CALIFORNIA INSTITUTE OF TECHNOLOGY
Division of the Humanities and Social Sciences

Alternative Linear Inequalities

KC Border
October 2007
v. 2011.10.13::14.27

1 Solutions of systems of equalities and inequalities

In this section we present some basic results on the existence of solutions to linear equalities and inequalities. These results 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" [17]. While it is possible to prove these results in inequalities in a purely algebraic fashion (cf. Gale [11, Chapter 2]), the geometric approach is illuminating.

> *In these notes I shall adopt David Gale's [11] notation, which does not distinguish between row and column vectors. This means that if A is an $n \times m$ matrix, and x is a vector, and I write Ax, you infer that that x is an m-dimensional column vector, and if I write yA, you infer that y is an n-dimensional row vector. The notation yAx means that x is an m-dimensional column vector, y is an n-dimensional row vector, and yAx is the scalar $yA \cdot x = y \cdot Ax$.*

The usual ordering on $\boldsymbol{R}$ is denoted $\geqslant$ or $\leqslant$. On $\boldsymbol{R}^n$, the ordering $x \geqq y$ means $x_i \geqslant y_i$, $i = 1, \ldots, n$, while $x \gg y$ means $x_i > y_i$, $i = 1, \ldots, n$. 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 $\boldsymbol{R}^n_+ = \{x \in \boldsymbol{R}^n : x \geqq 0\}$ and $\boldsymbol{R}^n_{++} = \{x \in \boldsymbol{R}^n : x \gg 0\}$, the **nonnegative orthant** and **strictly positive orthant** of $\boldsymbol{R}^n$ respectively.

$$x \geqq y \iff x_i \geqslant y_i,\ i = 1, \ldots, n$$
$$x > y \iff x_i \geqslant y_i,\ i = 1, \ldots, n \text{ and } x \neq y$$
$$x \gg y \iff x_i > y_i,\ i = 1, \ldots, n$$

Figure 1. Orderings on $\boldsymbol{R}^n$.

2 Basic nonnegative linear combinations

Let $A = \{x_1, \ldots, x_n\}$ 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_is 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.

1 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 $x_1, \ldots, x_n$ *are vectors and and* $y = \sum_{i=1}^{n} \lambda_i x_i$ *where each* λ_i *is nonnegative, then either* $y = 0$*, or there exist nonnegative* $\beta_1, \ldots, \beta_n$ *such that* $y = \sum_{i=1}^{n} \beta_i x_i$ *and* $\{x_i : \beta_i > 0\}$ *is 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 $\lambda_i > 0$, $i = 1, \ldots, n$. If $x_1, \ldots, x_n$ itself constitutes an independent set, there is nothing to prove, just set $\beta_i = \lambda_i$ for each i. On the other hand, if $x_1, \ldots, x_n$ are dependent, then there exist numbers $\alpha_1, \ldots, \alpha_n$, 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_is. Thus by the induction hypothesis $\mathbb{P}[n-1]$, we can express y as a linear combination that depends on a linearly independent subset. ∎

2 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 [13] did, that we did not really "eliminate" a variable, we just set its coefficient to zero.)

The first application of Lemma 1 is Carathéodory's theorem on convex hulls in finite dimensional spaces.

3 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 $j = 1, \dots, m$ 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 1, we can write $\hat{x} = \sum_{i=1}^{k} \alpha_i \hat{x}_i$, where $\hat{x}_1, \dots, \hat{x}_k$ 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. ∎

4 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.

5 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 $\big\{\alpha \in \boldsymbol{R}^{\mathrm{m}+1} : (\forall i = 0, \dots, m)\ [\alpha_i \geqslant 0]\ \&\ \sum_{i=0}^{m} \alpha_i = 1\big\}$) 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 1 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.[1]

[1] 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 $\boldsymbol{R}^{\mathrm{m}}$.

6 Lemma *Every finite 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 finite cone generated by A. Let y be the limit of some sequence y_n in C,

$$y_n \to y.$$

By Lemma 1 we can write each y_n as a nonnegative linear combination of an independent subset of the x_is. 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 $0 \to \alpha_j = 0$, so that y belongs to C. (For a proof that the coordinate mapping is continuous, something which most authors take for granted, see my on-line note at http://www.hss.caltech.edu/~kcb/Notes/Coordinates.pdf.) ∎

3 Solutions of systems of equalities

Consider the system of linear equations

$$Ax = b,$$

where $A = \begin{bmatrix} a_{i,j} \end{bmatrix}$ 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 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,

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

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 $p_1, \dots, p_m$, 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,

$$A^j \cdot p = c_j, \qquad j = 1, \dots, n$$

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.

Do I need this definition?

7 Definition *A vector $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$ is a* ***solution*** *of the system*

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

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 [19, p. 164] I shall refer to it as the Fredholm Alternative, as Fredholm [8] proved it in 1903 in the context of integral equations.

I'll bet there is an earlier proof.

8 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*

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

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

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

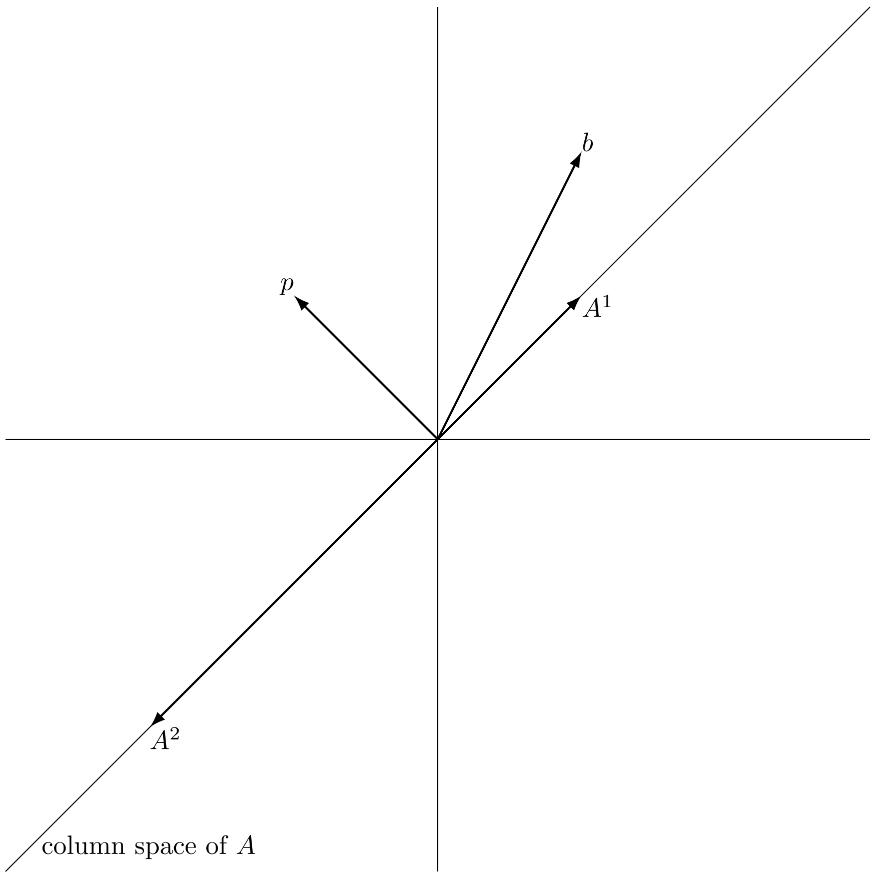

Figure 2. Geometry of the Fredholm Alternative

Proof: It is easy to see that both (1) and (2) 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 (1) 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 (2). ∎

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 $p \cdot b = p \cdot p > 0$. ∎

9 Remark There is another way to think about the Fredholm alternative, which was expounded by Kuhn [13]. Either the system $Ax = b$ has a solution, or we can find weights $p_1, \ldots, p_n$ such that if we weight the equations

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

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**.

10 Corollary *Let $p^0, p^1, \ldots, p^m \in \boldsymbol{R}^{\mathrm{n}}$ 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 (2) of Theorem 8 is false, so alternative (1) 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 12. Julius Farkas [6] proved them both in 1902.[2]

11 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*

$$\begin{aligned} Ax &= b \\ x &\geqq 0 \end{aligned} \tag{3}$$

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

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

Proof: (3) $\implies$ $\neg$(4) : 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, (4) fails.

$\neg$(3) $\implies$ (4) : Let $C = \{Ax : x \geqq 0\}$. If (3) fails, then b does not belong to C. By Lemma 6, the finite 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 $z \in X$ and $p \cdot b < 0$. Therefore (4). ∎

Note that there are many trivial variations on this result. For instance, multiplying p by -1, I could have written (4) as $pA \leqq 0$ & $p \cdot b > 0$. Or by replacing A by its transpose, we could rewrite (3) with $xA = b$ and (4) with $Ap \geqq 0$. Keep this in mind as you look at the coming theorems.

5 Solutions of systems of inequalities

Once we have a result on nonnegative solutions of equalities, we get on one 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

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

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

12 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*

$$\begin{aligned} Ax &\leqq b \\ x &\geqq 0 \end{aligned} \tag{5}$$

[2] 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.

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

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

13 Exercise *Prove the corollary by converting the inequalities to equalities as discussed above and apply the Farkas Lemma.*

Both versions of Farkas's Lemma are subsumed by the next result, which is also buried in Farkas [6].

14 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*

$$\begin{aligned} Ax &= b \\ Bx &\leqq c \\ x &\geqq 0 \end{aligned} \tag{7}$$

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

$$\begin{aligned} pA + qB &\geqq 0 \\ q &\geqq 0 \\ p \cdot b + q \cdot c &< 0. \end{aligned} \tag{8}$$

15 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 11.

We can reformulate the above as follows.

16 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*

$$\begin{aligned} (Ax)_i &= b_i \quad i \in E \\ (Ax)_i &\leqslant b_i \quad i \in I \\ x &\geqq 0 \end{aligned} \tag{9}$$

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

$$\begin{aligned} pA &\geqq 0 \\ p \cdot b &< 0 \\ p_i &\geqq 0 \quad i \in E. \end{aligned} \tag{10}$$

We can also handle strict inequalities. See Figure 3 for a geometric interpretation of the next theorem, which is due to Gordan [12] in 1873.

17 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*

$$Ax \gg 0. \tag{11}$$

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

$$\begin{aligned} pA &= 0 \\ p &> 0 \end{aligned} \tag{12}$$

There are two ways (11) can be satisfied. The first is that one column of A is zero, say column j—then $x = e^j$ satisfies (11). If no column of A is zero, then the finite cone $\langle A^1 \rangle + \cdots + \langle A^n \rangle$ 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, (11) fails, then the generators lie in the same open half space $[p > 0]$.

There is another, algebraic, interpretation of Gordan's Alternative in terms of consistency and solvability. It says that if (11) is not solvable, then we may multiply each equality by a multiplier p_i and add them so that the resulting coefficients on x_j, namely $p \cdot A^j$ are all strictly positive, but the right-hand side remains zero, showing that (11) is inconsistent.

18 Exercise *Prove Gordan's Alternative. Hint: If x satisfies (11), 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 (11) can be written as*

$$-A(u - v) \leqq -\mathbf{1}. \tag{11$'$}$$

Now use Corollary 12.

We can rewrite Corollary 17 as follows.

19 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*

$$Ax \gg 0 \tag{13}$$

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

$$\begin{aligned} pA &= 0 \\ p \cdot \mathbf{1} &= 1 \\ p &> 0. \end{aligned} \tag{14}$$

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

20 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*

$$\begin{aligned} Ax &\gg 0 \\ x &\geqq 0. \end{aligned} \tag{15}$$

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

$$\begin{aligned} pA &\leqq 0 \\ p &> 0 \end{aligned} \tag{16}$$

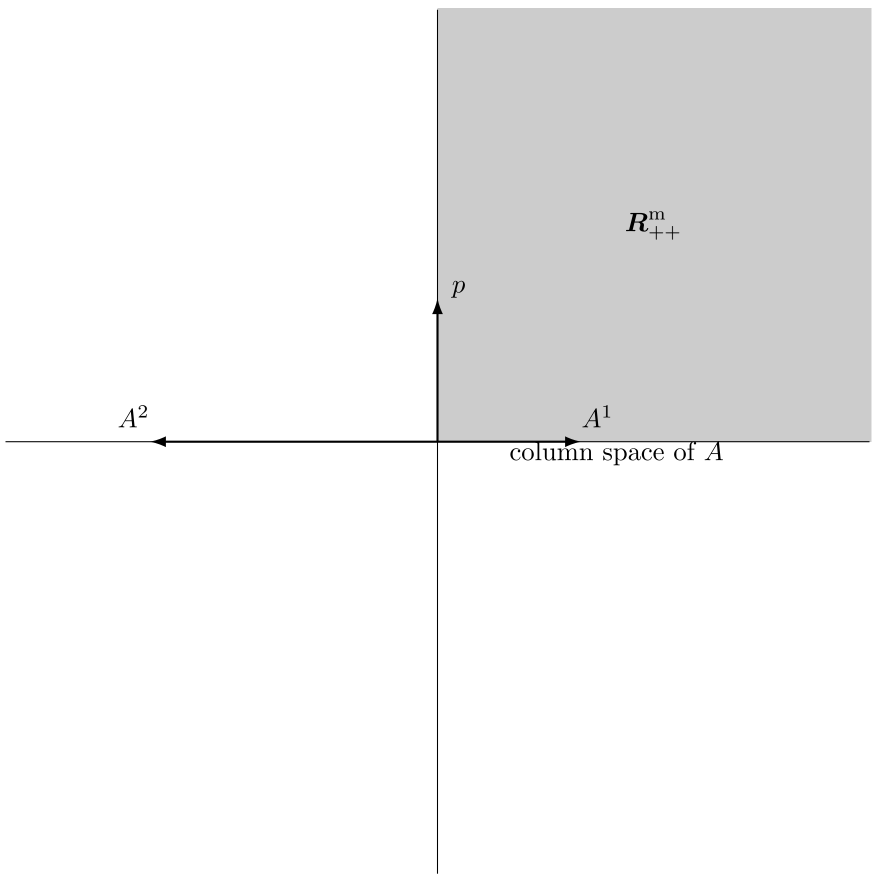

Figure 3. Geometry of Gordan's Alternative 17

21 Exercise *Prove the Corollary. Hint: If (15) holds we may normalize x so that $Ax \geqq \mathbf{1}$. Use Corollary 12.*

The following result was proved by Stiemke [20] in 1915.

22 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*

$$Ax > 0 \tag{17}$$

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

$$\begin{aligned} pA &= 0 \\ p &\gg 0. \end{aligned} \tag{18}$$

Proof: (17) $\implies \neg$(18) : Clearly both cannot be true, for then we must have both $pAx = 0$ (as $pA = 0$) and $pAx > 0$ (as $p \gg 0$ and $Ax > 0$).

$\neg$(17) $\implies$ (18) : Let $\Delta = \{z \in \boldsymbol{R}^{\mathrm{m}} : z \geqq 0 \text{ and } \sum_{j=1}^n z_j = 1\}$ be the the unit simplex in $\boldsymbol{R}^{\mathrm{m}}$, In geometric terms, (17) asserts that the span M of the columns $\{A^1, \ldots, A^n\}$ 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 (17) is equivalent to the disjointness of M and Δ,

So assume that (17) 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 (18), we could rescale p so that it is a strictly positive probability vector. Also not 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.

Alternate proof of Stiemke's Theorem: If (17) 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 12 to deduce the existence of p^i satisfying $pA = 0$, $p^i \cdot e^i > 0$, and $p^i > 0$.

Now observe that if (17) fails, then for each $i = 1, \ldots, m$, there must exist p^i as described. Now set $p = p^1 + \cdots + p^m$ to get p satisfying (18). ∎

Finally we come to another alternative, Motzkin's Transposition Theorem [15], proven in his 1934 Ph.D. thesis. This statement is take from his 1951 paper [16].[3]

[3] Motzkin [16] contains an unfortunate typo. The condition $Ax \gg 0$ is erroneously given as $Ax \ll 0$.

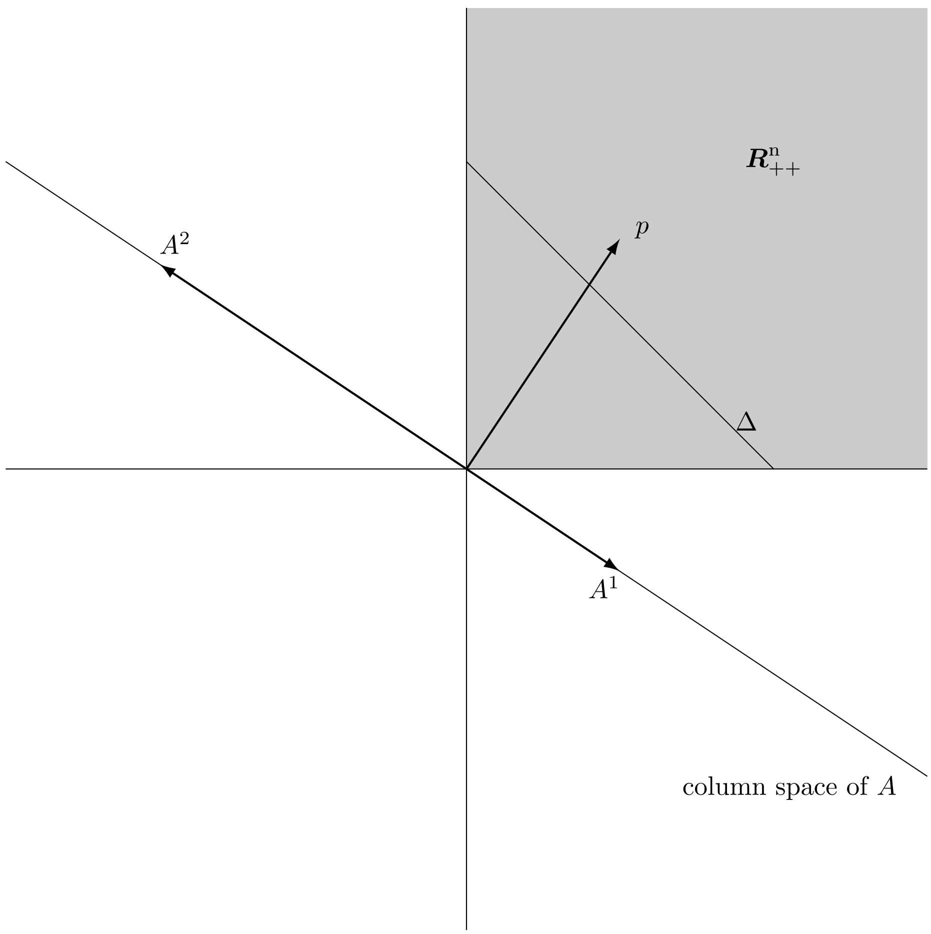

Figure 4. Geometry of the Stiemke Alternative

23 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*

$$\begin{aligned} Ax &\gg 0 \\ Bx &\geqq 0 \\ Cx &= 0 \end{aligned} \tag{19}$$

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

$$\begin{aligned} p^1A + p^2B + p^3C &= 0 \\ p^1 &> 0 \\ p^2 &\geqq 0. \end{aligned} \tag{20}$$

Motzkin expressed (20) in terms of the transpositions of A, B, and C.

24 Exercise *Prove the Transposition Theorem.*

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]?

25 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.[4]

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:

$$\left[\begin{array}{cc|c} 3 & 2 & 8 \\ 2 & 3 & 7 \end{array}\right]. \tag{21'}$$

[4]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 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:

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

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:

$$\left[\begin{array}{cc|c} 1 & \frac{2}{3} & \frac{8}{3} \\ 2-2\cdot 1 & 3-2\cdot\frac{2}{3} & 7-2\cdot\frac{8}{3} \end{array}\right] = \left[\begin{array}{cc|c} 1 & \frac{2}{3} & \frac{8}{3} \\ 0 & \frac{5}{3} & \frac{5}{3} \end{array}\right]. \tag{22$'$}$$

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

$$\left[\begin{array}{cc|c} 1-\frac{2}{3}\cdot 0 & \frac{2}{3}-\frac{2}{3}\cdot 1 & \frac{8}{3}-\frac{2}{3}\cdot 1 \\ 0 & 1 & 1 \end{array}\right] = \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \end{array}\right], \tag{23$'$}$$

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

7 A different look at the Gauss–Jordan method

David Gale [11] 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***.[5]

	a^1	a^2	b
e^1	3	2	8
e^2	2	3	7

(21)

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 (21), 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:

	a^1	a^2	b
a^1	?	?	?
e^2	?	?	?

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

	a^1	a^2	b
a^1	1	?	?
e^2	0	?	?

[5]The term *tableau*, a French word best translated as "picture" or "painting," harkens back to Quesnay's *Tableau économique* [18], which inspired Leontief [14], whose work spurred the Air Force's interest in linear programming [2, p. 17].

I claim the rest of the coefficients are

	a^1	a^2	b
a^1	1	$\frac{2}{3}$	$\frac{8}{3}$
e^2	0	$\frac{5}{3}$	$\frac{5}{3}$

(22)

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* (22)) is the same as (22′).

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

	a^1	a^2	b
a^1	1	0	2
a^2	0	1	1

(23)

This is the same as (23). 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 $\{b^1, \dots, b^m\}$ span $\mathcal{A}$. That is, each a^j can be written as a linear combination of b^is. Let $T = [t_{i,j}]$ be the $m \times n$ matrix of coordinates of the a^js with respect to the b^is.[6] That is,

$$a^j = \sum_{k=1}^{m} t_{k,j} b^k, \qquad j = 1, \dots, n. \tag{24}$$

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

	a^1	$\dots$	a^j	$\dots$	a^n
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	$t_{m,1}$	$\dots$	$t_{m,j}$	$\dots$	$t_{m,n}$

(24′)

[6]If the b^is are linearly dependent, T may not be unique.

- A *tableau* is actually a statement. It asserts that the equations (24) 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 (24) 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 $a^1, \ldots, a^n$, B is the matrix $m \times m$ matrix whose columns are $b^1, \ldots, b^m$, and T is the $m \times n$ matrix $[t_{i,j}]$.

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

26 Replacement Lemma *If $\{b^1, \ldots, b^m\}$ is a linearly independent set that spans $\mathcal{A}$, then*

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

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

	a^1	$\ldots$	$a^{\ell-1}$	a^ℓ	$a^{\ell+1}$	$\ldots$	a^n
b^1	$t'_{1,1}$	$\ldots$	$t'_{1,\ell-1}$	0	$t'_{1,\ell+1}$	$\ldots$	$t'_{1,n}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$
b^{k-1}	$t'_{k-1,1}$	$\ldots$	$t'_{k-1,\ell-1}$	0	$t'_{k-1,\ell+1}$	$\ldots$	$t'_{k-1,n}$
a^ℓ	$t'_{k,1}$	$\ldots$	$t'_{k,\ell-1}$	1	$t'_{k,\ell+1}$	$\ldots$	$t'_{k,n}$
b^{k+1}	$t'_{k+1,1}$	$\ldots$	$t'_{k+1,\ell-1}$	0	$t'_{k+1,\ell+1}$	$\ldots$	$t'_{k+1,n}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$
b^m	$t'_{m,1}$	$\ldots$	$t'_{m,\ell-1}$	0	$t'_{m,\ell+1}$	$\ldots$	$t'_{m,n}$

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

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

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

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

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 $\{b^1, \ldots, b^{k-1}, a^\ell, b^{k+1}, \ldots, b^m\}$ 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 $\{b^1, \ldots, b^m\}$ 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 $\{b^1, \ldots, b^{k-1}, a^\ell, b^{k+1}, \ldots, b^m\}$ 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.

27 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.

	a^1	$\dots$	a^j	$\dots$	a^n	e^1		$\dots$		e^m
b^1	$t_{1,1}$	$\dots$	$t_{1,j}$	$\dots$	$t_{1,n}$	$y_{1,1}$	$\dots$	$\dots$	$\dots$	$y_{1,m}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\ddots$	$\vdots$		$\vdots$
b^i	$t_{i,1}$	$\dots$	$t_{i,j}$	$\dots$	$t_{i,n}$	$y_{i,1}$	$\dots$	$\dots$	$\dots$	$y_{i,m}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\ddots$	$\vdots$
b^m	$t_{m,1}$	$\dots$	$t_{m,j}$	$\dots$	$t_{m,n}$	$y_{m,1}$	$\dots$	$\dots$	$\dots$	$y_{m,m}$

(25)

That is, for each j,

$$a^j = \sum_{i=1}^{m} t_{i,j} b^i \tag{26}$$

and

$$e^j = \sum_{i=1}^{m} y_{i,j} b^i. \tag{27}$$

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

$$y^i \cdot a^j = t_{i,j}. \tag{28}$$

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 (26) is just

$$A = BT$$

where and (27) is just

$$I = BY.$$

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

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

which is equivalent to (28). ∎

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

	a^1	$\dots$	a^m	e^1	$\dots$	e^m
a^1	1		0	$y_{1,1}$	$\dots$	$y_{1,m}$
$\vdots$		$\ddots$		$\vdots$		$\vdots$
a^m	0		1	$y_{m,1}$	$\dots$	$y_{m,m}$

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

10 The Fredholm Alternative revisited

Recall the Fredholm Alternative 8 that we previously proved using a separating hyperplane argument. We can now prove a stronger version using a purely algebraic argument.

29 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*

$$Ax = b \tag{29}$$

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

$$\begin{aligned} pA &= 0 \\ p \cdot b &> 0. \end{aligned} \tag{30}$$

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 26, and simultaneously compute x or p. Let A be the $m \times n$ with columns $A^1, \ldots, A^n$ in $\boldsymbol{R}^{\mathrm{m}}$. Then $x \in \boldsymbol{R}^{\mathrm{n}}$ and $b \in \boldsymbol{R}^{\mathrm{m}}$. Begin with this *tableau*.

	A^1	$\ldots$	A^n	b	e^1	$\ldots$	e^m
e^1	$\alpha_{1,1}$	$\ldots$	$\alpha_{1,n}$	β_1	1		0
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\ddots$	
e^m	$\alpha_{m,1}$	$\ldots$	$\alpha_{m,n}$	β_m	0		1

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:

	A^{j_1}	$\ldots$	A^{j_ℓ}	$A^{j_{\ell+1}}$	$\ldots$	A^{j_n}	b	e^1	$\ldots$	e^k	$\ldots$	e^m
A^{j_1}	1		0	$t_{1,\ell+1}$	$\ldots$	$t_{1,n}$	ξ_1	$p_{1,1}$	$\ldots$	$p_{1,k}$	$\ldots$	$p_{1,m}$
$\vdots$		$\ddots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
A^{j_ℓ}	0		1	$t_{\ell,\ell+1}$	$\ldots$	$t_{\ell,n}$	ξ_ℓ	$p_{\ell,1}$	$\ldots$	$p_{\ell,k}$	$\ldots$	$p_{\ell,m}$
e^{i_1}	0	$\ldots$	0	0	$\ldots$	0	$\xi_{\ell+1}$	$p_{\ell+1,1}$	$\ldots$	$p_{\ell+1,k}$	$\ldots$	$p_{\ell+1,m}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
e^{i_r}	0	$\ldots$	0	0	$\ldots$	0	$\xi_{\ell+r}$	$p_{\ell+r,1}$	$\ldots$	$p_{\ell+r,k}$	$\ldots$	$p_{\ell+r,m}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$e^{i_{m-\ell}}$	0	$\ldots$	0	0	$\ldots$	0	ξ_m	$p_{m,1}$	$\ldots$	$p_{m,k}$	$\ldots$	$p_{m,m}$

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 (29) holds, and we have found a solution. (We may have to rearrange the order of the coordinates of x.)

The Replacement Lemma 26 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 27 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 (30).

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. ∎

30 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.

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

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

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

[5] ——— . 1926. On positive solutions of a system of linear equations. *Annals of Mathematics* 28:386–392. www.jstor.org/stable/1968384.

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

[7] 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 [3, pp. 317–319].

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

[9] 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. cowles.econ.yale.edu/P/cm/m13/m13-17.pdf.

[10] ——— . 1969. How to solve linear inequalities. *American Mathematical Monthly* 76(6):589–599. www.jstor.org/stable/2316658
.

[11] ——— . 1989. *Theory of linear economic models*. Chicago: University of Chicago Press. Reprint of the 1960 edition published by McGraw-Hill.

[12] 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.

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

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

[15] T. S. Motzkin. 1934. *Beiträge zur Theorie der linearen Ungleichungen*. PhD thesis, Universität Basel. It appears from citations I have seen that this was published in Jerusalem in 1936. Motzkin [16] gives 1934 as the date.

[16] ——— . 1951. Two consequences of the transposition theorem on linear inequalities. *Econometrica* 19(2):184–185. www.jstor.org/stable/1905733
.

[17] 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.

[18] F. Quesnay. 1758. *Tableau économique*. Versailles.

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

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