Linear Programming Notes

Carl W. Lee
Department of Mathematics
University of Kentucky
Lexington, KY 40506
lee@ms.uky.edu

Fall 2009

Contents

1 References

The textbook for this course is Jon Lee, *A First Course in Combinatorial Optimization*, Cambridge, 2004. Four good references for linear programming are

1. Dimitris Bertsimas and John N. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific.

2. Vašek Chvátal, *Linear Programming*, W.H. Freeman.

3. George L. Nemhauser and Laurence A. Wolsey, *Integer and Combinatorial Optimization*, Wiley.

4. Christos H. Papadimitriou and Kenneth Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Prentice Hall.

I used some material from these sources in writing these notes. Also, some of the exercises were provided by Jon Lee and Francois Margot.

Exercise 1.1 Find as many errors in these notes as you can and report them to me. □

2 Exercises: Matrix Algebra

It is important to have a good understanding of the content of a typical one-semester undergraduate matrix algebra course. Here are some exercises to try. Note: Unless otherwise specified, all of my vectors are column vectors. If I want a row vector, I will transpose a column vector.

Exercise 2.1 Consider the product $C = AB$ of two matrices A and B. What is the formula for c_{ij}, the entry of C in row i, column j? Explain why we can regard the ith row of C as a linear combination of the rows of B. Explain why we can regard the jth column of C as a linear combination of the columns of A. Explain why we can regard the ith row of C as a sequence of inner products of the columns of B with a common vector. Explain why we can regard the jth column of C as a sequence of inner products of the rows of A with a common vector. Consider the block matrices

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \text{ and } \left[\begin{array}{c|c} E & F \\ \hline G & H \end{array}\right].$$

Assume that the number of columns of A and C equals the number of rows of E and F, and that the number of columns of B and D equals the number of rows of G and H. Describe the product of these two matrices. $\square$

Exercise 2.2 Associated with a matrix A are four vector spaces. What are they, how can you find a basis for each, and how are their dimensions related? Give a "natural" basis for the nullspace of the matrix $[A|I]$, where A is an $m \times n$ matrix and I is an $m \times m$ identity matrix concatenated onto A. $\square$

Exercise 2.3 Suppose V is a set of the form $\{Ax : x \in \mathbf{R}^k\}$, where A is an $n \times k$ matrix. Prove that V is also a set of the form $\{y \in \mathbf{R}^n : By = O\}$ where B is an $\ell \times n$ matrix, and explain how to find an appropriate matrix B. Conversely, suppose V is a set of the form $\{y \in \mathbf{R}^n : By = O\}$, where B is an $\ell \times n$ matrix. Prove that V is also a set of the form $\{Ax : x \in \mathbf{R}^k\}$, where A is an $n \times k$ matrix, and explain how to find an appropriate matrix A. $\square$

Exercise 2.4 Consider a linear system of equations, $Ax = b$. What are the various elementary row operations that can be used to obtain an equivalent system? What does it mean for two systems to be equivalent? $\square$

Exercise 2.5 Consider a linear system of equations, $Ax = b$. Describe the set of all solutions to this system. Explain how to use Gaussian elimination to determine this set. Prove that the system has no solution if and only if there is a vector y such that $y^T A = O^T$ and $y^T b \neq 0$. $\square$

Exercise 2.6 If $x \in \mathbf{R}^n$, what is the definition of $\|x\|_1$? Of $\|x\|_2$? Of $\|x\|_\infty$? For fixed matrix A (not necessarily square) and vector b, explain how to minimize $\|Ax - b\|_2$. Note: From now on in these notes, if no subscript appears in the notation $\|x\|$, then the norm $\|x\|_2$ is meant. □

Exercise 2.7 Consider a square $n \times n$ matrix A. What is the determinant of A? How can it be expressed as a sum with $n!$ terms? How can it be expressed as an expansion by cofactors along an arbitrary row or column? How is it affected by the application of various elementary row operations? How can it be determined by Gaussian elimination? What does it mean for A to be singular? Nonsingular? What can you tell about the determinant of A from the dimensions of each of the four vector spaces associated with A? The determinant of A describes the volume of a certain geometrical object. What is this object? □

Exercise 2.8 Consider a linear system of equations $Ax = b$ where A is square and nonsingular. Describe the set of all solutions to this system. What is Cramer's rule and how can it be used to find the complete set of solutions? □

Exercise 2.9 Consider a square matrix A. When does it have an inverse? How can Gaussian elimination be used to find the inverse? How can Gauss-Jordan elimination be used to find the inverse? Suppose e_j is a vector of all zeroes, except for a 1 in the jth position. What does the solution to $Ax = e_j$ have to do with A^{-1}? What does the solution to $x^T A = e_j^T$ have to do with A^{-1}? Prove that if A is a nonsingular matrix with integer entries and determinant ± 1, then A^{-1} is also a matrix with integer entries. Prove that if A is a nonsingular matrix with integer entries and determinant ± 1, and b is a vector with integer entries, then the solution to $Ax = b$ is an integer vector. □

Exercise 2.10 What is LU factorization? What is QR factorization, Gram-Schmidt orthogonalization, and their relationship? □

Exercise 2.11 What does it mean for a matrix to be orthogonal? Prove that if A is orthogonal and x and y are vectors, then $\|x - y\|_2 = \|Ax - Ay\|_2$; i.e., multiplying two vectors by A does not change the Euclidean distance between them. □

Exercise 2.12 What is the definition of an eigenvector and an eigenvalue of a square matrix? The remainder of the questions in this problem concern matrices over the real numbers, with real eigenvalues and eigenvectors. Find a square matrix with no eigenvalues. Prove that if A is a symmetric $n \times n$ matrix, there exists a basis for $\mathbf{R}^n$ consisting of eigenvectors of A. □

Exercise 2.13 What does it mean for a symmetric matrix A to be positive semi-definite? Positive definite? If A is positive definite, describe the set $\{x : x^T Ax \leq 1\}$. What is the geometrical interpretation of the eigenvectors and eigenvalues of A with respect to this set? $\square$

Exercise 2.14 Suppose E is a finite set of vectors in $\mathbf{R}^n$. Let V be the vector space spanned by the vectors in E. Let $\mathcal{I} = \{S \subseteq E : S \text{ is linearly independent}\}$. Let $\mathcal{C} = \{S \subseteq E : S \text{ is linearly dependent, but no proper subset of } S \text{ is linearly dependent}\}$. Let $\mathcal{B} = \{S \subseteq E : S \text{ is a basis for } V\}$. Prove the following:

1. $\emptyset \in \mathcal{I}$.
2. If $S_1 \in \mathcal{I}$, $S_2 \in \mathcal{I}$, and $\operatorname{card} S_2 > \operatorname{card} S_1$, then there exists an element $e \in S_2 \setminus S_1$ such that $S_1 \cup \{e\} \in \mathcal{I}$.
3. If $S \in \mathcal{I}$ and $S \cup \{e\}$ is dependent, then there is exactly one subset of $S \cup \{e\}$ that is in $\mathcal{C}$.
4. If $S_1 \in \mathcal{B}$ and $S_2 \in \mathcal{B}$, then $\operatorname{card} S_1 = \operatorname{card} S_2$.
5. If $S_1 \in \mathcal{B}$, $S_2 \in \mathcal{B}$, and $e_1 \in S_1$, then there exists an element $e_2 \in S_2$ such that $(S_1 \setminus \{e_1\}) \cup \{e_2\} \in \mathcal{B}$.
6. If $S_1 \in \mathcal{C}$, $S_2 \in \mathcal{C}$, $e \in S_1 \cap S_2$, and $e' \in S_1 \setminus S_2$, then there is a set $S_3 \in \mathcal{C}$ such that $S_3 \subseteq (S_1 \cup S_2) \setminus \{e\}$ and $e' \in S_3$.

$\square$

3 Polytopes

3.1 Convex Combinations and V-Polytopes

Definition 3.1 Let $v^1, \ldots, v^m$ be a finite set of points in $\mathbf{R}^n$ and $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$. Then

$$\sum_{j=1}^{m} \lambda_j v^j$$

is called a *linear combination* of $v^1, \ldots, v^m$. If $\lambda_1, \ldots, \lambda_m \geq 0$ then it is called a *nonnegative* (or *conical*) *combination*. If $\lambda_1 + \cdots + \lambda_m = 1$ then it is called an *affine combination*. If both $\lambda_1, \ldots, \lambda_m \geq 0$ and $\lambda_1 + \cdots + \lambda_m = 1$ then it is called *convex combination*. Note: We will regard an empty linear or nonnegative combination as equal to the point O, but will not consider empty affine or convex combinations.

Exercise 3.2 Give some examples of linear, nonnegative, affine, and convex combinations in $\mathbf{R}^1$, $\mathbf{R}^2$, and $\mathbf{R}^3$. Include diagrams. $\square$

Definition 3.3 The set $\{v^1, \ldots, v^m\} \subset \mathbf{R}^n$ is *linearly independent* if the only solution to $\sum_{j=1}^{m} \lambda_j v^j = O$ is the trivial one: $\lambda_j = 0$ for all j. Otherwise the set is *linearly dependent*.

Exercise 3.4 Prove that the set $S = \{v^1, \ldots, v^m\} \subset \mathbf{R}^n$ is linearly dependent if and only if there exists k such that v^k can be written as a linear combination of the elements of $S \setminus \{v^k\}$. $\square$

Solution: Assume S is linearly dependent. Then there is a nontrivial solution to $\sum_{j=1}^{m} \lambda_j v^j = O$. So there exists k such that $\lambda_k \neq 0$. Then

$$v^k = \sum_{j \neq k} \frac{-\lambda_j}{\lambda_k} v^j.$$

Therefore v^k can be written as a linear combination of the elements of $S \setminus \{v^k\}$.

Conversely, suppose there is a k such that v^k can be written as a linear combination of the elements of $S \setminus \{v^k\}$, say $v^k = \sum_{j \neq k} \lambda_j v^j$. Set $\lambda_k = -1$. Then $\sum_{j=1}^{m} \lambda_j v^j = O$ provides a nontrivial linear combination of the elements of S equaling O. Therefore S is linearly dependent.

Definition 3.5 The set $\{v^1, \ldots, v^m\} \subset \mathbf{R}^n$ is *affinely independent* if the only solution to $\sum_{j=1}^{m} \lambda_j v^j = O$, $\sum_{j=1}^{m} \lambda_j = 0$, is the trivial one: $\lambda_j = 0$ for all j. Otherwise the set is *affinely dependent*.

Exercise 3.6 Prove that the set $S = \{v^1, \ldots, v^m\} \subset \mathbf{R}^n$ is affinely dependent if and only if there exists k such that v^k can be written as an affine combination of the elements of $S \setminus \{v^k\}$. $\square$

Solution: Assume S is affinely dependent. Then there is a nontrivial solution to $\sum_{j=1}^m \lambda_j v^j = O$, $\sum_{j=1}^m \lambda_j = 0$. So there exists k such that $\lambda_k \neq 0$. Then

$$v^k = \sum_{j \neq k} \frac{-\lambda_j}{\lambda_k} v^j.$$

Note that

$$\begin{array}{rcl} \sum_{j \neq k} \frac{-\lambda_j}{\lambda_k} & = & \frac{1}{\lambda_k}(-\sum_{j \neq k} \lambda_j) \\ & = & \frac{1}{\lambda_k}\lambda_k \\ & = & 1. \end{array}$$

Therefore v^k can be written as an affine combination of the elements of $S \setminus \{v^k\}$.

Conversely, suppose there is a k such that v^k can be written as an affine combination of the elements of $S \setminus \{v^k\}$, say $v^k = \sum_{j \neq k} \lambda_j v^j$, where $\sum_{j \neq k} \lambda_j = 1$. Set $\lambda_k = -1$. Then $\sum_{j=1}^m \lambda_j v^j = O$ provides a nontrivial affine combination of the elements of S equaling O, since $\sum_{j=1}^m \lambda_j = 0$.

Exercise 3.7 Prove that the set $\{v^1, \ldots, v^m\}$ is affinely independent if and only if the set $\{v^1 - v^m, v^2 - v^m, \ldots, v^{m-1} - v^m\}$ is linearly independent. $\square$

Solution: Assume that the set $\{v^1, \ldots, v^m\}$ is affinely independent. Assume that $\sum_{j=1}^{m-1} \lambda_j (v^j - v^m) = O$. We need to prove that $\lambda_j = 0$, $j = 1, \ldots, m-1$. Now $\sum_{j=1}^{m-1} \lambda_j v^j - (\sum_{j=1}^{m-1} \lambda_j) v^m = O$. Set $\lambda_m = -\sum_{j=1}^{m-1} \lambda_j$. Then $\sum_{j=1}^m \lambda_j v^j = O$ and $\sum_{j=1}^m \lambda_j = 0$. Because the set $\{v^1, \ldots, v^m\}$ is affinely independent, we deduce $\lambda_j = 0$, $j = 1, \ldots, m$. Therefore the set $\{v^1 - v^m, \ldots, v^{m-1} - v^m\}$ is linearly independent.

Conversely, assume that the set $\{v^1 - v^m, \ldots, v^{m-1} - v^m\}$ is linearly independent. Assume that $\sum_{j=1}^m \lambda_j v^j = O$ and $\sum_{j=1}^m \lambda_j = 0$. We need to prove that $\lambda_j = 0$, $j = 1, \ldots, m$. Now $\lambda_m = -\sum_{j=1}^{m-1} \lambda_j$, so $\sum_{j=1}^{m-1} \lambda_j v^j - (\sum_{j=1}^{m-1} \lambda_j) v^m = O$. This is equivalent to $\sum_{j=1}^{m-1} \lambda_j (v^j - v^m) = O$. Because the set $\{v^1 - v^m, \ldots, v^{m-1} - v^m\}$ is linear independent, we deduce $\lambda_j = 0$, $j = 1, \ldots, m-1$. But $\lambda_m = -\sum_{j=1}^{m-1} \lambda_j$, so $\lambda_m = 0$ as well. Therefore the set $\{v^1, \ldots, v^m\}$ is affinely independent.

Definition 3.8 A subset $S \subseteq \mathbf{R}^n$ is a *subspace* (respectively, *cone, affine set, convex set*) if it is closed under all linear (respectively, nonnegative, affine, convex) combinations of its elements. Note: This implies that subspaces and cones must contain the point O.

Exercise 3.9 Give some examples of subspaces, cones, affine sets, and convex sets in $\mathbf{R}^1$, $\mathbf{R}^2$, and $\mathbf{R}^3$. $\square$

Exercise 3.10 Is the empty set a subspace, a cone, an affine set, a convex set? Is $\mathbf{R}^n$ a subspace, a cone, an affine set, a convex set? $\square$

Exercise 3.11 Prove that a subset $S \subseteq \mathbf{R}^n$ is affine if and only if it is a set of the form $L + x$, where L is a subspace and $x \in \mathbf{R}^n$. $\square$

Exercise 3.12 Are the following sets subspaces, cones, affine sets, convex sets?

1. $\{x \in \mathbf{R}^n : Ax = O\}$ for a given matrix A.

   Solution: This set is a subspace, hence also a cone, an affine set, and a convex set.

2. $\{x \in \mathbf{R}^n : Ax \leq O\}$ for a given matrix A.

   Solution: This is a cone, hence also convex. But it is not necessarily a subspace or an affine set.

3. $\{x \in \mathbf{R}^n : Ax = b\}$ for a given matrix A and vector b.

   Solution: this set is an affine set, hence also convex. But it is not necessarily a subspace or a cone.

4. $\{x \in \mathbf{R}^n : Ax \leq b\}$ for a given matrix A and vector b.

   Solution for this case: This set is convex. By Proposition 3.13 it suffices to show that it is closed under convex combinations of two elements. So assume $Av^1 \leq b$ and $Av^2 \leq b$, and $\lambda_1, \lambda_2 \geq 0$ such that $\lambda_1 + \lambda_2 = 1$. We need to verify that $Av \leq b$, where $v = \lambda_1 v^1 + \lambda_2 v^2$. But $Av = A(\lambda_1 v^1 + \lambda_2 v^2) = \lambda_1 Av^1 + \lambda_2 Av_2$. Knowing that $Av^1 \leq b$, $Av^2 \leq b$ and $\lambda_1, \lambda_2 \geq 0$, we deduce that $\lambda_1 Av^1 \leq \lambda_1 b$ and $\lambda_2 Av^2 \leq \lambda_2 b$. Thus $\lambda_1 Av^1 + \lambda_2 Av_2 \leq \lambda_1 b + \lambda_2 b = (\lambda_1 + \lambda_2)b = b$. Therefore $Av \leq b$ and the set is closed under pairwise convex combinations.

$\square$

Proposition 3.13 *A subset $S \subseteq \mathbf{R}^n$ is a subspace (respectively, a cone, an affine set, a convex set) if and only it is closed under all linear (respectively, nonnegative, affine, convex) combinations of pairs of elements.*

Proof. Exercise. $\Box$

Solution for convex combinations: It is immediate that if S is convex then it is closed under convex combinations of pairs of elements. So, conversely, assume that S is closed under convex combinations of pairs of elements. We will prove by induction on $m \geq 1$ that if $v^1, \ldots, v^m \in S$, $\lambda_1, \ldots, \lambda_m \geq 0$, and $\sum_{j=1}^m \lambda_j = 1$, then $v \in S$, where $v = \sum_{j=1}^m \lambda_j v^j$. If $m = 1$ then $\lambda_1 = 1$ and $v \in S$ trivially. If $m = 2$ then $v \in S$ by the assumption that S is closed under convex combinations of pairs of elements. So assume $m > 2$. First consider the case that $\lambda_m = 1$. Then v equals v^m and so $v \in S$. So now assume that $\lambda_m < 1$. Then

$$\begin{aligned} v &= \left(\sum_{j=1}^{m-1} \lambda_j v^j\right) + \lambda_m v^m \\ &= (1-\lambda_m)\left(\sum_{j=1}^{m-1} \frac{\lambda_j}{1-\lambda_m} v^j\right) + \lambda_m v^m \\ &= (1-\lambda_m) w + \lambda_m v^m, \end{aligned}$$

where

$$w = \sum_{j=1}^{m-1} \frac{\lambda_j}{1-\lambda_m} v^j.$$

Note that $1 - \lambda_m > 0$, and so $\frac{\lambda_j}{1-\lambda_m} \geq 0$, $j = 1, \ldots, m-1$. Also note that $\sum_{j=1}^{m-1} \frac{\lambda_j}{1-\lambda_m} = 1$. So $w \in X$ by the induction hypothesis. Hence $v \in S$ since it is convex combination of the pair w and v^m. Therefore S is closed under convex combinations of sets of m elements.

Proposition 3.14 *The intersection of any collection of subspaces (respectively, cones, affine sets, convex sets) is a subspace (respectively, cone, affine set, convex set).*

Proof. Exercise. $\Box$

Solution (for convex sets): Suppose each S_i, $i \in I$, is convex, and $S = \bigcap_{i \in I} S_i$. Let $v^1, \ldots, v^m \in S$ and $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$ such that $\lambda_1 + \cdots + \lambda_m = 1$ and $\lambda_1, \ldots, \lambda_m \geq 0$. Let $v = \sum_{j=1}^m \lambda_i v^i$. Then $v^1, \ldots, v^m \in S_i$ for all $i \in I$. Since S_i is convex, $v \in S_i$ for all i. Hence $v \in \bigcap_{i \in I} S_i = S$, and so S is convex.

Definition 3.15 Let $V \subseteq \mathbf{R}^n$. Define the *linear span* (respectively, *cone*, *affine span*, *convex hull*) of V, denoted $\operatorname{span} V$ (respectively, $\operatorname{cone} V$, $\operatorname{aff} V$, $\operatorname{conv} V$) to be the intersection of all subspaces (respectively, cones, affine sets, convex sets) containing V,

$$\begin{array}{rcl} \operatorname{span} V & = & \bigcap\{S : V \subseteq S,\ S \text{ is a subspace}\}, \\ \operatorname{cone} V & = & \bigcap\{S : V \subseteq S,\ S \text{ is a cone}\}, \\ \operatorname{aff} V & = & \bigcap\{S : V \subseteq S,\ S \text{ is an affine set}\}, \\ \operatorname{conv} V & = & \bigcap\{S : V \subseteq S,\ S \text{ is a convex set}\}. \end{array}$$

Lemma 3.16 *For all $V \subseteq \mathbf{R}^n$, the set $\operatorname{span} V$ (respectively, $\operatorname{cone} V$, $\operatorname{aff} V$, $\operatorname{conv} V$) is a subspace (respectively, cone, affine set, convex set).*

Proof. Exercise. □

Lemma 3.17 *"Linear/nonnegative/affine/convex combinations of linear/nonnegative/affine/convex combinations are linear/nonnegative/affine/convex combinations."*

Proof. Exercise. □

Solution (for convex combinations): Suppose $w^i = \sum_{j=1}^{m_i} \lambda_{ij} v^{ij}$, $i = 1, \ldots, \ell$, where $\lambda_{ij} \geq 0$ for all i, j, and $\sum_{j=1}^{m_i} \lambda_{ij} = 1$ for all i. Let $w = \sum_{i=1}^{\ell} \mu_i w^i$, where $\mu_i \geq 0$ and $\sum_{i=1}^{\ell} \mu_i = 1$. Then $w = \sum_{i=1}^{\ell} \sum_{j=1}^{m_i} \mu_i \lambda_{ij} v^{ij}$, where $\mu_i \lambda_{ij} \geq 0$ for all i, j, and

$$\begin{array}{rcl} \sum_{i=1}^{\ell} \sum_{j=1}^{m_i} \mu_i \lambda_{ij} & = & \sum_{i=1}^{\ell} \mu_i \sum_{j=1}^{m_i} \lambda_{ij} \\ & = & \sum_{i=1}^{\ell} \mu_i (1) \\ & = & 1. \end{array}$$

Proposition 3.18 *Let $V \subseteq \mathbf{R}^n$. Then $\operatorname{span} V$ (respectively, $\operatorname{cone} V$, $\operatorname{aff} V$, $\operatorname{conv} V$ equals the set of all linear (respectively, nonnegative, affine, convex) combinations of elements of V.*

Proof. Exercise. □

Solution (for convex sets): Let W be the set of all convex combinations of elements of V. By Lemma 3.17, W is closed under convex combinations, and hence is a convex set. Since $V \subseteq W$, we conclude $\operatorname{conv} V \subseteq W$. On the other hand, since $\operatorname{conv} V$ is a convex set containing V, $\operatorname{conv} V$ must contain all convex combinations of elements of V. Hence $W \subseteq \operatorname{conv} V$. Therefore $W = \operatorname{conv} V$.

Lemma 3.19 *Let $v^1, \ldots, v^m \in \mathbf{R}^n$. Let A be the matrix*

$$\begin{bmatrix} v^1 & \cdots & v^m \\ 1 & \cdots & 1 \end{bmatrix}$$

That is to say, A is created by listing the points v^i as columns and then appending a row of 1's. Let $v \in \mathbf{R}^n$. Then v equals the convex combination $\sum_{i=1}^m \lambda_i v^i$ if and only if $\lambda = [\lambda_1, \ldots, \lambda_m]^T$ is a solution of

$$A \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix} = \begin{bmatrix} v \\ 1 \end{bmatrix}$$
$$\lambda_1, \ldots, \lambda_m \geq 0$$

PROOF. Exercise. $\square$

Exercise 3.20 What can you say about the rank of the matrix A in the previous problem? $\square$

Theorem 3.21 (Carathéodory) *Suppose x is a convex combination of $v^1, \ldots, v^m \in \mathbf{R}^n$, where $m > n+1$. Then x is also a convex combination of a subset of $\{v^1, \ldots, v^m\}$ of cardinality at most $n+1$.*

PROOF. Suggestion: Think about the matrix A in the previous two problems. Assume that the columns associated with positive values of λ_i are linearly dependent. What can you do now? $\square$

Solution: Proof by induction on $m \geq n+1$. The statement is trivially true if $m = n+1$. Assume that $m > n+1$ and that x is a convex combination of $v^1, \ldots, v^m$. Define the matrix A as above and let λ^* be a solution to

$$A\lambda = \begin{bmatrix} x \\ 1 \end{bmatrix}$$
$$\lambda \geq O$$

If any $\lambda_k^* = 0$ then we actually have x written as a convex combination of the elements of $\{v^1, \ldots, v^m\} \setminus \{v^k\}$, a set of cardinality $m-1$, so the result is true by the induction hypothesis. Thus we assume that $\lambda_j^* > 0$ for all j. The matrix A has more columns than rows ($m > n+1$) so there is a nontrivial element of the nullspace of A, say, μ^*. So $A\mu^* = O$. That μ^* is nontrivial means that at least one μ_j^* is not zero. The last row of A implies that the sum of the μ_j^* equals 0. From this we conclude that at least one μ_j^* is negative. Now consider $\overline{\lambda} = \lambda^* + t\mu^*$, where t is a nonnegative real number. Start with $t = 0$ and increase t until you reach the first value, t^*, for which some component of $\lambda^* + t\mu^*$ becomes zero. In fact,

$$t^* = \min_{j:\mu_j^* < 0} \left\{ \frac{\lambda_j^*}{-\mu_j^*} \right\}.$$

Hence $\overline{\lambda}$ is a new solution to

$$\begin{aligned} A\lambda &= \begin{bmatrix} x \\ 1 \end{bmatrix} \\ \lambda &\geq O \end{aligned}$$

with fewer positive entries than λ^*. Let's assume $\overline{\lambda}_k = 0$. Then we have x written as a convex combination of the elements of $\{v^1, \ldots, v^m\} \setminus \{v^k\}$, a set of cardinality $m - 1$, so the result is true by the induction hypothesis.

Definition 3.22 A *V-polytope* is the convex hull of a finite collection of points in $\mathbf{R}^n$.

Exercise 3.23 Construct some examples of V-polytopes in $\mathbf{R}$, $\mathbf{R}^2$, and $\mathbf{R}^3$. □

Exercise 3.24 Is the empty set a V-polytope? □

Exercise 3.25 In each of the following cases describe $\operatorname{conv} V$.

1. $V = \{[\pm 1, \pm 1, \pm 1]\} \subset \mathbf{R}^3$.
2. $V = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\} \subset \mathbf{R}^3$.
3. $V = \{[\pm 1, 0, 0], [0, \pm 1, 0], [0, 0, \pm 1]\} \subset \mathbf{R}^3$.
4. $V = \{0, \pm\frac{1}{2}, \pm\frac{2}{3}, \pm\frac{3}{4}, \ldots\} \subset \mathbf{R}$.

□

Theorem 3.26 (Radon) *Let* $V = \{v^1, \ldots, v^m\} \subseteq \mathbf{R}^n$. *If* $m > n + 1$ *then there exists a partition* V_1, V_2 *of* V *such that* $\operatorname{conv} V_1 \cap \operatorname{conv} V_2 \neq \emptyset$. □

Proof. Exercise. □

Solution: Construct the matrix A as in the proof of Theorem 3.21. Because A has more columns than rows ($m > n + 1$) there is a nonzero element μ^* in its nullspace, $A\mu^* = O$. The last row of A then implies that at least one component of μ^* is negative and at least one component is positive. Let $I_\oplus = \{i : \mu_i^* \geq 0\}$ and $I_- = \{i : \mu_i < 0\}$. Note that $I_\oplus$ and I_- are each nonempty. Then

$$\sum_{i \in I_\oplus} \mu_i^* v^i + \sum_{i \in I_-} \mu_i^* v^i = O,$$

$$\sum_{i \in I_\oplus} \mu_i^* + \sum_{i \in I_-} \mu_i^* = 0.$$

So

$$\sum_{i \in I_\oplus} \mu_i^* v^i = \sum_{i \in I_-} (-\mu_i^*) v^i,$$

$$\sum_{i \in I_\oplus} \mu_i^* = \sum_{i \in I_-} (-\mu_i^*).$$

Define $c = \sum_{i \in I_\oplus} \mu_i^* = \sum_{i \in I_-} (-\mu_i^*)$. Note that c is positive. Dividing through by c yields

$$\sum_{i \in I_\oplus} \lambda_i^* v^i = \sum_{i \in I_-} \lambda_i^* v^i,$$

$$\sum_{i \in I_\oplus} \lambda_i^* = \sum_{i \in I_-} \lambda_i^* = 1,$$

$$\lambda \geq O,$$

where $\lambda_i^* = \frac{\mu_i^*}{c}$ if $i \in I_\oplus$, and $\lambda_i^* = \frac{-\mu_i^*}{c}$ if $i \in I_-$. Taking $x = \sum_{i \in I_\oplus} \lambda_i^* v^i = \sum_{i \in I_-} \lambda_i^* v^i$, $V_1 = \{v^i : i \in I_\oplus\}$, and $V_2 = \{v^i : i \in I_-\}$, we see that x is a convex combination of elements of V_1 as well as elements of V_2. So $\operatorname{conv} V_1 \cap \operatorname{conv} V_2 \neq \emptyset$.

Theorem 3.27 (Helly) *Let $\mathcal{V} = \{V_1, \ldots, V_m\}$ be a family of m convex subsets of $\mathbf{R}^n$ with $m \geq n+1$. If every subfamily of $n+1$ sets in $\mathcal{V}$ has a nonempty intersection, then $\cap_{i=1}^m V_i \neq \emptyset$.*
□

Proof. Exercise. □

3.2 Linear Inequalities and H-Polytopes

Definition 3.28 Let $a \in \mathbf{R}^n$ and $b_0 \in \mathbf{R}$. Then $a^T x = b_0$ is called a *linear equation*, and $a^T x \leq b_0$ and $a^T x \geq b_0$ are called *linear inequalities.*

Further, if $a \neq O$, then the set $\{x \in \mathbf{R}^n : a^T x = b_0\}$ is called a *hyperplane*, the sets $\{x \in \mathbf{R}^n : a^T x \leq b_0\}$ and $\{x \in \mathbf{R}^n : a^T x \geq b_0\}$ are called *closed halfspaces*, and the sets $\{x \in \mathbf{R}^n : a^T x < b_0\}$ and $\{x \in \mathbf{R}^n : a^T x > b_0\}$ are called *open halfspaces.*

Exercise 3.29 Why do we require $a \neq O$ in the definitions of hyperplanes and halfspaces?
□

Definition 3.30 A subset $S \subseteq \mathbf{R}^n$ is *bounded* if there exists a number $M \in \mathbf{R}$ such that $\|x\| \leq M$ for all $x \in S$.

Definition 3.31 We can represent systems of a finite collection of linear equations or linear inequalities in matrix form. For example, the system

$$\begin{gathered} a_{11}x_1 + \cdots + a_{1n}x_n \leq b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \leq b_m \end{gathered}$$

can be written compactly as

$$Ax \leq b$$

where

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

and

$$b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

A subset of $\mathbf{R}^n$ of the form $\{x \in \mathbf{R}^n : Ax \leq b\}$ is called an *H-polyhedron*. A bounded H-polyhedron is called an *H-polytope*.

Exercise 3.32 Construct some examples of H-polyhedra and H-polytopes in $\mathbf{R}$, $\mathbf{R}^2$, and $\mathbf{R}^3$. $\square$

Exercise 3.33 Is the empty set an H-polytope? Is $\mathbf{R}^n$ an H-polyhedron? $\square$

Exercise 3.34 Prove that a subset of $\mathbf{R}^n$ described by a finite collection of linear equations and inequalities is an H-polyhedron. $\square$

Exercise 3.35 In each case describe the H-polyhedron $P = \{x : Ax \leq b\}$ where A and b are as given.

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

2.

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

3.

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

4.

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

□

3.3 H-Polytopes are V-Polytopes

Definition 3.36 Suppose $P = \{x \in \mathbf{R}^n : Ax \le b\}$ is an H-polytope and $\overline{x} \in P$. Partition the linear inequalities into two sets: those that $\overline{x}$ satisfies with equality (the *tight* or *binding* inequalities) and those that $\overline{x}$ satisfies with strict inequality (the *slack* or *nonbinding* inequalities):

$$\begin{array}{c} A^1x \le b^1 \text{ where } A^1\overline{x} = b^1, \\ A^2x \le b^2 \text{ where } A^2\overline{x} < b^2. \end{array}$$

Define $N(\overline{x})$ to be the linear space that is the nullspace of the matrix A^1; i.e., the solution space to $A^1x = O$. Even though this is not an official term in the literature, we will call $N(\overline{x})$ the *nullspace* of $\overline{x}$ (with respect to the system defining P). □

Definition 3.37 Let P be an H-polyhedron and $\overline{x} \in P$ such that $\dim N(\overline{x}) = 0$. Then $\overline{x}$ is called a *vertex* of P. $\square$

Lemma 3.38 *No two different vertices of an H-polyhedron have the same set of tight inequalities.* $\square$

Solution: Assume that $\overline{x}$ and x' are each vertices of the $P = \{x \in \mathbf{R}^n : Ax \leq b\}$, and that they have the same associated submatrix A^1 for the set of tight inequalities. Then $A^1\overline{x} = b^1$ and $A^1x' = b^1$, so $A(\overline{x} - x') = O$. Because the nullspace of A^1 is trivial, we conclude $\overline{x} - x' = O$; i.e., $\overline{x} = x'$.

Proposition 3.39 *If P is an H-polyhedron then P has a finite number of vertices.* $\square$

Solution: There is only a finite number of choices for a submatrix A^1.

Lemma 3.40 *Let $P = \{x \in \mathbf{R}^n : Ax \leq b\}$ be an H-polytope and $\overline{x} \in P$ such that $\overline{x}$ is not a vertex. Choose any nonzero $\overline{w} \in N(\overline{x})$ and consider the line $\overline{x} + t\overline{w}$. Then there is a positive value of t for which $\overline{x} + t\overline{w}$ is in P and has more tight constraints than $\overline{x}$. Similarly, there is a negative value of t for which $\overline{x} + t\overline{w}$ is in P and has more tight constraints than $\overline{x}$.* $\square$

Theorem 3.41 (Minkowski) *Every H-polytope P is the convex hull of its vertices. Hence every H-polytope is a V-polytope. Suggestion: Prove this by induction on the number of slack inequalities for a point $\overline{x} \in P$.* $\square$

Solution: Let $\overline{x} \in P$ and assume $\overline{x}$ has k slack inequalities.

First assume $k = 0$. Then $A\overline{x} = b$. Claim: $\overline{x}$ is itself a vertex. If not, then there is a nontrivial $\overline{w} \neq O$ in the nullspace of A, and so $A(\overline{x} + t\overline{w}) = b$ for all real t. This implies P contains a line, contradicting that it is a polytope (and hence bounded). Therefore $\overline{x}$ is a vertex and trivially is a convex combination of vertices; namely, itself.

Now assume that $k > 0$. If $\overline{x}$ is itself a vertex, then we are done as before. So assume that $\overline{x}$ is not a vertex. By Lemma 3.40 there exist points $x^1 = \overline{x} + t_1\overline{w}$ with $t_1 > 0$, and $x^2 = \overline{x} + t_2\overline{w}$ with $t < 0$, each in P but having fewer slack inequalities. By the induction hypothesis, each can be written as a convex combination of vertices of P. But also $\overline{x} = \frac{-t_2}{t_1 - t_2}x^1 + \frac{t_1}{t_1 - t_2}x^2$. So $\overline{x}$ is a convex combination of x^1 and x^2. By Lemma 3.17, $\overline{x}$ is itself a convex combination of vertices of P, and we are done.

Exercise 3.42 Determine the vertices of the polytope in $\mathbf{R}^2$ described by the following inequalities:

$$
\begin{aligned}
x_1 + 2x_2 &\leq 120 \\
x_1 + x_2 &\leq 70 \\
2x_1 + x_2 &\leq 100 \\
x_1 &\geq 0 \\
x_2 &\geq 0
\end{aligned}
$$

$\square$

Exercise 3.43 Determine the vertices of the polytope in $\mathbf{R}^3$ described by the following inequalities:

$$
\begin{aligned}
x_1 + x_2 &\leq 1 \\
x_1 + x_3 &\leq 1 \\
x_2 + x_3 &\leq 1 \\
x_1 &\geq 0 \\
x_2 &\geq 0 \\
x_3 &\geq 0
\end{aligned}
$$

$\square$

Exercise 3.44 Consider the polytope in $\mathbf{R}^9$ described by the following inequalities:

$$
\begin{aligned}
x_{11} + x_{12} + x_{13} &= 1 \\
x_{21} + x_{22} + x_{23} &= 1 \\
x_{31} + x_{32} + x_{33} &= 1 \\
x_{11} + x_{21} + x_{31} &= 1 \\
x_{12} + x_{22} + x_{32} &= 1 \\
x_{13} + x_{23} + x_{33} &= 1 \\
x_{11} &\geq 0 \\
x_{12} &\geq 0 \\
x_{13} &\geq 0 \\
x_{21} &\geq 0 \\
x_{22} &\geq 0 \\
x_{23} &\geq 0 \\
x_{31} &\geq 0 \\
x_{32} &\geq 0 \\
x_{33} &\geq 0
\end{aligned}
$$

Find at least one vertex. $\square$

3.4 V-Polytopes are H-Polytopes

In order to prove the converse of the result in the previous section, i.e., to prove that every V-polytope is an H-polytope, we will need to invoke a procedure called Fourier-Motzkin elimination, which will be discussed later. What we need to know about this procedure for the moment is that whenever we have a polyhedron P described by a system of inequalities in the variables, say, $x_1, \ldots, x_n$, we can eliminate one or more variables of our choosing, say, $x_{k+1}, \ldots, x_n$, to obtain a new system of inequalities describing the projection of P onto the subspace associated with $x_1, \ldots, x_k$.

Theorem 3.45 (Weyl) *If P is a V-polytope, then it is an H-polytope.*

Proof. Assume $P = \operatorname{conv}\{v^1, \ldots, v^m\}$. Consider $P' = \{(r,x) : \sum_{i=1}^m r_i v^i - x = O,\ \sum_{i=1}^m r_i = 1,\ r \geq O\}$. Then $P' = \{(r,x) : \sum_{i=1}^m r_i v^i - x \leq O,\ \sum_{i=1}^m r_i v^i + x \geq O,\ \sum_{i=1}^m r_i \leq 1,\ \sum_{i=1}^m r_i \geq 1,\ r \geq O\}$. Then a description for P in terms of linear inequalities is obtained from that of P' by using Fourier-Motzkin elimination to eliminate the variables $r_1, \ldots, r_m$. Finally, we note that every V-polytope is necessarily a bounded set—we can, for example, bound the norm of any feasible point x in terms of the norms of $v^1, \ldots, v^m$: if $x = \sum_{i=1}^m r_i v^i$ with $\sum_{i=1}^m r_i = 1$ and $r_i \geq 0$ for all $i = 1, \ldots, m$, then

$$\begin{aligned}
\|x\| &= \left\| \sum_{i=1}^m r_i v^i \right\| \\
&\leq \sum_{i=1}^m r_i \|v^i\| \\
&\leq \sum_{i=1}^m \|v^i\|.
\end{aligned}$$

□

Exercise 3.46 Experiment with the online demo of "polymake," www.math.tu-berlin.de/polymake, which can convert between descriptions of polytopes as V-polytopes and as H-polytopes. □

4 Theorems of the Alternatives

4.1 Systems of Equations

Let's start with a system of linear equations:

$$Ax = b.$$

Suppose you wish to determine whether this system is feasible or not. One reasonable approach is to use Gaussian elimination. If the system has a solution, you can find a particular one, $\overline{x}$. (You remember how to do this: Use elementary row operations to put the system in row echelon form, select arbitrary values for the independent variables and use back substitution to solve for the dependent variables.) Once you have a feasible $\overline{x}$ (no matter how you found it), it is straightforward to convince someone else that the system is feasible by verifying that $A\overline{x} = b$.

If the system is infeasible, Gaussian elimination will detect this also. For example, consider the system

$$\begin{array}{c} x_1 + x_2 + x_3 + x_4 = 1 \\ 2x_1 - x_2 + 3x_3 = -1 \\ 8x_1 + 2x_2 + 10x_3 + 4x_4 = 0 \end{array}$$

which in matrix form looks like

$$\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & 0 & -1 \\ 8 & 2 & 10 & 4 & 0 \end{array}\right].$$

Perform elementary row operations to arrive at a system in row echelon form:

$$\left[\begin{array}{rrr} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{array}\right]\left[\begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -8 & 0 & 1 \end{array}\right]\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & 0 & -1 \\ 8 & 2 & 10 & 4 & 0 \end{array}\right] = \left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ 0 & -3 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 & -2 \end{array}\right],$$

which implies

$$\left[\begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -4 & -2 & 1 \end{array}\right]\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & 0 & -1 \\ 8 & 2 & 10 & 4 & 0 \end{array}\right] = \left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ 0 & -3 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 & -2 \end{array}\right].$$

Immediately it is evident that the original system is infeasible, since the resulting equivalent system includes the equation $0x_1 + 0x_2 + 0x_3 + 0x_4 = -2$.

This equation comes from multiplying the matrix form of the original system by the third row of the matrix encoding the row operations: $[-4,-2,1]$. This vector satisfies

$$\begin{bmatrix} -4 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & 0 \\ 8 & 2 & 10 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$\begin{bmatrix} -4 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = -2.$$

In matrix form, we have found a vector $\overline{y}$ such that $\overline{y}^T A = O$ and $\overline{y}^T b \neq 0$. Gaussian elimination will always produce such a vector if the original system is infeasible. Once you have such a $\overline{y}$ (regardless of how you found it), it is easy to convince someone else that the system is infeasible.

Of course, if the system is feasible, then such a vector $\overline{y}$ cannot exist, because otherwise there would also be a feasible $\overline{x}$, and we would have

$$0 = O^T \overline{x} = (\overline{y}^T A)\overline{x} = \overline{y}^T (A\overline{x}) = \overline{y}^T b \neq 0,$$

which is impossible. (Be sure you can justify each equation and inequality in the above chain.) We have established our first Theorem of the Alternatives:

Theorem 4.1 *Either the system*

$$(I) \quad Ax = b$$

has a solution, or the system

$$(II) \quad \begin{array}{c} y^T A = O^T \\ y^T b \neq 0 \end{array}$$

has a solution, but not both.

As a consequence of this theorem, the following question has a "good characterization": Is the system (I) feasible? I will not give an exact definition of this concept, but roughly speaking it means that whether the answer is yes or no, there exists a "short" proof. In this case, if the answer is yes, we can prove it by exhibiting any particular solution to (I). And if the answer is no, we can prove it by exhibiting any particular solution to (II).

Geometrically, this theorem states that precisely one of the alternatives occurs:

1. The vector b is in the column space of A.

2. There is a vector y orthogonal to each column of A (and hence to the entire column space of A) but not orthogonal to b.

4.2 Fourier-Motzkin Elimination — A Starting Example

Now let us suppose we are given a system of linear inequalities

$$Ax \le b$$

and we wish to determine whether or not the system is feasible. If it is feasible, we want to find a particular feasible vector $\overline{x}$; if it is not feasible, we want hard evidence!

It turns out that there is a kind of analog to Gaussian elimination that works for systems of linear inequalities: Fourier-Motzkin elimination. We will first illustrate this with an example:

$$(I) \quad \begin{array}{c} x_1 - 2x_2 \le -2 \\ x_1 + x_2 \le 3 \\ x_1 \le 2 \\ -2x_1 + x_2 \le 0 \\ -x_1 \le -1 \\ 8x_2 \le 15 \end{array}$$

Our goal is to derive a second system (II) of linear inequalities with the following properties:

1. It has one fewer variable.
2. It is feasible if and only if the original system (I) is feasible.
3. A feasible solution to (I) can be derived from a feasible solution to (II).

(Do you see why Gaussian elimination does the same thing for systems of linear equations?) Here is how it works. Let's eliminate the variable x_1. Partition the inequalities in (I) into three groups, (I_-), (I_+), and (I_0), according as the coefficient of x_1 is negative, positive, or zero, respectively.

$$(I_-) \ \begin{array}{c} -2x_1 + x_2 \le 0 \\ -x_1 \le -1 \end{array} \qquad (I_+) \ \begin{array}{c} x_1 - 2x_2 \le -2 \\ x_1 + x_2 \le 3 \\ x_1 \le 2 \end{array} \qquad (I_0) \ 8x_2 \le 15$$

For each pair of inequalities, one from (I_-) and one from (I_+), multiply by positive numbers and add to eliminate x_1. For example, using the first inequality in each group,

$$\begin{array}{c} (\frac{1}{2})(-2x_1 + x_2 \le 0) \\ +(1)(x_1 - 2x_2 \le -2) \\ \hline -\frac{3}{2}x_2 \le -2 \end{array}$$

System (II) results from doing this for all such pairs, and then also including the inequalities in (I_0):

$$
(II) \quad
\begin{array}{c}
-\frac{3}{2}x_2 \leq -2 \\
\frac{3}{2}x_2 \leq 3 \\
\frac{1}{2}x_2 \leq 2 \\
-2x_2 \leq -3 \\
x_2 \leq 2 \\
0x_2 \leq 1 \\
8x_2 \leq 15
\end{array}
$$

The derivation of (II) from (I) can also be represented in matrix form. Here is the original system:

$$
\left[\begin{array}{rr|r}
1 & -2 & -2 \\
1 & 1 & 3 \\
1 & 0 & 2 \\
-2 & 1 & 0 \\
-1 & 0 & -1 \\
0 & 8 & 15
\end{array}\right]
$$

Obtain the new system by multiplying on the left by the matrix that constructs the desired nonnegative combinations of the original inequalities:

$$
\left[\begin{array}{cccccc}
1 & 0 & 0 & 1/2 & 0 & 0 \\
0 & 1 & 0 & 1/2 & 0 & 0 \\
0 & 0 & 1 & 1/2 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]
\left[\begin{array}{rr|r}
1 & -2 & -2 \\
1 & 1 & 3 \\
1 & 0 & 2 \\
-2 & 1 & 0 \\
-1 & 0 & -1 \\
0 & 8 & 15
\end{array}\right]
$$

$$
= \left[\begin{array}{cc|c}
0 & -3/2 & -2 \\
0 & 3/2 & 3 \\
0 & 1/2 & 2 \\
0 & -2 & -3 \\
0 & 1 & 2 \\
0 & 0 & 1 \\
0 & 8 & 15
\end{array}\right].
$$

To see why the new system has the desired properties, let's break down this process a bit. First scale each inequality in the first two groups by positive numbers so that each coefficient of x_1 in (I_-) is -1 and each coefficient of x_1 in (I_+) is $+1$.

$$(I_-)\ \begin{array}{c} -x_1 + \frac{1}{2}x_2 \le 0 \\ -x_1 \le -1 \end{array} \quad (I_+)\ \begin{array}{c} x_1 - 2x_2 \le -2 \\ x_1 + x_2 \le 3 \\ x_1 \le 2 \end{array} \quad (I_0)\ 8x_2 \le 15$$

Isolate the variable x_1 in each of the inequalities in the first two groups.

$$(I_-)\ \begin{array}{c} \frac{1}{2}x_2 \le x_1 \\ 1 \le x_1 \end{array} \quad (I_+)\ \begin{array}{c} x_1 \le 2x_2 - 2 \\ x_1 \le -x_2 + 3 \\ x_1 \le 2 \end{array} \quad (I_0)\ 8x_2 \le 15$$

For each pair of inequalities, one from (I_-) and one from (I_+), create a new inequality by "sandwiching" and then eliminating x_1. Keep the inequalities in (I_0).

$$(IIa)\quad \begin{array}{ccc} \left\{ \begin{array}{c} \frac{1}{2}x_2 \\ 1 \end{array} \right\} & \le x_1 \le & \left\{ \begin{array}{c} 2x_2 - 2 \\ -x_2 + 3 \\ 2 \end{array} \right\} \\ 8x_2 & \le & 15 \end{array} \quad \longrightarrow \quad (IIb)\quad \begin{array}{c} \frac{1}{2}x_2 \le x_1 \le 2x_2 - 2 \\ \frac{1}{2}x_2 \le x_1 \le -x_2 + 3 \\ \frac{1}{2}x_2 \le x_1 \le 2 \\ 1 \le x_1 \le 2x_2 - 2 \\ 1 \le x_1 \le -x_2 + 3 \\ 1 \le x_1 \le 2 \\ 8x_2 \le 15 \end{array}$$

$$\longrightarrow \quad (IIc)\quad \begin{array}{c} \frac{1}{2}x_2 \le 2x_2 - 2 \\ \frac{1}{2}x_2 \le -x_2 + 3 \\ \frac{1}{2}x_2 \le 2 \\ 1 \le 2x_2 - 2 \\ 1 \le -x_2 + 3 \\ 1 \le 2 \\ 8x_2 \le 15 \end{array} \quad \longrightarrow \quad (II)\quad \begin{array}{c} -\frac{3}{2}x_2 \le -2 \\ \frac{3}{2}x_2 \le 3 \\ \frac{1}{2}x_2 \le 2 \\ -2x_2 \le -3 \\ x_2 \le 2 \\ 0x_2 \le 1 \\ 8x_2 \le 15 \end{array}$$

Observe that the system (II) does not involve the variable x_1. It is also immediate that if (I) is feasible, then (II) is also feasible. For the reverse direction, suppose that (II) is feasible. Set the variables (in this case, x_2) equal to any specific feasible values (in this case we choose a feasible value $\overline{x}_2$). From the way the inequalities in (II) were derived, it is evident that

$$\max \left\{ \begin{array}{c} \frac{1}{2}\overline{x}_2 \\ 1 \end{array} \right\} \le \min \left\{ \begin{array}{c} 2\overline{x}_2 - 2 \\ -\overline{x}_2 + 3 \\ 2 \end{array} \right\}.$$

So there exists a specific value $\overline{x}_1$ of x_1 such that

$$\begin{array}{ccc} \left\{ \begin{array}{c} \frac{1}{2}\overline{x}_2 \\ 1 \end{array} \right\} & \le \overline{x}_1 \le & \left\{ \begin{array}{c} 2\overline{x}_2 - 2 \\ -\overline{x}_2 + 3 \\ 2 \end{array} \right\} \\ 8\overline{x}_2 & \le & 15 \end{array}$$

We will then have a feasible solution to (I).

4.3 Showing our Example is Feasible

From this example, we now see how to eliminate one variable (but at the possible considerable expense of increasing the number of inequalities). If we have a solution to the new system, we can determine a value of the eliminated variable to obtain a solution of the original system. If the new system is infeasible, then so is the original system.

From this we can tackle any system of inequalities: Eliminate all of the variables one by one until a system with no variables remains! Then work backwards to determine feasible values of all of the variables.

In our previous example, we can now eliminate x_2 from system (II):

$$\left[\begin{array}{ccccccc} 2/3 & 2/3 & 0 & 0 & 0 & 0 & 0 \\ 2/3 & 0 & 2 & 0 & 0 & 0 & 0 \\ 2/3 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2/3 & 0 & 0 & 0 & 0 & 0 & 1/8 \\ 0 & 2/3 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 1/8 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array} \right] \left[\begin{array}{cc|c} 0 & -3/2 & -2 \\ 0 & 3/2 & 3 \\ 0 & 1/2 & 2 \\ 0 & -2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 8 & 15 \end{array} \right]$$

$$= \left[\begin{array}{cc|c} 0 & 0 & 2/3 \\ 0 & 0 & 8/3 \\ 0 & 0 & 2/3 \\ 0 & 0 & 13/24 \\ 0 & 0 & 1/2 \\ 0 & 0 & 5/2 \\ 0 & 0 & 1/2 \\ 0 & 0 & 3/8 \\ 0 & 0 & 1 \end{array} \right].$$

Each final inequality, such as $0x_1 + 0x_2 \leq 2/3$, is feasible, since the left-hand side is zero and the right-hand side is nonnegative. Therefore the original system is feasible. To find one specific feasible solution, rewrite (II) as

$$\{4/3, 3/2\} \leq x_2 \leq \{2, 4, 15/8\}\,.$$

We can choose, for example, $\overline{x}_2 = 3/2$. Substituting into (I) (or (IIa)), we require

$$\{3/4, 1\} \leq x_1 \leq \{1, 3/2, 2\}\,.$$

So we could choose $\overline{x}_1 = 1$, and we have a feasible solution $(1, 3/2)$ to (I).

4.4 An Example of an Infeasible System

Now let's look at the system:

$$(I) \qquad \begin{array}{c} x_1 - 2x_2 \leq -2 \\ x_1 + x_2 \leq 3 \\ x_1 \leq 2 \\ -2x_1 + x_2 \leq 0 \\ -x_1 \leq -1 \\ 8x_2 \leq 11 \end{array}$$

Multiplying by the appropriate nonnegative matrices to successively eliminate x_1 and x_2, we compute:

$$\begin{bmatrix} 1 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{rr|r} 1 & -2 & -2 \\ 1 & 1 & 3 \\ 1 & 0 & 2 \\ -2 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 8 & 11 \end{array}\right]$$

$$= \left[\begin{array}{cc|c} 0 & -3/2 & -2 \\ 0 & 3/2 & 3 \\ 0 & 1/2 & 2 \\ 0 & -2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 8 & 11 \end{array}\right] \quad (II)$$

and

$$\begin{bmatrix} 2/3 & 2/3 & 0 & 0 & 0 & 0 & 0 \\ 2/3 & 0 & 2 & 0 & 0 & 0 & 0 \\ 2/3 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2/3 & 0 & 0 & 0 & 0 & 0 & 1/8 \\ 0 & 2/3 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 1/8 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \left[\begin{array}{cc|c} 0 & -3/2 & -2 \\ 0 & 3/2 & 3 \\ 0 & 1/2 & 2 \\ 0 & -2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 8 & 11 \end{array}\right]$$

$$= \left[\begin{array}{cc|c} 0 & 0 & 2/3 \\ 0 & 0 & 8/3 \\ 0 & 0 & 2/3 \\ 0 & 0 & 1/24 \\ 0 & 0 & 1/2 \\ 0 & 0 & 5/2 \\ 0 & 0 & 1/2 \\ 0 & 0 & -1/8 \\ 0 & 0 & 1 \end{array}\right] \quad (III)$$

Since one inequality is $0x_1+0x_2 \leq -1/8$, the final system (III) is clearly infeasible. Therefore the original system (I) is also infeasible. We can go directly from (I) to (III) by collecting together the two nonnegative multiplier matrices:

$$\begin{bmatrix} 2/3 & 2/3 & 0 & 0 & 0 & 0 & 0 \\ 2/3 & 0 & 2 & 0 & 0 & 0 & 0 \\ 2/3 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2/3 & 0 & 0 & 0 & 0 & 0 & 1/8 \\ 0 & 2/3 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 1/8 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2/3 & 2/3 & 0 & 2/3 & 0 & 0 \\ 2/3 & 0 & 2 & 4/3 & 0 & 0 \\ 2/3 & 1 & 0 & 1/3 & 1 & 0 \\ 2/3 & 0 & 0 & 1/3 & 0 & 1/8 \\ 1/2 & 2/3 & 0 & 1/3 & 1/2 & 0 \\ 1/2 & 0 & 2 & 1 & 1/2 & 0 \\ 1/2 & 1 & 0 & 0 & 3/2 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 1/8 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} = M.$$

You can check that $M(I) = (III)$. Since M is a product of nonnegative matrices, it will itself be nonnegative. Since the infeasibility is discovered in the eighth inequality of (III), this comes from the eighth row of M, namely, $[1/2, 0, 0, 0, 1/2, 1/8]$. You can now demonstrate directly to anyone that (I) is infeasible using these nonnegative multipliers:

$$\begin{array}{c} (\frac{1}{2})(x_1 - 2x_2 \le -2) \\ +(\frac{1}{2})(-x_1 \le -1) \\ +(\frac{1}{8})(8x_2 \le 11) \\ \hline 0x_1 + 0x_2 \le -\frac{1}{8} \end{array}$$

In particular, we have found a nonnegative vector y such that $y^T A = O^T$ but $y^T b < 0$.

4.5 Fourier-Motzkin Elimination in General

Often I find that it is easier to understand a general procedure, proof, or theorem from a few good examples. Let's see if this is the case for you.

We begin with a system of linear inequalities

$$(I) \quad \sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = 1, \ldots, m.$$

Let's write this in matrix form as

$$Ax \le b$$

or

$$A^i x \le b_i, \quad i = 1, \ldots, m$$

where A^i represents the ith row of A.

Suppose we wish to eliminate the variable x_k. Define

$$\begin{array}{c} I_- = \{i : a_{ik} < 0\} \\ I_+ = \{i : a_{ik} > 0\} \\ I_0 = \{i : a_{ik} = 0\} \end{array}$$

For each $(p,q) \in I_- \times I_+$, construct the inequality

$$-\frac{1}{a_{pk}}(A^p x \le b_p) + \frac{1}{a_{qk}}(A^q x \le b_q).$$

By this I mean the inequality

$$\left(-\frac{1}{a_{pk}}A^p + \frac{1}{a_{qk}}A^q\right) x \le -\frac{1}{a_{pk}}b_p + \frac{1}{a_{qk}}b_q. \tag{1}$$

System (II) consists of all such inequalities, together with the original inequalities indexed by the set I_0.

It is clear that if we have a solution $(\overline{x}_1, \ldots, \overline{x}_n)$ to (I), then $(\overline{x}_1, \ldots, \overline{x}_{k-1}, \overline{x}_{k+1}, \ldots, \overline{x}_n)$ satisfies (II). Now suppose we have a solution $(\overline{x}_1, \ldots, \overline{x}_{k-1}, \overline{x}_{k+1}, \ldots, \overline{x}_n)$ to (II). Inequality (1) is equivalent to

$$\frac{1}{a_{pk}}(b_p - \sum_{j\ne k} a_{pj}x_j) \le \frac{1}{a_{qk}}(b_q - \sum_{j\ne k} a_{qj}x_j).$$

As this is satisfied by $(\overline{x}_1, \ldots, \overline{x}_{k-1}, \overline{x}_{k+1}, \ldots, \overline{x}_n)$ for all $(p,q) \in I_- \times I_+$, we conclude that

$$\max_{p\in I_-}\left\{\frac{1}{a_{pk}}(b_p - \sum_{j\ne k} a_{pj}\overline{x}_j)\right\} \le \min_{q\in I_+}\left\{\frac{1}{a_{qk}}(b_q - \sum_{j\ne k} a_{qj}\overline{x}_j)\right\}.$$

Choose $\overline{x}_k$ to be any value between these maximum and minimum values (inclusive). Then for all $(p,q) \in I_- \times I_+$,

$$\frac{1}{a_{pk}}(b_p - \sum_{j\ne k} a_{pj}\overline{x}_j) \le \overline{x}_k \le \frac{1}{a_{qk}}(b_q - \sum_{j\ne k} a_{qj}\overline{x}_j).$$

Now it is not hard to see that $(\overline{x}_1, \ldots, \overline{x}_{k-1}, \overline{x}_k, \overline{x}_{k+1}, \ldots, \overline{x}_n)$ satisfies all the inequalities in (I). Therefore (I) is feasible if and only if (II) is feasible.

Observe that each inequality in (II) is a nonnegative combination of inequalities in (I), so there is a nonnegative matrix M_k such that (II) is expressible as $M_k(Ax \le b)$. If we start with a system $Ax \le b$ and eliminate all variables sequentially via nonnegative matrices $M_1, \ldots, M_n$, then we will arrive at a system of inequalities of the form $0 \le b'_i$, $i = 1, \ldots, m'$. This system is expressible as $M(Ax \le b)$, where $M = M_n \cdots M_1$. If no b'_i is negative, then the final system is feasible and we can work backwards to obtain a feasible solution to the original system. If b'_i is negative for some i, then let $\overline{y}^T = M^i$ (the ith row of M), and we have a nonnegative vector $\overline{y}$ such that $\overline{y}^T A = O^T$ and $\overline{y}^T b < 0$.

This establishes a Theorem of the Alternatives for linear inequalities:

Theorem 4.2 *Either the system*

$$(I) \quad Ax \leq b$$

has a solution, or the system

$$(II) \quad \begin{array}{c} y^T A = O^T \\ y^T b < 0 \\ y \geq O \end{array}$$

has a solution, but not both.

Note that the "not both" part is the easiest to verify. Otherwise, we would have a feasible $\overline{x}$ and $\overline{y}$ satisfying

$$0 = O^T \overline{x} = (\overline{y}^T A)\overline{x} = \overline{y}^T (A\overline{x}) \leq \overline{y}^T b < 0,$$

which is impossible.

As a consequence of this theorem, we have a good characterization for the question: Is the system (I) feasible? If the answer is yes, we can prove it by exhibiting any particular solution to (I). If the answer is no, we can prove it by exhibiting any particular solution to (II).

4.6 More Alternatives

There are many Theorems of the Alternatives, and we shall encounter more later. Most of the others can be derived from the one of the previous section and each other. For example,

Theorem 4.3 *Either the system*

$$(I) \quad \begin{array}{c} Ax \leq b \\ x \geq O \end{array}$$

has a solution, or the system

$$(II) \quad \begin{array}{c} y^T A \geq O^T \\ y^T b < 0 \\ y \geq O \end{array}$$

has a solution, but not both.

PROOF. System (I) is feasible if and only if the following system is feasible:

$$(I') \quad \begin{bmatrix} A \\ -I \end{bmatrix} x \leq \begin{bmatrix} b \\ O \end{bmatrix}$$

System (II) is feasible if and only if the following system is feasible:

$$(II') \quad \begin{array}{c} \begin{bmatrix} y^T & w^T \end{bmatrix} \begin{bmatrix} A \\ -I \end{bmatrix} = O^T \\ \begin{bmatrix} y^T & w^T \end{bmatrix} \begin{bmatrix} b \\ O \end{bmatrix} < 0 \\ \begin{bmatrix} y^T & w^T \end{bmatrix} \geq \begin{bmatrix} O^T & O^T \end{bmatrix} \end{array}$$

Equivalently,

$$\begin{array}{c} y^T A - w^T = O^T \\ y^T b < O \\ y, w \geq O \end{array}$$

Now apply Theorem 4.2 to the pair (I'), (II'). □

4.7 Exercises: Systems of Linear Inequalities

Exercise 4.4 Discuss the consequences of having one or more of I_-, I_+, or I_0 being empty during the process of Fourier-Motzkin elimination. Does this create any problems? □

Exercise 4.5 Fourier-Motzkin elimination shows how we can start with a system of linear inequalities with n variables and obtain a system with $n-1$ variables. Explain why the set of all feasible solutions of the second system is a projection of the set of all feasible solutions of the first system. Consider a few examples where $n = 3$ and explain how you can classify the inequalities into types I_-, I_+, and I_0 geometrically (think about eliminating the third coordinate). Explain geometrically where the new inequalities in the second system are coming from. □

Exercise 4.6 Consider a given system of linear constraints. A subset of these constraints is called *irredundant* if it describes the same feasible region as the given system and no constraint can be dropped from this subset without increasing the set of feasible solutions.

Find an example of a system $Ax \leq b$ with three variables such that when x_3, say, is eliminated, the resulting system has a larger irredundant subset than the original system. That is to say, the feasible set of the resulting system requires more inequalities to describe than the feasible set of the original system. Hint: Think geometrically. Can you find such an example where the original system has two variables? □

Exercise 4.7 Use Fourier-Motzkin elimination to graph the set of solutions to the following system:

$$\begin{aligned}
+x_1 + x_2 + x_3 &\le 1\\
+x_1 + x_2 - x_3 &\le 1\\
+x_1 - x_2 + x_3 &\le 1\\
+x_1 - x_2 - x_3 &\le 1\\
-x_1 + x_2 + x_3 &\le 1\\
-x_1 + x_2 - x_3 &\le 1\\
-x_1 - x_2 + x_3 &\le 1\\
-x_1 - x_2 - x_3 &\le 1
\end{aligned}$$

What is this geometrical object called? □

Exercise 4.8 Prove the following Theorem of the Alternatives: Either the system

$$Ax \ge b$$

has a solution, or the system

$$\begin{aligned}
y^T A &= O^T\\
y^T b &> 0\\
y &\ge O
\end{aligned}$$

has a solution, but not both. □

Exercise 4.9 Prove the following Theorem of the Alternatives: Either the system

$$\begin{aligned}
Ax &\ge b\\
x &\ge O
\end{aligned}$$

has a solution, or the system

$$\begin{aligned}
y^T A &\le O^T\\
y^T b &> 0\\
y &\ge O
\end{aligned}$$

has a solution, but not both. □

Exercise 4.10 Prove or disprove: The system

$$(I) \quad Ax = b$$

has a solution if and only if each of the following systems has a solution:

$$(I') \quad Ax \le b \qquad (I'') \quad Ax \ge b$$

□

Exercise 4.11 (The Farkas Lemma). Derive and prove a Theorem of the Alternatives for the following system:

$$\begin{aligned} Ax &= b \\ x &\geq O \end{aligned}$$

Give a geometric interpretation of this theorem when A has two rows. When A has three rows. □

Exercise 4.12 Give geometric interpretations to other Theorems of the Alternatives that we have discussed. □

Exercise 4.13 Derive and prove a Theorem of the Alternatives for the system

$$\begin{gathered} \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i \in I_1 \\ \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i \in I_2 \\ x_j \geq 0, \quad j \in J_1 \\ x_j \text{ unrestricted}, \quad j \in J_2 \end{gathered}$$

where (I_1, I_2) is a partition of $\{1, \ldots, m\}$ and (J_1, J_2) is a partition of $\{1, \ldots, n\}$. □

Exercise 4.14 Derive and prove a Theorem of the Alternatives for the system

$$Ax < b.$$

□

5 More on Vertices

Proposition 5.1 *Suppose $P = \{x \in \mathbf{R}^n : Ax \le b\}$ is a polyhedron and $\overline{x} \in P$. Then $\overline{x}$ is a vertex of P if and only if the set of normal vectors of the binding inequalities for $\overline{x}$ spans $\mathbf{R}^n$ (i.e., $\operatorname{span}\{a^i : {a^i}^T\overline{x} = b_i\} = \mathbf{R}^n$, where a^i denotes row i of A).*

Solution: The normal vectors of the inequalities are precisely the rows of A. Let A^1 be the submatrix corresponding to the binding inequalities for $\overline{x}$. Note that A^1 has n columns. Then $\overline{x}$ is a vertex iff the nullspace of A^1 has dimension 0 iff the columns of A^1 are linearly independent iff the column rank of A^1 equals n iff the row rank of A^1 equals n iff the rows of A span $\mathbf{R}^n$.

Definition 5.2 Suppose $S \subseteq \mathbf{R}^n$, $\overline{x} \in S$, and there exists $c \in \mathbf{R}^n$ such that $\overline{x}$ is the unique maximizer of the linear function c^Tx over S. Then $\overline{x}$ is called an *exposed point* of of the set S.

Proposition 5.3 *Every vertex of a polyhedron P is an exposed point.*

Proof. Exercise. □

Solution: Suppose $P = \{x \in \mathbf{R}^n : Ax \le b\}$, and $\overline{x}$ is a vertex of P. Let A^1 be the submatrix of A corresponding to the tight inequalities for $\overline{x}$ and b^1 be the corresponding right-hand sides. Define $c^T = e^TA^1$, where e is a vector of 1's; i.e., c^T is obtained by adding the rows of A^1 together. Then $c^T\overline{x} = (e^TA^1)\overline{x} = e^T(A^1\overline{x}) = e^Tb^1$. Now suppose x' is any point in P. Then $c^Tx' = (e^TA^1)x' = e^T(A^1x') \le e^Tb^1$ with equality if and only if $A^1x' = b^1$. But this implies that $x' = \overline{x}$, since $\overline{x}$ is the unique solution to this system.

Lemma 5.4 *Suppose P is a polyhedron, $c \in \mathbf{R}^n$, and $\overline{x} \in P$ maximizes c^Tx over P. Assume that*

$$\overline{x} = \sum_{j=1}^{m} \lambda_j v^j,$$

where $v^1, \ldots, v^m$ are vertices of P and $\lambda_1, \ldots, \lambda_m$ are all strictly positive positive numbers summing to 1. Then v^j maximizes c^Tx over P, $j = 1, \ldots, m$.

Proof. Exercise. □

Solution: Let $M = c^T\overline{x}$. Then $c^Tv^j \le M$, $j = 1, \ldots, m$, and so $\lambda_jc^Tv^j \le \lambda_jM$. Hence $\sum_{j=1}^m \lambda_jc^Tv^j \le \sum_{j=1}^m \lambda_jM = M$ with equality if and only if $c^Tv^j = M$ for all $j = 1, \ldots, m$.

Hence

$$
\begin{aligned}
M &= c^T\overline{x} \\
&= \sum_{j=1}^{n} \lambda_j c^T v^j \\
&\leq \sum_{j=1}^{n} \lambda_j M \\
&= M
\end{aligned}
$$

forces $c^T v^j = M$, $j = 1, \ldots, m$.

Proposition 5.5 *Every exposed point of a polytope is a vertex.*

Proof. Exercise. □

Solution: Let $\overline{x}$ be an exposed point and $c^T x$ be a linear function having $\overline{x}$ as its unique maximizer. Write $\overline{x}$ as a convex combination of vertices of P, $\overline{x} = \sum_{j=1}^{m} \lambda_j v^j$. Discarding v^j for which $\lambda_j = 0$, if necessary, we may assume that $\lambda_j > 0$, $j = 1, \ldots, m$. By Lemma 5.4, $c^T x$ is also maximized at v^j, $j = 1, \ldots, m$. Since $\overline{x}$ is the unique maximizer, we conclude that $v^j = \overline{x}$, $j = 1, \ldots, m$; in particular, $\overline{x}$ is a vertex.

Definition 5.6 Suppose $S \subseteq \mathbf{R}^n$ and $\overline{x} \in S$ such that $\overline{x} \notin \operatorname{conv}(S \setminus \{\overline{x}\})$. Then $\overline{x}$ is called an *extreme point* of the set S.

Proposition 5.7 *Let $\overline{x}$ be a point in a polytope P. Then $\overline{x}$ is a vertex if and only if it is an extreme point.*

Proof. Exercise. □

Exercise 5.8 Give an example of a convex set S and a point $\overline{x} \in S$ such that $\overline{x}$ is an extreme point but not an exposed point. □

Proposition 5.9 *Let V be a finite set and $P = \operatorname{conv} V$. Then every vertex of P is in V.*

Proof. Exercise □

6 Introduction to Linear Programming

6.1 Example

Consider a hypothetical company that manufactures gadgets and gewgaws.

1. One kilogram of gadgets requires 1 hour of labor, 1 unit of wood, 2 units of metal, and yields a net profit of 5 dollars.

2. One kilogram of gewgaws requires 2 hours of labor, 1 unit of wood, 1 unit of metal, and yields a net profit of 4 dollars.

3. Available are 120 hours of labor, 70 units of wood, and 100 units of metal.

What is the company's optimal production mix? We can formulate this problem as the linear program

$$\begin{aligned} \max z &= 5x_1 + 4x_2 \\ \text{s.t. } x_1 + 2x_2 &\le 120 \\ x_1 + x_2 &\le 70 \\ 2x_1 + x_2 &\le 100 \\ x_1, x_2 &\ge 0 \end{aligned}$$

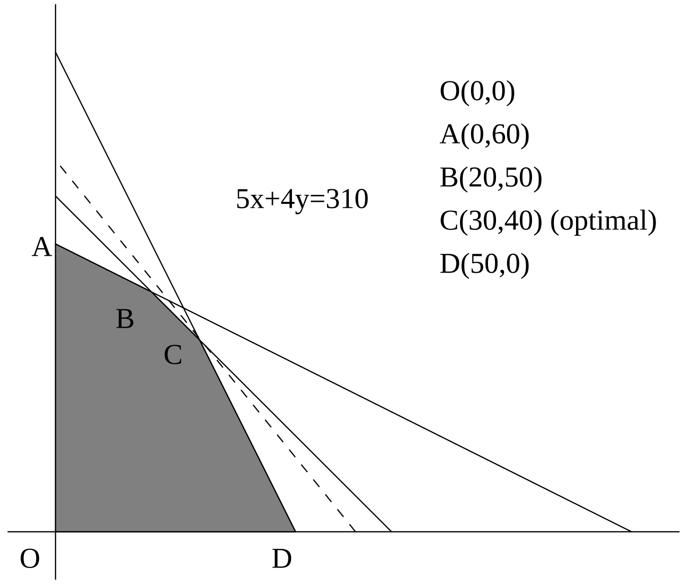
In matrix notation, this becomes

$$\begin{aligned} &\max \begin{bmatrix} 5 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ \text{s.t. } &\begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \le \begin{bmatrix} 120 \\ 70 \\ 100 \end{bmatrix} \\ &\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \ge \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

which is a problem of the form

$$\begin{aligned} &\max c^T x \\ \text{s.t. } &Ax \le b \\ &x \ge O \end{aligned}$$

We can determine the solution of this problem geometrically. Graph the set of all points that satisfy the constraints. Draw some lines for which the objective function assumes a constant value (note that these are all parallel). Find the line with the highest value of z that has nonempty intersection with the set of feasible points. In this case the optimal solution is $(30, 40)$ with optimal value 310.



6.2 Definitions

A *linear function* is a function of the form $a_1x_1 + \cdots + a_nx_n$, where $a_1, \ldots, a_n \in \mathbf{R}$. A *linear equation* is an equation of the form $a_1x_1 + \cdots + a_nx_n = \beta$, where $a_1, \ldots, a_n, \beta \in \mathbf{R}$. If there exists at least one nonzero a_j, then the set of solutions to a linear equation is called a *hyperplane*. A *linear inequality* is an inequality of the form $a_1x_1 + \cdots + a_nx_n \leq \beta$ or $a_1x_1 + \cdots + a_nx_n \geq \beta$, where $a_1, \ldots, a_n, \beta \in \mathbf{R}$. If there exists at least one nonzero a_j, then the set of solutions to a linear inequality is called a *halfspace*. A *linear constraint* is a linear equation or linear inequality.

A *linear programming problem* is a problem in which a linear function is to be maximized (or minimized), subject to a finite number of linear constraints. A *feasible solution* or *feasible point* is a point that satisfies all of the constraints. If such a point exists, the problem is *feasible*; otherwise, it is *infeasible* . The set of all feasible points is called the *feasible region* or *feasible set*. The *objective function* is the linear function to be optimized. An *optimal solution* or *optimal point* is a feasible point for which the objective function is optimized. The value of the objective function at an optimal point is the *optimal value* of the linear program. In the case of a maximization (minimization) problem, if arbitrarily large (small) values of the objective function can be achieved, then the linear program is said to be *unbounded*. More precisely, the maximization (minimization) problem is unbounded if for all $M \in \mathbf{R}$ there exists a feasible point x with objective function value greater than (less than) M. Note: It is possible to have a linear program that has bounded objective function value but unbounded feasible region, so don't let this confusing terminology confuse you. Also note that an infeasible linear program has a bounded feasible region.

Exercise 6.1 Graphically construct some examples of each of the following types of two-variable linear programs:

1. Infeasible.
2. With a unique optimal solution.
3. With more than one optimal solution.
4. Feasible with bounded feasible region.
5. Feasible and bounded but with unbounded feasible region.
6. Unbounded.

□

A linear program of the form

$$\begin{aligned} \max \sum_{j=1}^{n} c_j x_j & \\ \text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j \le b_i, & \quad i = 1, \ldots, m \\ x_j \ge 0, & \quad j = 1, \ldots, n \end{aligned}$$

which, in matrix form, is

$$\begin{aligned} \max c^T x & \\ \text{s.t.} \ Ax \le b & \\ x \ge O & \end{aligned}$$

is said to be in *standard form*. For every linear program there is an equivalent one in standard form (begin thinking about this).

6.3 Back to the Example

Suppose someone approached the Gadget and Gewgaw Manufacturing Company (GGMC), offering to purchase the company's available labor hours, wood, and metal, at \$1.50 per hour of labor, \$1 per unit of wood, and \$1 per unit of metal. They are willing to buy whatever amount GGMC is willing to sell. Should GGMC sell everything? This is mighty tempting, because they would receive \$350, more than what they would gain by their current manufacturing plan. However, observe that if they manufactured some gadgets instead, for each kilogram of gadgets they would lose \$4.50 from the potential sale of their resources but gain \$5 from the sale of the gadgets. (Note, however, that it would be better to sell their resources than make gewgaws.) So they should not accept the offer to sell *all* of their resources at these prices.

Exercise 6.2 In the example above, GGMC wouldn't want to sell all of their resources at those prices. But they might want to sell some. What would be their best strategy? □

Exercise 6.3 Suppose now that GGMC is offered \$3 for each unit of wood and \$1 for each unit of metal that they are willing to sell, but no money for hours of labor. Explain why they would do just as well financially by selling all of their resources as by manufacturing their products. □

Exercise 6.4 In general, what conditions would proposed prices have to satisfy to induce GGMC to sell all of their resources? If you were trying to buy all of GGMC's resources as cheaply as possible, what problem would you have to solve? □

Exercise 6.5 If you want to purchase just one hour of labor, or just one unit of wood, or just one unit of metal, from GGMC, what price in each case must you offer to induce GGMC to sell? □

6.4 Exercises: Linear Programs

Exercise 6.6 Consider the following linear program (P):

$$\begin{array}{rll} \max z = x_1 + 2x_2 & \\ \text{s.t. } 3x_1 + x_2 \leq 3 & (1) \\ x_1 + x_2 \leq 3/2 & (2) \\ x_1 \geq 0 & (3) \\ x_2 \geq 0 & (4) \end{array}$$

1. Graph the feasible region.
2. Locate the optimal point(s).
3. Explain why the four constraints have the following respective outer normal vectors (an outer normal vector to a constraint is perpendicular to the defining line of the constraint and points in the opposite direction of the shaded side of the constraint):

(1) $[3, 1]^T$.

(2) $[1, 1]^T$.

(3) $[-1, 0]^T$.

(4) $[0, -1]^T$.

Explain why the gradient of the objective function is the vector $[1, 2]^T$. For each corner point of the feasible region, compare the outer normals of the binding constraints at that point (the constraints satisfied with equality by that point) with the gradient of z. From this comparison, how can you tell geometrically if a given corner point is optimal or not?

4. Vary the objective function coefficients and consider the following linear program:

$$\begin{aligned} \max z &= c_1x_1 + c_2x_2 \\ \text{s.t. } 3x_1 + x_2 &\leq 3 \\ x_1 + x_2 &\leq 3/2 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Carefully and completely describe the optimal value $z^*(c_1, c_2)$ as a function of the pair (c_1, c_2). What kind of function is this? Optional: Use some software such as Maple to plot this function of two variables.

5. Vary the right hand sides and consider the following linear program:

$$\begin{aligned} \max z &= x_1 + 2x_2 \\ \text{s.t. } 3x_1 + x_2 &\leq b_1 \\ x_1 + x_2 &\leq b_2 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Carefully and completely describe the optimal value $z^*(b_1, b_2)$ as a function of the pair (b_1, b_2). What kind of function is this? Optional: Use some software such as Maple to plot this function of two variables.

6. Find the best nonnegative integer solution to (P). That is, of all feasible points for (P) having integer coordinates, find the one with the largest objective function value.

□

Exercise 6.7 Consider the following linear program (P):

$$\begin{array}{rl} \max z = -x_1 - x_2 & \\ \text{s.t. } x_1 \leq 1/2 & (1) \\ x_1 - x_2 \leq -1/2 & (2) \\ x_1 \geq 0 & (3) \\ x_2 \geq 0 & (4) \end{array}$$

Answer the analogous questions as in Exercise 6.6. □

Exercise 6.8

1. Consider the following linear program (P):

$$\begin{array}{rl} \max z = 2x_1 + x_2 & \\ \text{s.t. } x_1 \leq 2 & (1) \\ x_2 \leq 2 & (2) \\ x_1 + x_2 \leq 4 & (3) \\ x_1 - x_2 \leq 1 & (4) \\ x_1 \geq 0 & (5) \\ x_2 \geq 0 & (6) \end{array}$$

   Associated with each of the 6 constraints is a line (change the inequality to equality in the constraint). Consider each pair of constraints for which the lines are not parallel, and examine the point of intersection of the two lines. Call this pair of constraints a *primal feasible pair* if the intersection point falls in the feasible region for (P). Call this pair of constraints a *dual feasible pair* if the gradient of the objective function can be expressed as a nonnegative linear combination of the two outer normal vectors of the two constraints. (The motivation for this terminology will become clearer later on.) List all primal-feasible pairs of constraints, and mark the intersection point for each pair. List all dual-feasible pairs of constraints (whether primal-feasible or not), and mark the intersection point for each pair. What do you observe about the optimal point(s)?

2. Repeat the above exercise for the GGMC problem.

□

Exercise 6.9 We have observed that any two-variable linear program appears to fall into exactly one of three categories: (1) those that are infeasible, (2) those that have unbounded objective function value, and (3) those that have a finite optimal objective function value. Suppose (P) is any two-variable linear program that falls into category (1). Into which of the other two categories can (P) be changed if we only alter the right hand side vector b? The objective function vector c? Both b and c? Are your answers true regardless of the initial choice of (P)? Answer the analogous questions if (P) is initially in category (2). In category (3). □

Exercise 6.10 Find a two-variable linear program

$$(P) \quad \begin{array}{c} \max c^T x \\ \text{s.t. } Ax \leq b \\ x \geq O \end{array}$$

with associated integer linear program

$$(IP) \quad \begin{array}{c} \max c^T x \\ \text{s.t. } Ax \leq b \\ x \geq O \text{ and integer} \end{array}$$

such that (P) has unbounded objective function value, but (IP) has a finite optimal objective function value. Note: "x integer" means that each coordinate x_j of x is an integer. □

Exercise 6.11 Prove the following: For each positive real number d there exists a two-variable linear program (P) with associated integer linear program (IP) such that the entries of A, b, and c are rational, (P) has a unique optimal solution x^*, (IP) has a unique optimal solution $\overline{x}^*$, and the Euclidean distance between x^* and $\overline{x}^*$ exceeds d. Can you do the same with a one-variable linear program? □

Exercise 6.12 Find a subset S of $\mathbf{R}^2$ and a linear objective function $c^T x$ such that the optimization problem

$$\begin{array}{c} \max c^T x \\ \text{s.t. } x \in S \end{array}$$

is feasible, has no optimal objective function value, but yet does not have unbounded objective function value. □

Exercise 6.13 Find a quadratic objective function $f(x)$, a matrix A with two columns, and a vector b such that the optimization problem

$$\begin{array}{c} \max f(x) \\ \text{s.t. } Ax \leq b \\ x \geq O \end{array}$$

has a unique optimal solution, but not at a corner point. □

Exercise 6.14 (Chvátal problem 1.5.) Prove or disprove: If the linear program

$$\begin{array}{rl} & \max c^T x \\ (P) & \text{s.t. } Ax \le b \\ & x \ge O \end{array}$$

is unbounded, then there is a subscript k such that the linear program

$$\begin{array}{c} \max x_k \\ \text{s.t. } Ax \le b \\ x \ge O \end{array}$$

is unbounded. □

Exercise 6.15 (Bertsimas and Tsitsiklis problem 1.12.) Consider a set $S \subseteq \mathbf{R}^n$ described by the constraints $Ax \le b$. The ball with center $y \in \mathbf{R}^n$ and radius $r \in \mathbf{R}_+$ is defined as $\{x \in \mathbf{R}^n : \|x - y\| \le r\}$. Construct a linear program to solve the problem of finding a ball with the largest possible radius that is entirely contained within the set S. □

7 Duality

In this section we will learn that associated with a given linear program is another one, its dual, which provides valuable information about the nature of the original linear program.

7.1 Economic Motivation

The dual linear program can be motivated economically, algebraically, and geometrically. You have already seen an economic motivation in Section 6.3. Recall that GGMC was interested in producing gadgets and gewgaws and wanted to solve the linear program

$$\begin{aligned} \max z = 5x_1 + 4x_2 \\ \text{s.t. } x_1 + 2x_2 \leq 120 \\ x_1 + x_2 \leq 70 \\ 2x_1 + x_2 \leq 100 \\ x_1, x_2 \geq 0 \end{aligned}$$

Another company (let's call it the Knickknack Company, KC) wants to offer money for GGMC's resources. If they are willing to buy whatever GGMC is willing to sell, what prices should be set so that GGMC will end up selling all of its resources? What is the minimum that KC must spend to accomplish this? Suppose y_1, y_2, y_3 represent the prices for one hour of labor, one unit of wood, and one unit of metal, respectively. The prices must be such that GGMC would not prefer manufacturing any gadgets or gewgaws to selling all of their resources. Hence the prices must satisfy $y_1 + y_2 + 2y_3 \geq 5$ (the income from selling the resources needed to make one kilogram of gadgets must not be less than the net profit from making one kilogram of gadgets) and $2y_1 + y_2 + y_3 \geq 4$ (the income from selling the resources needed to make one kilogram of gewgaws must not be less than the net profit from making one kilogram of gewgaws). KC wants to spend as little as possible, so it wishes to minimize the total amount spent: $120y_1 + 70y_2 + 100y_3$. This results in the linear program

$$\begin{aligned} \min 120y_1 + 70y_2 + 100y_3 \\ \text{s.t. } y_1 + y_2 + 2y_3 \geq 5 \\ 2y_1 + y_2 + y_3 \geq 4 \\ y_1, y_2, y_3 \geq 0 \end{aligned}$$

In matrix form, this is

$$
\begin{aligned}
\min \begin{bmatrix} 120 & 70 & 100 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \\
\text{s.t.} \quad \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \geq \begin{bmatrix} 5 \\ 4 \end{bmatrix} \\
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}
\end{aligned}
$$

or

$$
\begin{aligned}
\min \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} 120 \\ 70 \\ 100 \end{bmatrix} \\
\text{s.t.} \quad \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \geq \begin{bmatrix} 5 & 4 \end{bmatrix} \\
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}
\end{aligned}
$$

If we represent the GGMC problem and the KC problem in the following compact forms, we see that they are "transposes" of each other.

1	2	120
1	1	70
2	1	100
5	4	max

GGMC

1	1	2	5
2	1	1	4
120	70	100	min

KC

7.2 The Dual Linear Program

Given any linear program (P) in standard form

$$
(P) \quad \begin{aligned} & \max c^T x \\ & \text{s.t. } Ax \leq b \\ & x \geq O \end{aligned}
$$

or

$$\begin{aligned} \max \sum_{j=1}^{n} c_j x_j \\ \text{s.t.} \ \sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = 1, \ldots, m \\ x_j \ge 0, \quad j = 1, \ldots, n \end{aligned}$$

its *dual* is the LP

$$(D) \quad \begin{aligned} &\min y^T b \\ \text{s.t.} \ &y^T A \ge c^T \\ &y \ge O \end{aligned}$$

or

$$\begin{aligned} &\min b^T y \\ \text{s.t.} \ &A^T y \ge c \\ &y \ge O \end{aligned}$$

or

$$\begin{aligned} \min \sum_{i=1}^{m} b_i y_i \\ \text{s.t.} \ \sum_{i=1}^{m} a_{ij} y_i \ge c_j, \quad j = 1, \ldots, n \\ y_i \ge 0, \quad i = 1, \ldots, m \end{aligned}$$

Note the change from maximization to minimization, the change in the direction of the inequalities, the interchange in the roles of objective function coefficients and right-hand sides, the one-to-one correspondence between the inequalities in $Ax \le b$ and the variables in (D), and the one-to-one correspondence between the inequalities in $y^T A \ge c^T$ and the variables in (P). In compact form, the two problems are transposes of each other:

A	b
c^T	max

(P)

A^T	c
b^T	min

(D)

By the way, the problem (P) is called the *primal* problem. It has been explained to me that George Dantzig's father made two contributions to the theory of linear programming: the word "primal," and George Dantzig. Dantzig had already decided to use the word "dual" for the second LP, but needed a term for the original problem.

7.3 The Duality Theorems

One algebraic motivation for the dual is given by the following theorem, which states that any feasible solution for the dual LP provides an upper bound for the value of the primal LP:

Theorem 7.1 (Weak Duality) *If $\overline{x}$ is feasible for (P) and $\overline{y}$ is feasible for (D), then $c^T\overline{x} \le \overline{y}^T b$.*

PROOF. $c^T\overline{x} \le (\overline{y}^T A)\overline{x} = \overline{y}^T(A\overline{x}) \le \overline{y}^T b$. □

Example 7.2 The prices $(1, 2, 3)$ are feasible for KC's problem, and yield an objective function value of 560, which is ≥ 310. □

As an easy corollary, if we are fortunate enough to be given $\overline{x}$ and $\overline{y}$ feasible for (P) and (D), respectively, with equal objective function values, then they are each optimal for their respective problems:

Corollary 7.3 *If $\overline{x}$ and $\overline{y}$ are feasible for (P) and (D), respectively, and if $c^T\overline{x} = \overline{y}^T b$, then $\overline{x}$ and $\overline{y}$ are optimal for (P) and (D), respectively.*

PROOF. Suppose $\hat{x}$ is any feasible solution for (P). Then $c^T\hat{x} \le \overline{y}^T b = c^T\overline{x}$. Similarly, if $\hat{y}$ is any feasible solution for (D), then $\hat{y}^T b \ge \overline{y}^T b$. □

Example 7.4 The prices $(0, 3, 1)$ are feasible for KC's problem, and yield an objective function value of 310. Therefore, $(30, 40)$ is an optimal solution to GGMC's problem, and $(0, 3, 1)$ is an optimal solution to KC's problem. □

Weak Duality also immediately shows that if (P) is unbounded, then (D) is infeasible:

Corollary 7.5 *If (P) has unbounded objective function value, then (D) is infeasible. If (D) has unbounded objective function value, then (P) is infeasible.*

PROOF. Suppose (D) is feasible. Let $\overline{y}$ be a particular feasible solution. Then for all $\overline{x}$ feasible for (P) we have $c^T\overline{x} \le \overline{y}^T b$. So (P) has bounded objective function value if it is feasible, and therefore cannot be unbounded. The second statement is proved similarly. □

Suppose (P) is feasible. How can we verify that (P) is unbounded? One way is if we discover a vector $\overline{w}$ such that $A\overline{w} \le O$, $\overline{w} \ge O$, and $c^T\overline{w} > 0$. To see why this is the case, suppose that $\overline{x}$ is feasible for (P). Then we can add a positive multiple of $\overline{w}$ to $\overline{x}$ to get another feasible solution to (P) with objective function value as high as we wish.

Perhaps surprisingly, the converse is also true, and the proof shows some of the value of Theorems of the Alternatives.

Theorem 7.6 *Assume (P) is feasible. Then (P) is unbounded (has unbounded objective function value) if and only if the following system is feasible:*

$$(UP) \quad \begin{array}{c} Aw \leq O \\ c^T w > 0 \\ w \geq O \end{array}$$

PROOF. Suppose $\overline{x}$ is feasible for (P).

First assume that $\overline{w}$ is feasible for (UP) and $t \geq 0$ is a real number. Then

$$\begin{array}{c} A(\overline{x} + t\overline{w}) = A\overline{x} + tA\overline{w} \leq b + O = b \\ \overline{x} + t\overline{w} \geq O + tO = O \\ c^T(\overline{x} + t\overline{w}) = c^T\overline{x} + tc^T\overline{w} \end{array}$$

Hence $\overline{x}+t\overline{w}$ is feasible for (P), and by choosing t appropriately large, we can make $c^T(\overline{x}+t\overline{w})$ as large as desired since $c^T\overline{w}$ is a positive number.

Conversely, suppose that (P) has unbounded objective function value. Then by Corollary 7.5, (D) is infeasible. That is, the following system has no solution:

$$\begin{array}{c} y^T A \geq c^T \\ y \geq O \end{array}$$

or

$$\begin{array}{c} A^T y \geq c \\ y \geq O \end{array}$$

By the Theorem of the Alternatives proved in Exercise 4.9, the following system is feasible:

$$\begin{array}{c} w^T A^T \leq O^T \\ w^T c > 0 \\ w \geq O \end{array}$$

or

$$\begin{array}{c} Aw \leq O \\ c^T w > 0 \\ w \geq O \end{array}$$

Hence (UP) is feasible. $\square$

Example 7.7 Consider the LP:

$$(P) \quad \begin{array}{c} \max 100x_1 + x_2 \\ \text{s.t.} \; -2x_1 + 3x_2 \leq 1 \\ x_1 - 2x_2 \leq 2 \\ x_1, x_2 \geq 0 \end{array}$$

The system (UP) in this case is:

$$\begin{gathered}-2w_1 + 3w_2 \le 0\\ w_1 - 2w_2 \le 0\\ 100w_1 + w_2 > 0\\ w_1, w_2 \ge 0\end{gathered}$$

One feasible point for (P) is $\overline{x} = (1, 0)$. One feasible solution to (UP) is $\overline{w} = (2, 1)$. So (P) is unbounded, and we can get points with arbitrarily high objective function values by $\overline{x} + t\overline{w} = (1 + 2t, t)$, $t \ge 0$, which has objective function value $100 + 201t$. $\square$

There is an analogous theorem for the unboundedness of (D) that is proved in the obviously similar way:

Theorem 7.8 *Assume (D) is feasible. Then (D) is unbounded if and only if the following system is feasible:*

$$(UD)\qquad \begin{gathered} v^T A \ge O^T\\ v^T b < 0\\ v \ge O\end{gathered}$$

The following highlights an immediate corollary of the proof:

Corollary 7.9 *(P) is feasible if and only if (UD) is infeasible. (D) is feasible if and only if (UP) is infeasible.*

Let's summarize what we now know in a slightly different way:

Corollary 7.10 *If (P) is infeasible, then either (D) is infeasible or (D) is unbounded. If (D) is infeasible, then either (P) is infeasible or (P) is unbounded.*

We now turn to a very important theorem, which is part of the strong duality theorem, that lies at the heart of linear programming. This shows that the bounds on each other's objective function values that the pair of dual LP's provides are always tight.

Theorem 7.11 *Suppose (P) and (D) are both feasible. Then (P) and (D) each have finite optimal objective function values, and moreover these two values are equal.*

Proof. We know by Weak Duality that if $\overline{x}$ and $\overline{y}$ are feasible for (P) and (D), respectively, then $c^T\overline{x} \leq \overline{y}^T b$. In particular, neither (P) nor (D) is unbounded. So it suffices to show that the following system is feasible:

$$(I)\quad \begin{array}{c} Ax \leq b \\ x \geq O \\ y^T A \geq c^T \\ y \geq O \\ c^T x \geq y^T b \end{array}$$

For if $\overline{x}$ and $\overline{y}$ are feasible for this system, then by Weak Duality in fact it would have to be the case that $c^T\overline{x} = \overline{y}^T b$.

Let's rewrite this system in matrix form:

$$\begin{bmatrix} A & O \\ O & -A^T \\ -c^T & b^T \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} b \\ -c \\ 0 \end{bmatrix}$$
$$x, y \geq O$$

We will assume that this system is infeasible and derive a contradiction. If it is not feasible, then by Theorem 4.3 the following system has a solution $\overline{v}, \overline{w}, \overline{t}$:

$$(II)\quad \begin{array}{c} \begin{bmatrix} v^T & w^T & t \end{bmatrix} \begin{bmatrix} A & O \\ O & -A^T \\ -c^T & b^T \end{bmatrix} \geq \begin{bmatrix} O^T & O^T \end{bmatrix} \\ \begin{bmatrix} v^T & w^T & t \end{bmatrix} \begin{bmatrix} b \\ -c \\ 0 \end{bmatrix} < 0 \\ v, w, t \geq O \end{array}$$

So we have

$$\begin{array}{c} \overline{v}^T A - \overline{t} c^T \geq O^T \\ -\overline{w}^T A^T + \overline{t} b^T \geq O^T \\ \overline{v}^T b - \overline{w}^T c < 0 \\ \overline{v}, \overline{w}, \overline{t} \geq O \end{array}$$

Case 1: Suppose $\overline{t} = 0$. Then

$$\begin{array}{c} \overline{v}^T A \geq O^T \\ A\overline{w} \leq O \\ \overline{v}^T b < c^T \overline{w} \\ \overline{v}, \overline{w} \geq O \end{array}$$

Now we cannot have both $c^T\overline{w} \leq 0$ and $\overline{v}^T b \geq 0$; otherwise $0 \leq \overline{v}^T b < c^T\overline{w} \leq 0$, which is a contradiction.

Case 1a: Suppose $c^T\overline{w} > 0$. Then $\overline{w}$ is a solution to (UP), so (D) is infeasible by Corollary 7.9, a contradiction.

Case 1b: Suppose $\overline{v}^T b < 0$. Then $\overline{v}$ is a solution to (UD), so (P) is infeasible by Corollary 7.9, a contradiction.

Case 2: Suppose $\overline{t} > 0$. Set $\overline{x} = \overline{w}/\overline{t}$ and $\overline{y} = \overline{v}/\overline{t}$. Then

$$\begin{array}{c} A\overline{x} \leq b \\ \overline{x} \geq O \\ \overline{y}^T A \geq c^T \\ \overline{y} \geq O \\ c^T\overline{x} > \overline{y}^T b \end{array}$$

Hence we have a pair of feasible solutions to (P) and (D), respectively, that violates Weak Duality, a contradiction.

We have now shown that (II) has no solution. Therefore, (I) has a solution. □

Corollary 7.12 *Suppose (P) has a finite optimal objective function value. Then so does (D), and these two values are equal. Similarly, suppose (D) has a finite optimal objective function value. Then so does (P), and these two values are equal.*

Proof. We will prove the first statement only. If (P) has a finite optimal objective function value, then it is feasible, but not unbounded. So (UP) has no solution by Theorem 7.6. Therefore (D) is feasible by Corollary 7.9. Now apply Theorem 7.11. □

We summarize our results in the following central theorem, for which we have already done all the hard work:

Theorem 7.13 (Strong Duality) *Exactly one of the following holds for the pair (P) and (D):*

1. *They are both infeasible.*
2. *One is infeasible and the other is unbounded.*
3. *They are both feasible and have equal finite optimal objective function values.*

Corollary 7.14 *If $\overline{x}$ and $\overline{y}$ are feasible for (P) and (D), respectively, then $\overline{x}$ and $\overline{y}$ are optimal for (P) and (D), respectively, if and only if $c^T\overline{x} = \overline{y}^T b$.*

Corollary 7.15 *Suppose $\overline{x}$ is feasible for (P). Then $\overline{x}$ is optimal for (P) if and only if there exists $\overline{y}$ feasible for (D) such that $c^T\overline{x} = \overline{y}^T b$. Similarly, suppose $\overline{y}$ is feasible for (D). Then $\overline{y}$ is optimal for (D) if and only if there exists $\overline{x}$ feasible for (P) such that $c^T\overline{x} = \overline{y}^T b$.*

7.4 Comments on Good Characterization

The duality theorems show that the following problems for (P) have "good characterizations." That is to say, whatever the answer, there exists a "short" proof.

1. Is (P) feasible? If the answer is yes, you can prove it by producing a particular feasible solution to (P). If the answer is no, you can prove it by producing a particular feasible solution to (UD).

2. Assume that you know that (P) is feasible. Is (P) unbounded? If the answer is yes, you can prove it by producing a particular feasible solution to (UP). If the answer is no, you can prove it by producing a particular feasible solution to (D).

3. Assume that $\overline{x}$ is feasible for (P). Is $\overline{x}$ optimal for (P)? If the answer is yes, you can prove it by producing a particular feasible solution to (D) with the same objective function value. If the answer is no, you can prove it by producing a particular feasible solution to (P) with higher objective function value.

7.5 Complementary Slackness

Suppose $\overline{x}$ and $\overline{y}$ are feasible for (P) and (D), respectively. Under what conditions will $c^T\overline{x}$ equal $\overline{y}^Tb$? Recall the chain of inequalities in the proof of Weak Duality:

$$c^T\overline{x} \le (\overline{y}^TA)\overline{x} = \overline{y}^T(A\overline{x}) \le \overline{y}^Tb.$$

Equality occurs if and only if both $c^T\overline{x} = (\overline{y}^TA)\overline{x}$ and $\overline{y}^T(A\overline{x}) = \overline{y}^Tb$. Equivalently,

$$\overline{y}^T(b - A\overline{x}) = 0$$

and

$$(\overline{y}^TA - c^T)\overline{x} = 0.$$

In each case, we are requiring that the inner product of two nonnegative vectors (for example, $\overline{y}$ and $b - A\overline{x}$) be zero. The only way this can happen is if these two vectors are never both positive in any common component. This motivates the following definition: Suppose $\overline{x} \in \mathbf{R}^n$ and $\overline{y} \in \mathbf{R}^m$. Then $\overline{x}$ and $\overline{y}$ satisfy *complementary slackness* if

1. For all j, either $\overline{x}_j = 0$ or $\sum_{i=1}^m a_{ij}\overline{y}_i = c_j$ or both; and

2. For all i, either $\overline{y}_i = 0$ or $\sum_{j=1}^n a_{ij}\overline{x}_j = b_i$ or both.

Theorem 7.16 *Suppose $\overline{x}$ and $\overline{y}$ are feasible for (P) and (D), respectively. Then $c^T\overline{x} = \overline{y}^T b$ if and only if $\overline{x}, \overline{y}$ satisfy complementary slackness.*

Corollary 7.17 *If $\overline{x}$ and $\overline{y}$ are feasible for (P) and (D), respectively, then $\overline{x}$ and $\overline{y}$ are optimal for (P) and (D), respectively, if and only if they satisfy complementary slackness.*

Corollary 7.18 *Suppose $\overline{x}$ is feasible for (P). Then $\overline{x}$ is optimal for (P) if and only if there exists $\overline{y}$ feasible for (D) such that $\overline{x}, \overline{y}$ satisfy complementary slackness. Similarly, suppose $\overline{y}$ is feasible for (D). Then $\overline{y}$ is optimal for (D) if and only if there exists $\overline{x}$ feasible for (P) such that $\overline{x}, \overline{y}$ satisfy complementary slackness.*

Example 7.19 Consider the optimal solution $(30, 40)$ of GGMC's problem, and the prices $(0, 3, 1)$ for KC's problem. You can verify that both solutions are feasible for their respective problems, and that they satisfy complementary slackness. But let's exploit complementary slackness a bit more. Suppose you only had the feasible solution $(30, 40)$ and wanted to verify optimality. Try to find a feasible solution to the dual satisfying complementary slackness. Because the constraint on hours is not satisfied with equality, we must have $y_1 = 0$. Because both x_1 and x_2 are positive, we must have both dual constraints satisfied with equality. This results in the system:

$$\begin{aligned} y_1 &= 0 \\ y_2 + 2y_3 &= 5 \\ y_2 + y_3 &= 4 \end{aligned}$$

which has the unique solution $(0, 3, 1)$. Fortunately, all values are also nonnegative. Therefore we have a feasible solution to the dual that satisfies complementary slackness. This proves that $(30, 40)$ is optimal and produces a solution to the dual in the bargain. $\square$

7.6 Duals of General LP's

What if you want a dual to an LP not in standard form? One approach is first to transform it into standard form somehow. Another is to come up with a definition of a general dual that will satisfy all of the duality theorems (weak and strong duality, correspondence between constraints and variables, complementary slackness, etc.). Both approaches are related.

Here are some basic transformations to convert an LP into an equivalent one:

1. Multiply the objective function by -1 and change "max" to "min" or "min" to "max."

2. Multiply an inequality constraint by -1 to change the direction of the inequality.

3. Replace an equality constraint
$$\sum_{j=1}^{n} a_{ij}x_j = b_i$$
with two inequality constraints
$$\begin{aligned} \sum_{j=1}^{n} a_{ij}x_j &\le b_i \\ -\sum_{j=1}^{n} a_{ij}x_j &\le -b_i \end{aligned}$$

4. Replace a variable that is nonpositive with a variable that is its negative. For example, if x_j is specified to be nonpositive by $x_j \le 0$, replace every occurrence of x_j with $-\hat{x}_j$ and require $\hat{x}_j \ge 0$.

5. Replace a variable that is unrestricted in sign with the difference of two nonnegative variables. For example, if x_j is unrestricted (sometimes called *free*), replace every occurrence of x_j with $x_j^+ - x_j^-$ and require that x_j^+ and x_j^- be nonnegative variables.

Using these transformations, every LP can be converted into an equivalent one in standard form. By *equivalent* I mean that a feasible (respectively, optimal) solution to the original problem can be obtained from a feasible (respectively, optimal) solution to the new problem. The dual to the equivalent problem can then be determined. But you can also apply the inverses of the above transformations to the dual and get an appropriate dual to the original problem.

Try some concrete examples for yourself, and then dive into the proof of the following theorem:

Theorem 7.20 *The following is a pair of dual LP's:*

$$(P)\quad \begin{aligned} &\max \sum_{j=1}^{n} c_j x_j \\ \text{s.t. } &\sum_{j=1}^{n} a_{ij}x_j \le b_i, \quad i \in I_1 \\ &\sum_{j=1}^{n} a_{ij}x_j \ge b_i, \quad i \in I_2 \\ &\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i \in I_3 \\ &x_j \ge 0, \quad j \in J_1 \\ &x_j \le 0, \quad j \in J_2 \\ &x_j \text{ unrestricted in sign}, \quad j \in J_3 \end{aligned} \qquad (D)\quad \begin{aligned} &\min \sum_{i=1}^{m} b_i y_i \\ \text{s.t. } &\sum_{i=1}^{m} a_{ij}y_i \ge c_j, \quad j \in J_1 \\ &\sum_{i=1}^{m} a_{ij}y_i \le c_j, \quad j \in J_2 \\ &\sum_{i=1}^{m} a_{ij}y_i = c_j, \quad j \in J_3 \\ &y_i \ge 0, \quad i \in I_1 \\ &y_i \le 0, \quad i \in I_2 \\ &y_i \text{ unrestricted in sign}, \quad i \in I_3 \end{aligned}$$

where (I_1, I_2, I_3) *is a partition of* $\{1, \ldots, m\}$ *and* (J_1, J_2, J_3) *is a partition of* $\{1, \ldots, n\}$.

Proof. Rewrite (P) in matrix form:

$$\begin{array}{c}
\max c^{1T}x^1 + c^{2T}x^2 + c^{3T}x^3 \\
\text{s.t.} \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} \begin{array}{c} \le \\ \ge \\ = \end{array} \begin{bmatrix} b^1 \\ b^2 \\ b^3 \end{bmatrix} \\
x^1 \ge O \\
x^2 \le O \\
x^3 \text{ unrestricted}
\end{array}$$

Now make the substitutions $\hat{x}^1 = x^1$, $\hat{x}^2 = -x^2$ and $\hat{x}^3 - \hat{x}^4 = x^3$:

$$\begin{array}{c}
\max c^{1T}\hat{x}^1 - c^{2T}\hat{x}^2 + c^{3T}\hat{x}^3 - c^{3T}\hat{x}^4 \\
\text{s.t.} \begin{bmatrix} A_{11} & -A_{12} & A_{13} & -A_{13} \\ A_{21} & -A_{22} & A_{23} & -A_{23} \\ A_{31} & -A_{32} & A_{33} & -A_{33} \end{bmatrix} \begin{bmatrix} \hat{x}^1 \\ \hat{x}^2 \\ \hat{x}^3 \\ \hat{x}^4 \end{bmatrix} \begin{array}{c} \le \\ \ge \\ = \end{array} \begin{bmatrix} b^1 \\ b^2 \\ b^3 \end{bmatrix} \\
\hat{x}^1, \hat{x}^2, \hat{x}^3, \hat{x}^4 \ge O
\end{array}$$

Transform the constraints:

$$\begin{array}{c}
\max c^{1T}\hat{x}^1 - c^{2T}\hat{x}^2 + c^{3T}\hat{x}^3 - c^{3T}\hat{x}^4 \\
\text{s.t.} \begin{bmatrix} A_{11} & -A_{12} & A_{13} & -A_{13} \\ -A_{21} & A_{22} & -A_{23} & A_{23} \\ A_{31} & -A_{32} & A_{33} & -A_{33} \\ -A_{31} & A_{32} & -A_{33} & A_{33} \end{bmatrix} \begin{bmatrix} \hat{x}^1 \\ \hat{x}^2 \\ \hat{x}^3 \\ \hat{x}^4 \end{bmatrix} \le \begin{bmatrix} b^1 \\ -b^2 \\ b^3 \\ -b^3 \end{bmatrix} \\
\hat{x}^1, \hat{x}^2, \hat{x}^3, \hat{x}^4 \ge O
\end{array}$$

Take the dual:

$$\begin{array}{c}
\min b^{1T}\hat{y}^1 - b^{2T}\hat{y}^2 + b^{3T}\hat{y}^3 - b^{3T}\hat{y}^4 \\
\text{s.t.} \begin{bmatrix} A_{11}^T & -A_{21}^T & A_{31}^T & -A_{31}^T \\ -A_{12}^T & A_{22}^T & -A_{32}^T & A_{32}^T \\ A_{13}^T & -A_{23}^T & A_{33}^T & -A_{33}^T \\ -A_{13}^T & A_{23}^T & -A_{33}^T & A_{33}^T \end{bmatrix} \begin{bmatrix} \hat{y}^1 \\ \hat{y}^2 \\ \hat{y}^3 \\ \hat{y}^4 \end{bmatrix} \ge \begin{bmatrix} c^1 \\ -c^2 \\ c^3 \\ -c^3 \end{bmatrix} \\
\hat{y}^1, \hat{y}^2, \hat{y}^3, \hat{y}^4 \ge O
\end{array}$$

Transform the constraints:

$$\min b^{1T}\hat{y}^1 - b^{2T}\hat{y}^2 + b^{3T}\hat{y}^3 - b^{3T}\hat{y}^4$$
$$\text{s.t.} \quad \begin{bmatrix} A_{11}^T & -A_{21}^T & A_{31}^T & -A_{31}^T \\ A_{12}^T & -A_{22}^T & A_{32}^T & -A_{32}^T \\ A_{13}^T & -A_{23}^T & A_{33}^T & -A_{33}^T \end{bmatrix} \begin{bmatrix} \hat{y}^1 \\ \hat{y}^2 \\ \hat{y}^3 \\ \hat{y}^4 \end{bmatrix} \begin{matrix} \geq \\ \leq \\ = \end{matrix} \begin{bmatrix} c^1 \\ c^2 \\ c^3 \end{bmatrix}$$
$$\hat{y}^1, \hat{y}^2, \hat{y}^3, \hat{y}^4 \geq O$$

Transform the variables by setting $y^1 = \hat{y}^1$, $y^2 = -\hat{y}^2$, and $y^3 = \hat{y}^3 - \hat{y}^4$:

$$\min b^{1T}y^1 + b^{2T}y^2 + b^{3T}y^3$$
$$\text{s.t.} \quad \begin{bmatrix} A_{11}^T & A_{21}^T & A_{31}^T \\ A_{12}^T & A_{22}^T & A_{32}^T \\ A_{13}^T & A_{23}^T & A_{33}^T \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \\ y^3 \end{bmatrix} \begin{matrix} \geq \\ \leq \\ = \end{matrix} \begin{bmatrix} c^1 \\ c^2 \\ c^3 \end{bmatrix}$$
$$y^1 \geq O$$
$$y^2 \leq O$$
$$y^3 \text{ unrestricted}$$

Write this in summation form, and you have (D). □

Whew! Anyway, this pair of dual problems will satisfy all of the duality theorems, so it was probably worth working through this generalization at least once. We say that (D) is the dual of (P), and also that (P) is the dual of (D). Note that there is still a one-to-one correspondence between the variables in one LP and the "main" constraints (not including the variable sign restrictions) in the other LP. Hillier and Lieberman (*Introduction to Operations Research*) suggest the following mnemonic device. Classify variables and constraints of linear programs as *standard* (S), *opposite* (O), or *bizarre* (B) as follows:

Maximization Problems

	Variables	Constraints
S	≥ 0	$\leq$
O	≤ 0	$\geq$
B	unrestricted in sign	$=$

Minimization Problems

	Variables	Constraints
S	≥ 0	$\geq$
O	≤ 0	$\leq$
B	unrestricted in sign	$=$

Then in the duality relationship, standard variables are paired with standard constraints, opposite variables are paired with opposite constraints, and bizarre variables are paired with bizarre constraints. If we express a pair of dual linear programs in compact form, labeling columns according to the type of variable and rows according to the type of constraint, we see that they are still transposes of each other:

	S	O	B	
S	A_{11}	A_{12}	A_{13}	b^1
O	A_{21}	A_{22}	A_{23}	b^2
B	A_{31}	A_{32}	A_{33}	b^3
	c^{1T}	c^{2T}	c^{3T}	max

(P)

	S	O	B	
S	A_{11}	A_{21}	A_{31}	c^1
O	A_{12}	A_{22}	A_{32}	c^2
B	A_{13}	A_{23}	A_{33}	c^3
	b^{1T}	b^{2T}	b^{3T}	min

(D)

Example 7.21 The following is a pair of dual linear programs:

(P)

$$\begin{array}{lrrrrrcr}
\max & 3x_1 & -2x_2 & & +4x_4 & +5x_5 & & \\
\text{s.t.} & x_1 & +x_2 & & & & \geq & 3 \\
& x_1 & -x_2 & +x_3 & -x_4 & +x_5 & = & 10 \\
& -6x_1 & & +2x_3 & +4x_4 & +x_5 & \leq & 2 \\
& & 9x_2 & & -11x_4 & & \geq & 0
\end{array}$$

$$x_1, x_5 \geq 0$$
$$x_2, x_3 \leq 0$$
$$x_4 \text{ unrestricted in sign}$$

(D)

$$\begin{array}{lrrrrcr}
\min & 3y_1 & +10y_2 & +2y_3 & & & \\
\text{s.t.} & y_1 & +y_2 & -6y_3 & & \geq & 3 \\
& y_1 & -y_2 & & +9y_4 & \leq & -2 \\
& & y_2 & +2y_3 & & \leq & 0 \\
& & -y_2 & +4y_3 & -11y_4 & = & 4 \\
& & y_2 & +y_3 & & \geq & 5
\end{array}$$

$$y_1, y_4 \leq 0$$
$$y_2 \text{ unrestricted in sign}$$
$$y_3 \geq 0$$

	S	O	O	B	S	
O	1	1	0	0	0	3
B	1	−1	1	−1	1	10
S	−6	0	2	4	1	2
O	0	9	0	−11	0	0
	3	−2	0	4	5	max

	O	B	S	O	
S	1	1	−6	0	3
O	1	−1	0	9	−2
O	0	1	2	0	0
B	0	−1	4	−11	4
S	0	1	1	0	5
	3	10	2	0	min

□

Here are some special cases of pairs of dual LP's:

$$(P)\quad \begin{array}{c} \max c^T x \\ \text{s.t. } Ax \leq b \end{array} \qquad (D)\quad \begin{array}{c} \min y^T b \\ \text{s.t. } y^T A = c^T \\ y \geq O \end{array}$$

and

$$(P) \quad \begin{array}{c} \max c^T x \\ \text{s.t. } Ax = b \\ x \geq O \end{array} \qquad (D) \quad \begin{array}{c} \min y^T b \\ \text{s.t. } y^T A \geq c^T \end{array}$$

Exercise 7.22 Suppose (P) and (D) are as given in Theorem 7.20. Show that the appropriate general forms of (UP) and (UD) are:

$$(UP) \quad \begin{array}{c} \sum_{j=1}^{n} a_{ij} w_j \leq 0, \quad i \in I_1 \\ \sum_{j=1}^{n} a_{ij} w_j \geq 0, \quad i \in I_2 \\ \sum_{j=1}^{n} a_{ij} w_j = 0, \quad i \in I_3 \\ \sum_{j=1}^{n} c_j w_j > 0 \\ w_j \geq 0, \quad j \in J_1 \\ w_j \leq 0, \quad j \in J_2 \\ w_j \text{ unrestricted in sign}, \quad j \in J_3 \end{array} \qquad (UD) \quad \begin{array}{c} \sum_{i=1}^{m} a_{ij} v_i \geq 0, \quad j \in J_1 \\ \sum_{i=1}^{m} a_{ij} v_i \leq 0, \quad j \in J_2 \\ \sum_{i=1}^{m} a_{ij} v_i = 0, \quad j \in J_3 \\ \sum_{i=1}^{m} b_i v_i < 0 \\ v_i \geq 0, \quad i \in I_1 \\ v_i \leq 0, \quad i \in I_2 \\ v_i \text{ unrestricted in sign}, \quad i \in I_3 \end{array}$$

□

7.7 Geometric Motivation of Duality

We mentioned in the last section that the following is a pair of dual LP's:

$$(P) \quad \begin{array}{c} \max c^T x \\ \text{s.t. } Ax \leq b \end{array} \qquad (D) \quad \begin{array}{c} \min y^T b \\ \text{s.t. } y^T A = c^T \\ y \geq O \end{array}$$

What does it mean for $\overline{x}$ and $\overline{y}$ to be feasible and satisfy complementary slackness for this pair of LP's? The solution $\overline{y}$ to (D) gives a way to write the objective function vector of (P) as a nonnegative linear combination of the outer normals of the constraints of (P). In effect, (D) is asking for the "cheapest" such expression. If $\overline{x}$ does not satisfy a constraint of (P) with equality, then the corresponding dual variable must be zero by complementary slackness. So the only outer normals used in the nonnegative linear combination are those for the binding constraints (the constraints satisfied by $\overline{x}$ with equality).

We have seen this phenomenon when we looked at two-variable linear programs earlier. For example, look again at Exercise 6.8. Every dual-feasible pair of constraints corresponds to a particular solution to the dual problem (though there are other solutions to the dual as well), and a pair of constraints that is both primal-feasible and dual feasible corresponds to a pair of solutions to (P) and (D) that satisfy complementary slackness and hence are optimal.

7.8 Exercises: Duality

Note: By e is meant a vector consisting of all 1's.

Exercise 7.23 Consider the classic diet problem: Various foods are available, each unit of which contributes a certain amount toward the minimum daily requirements of various nutritional needs. Each food has a particular cost. The goal is to choose how many units of each food to purchase to meet the minimum daily nutritional requirements, while minimizing the total cost. Formulate this as a linear program, and give an "economic interpretation" of the dual problem. □

Exercise 7.24 Find a linear program (P) such that both (P) and its dual (D) are infeasible. □

Exercise 7.25 Prove that the *set* $S = \{x : Ax \le b,\ x \ge O\}$ is unbounded if and only if $S \ne \emptyset$ and the following system is feasible:

$$\begin{gathered} Aw \le O \\ w \ge O \\ w \ne O \end{gathered}$$

Note: By $w \ge O$, $w \ne O$ is meant that every component of w is nonnegative, and at least one component is positive. A solution to the above system is called a *feasible direction* for S. Draw some examples of two variable regions to illustrate how you can find the set of feasible directions geometrically. □

Exercise 7.26 Prove that if the LP

$$\begin{gathered} \max c^T x \\ \text{s.t. } Ax \le b \\ x \ge O \end{gathered}$$

is unbounded, then the LP

$$\begin{gathered} \max e^T x \\ \text{s.t. } Ax \le b \\ x \ge O \end{gathered}$$

is unbounded. What can you say about the converse of this statement? □

Exercise 7.27 Suppose you use Lagrange multipliers to solve the following problem:

$$\begin{array}{c}\max c^T x \\ \text{s.t. } Ax = b\end{array}$$

What is the relationship between the Lagrange multipliers and the dual problem? □

Exercise 7.28 Suppose that the linear program

$$\begin{array}{c}\max c^T x \\ \text{s.t. } Ax \leq b \\ x \geq O\end{array}$$

is unbounded. Prove that, for any $\hat{b}$, the following linear program is either infeasible or unbounded:

$$\begin{array}{c}\max c^T x \\ \text{s.t. } Ax \leq \hat{b} \\ x \geq O\end{array}$$

□

Exercise 7.29 Consider the following linear programs:

$$(P)\ \begin{array}{c}\max c^T x \\ \text{s.t. } Ax \leq b \\ x \geq O\end{array} \qquad (\overline{P})\ \begin{array}{c}\max c^T x \\ \text{s.t. } Ax \leq b + u \\ x \geq O\end{array} \qquad (D)\ \begin{array}{c}\min y^T b \\ \text{s.t. } y^T A \geq c^T \\ y \geq O\end{array}$$

Here, u is a vector the same size as b. (u is a vector of real numbers, not variables.) Assume that (P) has a finite optimal objective function value z^*. Let y^* be any optimal solution to (D). Prove that $c^T x \leq z^* + u^T y^*$ for every feasible solution x of $(\overline{P})$. What does this mean economically when applied to the GGMC problem? □

Exercise 7.30 Consider the following pair of linear programs:

$$(P)\ \begin{array}{c}\max c^T x \\ \text{s.t. } Ax \leq b \\ x \geq O\end{array} \qquad (D)\ \begin{array}{c}\min y^T b \\ \text{s.t. } y^T A \geq c^T \\ y \geq O\end{array}$$

For all nonnegative x and y, define the function $\phi(x, y) = c^T x + y^T b - y^T Ax$. Assume that $\overline{x}$ and $\overline{y}$ are nonnegative. Prove that $\overline{x}$ and $\overline{y}$ are feasible and optimal for the above two linear programs, respectively, if and only if

$$\phi(\overline{x}, y) \geq \phi(\overline{x}, \overline{y}) \geq \phi(x, \overline{y})$$

for all nonnegative x and y (whether x and y are feasible for the above linear programs or not). (This says that $(\overline{x}, \overline{y})$ is a *saddlepoint* of ϕ.) □

Exercise 7.31 Consider the *fractional linear program*

$$(FP)\quad \begin{array}{c} \max \dfrac{c^T x + \alpha}{d^T x + \beta} \\ \text{s.t. } Ax \le b \\ x \ge O \end{array}$$

and the associated linear program

$$(P)\quad \begin{array}{c} \max c^T w + \alpha t \\ \text{s.t. } Aw - bt \le O \\ d^T w + \beta t = 1 \\ w \ge O, t \ge 0 \end{array}$$

where A is an $m \times n$ matrix, b is an $m \times 1$ vector, c and d are $n \times 1$ vectors, and α and β are scalars. The variables x and w are $n \times 1$ vectors, and t is a scalar variable.

Suppose that the feasible region for (FP) is nonempty, and that $d^T x + \beta > 0$ for all x that are feasible to (FP). Let (w^*, t^*) be an optimal solution to (P).

1. Suppose that the feasible region of (FP) is a bounded set. Prove that $t^* > 0$.

2. Given that $t^* > 0$, demonstrate how an optimal solution of (FP) can be recovered from (w^*, t^*) and prove your assertion.

□

Exercise 7.32

1. Give a geometric interpretation of complementary slackness for the LP

$$\begin{array}{c} \max c^T x \\ \text{s.t. } Ax \le b \\ x \ge O \end{array}$$

   and its dual.

2. Now give an economic interpretation of complementary slackness.

□

Exercise 7.33 Consider the linear program

$$(P)\quad \begin{array}{l} \min c^T x \\ \text{s.t. } Ax = b \\ \ell \le x \le u \end{array}$$

where ℓ and u are vectors of constants and $\ell_i < u_i$ for all i. Suppose that x is feasible for (P). Prove that x is optimal for (P) if and only if there exists a vector y such that, for all i,

$$\begin{array}{ll} (A^T y)_i \ge c_i & \text{if } x_i > \ell_i \\ (A^T y)_i \le c_i & \text{if } x_i < u_i. \end{array}$$

$\square$

Exercise 7.34 There are algorithmic proofs using the simplex method of Theorem 7.13 that do not explicitly rely upon Theorem 4.3. Assume that Theorem 7.13 has been proved some other way. Now reprove Theorem 4.3 using Theorem 7.13 and the fact that (I) is feasible if and only if the following LP is feasible (and thus has optimal value 0):

$$(P)\quad \begin{array}{l} \max O^T x \\ \text{s.t. } Ax \le b \\ x \ge O \end{array}$$

$\square$

Exercise 7.35 Derive and prove a Theorem of the Alternatives for the system

$$(I)\quad Ax < b$$

in the following way: Introduce a scalar variable t and a vector e of 1's, and consider the LP

$$(P)\quad \begin{array}{l} \max t \\ \text{s.t. } Ax + et \le b \end{array}$$

Begin by noting that (P) is always feasible, and proving that (I) is infeasible if and only if (P) has a nonpositive optimal value. $\square$

Exercise 7.36 Consider the pair of dual LP's

$$(P)\quad \begin{array}{l} \max c^T x \\ \text{s.t. } Ax \le b \\ x \ge O \end{array} \qquad (D)\quad \begin{array}{l} \min y^T b \\ \text{s.t. } y^T A \ge c^T \\ y \ge O \end{array}$$

Suppose $\overline{x}$ and $\overline{y}$ are feasible for (P) and (D), respectively. Then $\overline{x}$ and $\overline{y}$ satisfy *strong complementary slackness* if

1. For all j, either $\overline{x}_j = 0$ or $\sum_{i=1}^m a_{ij}\overline{y}_i = c_j$, but not both; and

2. For all i, either $\overline{y}_i = 0$ or $\sum_{j=1}^n a_{ij}\overline{x}_j = b_i$, but not both.

Prove that if (P) and (D) are both feasible, then there exists a pair $\overline{x}, \overline{y}$ of optimal solutions to (P) and (D), respectively, that satisfies strong complementary slackness. Illustrate with some examples of two variable LP's. Hint: One way to do this is to consider the following LP:

$$\begin{aligned} &\max t \\ &\text{s.t. } Ax \le b \\ &Ax - Iy + et \le b \\ &-A^T y \le -c \\ &-Ix - A^T y + ft \le -c \\ &-c^T x + b^T y \le 0 \\ &x, y, t \ge O \end{aligned}$$

Here, both e and f are vectors of all 1's, and t is a scalar variable. $\square$

Exercise 7.37 Consider the quadratic programming problem

$$(P) \qquad \begin{aligned} &\min Q(x) = c^T x + \frac{1}{2} x^T D x \\ &\text{s.t. } Ax \ge b \\ &x \ge O \end{aligned}$$

where A is an $m \times n$ matrix and D is a symmetric $n \times n$ matrix.

1. Assume that $\overline{x}$ is an optimal solution of (P). Prove that $\overline{x}$ is an optimal solution of the following linear program:

$$(P') \qquad \begin{aligned} &\min (c^T + \overline{x}^T D)x \\ &\text{s.t. } Ax \ge b \\ &x \ge O \end{aligned}$$

   Suggestion: Let $\hat{x}$ be any other feasible solution to (P'). Then $\lambda\hat{x} + (1-\lambda)\overline{x}$ is also a feasible solution to (P') for any $0 < \lambda < 1$.

2. Assume that $\overline{x}$ is an optimal solution of (P). Prove that there exist nonnegative vectors $\overline{y} \in \mathbf{R}^m$, $\overline{u} \in \mathbf{R}^n$, and $\overline{v} \in \mathbf{R}^m$ such that

$$\begin{bmatrix} \overline{u} \\ \overline{v} \end{bmatrix} - \begin{bmatrix} D & -A^T \\ A & O \end{bmatrix} \begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix} = \begin{bmatrix} c \\ -b \end{bmatrix}$$

   and such that $\overline{u}^T\overline{x} + \overline{v}^T\overline{y} = 0$.

□

Exercise 7.38 Consider a $p \times q$ chessboard. Call a subset of cells *independent* if no pair of cells are adjacent to each other via a single knight's move. Call any line segment joining the centers of two cells that are adjacent via a single knight's move a *knight line*. A knight line is said to *cover* its two endpoint cells. A *knight line cover* is a set of knight lines such that every cell on a chessboard is covered by at least one knight line. Consider the problem (P) of finding the maximum size k^* of an independent set. Consider the problem (D) of finding the minimum size ℓ^* of a knight lines cover. Prove that if k is the size of any independent set and ℓ is the size of any knight line cover, then $k \leq \ell$. Conclude that $k^* \leq \ell^*$. Use this result to solve both (P) and (D) for the 8×8 chessboard. For the 2×6 chessboard. □

Exercise 7.39 Look up the definitions and some theorems about Eulerian graphs. Explain why the question: "Is a given graph G Eulerian?" has a good characterization. □

8 Solving Linear Programs

8.1 Matrices

Suppose A is an $m \times n$ matrix, B is an $n \times p$ matrix, and $C = AB$. Then

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}, \ i = 1, \ldots, m, \ k = 1, \ldots, p.$$

We can recognize this as the inner product of the ith row of A with the kth column of B:

$$c_{ik} = [a_{i1}, \ldots, a_{in}] \begin{bmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{bmatrix}.$$

We can also see that the ith row of C is a linear combination of the rows of B using as coefficients the entries in the ith row of A:

$$[c_{i1}, \ldots, c_{ip}] = a_{i1} [b_{11}, \ldots, b_{1p}] + \cdots + a_{in} [b_{n1}, \ldots, b_{np}],$$

and the kth column of C is a linear combination of the columns of A using as coefficients the entries in the kth column of B:

$$\begin{bmatrix} c_{1k} \\ \vdots \\ c_{mk} \end{bmatrix} = b_{1k} \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + b_{nk} \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

8.2 Four Vector Spaces Associated with a Matrix

Definition 8.1 Let A be an $m \times n$ matrix. The four vector spaces associated with A are:

1. The *column space* of A. This is the space of all linear combinations of the columns of A, $\text{columnspace}(A) = \{Ax : x \in \mathbf{R}^n\}$.

2. The *row space* of A. This is the space of all linear combinations of the rows of A, $\text{rowspace}(A) = \{y^T A : y \in \mathbf{R}^m\}$.

3. The *nullspace* of A, $\text{nullspace}(A) = \{x \in \mathbf{R}^n : Ax = O\}$.

4. The *left nullspace* of A, $\text{leftnullspace}(A) = \{y \in \mathbf{R}^m : y^T A = O^T\}$.

To find bases for each of the four spaces, perform Gaussian elimination on A to obtain a matrix A' in row-reduced form. By multiplying the matrices corresponding to the various row operations, determine the square invertible matrix M such that $MA = A'$. The leading nonzero entries in each nonzero row of A' are called the *pivot entries*. The rows in which they appear are called the *pivot rows*, and the columns in which they appear are called the *pivot columns*. You should be able to verify the following assertions:

1. To obtain a basis for the column space of A select the columns of A corresponding to the pivot columns of A'. Note that A and A' do not necessarily have the same column space, but their column spaces have the same dimension, namely, the number of pivot entries.

2. The nonzero rows of A' form a basis for the row space of A. Hence they have the same dimension, namely, the number of pivot entries.

3. Matrices A and A' have the same nullspace. There is one basis vector for each nonpivot column A'_s of A': Set $w_s = 1$ and $w_j = 0$ for all other nonpivot columns A_j. Then determine the unique multipliers w_j for the pivot columns of A'_j to solve $A'w = O$. So the dimension of the nullspace equals the number of nonpivot columns.

4. Suppose the zero rows of A' are precisely the last k rows of A'. Then the last k rows of M form a basis for the left nullspace of A. In particular, the dimension of the left nullspace equals the number of nonpivot rows.

As an immediate consequence we have:

Theorem 8.2

1. *The dimension of the row space equals the dimension of the column space. This number is also called the* rank *of A.*

2. *The dimension of the column space plus the dimension of the nullspace equals the number of columns.*

3. *The dimension of the row space plus the dimension of the left nullspace equals the number of rows.*

8.3 Graphs and Digraphs

Definition 8.3 A *directed graph (digraph)* $G = (V, E)$ is a finite set $V = V(G)$ of *vertices* and a finite set $E = E(G)$ of *edges*. Associated with each edge e is an ordered pair (u, v) of (usually distinct) vertices. We way u is the *tail* of e, v is the *head* of e, and u and v are the *endpoints* of e. If u and v are distinct, we may sometimes write $e = uv$ if there is no other edge having the same tail as e and the same head as e. If $u = v$ then e is called a *loop*. In this course, unless otherwise indicated, we will consider only digraphs without loops. We can represent a directed graph by a drawing, using points for vertices and arrows for edges, with the arrow pointing from the tail to the head.

If each edge is associated with an unordered (instead of ordered) pair of endpoints, then we say that we have an *(undirected) graph*.

Definition 8.4 A *subgraph* of a digraph $G = (V, E)$ is a digraph $G' = (V', E')$ where $V' \subseteq V$, $E' \subseteq E$.

A *path* in a digraph is an alternating sequence $v_0, e_1, v_1, e_2, v_2, \ldots, e_k, v_k$ of vertices and edges, where the vertices v_i are all distinct, the edges e_i are all distinct, and v_{i-1} and v_i are the two endpoints of e_i (either one of v_i, v_j could be the tail, the other being the head). If you have an alternating sequence in which $k \geq 1$, $v_0 = v_k$ and otherwise the vertices and edges are distinct, the sequence is called a *cycle*. We also often identify paths and cycles with just the sets of edges in them.

Two vertices are *connected* if there is a path from one to the other. The set of all vertices connected to a given vertex is a *component* of the digraph. A digraph is *connected* if it has only one component.

A *forest* is a graph containing no cycle (i.e., *acylic*). A *tree* is a connected forest. A *spanning forest* of a digraph is a subgraph that is a forest containing every vertex of the digraph. A connected spanning forest of a digraph is a *spanning tree*.

A *twig* of a digraph is an edge that is not a loop and has an endpoint that is not the endpoint of any other edge in the digraph.

Theorem 8.5 *Forests with at least one edge have twigs.*

Proof. Assume the contrary. Choose any vertex in the forest that is the endpoint of some edge in the forest. Begin making a path. If every vertex encountered is new, at every new vertex there will be another vertex having that vertex as an endpoint that can be used to continue the path. Since there is only a finite number of vertices, this cannot continue forever. So at some point a vertex will be encountered for a second time. But then a cycle will be detected, which is also impossible. □

Definition 8.6 Associated with a digraph having no loops is the *vertex-edge incidence matrix* A. Rows are indexed by vertices, columns by edges. The entry in row v column e is -1 if v is the tail of e, $+1$ if v is the head of e, and zero otherwise. (Note that other texts may interchange -1 and 1 in the definition.)

Theorem 8.7 *For a digraph, the dimension of the left nullspace of A equals the number of components of G.*

Proof. An element of the left nullspace is an assignment of numbers y_v to vertices v such that for every edge $e = uv$, $y_v - y_u = 0$; i.e., $y_v = y_u$. By connectivity, the same number must be assigned to every vertex in a given component. But different components may have vertices with different numbers. □

For a digraph, an element of the nullspace of A corresponds to an assignment of numbers x_e to edges $e \in E(G)$ such that for every vertex v,

$$\sum_{e\in\delta^-(v)} x_e - \sum_{e\in\delta^+(v)} x_e = 0.$$

We are using the notation $\delta^+(v)$ to denote the set of edges whose tails are v (the set of edges leaving v), and $\delta^-(v)$ to denote the set of edges whose heads are v (the set of edges entering v). That is to say, for every vertex v, the sum of the numbers on edges entering v equals the sum of the numbers on edges leaving v. The above equations are known as the *flow-conservation* equations because by interpreting the numbers on the edges as flows, the equations state that what flows in at each vertex must flow out. (Regard negative flow as positive flow in the opposite direction.)

Theorem 8.8 *If a subset of edges contains a cycle, then the corresponding subset of columns of A is dependent.*

Proof. Trace the cycle in some direction, assigning $+1$ to edges traversed in the forward direction, and -1 to edges traversed in the reverse direction. □

Theorem 8.9 *If a subset S of columns of A is dependent, then the corresponding subset of edges contains a cycle.*

Proof. Let x be a nonzero solution to $Ax = O$. Without loss of generality assume that $x_e \neq 0$ precisely when $e \in S$. From the flow-conservations equations we can deduce that there can be no twig in the subgraph of G determined by the edges in S. So the set of edges of S cannot be a forest, and hence contains a cycle. □

Exercise 8.10 Prove that the following are equivalent for a digraph G with at least one edge. Try using some of the properties of the dimensions of the vector spaces associated with the vertex-edge incidence matrix A of G.

1. G is a tree.
2. G is minimally connected; i.e., G is connected, but no subgraph with the same vertex set and fewer edges is connected.
3. G is maximally acyclic; i.e, G is acyclic, but no supergraph with the same vertex set and more edges is acyclic.
4. $|V(G)| = |E(G)| + 1$ and G is connected.
5. $|V(G)| = |E(G)| + 1$ and G is acyclic.

Exercise 8.11 Let A be the vertex-edge incidence matrix of a digraph G with at least one edge. Let M be any square submatrix of A, determined by selecting any subsets of equal numbers of rows and columns of A, not necessarily adjacent. Prove that the determinant of M is 0, 1, or -1. Suggestion: Recall how to calculate a determinant by expansion along a column.

8.4 Systems of Equations

Definition 8.12 Let A be an $m \times n$ matrix. For a subset $B \subseteq \{1, \ldots, n\}$ we let A_B denote the submatrix of A consisting only of those columns indexed by B. Similarly, if $x \in \mathbf{R}^n$ we let x_B denote the subvector of x consisting only of those components of x indexed by B.

Given an $m \times n$ matrix A and a vector $b \in \mathbf{R}^m$ we can determine whether or not b is in the column space of A and, if so, find a particular solution $\overline{x}$ to $Ax = b$ using Gaussian elimination. The solution $\overline{x}$ will satisfy $A_B\overline{x}_B = b$ and $\overline{x}_N = O$, where $N = \{1, \ldots, n\} \setminus B$. We can also find a basis for $\{\overline{w}^1, \ldots, \overline{w}^p\}$ for the nullspace of A. Then the complete set of solutions to $Ax = b$ is given by

$$\{\overline{x} + t_1\overline{w}^1 + \cdots + t_p\overline{w}^p : t_1, \ldots, t_p \in \mathbf{R}\}.$$

Actually, we can do the above for any set B of indices of columns forming a basis for the column space of A. Such a solution $\overline{x}$ is called the *basic solution* associated with B, and the vectors $w^1, \ldots, w^p$ are called the *basic directions* associated with B.

Example 8.13 Consider the system $Ax = b$ below:

$$\begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 120 \\ 70 \\ 100 \end{bmatrix}.$$

Here are some choices for B and the corresponding basic solutions and basic directions, each of which can easily be computed by hand in this small example:

1. $B = \{3,4,5\}$, $\overline{x} = (0,0,120,70,100)$, $\overline{w}^1 = (1,0,-1,-1,-2)$, $\overline{w}^2 = (0,1,-2,-1,-1)$.
2. $B = \{2,3,5\}$, $\overline{x} = (0,70,-20,0,30)$, $\overline{w}^1 = (1,-1,1,0,-1)$, $\overline{w}^2 = (0,-1,2,1,1)$.
3. $B = \{1,2,3\}$, $\overline{x} = (30,40,10,0,0)$, $\overline{w}^1 = (1,-2,3,1,0)$, $\overline{w}^2 = (-1,1,-1,0,1)$.
4. $B = \{1,2,4\}$, $\overline{x} = (\frac{80}{3},\frac{140}{3},0,-\frac{10}{3},0)$, $\overline{w}^1 = (\frac{1}{3},-\frac{2}{3},1,\frac{1}{3},0)$, $\overline{w}^2 = (-\frac{2}{3},\frac{1}{3},0,\frac{1}{3},1)$.
5. $B = \{1,2,5\}$, $\overline{x} = (20,50,0,0,10)$, $\overline{w}^1 = (1,-1,1,0,-1)$, $\overline{w}^2 = (-2,1,0,1,3)$.

So, for example, from $B = \{1,2,5\}$ we obtain the complete set of solutions

$$\left\{ \begin{bmatrix} 20 \\ 50 \\ 0 \\ 0 \\ 10 \end{bmatrix} + t_1 \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ -1 \end{bmatrix} + t_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \\ 3 \end{bmatrix} : t_1, t_2 \in \mathbf{R} \right\}.$$

In the special case that the matrix A has full row rank (the rows are linear independent), then for every choice of indices B for the column space, the matrix A_B will be square and invertible. In this case we can derive formulas for $\overline{x}$ and the vectors $\overline{w}^i$.

Since $A_B\overline{x}_B = b$, we have

$$\begin{aligned} \overline{x}_B &= A_B^{-1}b \\ \overline{x}_N &= O \end{aligned}$$

Since each basic direction $\overline{w}$ is obtained by setting $\overline{w}_s = 1$ for some $s \in N$, $\overline{w}_j = 0$ for $j \in N \setminus \{s\}$, and then solving $A\overline{w} = O$, we have $A_B\overline{w}_B + A_s = O$ so

$$\begin{aligned} \overline{w}_B &= -A_B^{-1}A_s \\ \overline{w}_s &= 1 \\ \overline{w}_j &= 0,\ j \in N \setminus \{s\} \end{aligned}$$

8.5 Solving Linear Programs

In this section we finally begin to discuss how to solve linear programs. Let's start with a linear program in standard form

$$(\hat{P}) \quad \begin{array}{c} \max z = \hat{c}^T\hat{x} \\ \text{s.t. } \hat{A}\hat{x} \le b \\ \hat{x} \ge O \end{array}$$

where $\hat{A}$ is an $m \times n$ matrix.

In summation notation, $(\hat{P})$ is of the form

$$\begin{array}{c} \max z = \sum_{j=1}^{n} c_j x_j \\ \text{s.t. } \sum_{j=1}^{n} a_{ij}x_j \le b_i, \quad i = 1, \ldots, m \\ x_j \ge 0, \quad j = 1, \ldots, n \end{array}$$

The first step will be to turn this system into a system of equations by introducing m nonnegative *slack* variables, one for each inequality in $\hat{A}\hat{x} \le b$:

$$\begin{array}{c} \max z = \sum_{j=1}^{n} c_j x_j \\ \text{s.t. } (\sum_{j=1}^{n} a_{ij}x_j) + x_{n+i} = b_i, \quad i = 1, \ldots, m \\ x_j \ge 0, \quad j = 1, \ldots, n+m \end{array}$$

Now we have a problem of the form

$$(P) \quad \begin{array}{c} \max c^T x \\ \text{s.t. } Ax = b \\ x \ge O \end{array}$$

where $x = (\hat{x}, x_{n+1}, \ldots, x_{n+m})$, $c = (\hat{c}, 0, \ldots, 0)$, and $A = [\hat{A}|I]$. In particular, the rows of A are linearly independent (A has *full row rank*).

Example 8.14 With the addition of slack variables, the GGMC problem becomes

$$\begin{array}{c} \max z = 5x_1 + 4x_2 \\ \text{s.t. } x_1 + 2x_2 + x_3 = 120 \\ x_1 + x_2 + x_4 = 70 \\ 2x_1 + x_2 + x_5 = 100 \\ x_1, \ldots, x_5 \ge 0 \end{array}$$

We are going to solve this problem by moving through a sequence of basic solutions by following a sequence of basic directions.

1. Begin with the easy basis $\{3,4,5\}$. Compute the associated basic solution and basic directions: $\overline{x} = (0,0,120,70,100)$, $\overline{w}^1 = (1,0,-1,-1,-2)$, $\overline{w}^2 = (0,1,-2,-1,-1)$. Notice that $\overline{x}$ is a *nonnegative* basic solution. We call such solutions *basic feasible solutions*, and such a basis a *(primal) feasible basis*. Using the objective function, determine the values of each of these: $c^T\overline{x} = \$0$, $c^T\overline{w}^1 = \$5$, $c^T\overline{w}^2 = \$4$. So if we start with the solution $\overline{x}$ and move in the direction of $\overline{w}^1$ along the ray $\overline{x} + t\overline{w}^1 = (0+t, 0, 120-t, 70-t, 100-2t)$, we gain \$5 for each unit increase in t. How high can we make t without having any of our variables become negative? We clearly need to be only concerned with the variables that are dropping in value, and in this case we see that t must be the minimum of $\{120, 70, \frac{100}{2}\}$; namely, 50.

   Setting $t = 50$ we get the new solution $(50,0,70,20,0)$. So we went from a solution involving basis $\{3,4,5\}$, used the linear relation among columns $\{1,3,4,5\}$, and noticed that dropping column 5 leaves a new basis $\{1,3,4\}$ with associated basic feasible solution $(50,0,70,20,0)$. At this point it would be wise to take a second look at Exercise 2.14(3).

2. The basis $\{1,3,4\}$ has the associated basic feasible solution and basic directions: $\overline{x} = (50,0,70,20,0)$ with value \$250, $\overline{w}^1 = (-\frac{1}{2}, 1, -\frac{3}{2}, -\frac{1}{2}, 0)$ with value $\$\frac{3}{2}$, $\overline{w}^2 = (-\frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, 1)$ with value $\$-\frac{5}{2}$. If we move from $\overline{x}$ in the direction of $\overline{w}^1$ along the ray $\overline{x} + t\overline{w}^1 = (50 - \frac{1}{2}t, 0+t, 70 - \frac{3}{2}t, 20 - \frac{1}{2}t, 0)$, we gain $\$\frac{3}{2}$ for each unit increase in t. How high can we make t without having any of our variables become negative? In this case it is the minimum of $\{50/\frac{1}{2}, 70/\frac{3}{2}, 20/\frac{1}{2}\}$; namely, 40.

   Setting $t = 40$ we get the new solution $(30,40,10,0,0)$. So we went from a solution involving basis $\{1,3,4\}$, used the linear relation among columns $\{1,2,3,4\}$, and noticed that dropping column 4 leaves a new linearly independent set of columns $\{1,2,3\}$ with associated basic feasible solution $(30,40,10,0,0)$.

3. The basis $\{1,2,3\}$ has the associated basic feasible solution and basic directions: $\overline{x} = (30,40,10,0,0)$ with value \$310, $\overline{w}^1 = (1,-2,3,1,0)$ with value $\$-3$, $\overline{w}^2 = (-1,1,-1,0,1)$ with value $\$-1$. If we move from $\overline{x}$ in either direction $\overline{w}^1$ or $\overline{w}^2$ or any nonnegative combination of these, we cannot get a higher value. On the other hand, since every solution to the set of equations is of the form $\overline{x} + t_1\overline{w}^1 + t_2\overline{w}^2$ and we require nonnegative variables, we cannot let t_1 become negative without making x_4 negative, and we cannot let t_2 become negative without making x_5 negative. So we are

restricted to considering only nonnegative t_1, t_2. Therefore there is no solution with higher value than the one we have. $\square$

This gives an idea of the *simplex method* to solve linear programs. If you have a basis B with an associated basic feasible solution, determine the basic directions and their values. If no value is positive, you have arrived at an optimal solution. Otherwise move in the direction of one of the basic directions with positive value until some basic variable is forced to zero. If no variable decreases to zero, then the basic direction $\overline{w}$ is nonnegative. Such a basic direction is called a *basic feasible direction*. Then you have found a ray of feasible solutions with ever-increasing objective function value; hence the linear program is unbounded. If, on the other hand, some variable decreases to zero, drop the index of that column from the basis and add the index of the nonbasic column associated with the basic direction.

There are still some details to be worked out, including how to make choices to insure that the algorithm will terminate, and how to get a basis with a basic feasible solution in the first place.

8.6 The Revised Simplex Method

Another question has to do with the computations involved. Let us suppose that our matrix A has full row rank. Given a basis B, we already have formulas for the basic solution and the basic directions. Let's think about the costs of the basic directions. Suppose we have basic direction $\overline{w}$ associated with basis B and nonbasic column s. Recall that $\overline{w}_B = -A_B^{-1}A_s$, $\overline{w}_s = 1$, and $\overline{w}_j = 0$ otherwise. So its cost is $c^T\overline{w} = c_s\overline{w}_s + c_B^T\overline{w}_B = c_s - c_B^T A_B^{-1}A_s$.

Define $\overline{y}^T = c_B^T A_B^{-1}$. Thus $\overline{y}$ is the unique solution to the equation $y^T A_B = c_B^T$. Then the cost of $\overline{w}$ becomes $c_s - \overline{y}^T A_s$. By computing $\overline{y}$ first, we can screen our possible directions to determine if any have positive value. If we find that there is a basic direction with positive value, we can then compute it and use it to move to another basic feasible solution.

The general procedure to make a step (also known as a *pivot*) in the revised simplex method can be described as follows:

At some stage in the simplex method we have a feasible basis B and an associated basic feasible solution $\overline{x}$. To get $\overline{y}^T = c_B^T A_B^{-1}$, we solve $y^T A_B = c_B^T$. Then we can calculate $\overline{c}_N = c_N^T - \overline{y}^T A_N$. These numbers are known as the *reduced costs* of the nonbasic variables. If $\overline{c}_N \leq O$, then $\overline{x}$ is optimal. If not, then $\overline{c}_s > 0$ for some $s \in N$. Find $\overline{w}$ by solving $A_B\overline{w}_B = -A_s$, and setting $\overline{w}_s = 1$ and $\overline{w}_j = 0$ for all other $j \in N \setminus s$. If $\overline{w} \geq O$ then (P) is unbounded. Otherwise find the largest value of t for which $\overline{x} + t\overline{w}$ remains nonnegative. Calculate the new basic feasible solution with this value of t. Suppose the previously basic variable x_r drops to zero. Change B by dropping r and including s.

Example 8.15 Solving GGMC using the revised simplex method.

1. Begin with the basis $B = \{3,4,5\}$ and associated basic feasible solution $\overline{x} = (0,0,120,70,100)$.

   Our starting basis is $(3,4,5)$, So

$$A_B = \begin{array}{c} \begin{array}{ccc} 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

   Find $\overline{y}$ by solving $y^T A_B = c_B^T$:

$$\begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{array}{c} \begin{array}{ccc} 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} = \begin{array}{c} \begin{array}{ccc} 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \end{array}$$

   Thus $\overline{y} = (0,0,0)^T$.

   Calculate $\overline{c}_N^T = c_N^T - \overline{y}^T A_N$.

$$\begin{bmatrix} \overline{c}_1 & \overline{c}_2 \end{bmatrix} = \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{bmatrix} 5 & 4 \end{bmatrix} \end{array} - \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \end{array} = \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{bmatrix} 5 & 4 \end{bmatrix} \end{array}$$

   Since 5 is positive we can choose the basic direction $\overline{w}$ associated with column $s = 1$; i.e., x_1 will be the entering variable. Set $\overline{w}_1 = 1$ and $\overline{w}_2 = 0$. Solve $A_B w_B = -A_s$:

$$\begin{array}{c} \begin{array}{ccc} 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \begin{bmatrix} w_3 \\ w_4 \\ w_5 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -2 \end{bmatrix}$$

   Thus $\overline{w} = (1,0,-1,-1,-2)$.

To find the new solution, write $\hat{x} = \overline{x} + t\overline{w}$:

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \\ \hat{x}_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 120 \\ 70 \\ 100 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ -1 \\ -1 \\ -2 \end{bmatrix}$$

Therefore $t = 50$, x_5 is the leaving variable, $\{1,3,4\}$ is the new basis, $\overline{x} = (50,0,70,20,0)$ is the new basic feasible solution, and

$$B = \begin{array}{c} \begin{matrix} 1 & 3 & 4 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix} \end{array}$$

is the new basis matrix A_B.

2. Find $\overline{y}$ by solving $y^T A_B = c_B^T$:

$$\begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{array}{c} \begin{matrix} 1 & 3 & 4 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix} \end{array} = \begin{array}{c} \begin{matrix} 1 & 3 & 4 \end{matrix} \\ \begin{bmatrix} 5 & 0 & 0 \end{bmatrix} \end{array}$$

Thus $\overline{y} = (0,0,2.5)^T$.
Calculate $\overline{c}_N^T = c_N - \overline{y}^T A_N$.

$$\begin{bmatrix} \overline{c}_2 & \overline{c}_5 \end{bmatrix} = \begin{array}{c} \begin{matrix} 2 & 5 \end{matrix} \\ \begin{bmatrix} 4 & 0 \end{bmatrix} \end{array} - \begin{bmatrix} 0 & 0 & 2.5 \end{bmatrix} \begin{array}{c} \begin{matrix} 2 & 5 \end{matrix} \\ \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \end{array} = \begin{array}{c} \begin{matrix} 2 & 5 \end{matrix} \\ \begin{bmatrix} 1.5 & -2.5 \end{bmatrix} \end{array}$$

Since 1.5 is positive we can choose the basic direction $\overline{w}$ associated with column $s = 2$; i.e., x_2 will be the entering variable. Set $\overline{w}_2 = 1$ and $\overline{w}_5 = 0$. Solve $A_B w_B = -A_s$:

$$\begin{array}{c} \begin{matrix} 1 & 3 & 4 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix} \end{array} \begin{bmatrix} w_1 \\ w_3 \\ w_4 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ -1 \end{bmatrix}$$

Thus $\overline{w} = (-0.5, 1, -1.5, -0.5, 0)$.

To find the new solution, write $\hat{x} = \overline{x} + t\overline{w}$:

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \\ \hat{x}_5 \end{bmatrix} = \begin{bmatrix} 50 \\ 0 \\ 70 \\ 20 \\ 0 \end{bmatrix} + t \begin{bmatrix} -0.5 \\ 1 \\ -1.5 \\ -0.5 \\ 0 \end{bmatrix}$$

Therefore $t = 40$, x_4 is the leaving variable, $\{1, 2, 3\}$ is the new basis, $\overline{x} = (30, 40, 10, 0, 0)$ is the new basic feasible solution, and

$$B = \begin{array}{c} \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{bmatrix} \end{array}$$

is the new basis matrix A_B.

3. Find $\overline{y}$ by solving $y^T A_B = c_B^T$:

$$\begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{array}{c} \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{bmatrix} \end{array} = \begin{array}{c} \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{bmatrix} 5 & 4 & 0 \end{bmatrix} \end{array}$$

Thus $\overline{y} = (0, 3, 1)^T$.

Calculate $\overline{c}_N^T = c_N - \overline{y}^T A_N$.

$$\begin{bmatrix} \overline{c}_4 & \overline{c}_5 \end{bmatrix} = \begin{array}{c} \begin{matrix} 4 & 5 \end{matrix} \\ \begin{bmatrix} 0 & 0 \end{bmatrix} \end{array} - \begin{bmatrix} 0 & 3 & 1 \end{bmatrix} \begin{array}{c} \begin{matrix} 4 & 5 \end{matrix} \\ \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array} = \begin{array}{c} \begin{matrix} 4 & 5 \end{matrix} \\ \begin{bmatrix} -3 & -1 \end{bmatrix} \end{array}$$

Since $\overline{c}$ is nonpositive, our current solution $\overline{x}$ is optimal.

$\square$

Index