Topic 25: Inequalities and alternatives

This chapter discusses both the solution of linear inequalities and the geometry of polyhedra. The two are intimately related, since the set of solutions to a system of linear inequalities is a polyhedron and vice versa.

Some of these results are in the form of an alternative, that is, "an opportunity for choice between two things, courses, or propositions, either of which may be chosen, but not both" [18]. These theorems will be proven twice. First we shall prove them geometrically using simple separating hyperplane arguments. Then we shall prove them algebraically and explore algorithms for solution. The latter approach also provides results on rational solutions.

25.1 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 $p_1, \ldots, p_m$, 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,

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

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.

Do I need this definition?

25.1.1 Definition *A vector $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)$ 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

I'll bet there is an earlier proof in the finite-dimensional case.

the question of how to characterize the existence of a solution. The answer is given by the following theorem. Following Riesz–Sz.-Nagy [21, p. 164] and wikipedia, I shall refer to it as the Fredholm Alternative, as Fredholm [8] proved it in 1903 in the context of integral equations. But I do note that Marlow [14, p. 86] refers to it as Gale's Theorem. It does appear in Gale [10, Theorem 2.5, p. 41].

25.1.2 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

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

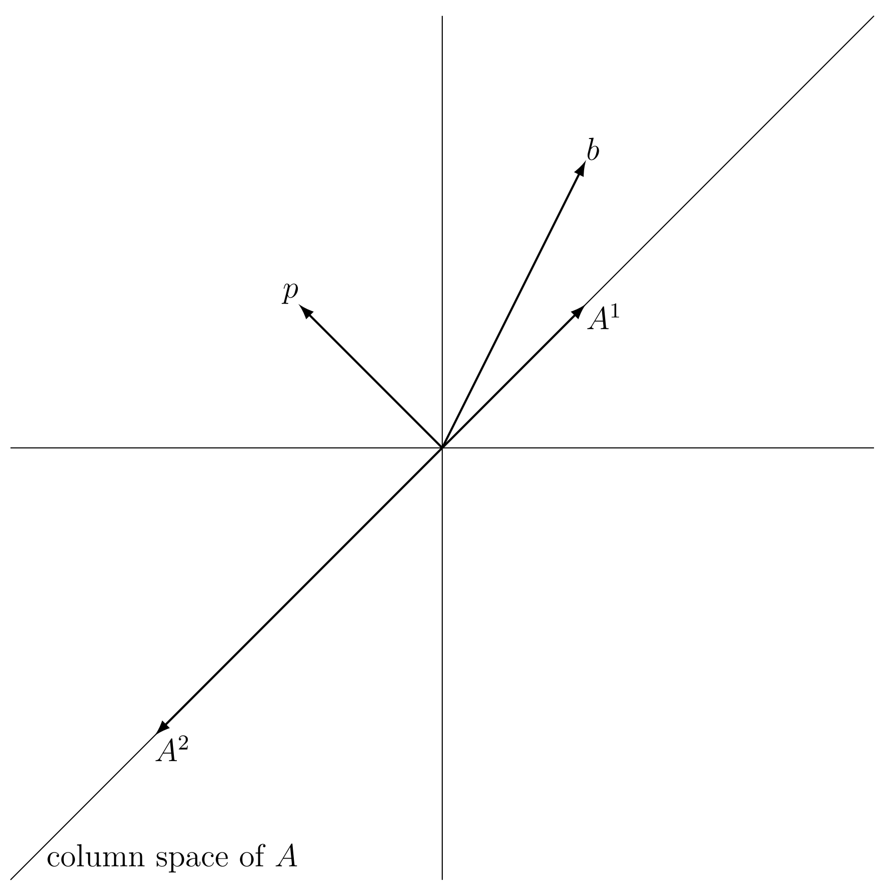

Figure 25.1.1. Geometry of the Fredholm Alternative

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 $p \cdot b = p \cdot p > 0$. ∎

25.1.3 Remark There is another way to think about the Fredholm alternative, which was expounded by Kuhn [12]. Either the system $Ax = b$ has a solution, or we can find weights $p_1, \dots, p_n$ 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**.

25.1.4 Corollary *Let $p^0, p^1, \dots, p^m \in \boldsymbol{R}^{\mathrm{m}}$ 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 $p^0 = \sum_{i=1}^m \lambda_i p^i$.*

Proof: Consider the matrix A whose columns are $p^1, \dots, p^m$, and set $b = p^0$. By hypothesis alternative (2) of Theorem 25.1.2 is false, so alternative (1) must hold. But that is precisely the conclusion of this theorem. ∎

Proof using orthogonal decomposition: Let $M = \operatorname{span}\{p^1, \dots, p^m\}$ 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, $p^i \cdot p^0_\perp = 0$, $i = 1, \dots, m$. 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 $p^1, \dots, p^m$. ∎

25.2 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' Lemma, and so is Corollary 25.3.1. What Julius Farkas [6][1] did prove in 1902 is a hybrid result of these two, which I present as Theorem 25.3.7 below.

25.2.1 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 $C = \{Ax : x \geqq 0\}$ be the cone generated by the columns of A. If (3) fails, then b does not belong to C. By Lemma 3.1.7, the finitely generated convex 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 (4). ∎

[1] 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.

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.

25.3 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 $p \neq 0$.

25.3.1 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*

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

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

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

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

There is a corollary of this where we do not impose the condition $x \geqq 0$. One way to prove the following is to note that $Ax \leqq b$ (with no restriction on x) is equivalent to $Ay - Az \leqq b$, $y \geqq 0$, $z \geqq 0$.

25.3.3 Corollary *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*

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

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

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

25.3.4 Exercise Prove Corollary 25.3.3. □

We can of course reverse the inequality.

25.3.5 Corollary *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*

$$Ax \geqq b \tag{9}$$

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

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

25.3.6 Exercise Prove Corollary 25.3.5. □

Now we come to another theorem articulated by Farkas [6].

25.3.7 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*

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

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

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

25.3.8 Exercise Prove this version of Farkas's Alternative. Hint: Rewrite (11) as

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

□

We can also handle strict inequalities. See Figure 25.3.1 for a geometric interpretation of the next theorem.

The next alternative is from Gordan [11] in 1873.

25.3.9 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*

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

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

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

There are two ways (14) can be satisfied. The first is that some row of A is zero, say row i. Then $p = e^i$ satisfies (14). If no row of A is zero, then the finitely generated convex cone $\langle A_1 \rangle + \cdots + \langle A_m \rangle$, where a_i is the i^{th} row of A, must contain a nonzero point and its negative. That is, the cone is not pointed. Gordan's Alternative says that if the cone is pointed, that is, (14) 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 (13) 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 (13) is inconsistent.

25.3.10 Exercise Prove Gordan's Alternative. Hint: If x satisfies (13), 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 (13) can be written as

$$\begin{bmatrix} A & -A \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \geqq \mathbf{1}. \tag{13$'$}$$

Now use Corollary 25.3.5 to Farkas's Alternative. □

Dantzig [5, p. 139] attributes the nest result to Jean Ville [27]. It may also be found in Gale [10, Theorem 2.10, p. 49].

25.3.11 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*

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

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

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

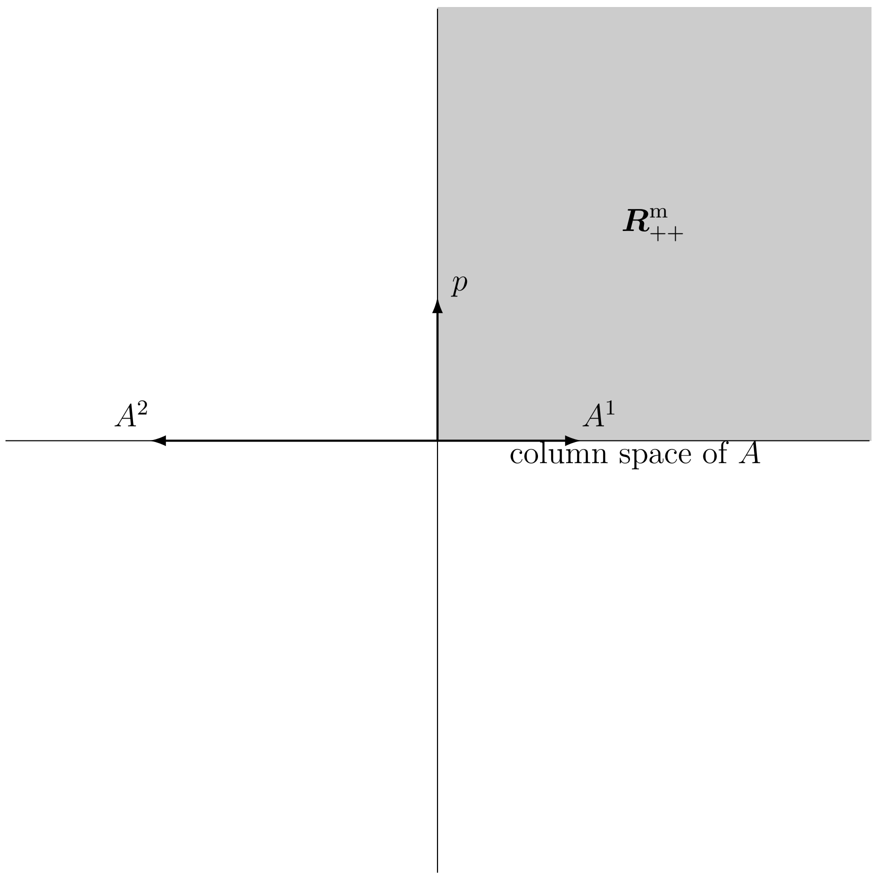

Figure 25.3.1. Geometry of Farkas's Alternative 25.3.1

25.3.12 Exercise Prove the Ville's Alternative. □

The following result was proved by Stiemke [23] in 1915.

25.3.13 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 unit simplex in $\boldsymbol{R}^{\mathrm{m}}$. In geometric terms, (17) 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 (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 $p \in \boldsymbol{R}^{\mathrm{m}}$ 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 column $A^j \in M$, we have that $p \cdot A^j = 0$, that is,

$$pA = 0.$$

This completes the proof. ∎

Note that in (18), 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.

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,

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

Fixing i for the moment, if this fails, then we can use the Corollary to Farkas's Alternative 25.3.5 to deduce the existence of p^i satisfying $p^i A = 0$, $p^i \cdot e^i > 0$, and $p^i > 0$.

Now observe that if (17) 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 (18). ∎

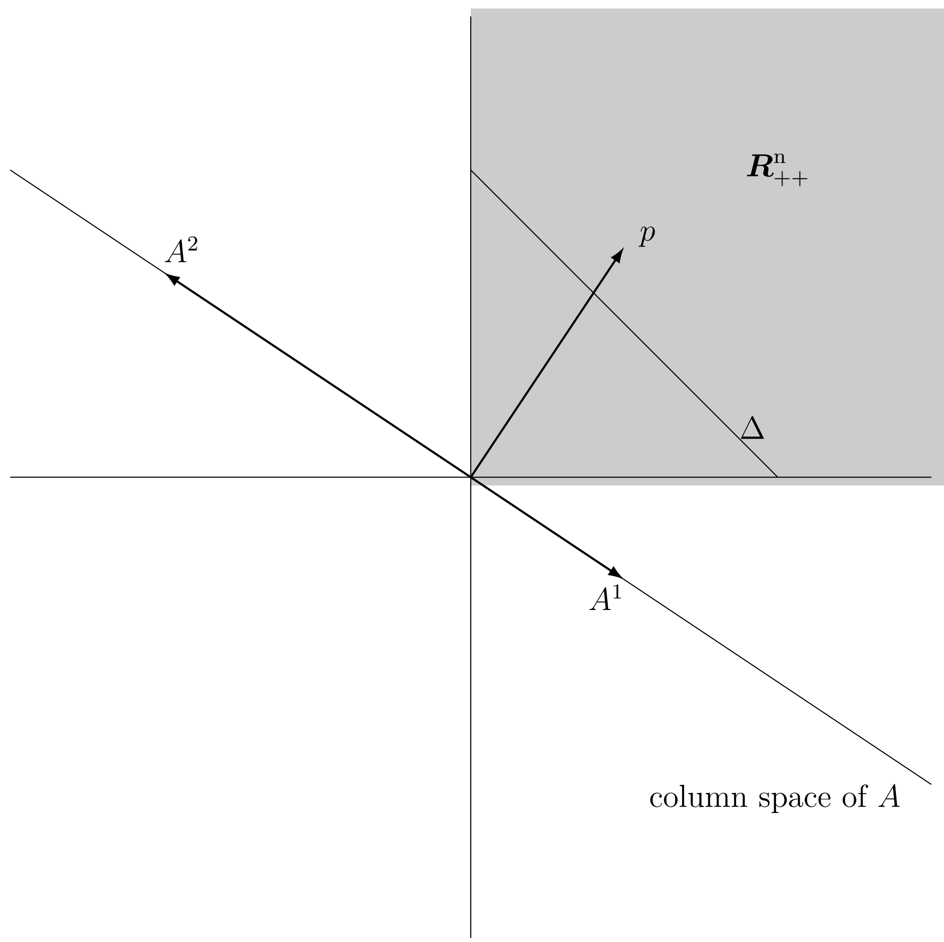

Figure 25.3.2. Geometry of the Stiemke Alternative

Finally we come to another alternative, Motzkin's Transposition Theorem [16], proven in his 1934 Ph.D. thesis. This statement is take from his 1951 paper [17].[2]

25.3.14 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.

25.3.15 Exercise Prove the Transposition Theorem. Hint: If x satisfies (19), it can be scaled to satisfy

$$\begin{bmatrix} A \\ B \\ C \\ -C \end{bmatrix} x \geqq \begin{bmatrix} \mathbf{1} \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Apply Corollary 25.3.5. □

Stoer and Witzgall [24] also provide a rational version of Motzkin's theorem, which can be recast as follows.

25.3.16 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*

$$\begin{aligned} Ax &\gg 0 \\ Bx &\geqq 0 \\ Cx &= 0 \end{aligned} \tag{21}$$

[2] Motzkin [17] contains an unfortunate typo. The condition $Ax \gg 0$ is erroneously given as $Ax \ll 0$.

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*

$$\begin{aligned} p^1A + p^2B + p^3C &= 0 \\ p^1 &> 0 \\ p^2 &\geqq 0. \end{aligned} \tag{22}$$

The alternate proof of Stiemke's Alternative points to another whole class of theorems that I first encountered in Morris [15]. Here is Corollary A1 from his paper.

25.3.17 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*

$$\begin{aligned} Ax &\geqq 0 \\ A_i \cdot x &> 0 \quad \text{for all } i \in S \end{aligned} \tag{23}$$

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

$$\begin{aligned} pA &= 0 \\ p &\geqq 0 \\ \sum_{i\in S} p_i &> 0 \quad \text{for all } S \in \mathcal{S}. \end{aligned} \tag{24}$$

25.3.18 Remark Observe that Stiemke's Alternative corresponds to the case where $\mathcal{S}$ is the set of all singletons: (23) reduces to Ax being semipositive. And $\sum_{i\in S} p_i > 0$ for the singleton $S = \{i\}$ (given that $p \geqq 0$) simply says $p_i > 0$. Requiring this for all singletons asserts that $p \gg 0$.

Proof of Theorem 25.3.17: It is clear that (23) and (24) cannot both be true.

So assume that (23) 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 25.3.14 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 (24). ∎

25.4 Tucker's Theorem

Tucker [25, Lemma, p. 5] proves the following theorem that is related to Theorems of the Alternative, but not stated as an alternative. See Nikaidô [19, Theorem 3.7, pp. 36–37] for a proof of Tucker's Theorem using the Stiemke's Alternative, and vice-versa.

25.4.1 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*

$$\begin{aligned} Ax &= 0 \\ x &\geqq 0 \\ A'p &\geqq 0 \\ A'p + x &\gg 0, \end{aligned} \tag{25}$$

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 25.3.13. It has two "dual" systems of inequalities

$$A'p > 0 \tag{17$'$}$$

and

$$Ax = 0, \qquad x \gg 0 \tag{18$'$}$$

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

$$A'p \geqq 0, \tag{17$''$}$$

$$Ax = 0, \qquad x \geqq 0. \tag{18$''$}$$

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 ((17'')–(18'')) 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., § 6.2 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.

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

$$x \in M, \quad x \geqq 0, \quad x_i > 0 \tag{26}$$

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

$$y \in M^\perp, \quad y \geqq 0, \quad y_i > 0. \tag{27}$$

25.5 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]?

25.5.1 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.[3]

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:

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

[3] 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 25.9.) 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:

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

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

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

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

25.6 A different look at the Gauss–Jordan method

David Gale [10] 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***.[4]

$$\begin{array}{c||cc|c} & a^1 & a^2 & b \\ \hline e^1 & 3 & 2 & 8 \\ e^2 & 2 & 3 & 7 \end{array} \tag{28}$$

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 (1′), 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

[4] The term *tableau*, a French word best translated as "picture" or "painting," harkens back to Quesnay's *Tableau économique* [20], which inspired Leontief [13], whose work spurred the Air Force's interest in linear programming [5, p. 17].

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	?	?

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	$\frac{5}{3}$	$\frac{5}{3}$

(29)

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* (29) is the same as (2′).

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

(30)

This is the same as (3′). 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.

25.7 *Tableaux* and 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^i's. Let $T = \begin{bmatrix} t_{i,j} \end{bmatrix}$ be the $m \times n$ matrix of coordinates of the a^j's with respect to the b^i's.[5] That is,

$$a^j = \sum_{k=1}^{m} t_{k,j} b^k, \qquad j = 1, \dots, n. \tag{31}$$

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

$$\begin{array}{c||ccccc} & a^1 & \dots & a^j & \dots & a^n \\ \hline b^1 & t_{1,1} & \dots & t_{1,j} & \dots & t_{1,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b^i & t_{i,1} & \dots & t_{i,j} & \dots & t_{i,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b^m & t_{m,1} & \dots & t_{m,j} & \dots & t_{m,n} \end{array} \tag{31$'$}$$

- A *tableau* is actually a statement. It asserts that the equations (31) 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 (31) in terms of the coordinates of the vectors as

  $$a^j_i = \sum_{k=1}^{m} t_{k,j} b^k_i$$

  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 $\begin{bmatrix} t_{i,j} \end{bmatrix}$.

[5] 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.

25.7.1 Replacement Lemma *If* $\{b^1, \dots, b^m\}$ *is a basis for* $\mathcal{A}$*, then*

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

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 $\{b^1,\dots,b^m\}$ 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 $\{b^1,\dots,b^{k-1},a^\ell,b^{k+1},\dots,b^m\}$ 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 $\mathcal{A}$. See Example 25.9.4.

- Note that elementary row operations preserve the scalar 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.

25.8 More on *tableaux*

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

25.8.1 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.

$$\begin{array}{c||ccccc||ccccc}
 & a^1 & \dots & a^j & \dots & a^n & e^1 & & \dots & & e^m \\ \hline
b^1 & t_{1,1} & \dots & t_{1,j} & \dots & t_{1,n} & y_{1,1} & \dots & \dots & \dots & y_{1,m} \\
\vdots & \vdots & & \vdots & & \vdots & \vdots & \ddots & \vdots & & \vdots \\
b^i & t_{i,1} & \dots & \boxed{t_{i,j}} & \dots & t_{i,n} & \boldsymbol{y_{i,1}} & \dots & \dots & \dots & \boldsymbol{y_{i,m}} \\
\vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots & \ddots & \vdots \\
b^m & t_{m,1} & \dots & t_{m,j} & \dots & t_{m,n} & y_{m,1} & \dots & \dots & \dots & y_{m,m}
\end{array} \tag{32}$$

That is, for each j,

$$a^j = \sum_{i=1}^m t_{i,j} b^i \tag{33}$$

and

$$e^j = \sum_{i=1}^m y_{i,j} b^i. \tag{34}$$

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

$$y^i \cdot a^j = t_{i,j}. \tag{35}$$

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 (33) is just

$$A = BT$$

where and (34) is just

$$I = BY.$$

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

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

which is equivalent to (35). ∎

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

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

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

25.9 The Fredholm Alternative revisited

Recall the Fredholm Alternative 25.1.2 that we previously proved using a separating hyperplane argument. We can now prove a stronger version using a purely algebraic argument.

25.9.1 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*

$$Ax = b \tag{36}$$

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

$$\begin{aligned} pA &= 0 \\ p \cdot b &> 0. \end{aligned} \tag{37}$$

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 25.7.1, 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:

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

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

The Replacement Lemma 25.7.1 guarantees that $A^{j_1}, \dots, A^{j_\ell}, e^{i_1}, \dots, e^{i_{m-\ell}}$ is a basis for $\boldsymbol{R}^{\mathrm{m}}$. So if some ξ_k is not zero for $m \geqslant k > \ell$, then Proposition 25.8.1 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 (37).

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. ∎

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

25.9.3 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.

25.9.4 Example Find a basis for the column space of

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 3 & 2 & -4 & 1 & 1 & 1 \\ 4 & 5 & -3 & 3 & 2 & 3 \end{bmatrix}$$

and a basis for its orthogonal complement. Note that the last row is the sum of the first three rows, so the rows are not independent.

Start with a basis of unit coordinate vectors.

Initial *tableau*:

	a^1	a^2	a^3	a^4	a^5	a^6	e^1	e^2	e^3	e^4
e^1	$\boxed{1}$	2	0	1	1	1	1	0	0	0
e^2	0	1	1	1	0	1	0	1	0	0
e^3	3	2	−4	1	1	1	0	0	1	0
e^4	4	5	−3	3	2	3	0	0	0	1

Replace e^1 by a^1 to get:

	a^1	a^2	a^3	a^4	a^5	a^6	e^1	e^2	e^3	e^4
a^1	1	2	0	1	1	1	1	0	0	0
e^2	0	$\boxed{1}$	1	1	0	1	0	1	0	0
e^3	0	−4	−4	−2	−2	−2	−3	0	1	0
e^4	0	−3	−3	−1	−2	−1	−4	0	0	1

Replace e^2 by a^2 to get:

	a^1	a^2	a^3	a^4	a^5	a^6	e^1	e^2	e^3	e^4
a^1	1	0	−2	−1	1	−1	1	−2	0	0
a^2	0	1	1	1	0	1	0	1	0	0
e^3	0	0	0	$\boxed{2}$	−2	2	−3	4	1	0
e^4	0	0	0	2	−2	2	−4	3	0	1

Replace e^3 by a^4 to get:

	a^1	a^2	a^3	a^4	a^5	a^6	e^1	e^2	e^3	e^4
a^1	1	0	−2	0	0	0	$-\frac{1}{2}$	0	$\frac{1}{2}$	0
a^2	0	1	1	0	1	0	$1\frac{1}{2}$	−1	$-\frac{1}{2}$	0
a^4	0	0	0	1	−1	1	$-1\frac{1}{2}$	2	$\frac{1}{2}$	0
e^4	0	0	0	0	0	0	−1	−1	−1	1

We are unable to replace the unit coordinate vector e^4, but none of the columns of A depend on it. That is, $\{a^1, a^2, a^4\}$ is a basis for the column space (there are others). Also observe that the vector $(-1, -1, -1, 1)$ is orthogonal to each column, and so is a basis vector for the one-dimensional orthogonal complement. □

25.10 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, so I will postpone the discussion until we discuss the Simplex Algorithm in Lecture 29. But that discussion will make use of the following variation of the Replacement Lemma.

25.10.1 Replacement Lemma with Nonnegativity *If the* tableau

$$\begin{array}{c||cccccccc||c}
 & a^1 & \ldots & a^{\ell-1} & a^\ell & a^{\ell+1} & \ldots & a^n & & z \\
\hline
b^1 & t_{1,1} & \ldots & t_{1,\ell-1} & t_{1,\ell} & t_{1,\ell+1} & \ldots & t_{1,n} & & \zeta_1 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\
b^{k-1} & t_{k-1,1} & \ldots & t_{k-1,\ell-1} & t_{k-1,\ell} & t_{k-1,\ell+1} & \ldots & t_{k-1,n} & & \zeta_{k-1} \\
b^k & t_{k,1} & \ldots & t_{k,\ell-1} & \boxed{t_{k,\ell}} & t_{k,\ell+1} & \ldots & t_{k,n} & & \zeta_k \\
b^{k+1} & t_{k+1,1} & \ldots & t_{k+1,\ell-1} & t_{k+1,\ell} & t_{k+1,\ell+1} & \ldots & t_{k+1,n} & & \zeta_{k+1} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\
b^m & t_{m,1} & \ldots & t_{m,\ell-1} & t_{m,\ell} & t_{m,\ell+1} & \ldots & t_{m,n} & & \zeta_m
\end{array}$$

satisfies

$$\zeta_i \geqslant 0, \quad i = 1, \ldots, m,$$

then the replacement operation with pivot $t_{k,\ell}$ yields a new tableau *with all $\zeta'_i \geqslant 0$ if and only if*

$$t_{k,\ell} > 0$$

and

$$\frac{\zeta_k}{t_{k,\ell}} = \min\left\{\frac{\zeta_i}{t_{i,\ell}} : t_{i,\ell} > 0\right\}.$$

That is, in order to keep the ζ's nonnegative the pivot $t_{k,\ell}$ must be chosen to be strictly positive and also to minimize the ratio $\zeta_i/t_{i,\ell}$ for strictly positive $t_{i,\ell}$ in column ℓ.

Proof: Since $\zeta'_k = \zeta_k/t_{k,\ell}$ we need $t_{k,\ell} > 0$. For the other rows

$$\zeta'_i = \zeta_i - \frac{t_{i,\ell}}{t_{k,\ell}}\zeta_k.$$

So if $t_{i,\ell} \leqslant 0$ we must have $\zeta'_i \geqslant 0$ given that we choose $t_{k,\ell} > 0$ since $\zeta'_i \geqslant 0$ by hypothesis. Otherwise, if $t_{i,\ell} > 0$, then $\zeta'_i \geqslant 0$ if and only if

$$\frac{\zeta_i}{t_{i,\ell}} \geqslant \frac{\zeta_k}{t_{k,\ell}}.$$

This completes the proof. ∎

25.11 Application to saddlepoint theorems

Recall the Saddlepoint Theorem 10.3.6 with Karlin's condition.

25.11.1 Saddlepoint Theorem *Let $f, g_1, \ldots, g_m \colon C \to \boldsymbol{R}$ be concave, where $C \subset \boldsymbol{R}^n$ is convex. Assume in addition that **Karlin's Condition**,*

$$(\forall \lambda > 0)\ (\exists \bar{x}(\lambda) \in C)\ \Big[\lambda \cdot g\big(\bar{x}(\lambda)\big) > 0\Big], \tag{K}$$

is satisfied.

The following are equivalent.

1. *The point x^* maximizes f over C subject to the constraints $g_j(x) \geqslant 0$, $j = 1, \ldots, m$.*

2. *Then there exist real numbers $\lambda_1^*, \ldots, \lambda_m^* \geqslant 0$ such that $(x^*; \lambda^*)$ is a saddlepoint of the Lagrangean L. That is,*

$$f(x) + \sum_{j=1}^{m} \lambda_j^* g_j(x) \leqslant f(x^*) + \sum_{j=1}^{m} \lambda_j^* g_j(x^*) \leqslant f(x^*) + \sum_{j=1}^{m} \lambda_j g_j(x^*)$$

for all $x \in C$ and all $\lambda_1, \ldots, \lambda_m \geqslant 0$. Furthermore $\lambda_j^* g_j(x^*) = 0$, $j = 1, \ldots, m$.

Uzawa [26] proposed the following modification.
****

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] B. D. Craven and J. J. Koliha. 1977. Generalizations of Farkas' theorem. *SIAM Journal on Mathematical Analysis* 8(6):983–997.
DOI: 10.1137/0508076

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

[6] 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

[7] W. Fenchel. 1953. Convex cones, sets, and functions. Lecture notes, Princeton University, Department of Mathematics. From notes taken by D. W. Blackett, Spring 1951.

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

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

[10] ——— . 1989. *Theory of linear economic models*. Chicago: University of Chicago Press. Reprint of the 1960 edition published by McGraw-Hill.

[11] 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

[12] 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

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

[14] 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.

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

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

[17] ——— . 1951. Two consequences of the transposition theorem on linear inequalities. *Econometrica* 19(2):184–185. http://www.jstor.org/stable/1905733

[18] 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.

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

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

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

[22] R. T. Rockafellar. 1970. *Convex analysis*. Number 28 in Princeton Mathematical Series. Princeton: Princeton University Press.

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

[24] 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.

[25] A. W. Tucker. 1956. Dual systems of homogeneous linear relations. In H. W. Kuhn and A. W. Tucker, eds., *Linear Inequalities and Related Systems*, number 38 in Annals of Mathematics Studies, pages 3–18. Princeton: Princeton University Press.

[26] H. Uzawa. 1958. The Kuhn–Tucker conditions in concave programming. In K. J. Arrow, L. Hurwicz, and H. Uzawa, eds., *Studies in Linear and Non-linear Programming*, number 2 in Stanford Mathematical Studies in the Social Sciences, chapter 3, pages 32–37. Stanford, California: Stanford University Press.

[27] 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.