{5}{3} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 8 \\ 7 \end{bmatrix} = \frac{8}{3} \begin{bmatrix} 3 \\ 2 \end{bmatrix} + \frac{5}{3} \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

which is correct. Now observe that the *tableau* (67)) is the same as (67').

Now we proceed to replace e^2 in our basis by a^1. The resulting *tableau* is

$$\begin{array}{c||cc|c} & a^1 & a^2 & b \\ \hline a^1 & 1 & 0 & 2 \\ a^2 & 0 & 1 & 1 \end{array} \tag{68}$$

[6] The term *tableau*, a French word best translated as "picture" or "painting," harkens back to Quesnay's *Tableau économique* [25], which inspired Leontief [17], whose work spurred the Air Force's interest in linear programming [4, p. 17].

This is the same as (68). In other words, in terms of our original problem $x_1 = 2$ and $x_2 = 1$.

So far we have done nothing that we would not have done in the standard method of solving linear equations. The only difference is in the description of what we are doing.

> Instead of describing our steps as eliminating variables from equations one by one, we say that we are replacing one basis by another, one vector at a time.

We now formalize this notion more generally.

8 The replacement operation

Let $\mathcal{A} = \{a^1, \dots, a^n\}$ be a set of vectors in some vector space, and let $p^1, \dots, p^m\}$ span $\mathcal{A}$. That is, each a^j can be written as a linear combination of b^i's. Let $T = [t_{i,j}]$ be the $m \times n$ matrix of coordinates of the a^j's with respect to the b^i's.[7] That is,

$$a^j = \sum_{k=1}^{m} t_{k,j} b^k, \qquad j = 1, \dots, n. \tag{69}$$

We express this as the following ***tableau***:

$$\begin{array}{c||ccccc} & a^1 & \dots & a^j & \dots & a^n \\ \hline b^1 & t_{1,1} & \dots & t_{1,j} & \dots & t_{1,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b^i & t_{i,1} & \dots & t_{i,j} & \dots & t_{i,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b^m & t_{m,1} & \dots & t_{m,j} & \dots & t_{m,n} \end{array} \tag{69$'$}$$

- A *tableau* is actually a statement. It asserts that the equations (69) hold. In this sense a *tableau* may be true or false, but we shall only consider true *tableaux*.
- It is obvious that interchanging any two rows or interchanging any two columns represents the same information, namely that each vector listed in the top margin is a linear combination of the vectors in the left margin, with the coefficients being displayed in the *tableau*'s matrix.
- We can rewrite (69) in terms of the coordinates of the vectors as

  $$a_i^j = \sum_{k=1}^{m} t_{k,j} b_i^k$$

  or perhaps more familiarly as the matrix equation

  $$BT = A,$$

  where A is the matrix $m \times n$ matrix whose columns are $a^1, \dots, a^n$, B is the matrix $m \times m$ matrix whose columns are $b^1, \dots, b^m$, and T is the $m \times n$ matrix $[t_{i,j}]$.

[7]If the b^i's are linearly dependent, T may not be unique.

The usefulness of the *tableau* is the ease with which we can change the basis of a subspace. The next lemma is the key.

51 Replacement Lemma *If $p^1, \dots, p^m\}$ is a linearly independent set that spans $\mathcal{A}$, then*

$$t_{k,\ell} \neq 0 \quad \text{if and only if} \quad \{b^1, \dots, b^{k-1}, a^\ell, b^{k+1}, \dots, b^m\} \text{ is independent and spans } \mathcal{A}.$$

Moreover, in the latter case ***the new*** **tableau** ***is derived from the old one by applying elementary row operations that transform the ℓ^{th} column into the k^{th} unit coordinate vector****. That is, the* tableau

$$\begin{array}{c||ccccccc}
 & a^1 & \dots & a^{\ell-1} & a^\ell & a^{\ell+1} & \dots & a^n \\
\hline
b^1 & t'_{1,1} & \dots & t'_{1,\ell-1} & 0 & t'_{1,\ell+1} & \dots & t'_{1,n} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
b^{k-1} & t'_{k-1,1} & \dots & t'_{k-1,\ell-1} & 0 & t'_{k-1,\ell+1} & \dots & t'_{k-1,n} \\
a^\ell & t'_{k,1} & \dots & t'_{k,\ell-1} & 1 & t'_{k,\ell+1} & \dots & t'_{k,n} \\
b^{k+1} & t'_{k+1,1} & \dots & t'_{k+1,\ell-1} & 0 & t'_{k+1,\ell+1} & \dots & t'_{k+1,n} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
b^m & t'_{m,1} & \dots & t'_{m,\ell-1} & 0 & t'_{m,\ell+1} & \dots & t'_{m,n}
\end{array}$$

is obtained by dividing the k^{th} row by $t_{k,\ell}$,

$$t'_{k,j} = \frac{t_{k,j}}{t_{k,\ell}}, \qquad j = 1, \dots, n,$$

and adding $-\dfrac{t_{i,\ell}}{t_{k,\ell}}$ times row k to row i for $i \neq k$,

$$t'_{i,j} = t_{i,j} - \frac{t_{i,\ell}}{t_{k,\ell}} t_{k,j} \quad \left(= t_{i,j} - t_{i,\ell} t'_{k,j}\right), \qquad \begin{array}{c} i = 1, \dots, m,\ i \neq k \\ j = 1, \dots, n \end{array}.$$

Proof: If $t_{k,\ell} = 0$, then

$$a^\ell = \sum_{i:i\neq k} t_{i,\ell} b^i,$$

or

$$\sum_{i:i\neq k} t_{i,\ell} b^i - 1a^\ell = 0,$$

so $\{b^1, \dots, b^{k-1}, a^\ell, b_{k+1}, \dots, b_m\}$ is dependent.

For the converse, assume $t_{k,\ell} \neq 0$, and that

$$\begin{aligned}
0 &= \alpha a^\ell + \sum_{i:i\neq k} \beta_i b^i \\
&= \alpha \left(\sum_{i=1}^m t_{i,\ell} b^i \right) + \sum_{i:i\neq k} \beta_i b^i \\
&= \alpha t_{k,\ell} b^k + \sum_{i:i\neq k} (\alpha t_{i,\ell} + \beta_i) b^i.
\end{aligned}$$

Since $p^1, \dots, p^m\}$ is independent by hypothesis, we must have (i) $\alpha t_{k,\ell} = 0$ and (ii) $\alpha t_{i,\ell} + \beta_i = 0$ for $i \neq k$. Since $t_{k,\ell} \neq 0$, (i) implies that $\alpha = 0$. But then (ii) implies that each $\beta_i = 0$, too, which shows that the set $\{b^1, \dots, b^{k-1}, a^\ell, b_{k+1}, \dots, b_m\}$ is linearly independent.

To show that this set spans $\mathcal{A}$, and to verify the *tableau*, we must show that for each $j \neq \ell$,

$$a^j = \sum_{i:i\neq k} t'_{i,j} b^i + t'_{k,j} a^\ell.$$

But the right-hand side is just

$$\begin{aligned}
&= \sum_{i:i\neq k} \underbrace{\left(t_{i,j} - \frac{t_{i,\ell}}{t_{k,\ell}} t_{k,j}\right)}_{t'_{i,j}} b^i + \underbrace{\frac{t_{k,j}}{t_{k,\ell}}}_{t'_{k,j}} \underbrace{\sum_{i=1}^m t_{i,\ell} b^i}_{a^\ell} \\
&= \sum_{i=1}^m t_{i,j} b^i \\
&= a^j,
\end{aligned}$$

which completes the proof. ∎

Thus whenever $t_{k,\ell} \neq 0$, we can replace b^k by a^ℓ, and get a valid new *tableau*. We call this the **replacement operation** and the entry $t_{k,\ell}$, the **pivot**. Note that one replacement operation is actually m elementary row operations.

Here are some observations.

- If at some point, an entire row of the *tableau* becomes 0, then any replacement operation leaves the row unchanged. This means that the dimension of the span of $\mathcal{A}$ is less than m, and that row may be omitted.
- We can use this method to select a basis from $\mathcal{A}$. Replace the standard basis with elements of $\mathcal{A}$ until no additional replacements can be made. By construction, the set $\mathcal{B}$ of elements of $\mathcal{A}$ appearing in the left-hand margin of the *tableau* will constitute a linearly independent set. If no more replacements can be made, then each row i associated with a vector not in $\mathcal{A}$ must have $t_{i,j} = 0$, for $j \notin \mathcal{B}$ (otherwise we could make another replacement with $t_{i,j}$ as pivot.) Thus $\mathcal{B}$ must be a basis for $\mathcal{A}$.
- Note that the elementary row operations used preserve the field to which the coefficients belong. In particular, if the original coefficients belong to the field of rational numbers, the coefficients after a replacement operation also belong to the field of rational numbers.

9 More on *tableaux*

An important feature of *tableaux* is given in the following proposition.

52 Proposition *Let $b^1, \dots, b^m$ be a basis for $\boldsymbol{R}^{\mathrm{m}}$ and let $a^1, \dots, a^n$ be vectors in $\boldsymbol{R}^{\mathrm{m}}$. Consider* *the following* tableau.

$$
\begin{array}{c||c c|c|c c||c c c c c}
 & a^1 & \dots & a^j & \dots & a^n & e^1 & & \dots & & e^m \\
\hline
b^1 & t_{1,1} & \dots & t_{1,j} & \dots & t_{1,n} & y_{1,1} & \dots & \dots & \dots & y_{1,m} \\
\vdots & \vdots & & \vdots & & \vdots & \vdots & \ddots & \vdots & & \vdots \\
\cline{7-11}
b^i & t_{i,1} & \dots & \boxed{t_{i,j}} & \dots & t_{i,n} & y_{i,1} & \dots & \dots & \dots & y_{i,m} \\
\cline{7-11}
\vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots & \ddots & \vdots \\
b^m & t_{m,1} & \dots & t_{m,j} & \dots & t_{m,n} & y_{m,1} & \dots & \dots & \dots & y_{m,m}
\end{array}
\tag{70}
$$

That is, for each j*,*

$$a^j = \sum_{i=1}^m t_{i,j} b^i \tag{71}$$

and

$$e^j = \sum_{i=1}^m y_{i,j} b^i. \tag{72}$$

Let y^i *be the (row) vector made from the last* m *elements of the* i^{th} *row. Then*

$$y^i \cdot a^j = t_{i,j}. \tag{73}$$

Proof: Let B be the $m \times m$ matrix whose j^{th} column is b^j, let A be the $m \times n$ matrix with column j equal to a^j, let T be the $m \times n$ matrix with (i,j) element $t_{i,j}$, and let Y be the $m \times m$ matrix with (i,j) element $y_{i,j}$ (that is, y^i is the i^{th} row of Y). Then (71) is just

$$A = BT$$

where and (72) is just

$$I = BY.$$

Thus $Y = B^{-1}$, so

$$YA = B^{-1}(BT) = (B^{-1}B)T = T,$$

which is equivalent to (73). ∎

53 Corollary *Let* A *be an* $m \times m$ *matrix with columns* $a^1, \dots, a^m$*. If the* tableau

$$
\begin{array}{c||c c c||c c c}
 & a^1 & \dots & a^m & e^1 & \dots & e^m \\
\hline
a^1 & 1 & & 0 & y_{1,1} & \dots & y_{1,m} \\
\vdots & & \ddots & & \vdots & & \vdots \\
a^m & 0 & & 1 & y_{m,1} & \dots & y_{m,m}
\end{array}
$$

is true, then the matrix Y *is the inverse of* A*.*

10 The Fredholm Alternative revisited

Recall the Fredholm Alternative 31 that we previously proved using a separating hyperplane argument. We can now prove a stronger version using a purely algebraic argument.

54 Theorem (Fredholm Alternative) *Let A be an $m \times n$ matrix and let $b \in \boldsymbol{R}^{\mathrm{m}}$. Exactly one of the following alternatives holds. Either there exists an $x \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$Ax = b \tag{74}$$

or else there exists $p \in \boldsymbol{R}^{\mathrm{m}}$ satisfying

$$\begin{aligned} pA &= 0 \\ p \cdot b &> 0. \end{aligned} \tag{75}$$

Moreover, if A and b have all rational entries, then x or p may be taken to have rational entries.

Proof: We prove the theorem based on the Replacement Lemma 51, and simultaneously compute x or p. Let A be the $m \times n$ with columns $A^1, \dots, A^n$ in $\boldsymbol{R}^{\mathrm{m}}$. Then $x \in \boldsymbol{R}^{\mathrm{n}}$ and $b \in \boldsymbol{R}^{\mathrm{m}}$. Begin with this *tableau*.

$$\begin{array}{c||ccc|c|ccc}
 & A^1 & \dots & A^n & b & e^1 & \dots & e^m \\
\hline
e^1 & \alpha_{1,1} & \dots & \alpha_{1,n} & \beta_1 & 1 & & 0 \\
\vdots & \vdots & & \vdots & \vdots & & \ddots & \\
e^m & \alpha_{m,1} & \dots & \alpha_{m,n} & \beta_m & 0 & & 1
\end{array}$$

Here $\alpha_{i,j}$ is the i^{th} row, j^{th} column element of A and β_i is the i^{th} coordinate of b with respect to the standard ordered basis. Now use the replacement operation to replace as many non-column vectors as possible in the left-hand margin basis. Say that we have replaced ℓ members of the standard basis with columns of A. Interchange rows and columns as necessary to bring the *tableau* into this form:

$$\begin{array}{c||ccc|ccc|c|ccccc}
 & A^{j_1} & \dots & A^{j_\ell} & A^{j_{\ell+1}} & \dots & A^{j_n} & b & e^1 & \dots & e^k & \dots & e^m \\
\hline
A^{j_1} & 1 & & 0 & t_{1,\ell+1} & \dots & t_{1,n} & \xi_1 & p_{1,1} & \dots & p_{1,k} & \dots & p_{1,m} \\
\vdots & & \ddots & & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\
A^{j_\ell} & 0 & & 1 & t_{\ell,\ell+1} & \dots & t_{\ell,n} & \xi_\ell & p_{\ell,1} & \dots & p_{\ell,k} & \dots & p_{\ell,m} \\
\hline
e^{i_1} & 0 & \dots & 0 & 0 & \dots & 0 & \xi_{\ell+1} & p_{\ell+1,1} & \dots & p_{\ell+1,k} & \dots & p_{\ell+1,m} \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\
e^{i_r} & 0 & \dots & 0 & 0 & \dots & 0 & \boxed{\xi_{\ell+r}} & p_{\ell+r,1} & \dots & p_{\ell+r,k} & \dots & p_{\ell+r,m} \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\
e^{i_{m-\ell}} & 0 & \dots & 0 & 0 & \dots & 0 & \xi_m & p_{m,1} & \dots & p_{m,k} & \dots & p_{m,m}
\end{array}$$

The $\ell \times \ell$ block in the upper left is an identity matrix, with an $(m-\ell) \times \ell$ block of zeroes below it. This comes from the fact that the representation of columns of A in the left-hand margin basis puts coefficient 1 on the basis element and 0 elsewhere. The $(m-\ell) \times (n-\ell)$ block to the right is zero since no additional replacements can be made. The middle column indicates that

$$b = \sum_{k=1}^{\ell} \xi_k A^{j_k} + \sum_{r=1}^{m-\ell} \xi_{\ell+r} e^{i_r}.$$

If $\xi_{\ell+1} = \cdots = \xi_m = 0$ (which must be true if $\ell = m$), then b is a linear combination only of columns of A, so alternative (74) holds, and we have found a solution. (We may have to rearrange the order of the coordinates of x.)

The Replacement Lemma 51 guarantees that $A^{j_1}, \ldots, A^{j_\ell}, e^{i_1}, \ldots, e^{i_{m-\ell}}$ is a basis for $\boldsymbol{R}^{\mathrm{m}}$. So if some ξ_k is not zero for $m \geqslant k > \ell$, then Proposition 52 implies that the corresponding p^k row vector satisfies $p^k \cdot b = \xi_k \neq 0$, and $p^k \cdot A^j = 0$ for all j. Multiplying by -1 if necessary, p_k satisfies alternative (75).

As for the rationality of x and p, if all the elements of A are rational, then all the elements of the original *tableau* are rational, and the results of pivot operation are all rational, so the final *tableau* is rational. ∎

55 Remark As an aside, observe that $A^{j_1}, \ldots, A^{j_\ell}$ is a basis for the column space of A, and $p^{\ell+1}, \ldots, p^m$ is a basis for its orthogonal complement.

56 Remark Another corollary is that if all the columns of A are used in the basis, the matrix P is the inverse of A. This is the well-known result that the Gauss–Jordan method can be used to invert a matrix.

11 Farkas' Lemma Revisited

The Farkas Lemma concerns nonnegative solutions to linear inequalities. You would think that we can apply the Replacement Lemma here to a constructive proof of the Farkas Lemma, and indeed we can. But the choice of replacements is more complicated when we are looking for nonnegative solutions to systems of inequalities, and uses the Simplex Algorithm of Linear Programming.

References

[1] T. M. Apostol. 1969. *Calculus*, 2d. ed., volume 2. Waltham, Massachusetts: Blaisdell.

[2] A. Bachem and W. Kern. 1992. *Linear programming duality: An introduction to oriented matroids*. Berlin: Springer–Verlag.

[3] K. C. Border. Quick review of matrix and linear algebra. http://www.its.caltech.edu/~kcborder/Notes/LinearAlgebra.pdf

[4] G. B. Dantzig. 1963. *Linear programming and extensions*. Princeton: Princeton University Press.

[5] G. Darboux, ed. 1890. *Oeuvres de Jean-Baptiste Joseph Fourier*, volume 2. Paris: Gauthier–Villars.

[6] L. L. Dines. 1919. Systems of linear inequalities. *Annals of Mathematics* 20(3):191–199. http://www.jstor.org/stable/1967869

[7] ——— . 1926. On positive solutions of a system of linear equations. *Annals of Mathematics* 28:386–392. http://www.jstor.org/stable/1968384

[8] J. Farkas. 1902. Über die Theorie der einfachen Ungleichungen. *Journal für die Reine und Angewandte Mathematik* 124:1–27. http://www.digizeitschriften.de/main/dms/img/?PPN=GDZPPN002165023

[9] J.-B. J. Fourier. 1826. Solution d'une question particuliére du calcul des inégalités [Solution of a particular problem in the calculus of inequalities]. *Nouveau Bulletin des Sciences par la Société Philomathique de Paris* pages 99–101. Reprinted in [5, pp. 317–319].

[10] I. Fredholm. 1903. Sur une classe d'Équations fonctionelles. *Acta Mathematica* 27(1):365–390. DOI: 10.1007/BF02421317

[11] D. Gale. 1951. Convex polyhedral cones and linear inequalities. In T. C. Koopmans, ed., *Activity Analysis of Production and Allocation*, number 13 in Cowles Commission for Research in Economics Monographs, chapter 17, pages 287–297. New York: John Wiley and Sons. `http://cowles.econ.yale.edu/P/cm/m13/m13-17.pdf`

[12] ——— . 1969. How to solve linear inequalities. *American Mathematical Monthly* 76(6):589–599. `http://www.jstor.org/stable/2316658`

[13] ——— . 1989. *Theory of linear economic models*. Chicago: University of Chicago Press. Reprint of the 1960 edition published by McGraw-Hill.

[14] P. Gordan. 1873. Über die auflösung linearer Gleichungen mit reelen Coefficienten [On the solution of linear inequalities with real coefficients]. *Mathematische Annalen* 6(1):23–28. DOI: 10.1007/BF01442864

[15] H. W. Kuhn. 1956. Solvability and consistency for linear equations and inequalities. *American Mathematical Monthly* 63(4):217–232. `http://www.jstor.org/stable/2310345`

[16] H. W. Kuhn and A. W. Tucker, eds. 1956. *Linear inequalities and related systems*. Number 38 in Annals of Mathematics Studies. Princeton: Princeton University Press.

[17] W. W. Leontief. 1941. *The structure of the American economy, 1919–1939*. New York: Oxford University Press.

[18] O. L. Mangasarian. 1994. *Nonlinear programming*. Classics in Applied Mathematics. Philadelphia: SIAM. Reprint of the 1969 edition published by McGraw–Hill.

[19] W. H. Marlow. 1993. *Mathematics for operations research*. New York: Dover Publications. Reprint of the 1978 edition published by John Wiley and Sons, New York.

[20] S. Morris. 1994. Trade with heterogeneous prior beliefs and asymmetric information. *Econometrica* 62(6):1327–1347. `http://www.jstor.org/stable/2951751`

[21] T. S. Motzkin. 1934. *Beiträge zur Theorie der linearen Ungleichungen*. PhD thesis, Universität Basel.

[22] ——— . 1951. Two consequences of the transposition theorem on linear inequalities. *Econometrica* 19(2):184–185. `http://www.jstor.org/stable/1905733`

[23] W. A. Neilson, T. A. Knott, and P. W. Carhart, eds. 1944. *Webster's new international dictionary of the English language*, second unabridged ed. Springfield, Massachusetts: G. & C. Merriam Company.

[24] H. Nikaidô. 1968. *Convex structures and economic theory*. Mathematics in Science and Engineering. New York: Academic Press.

[25] F. Quesnay. 1758. *Tableau Économique*. Versailles.

[26] F. Riesz and B. Sz.-Nagy. 1955. *Functional analysis*. New York: Frederick Ungar. Translated from the 1953 second French edition by Leo T. Boron.

[27] E. Stiemke. 1915. Über positive Lösungen homogener linearer Gleichungen. *Mathematische Annalen* 76(2–3):340–342. DOI: 10.1007/BF01458147

[28] J. Stoer and C. Witzgall. 1970. *Convexity and optimization in finite dimensions I*. Number 163 in Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berüsichtigung der Anwendungsgebiete. Berlin: Springer–Verlag.

[29] A. W. Tucker. 1956. Dual systems of homogeneous linear relations. In Kuhn and Tucker [16], pages 3–18.

[30] J. A. Ville. 1938. Sur la théorie générale des jeux où intervient l'habileté des joueurs. In . Borel, ed., *Traité du Calcul des Probabilités et de ses Applications, Applications aux Jeux de Hasard*, volume IV, 2, pages 105–113. Paris: Gauthier-Villars.

[31] H. Weyl. 1935. Elementare Theorie der konvexen Polyeder. *Commentarii Mathematici Helvetici* 7:290–306. Translated in [32].

[32] ——— . 1950. The elementary theory of convex polyhedra. In H. W. Kuhn and A. W. Tucker, eds., *Contributions to the Theory of Games, I*, number 24 in Annals of Mathematics Studies, chapter 1, pages 3–18. Princeton: Princeton University Press. Translation by H. W. Kuhn of [31].

[33] G. M. Ziegler. 1995. *Lectures on polytopes*. Number 152 in Graduate Texts in Mathematics. New York: Springer–Verlag.