Chapter 4

Duality

Given any linear program, there is another related linear program called the *dual*. In this chapter, we will develop an understanding of the dual linear program. This understanding translates to important insights about many optimization problems and algorithms. We begin in the next section by exploring the main concepts of duality through the simple graphical example of building cars and trucks that was introduced in Section 3.1.1. Then, we will develop the theory of duality in greater generality and explore more sophisticated applications.

4.1 A Graphical Example

Recall the linear program from Section 3.1.1, which determines the optimal numbers of cars and trucks to build in light of capacity constraints. There are two decision variables: the number of cars x_1 in thousands and the number of trucks x_2 in thousands. The linear program is given by

$$
\begin{array}{lll}
\text{maximize} & 3x_1 + 2.5x_2 & \text{(profit in thousands of dollars)} \\
\text{subject to} & 4.44x_1 \leq 100 & \text{(car assembly capacity)} \\
 & 6.67x_2 \leq 100 & \text{(truck assembly capacity)} \\
 & 4x_1 + 2.86x_2 \leq 100 & \text{(metal stamping capacity)} \\
 & 3x_1 + 6x_2 \leq 100 & \text{(engine assembly capacity)} \\
 & x \geq 0 & \text{(nonnegative production).}
\end{array}
$$

The optimal solution is given approximately by $x_1 = 20.4$ and $x_2 = 6.5$, generating a profit of about \$77.3 million. The constraints, feasible region, and optimal solution are illustrated in Figure 4.1.

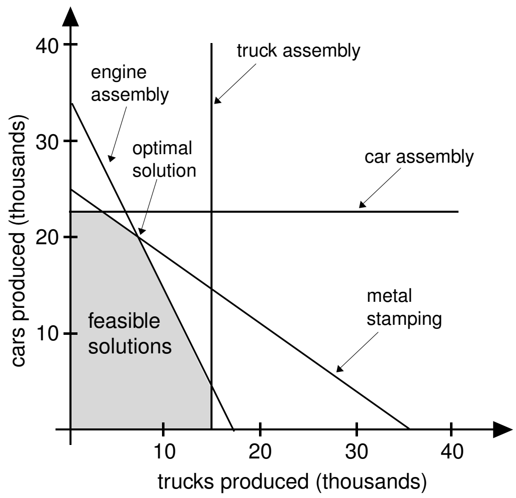

Figure 4.1: The constraints, feasible region, and optimal solution of the linear program associated with building cars and trucks.

Written in matrix notation, the linear program becomes

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \geq 0, \end{array}$$

where

$$c = \left[\begin{array}{c} 3 \\ 2.5 \end{array}\right], \qquad A = \left[\begin{array}{cc} 4.44 & 0 \\ 0 & 6.67 \\ 4 & 2.86 \\ 3 & 6 \end{array}\right] \quad \text{and} \quad b = \left[\begin{array}{c} 100 \\ 100 \\ 100 \\ 100 \end{array}\right].$$

The optimal solution of our problem is a basic feasible solution. Since there are two decision variables, each basic feasible solution is characterized by a set of two linearly independent binding constraints. At the optimal solution, the two binding constraints are those associated with metal stamping and engine assembly capacity. Hence, the optimal solution is the unique solution to a pair of linear equations:

$$\begin{array}{ll} 4x_1 + 2.86x_2 = 100 & \text{(metal stamping capacity is binding)} \\ 3x_1 + 6x_2 = 100 & \text{(engine assembly capacity is binding).} \end{array}$$

In matrix form, these equations can be written as $\overline{A}x = \overline{b}$, where

$$\overline{A} = \left[\begin{array}{c} (A_{3*})^T \\ (A_{4*})^T \end{array}\right] \quad \text{and} \quad \overline{b} = \left[\begin{array}{c} b_3 \\ b_4 \end{array}\right].$$

Note that the matrix $\overline{A}$ has full rank. Therefore, it has an inverse $\overline{A}^{-1}$. Through some calculations, we get (approximately)

$$\overline{A}^{-1} = \begin{bmatrix} 0.389 & -0.185 \\ -0.195 & 0.259 \end{bmatrix}.$$

The optimal solution of the linear program is given by $x = \overline{A}^{-1}b$, and therefore, the optimal profit is $c^T\overline{A}^{-1}\overline{b} = 77.3$.

4.1.1 Sensitivity Analysis

Suppose we wish to increase profit by expanding manufacturing capacities. In such a situation, it is useful to think of profit as a function of a vector $\Delta \in \Re^4$ of changes to capacity. We denote this profit by $z(\Delta)$, defined to be the maximal objective value associated with the linear program

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b + \Delta \\ & x \geq 0. \end{array} \tag{4.1}$$

Hence, the maximal profit in our original linear program is equal to $z(0)$. In this section, we will examine how incremental changes in capacities influence the optimal profit $z(\Delta)$. The study of such changes is called *sensitivity analysis.*

Consider a situation where the metal stamping and engine assembly capacity constraints are binding at the optimal solution to the linear program (4.1). Then, this optimal solution must be given by $x = \overline{A}^{-1}(\overline{b} + \Delta)$, and the optimal profit must be $z(\Delta) = c^T\overline{A}^{-1}(\overline{b} + \overline{\Delta})$, where

$$\overline{\Delta} = \begin{bmatrix} \Delta_3 \\ \Delta_4 \end{bmatrix}.$$

Furthermore, the difference in profit is $z(\Delta) - z(0) = c^T\overline{A}^{-1}\overline{\Delta}$.

This matrix equation provides a way to gauge the impact of changes in capacities on optimal profit in the event that the set of binding constraints does not change. It turns out that this also gives us the information required to conduct sensitivity analysis. This is because small changes in capacities will not change which constraints are binding. To understand why, consider the illustration in Figure 4.2, where the engine assembly capacity is increased by a small amount. Clearly, the new optimal solution is still at the intersection where metal stamping and engine assembly capacity constraints are binding. Similarly, though not illustrated in the figure, one can easily see that incremental changes in any of the other capacity constraints will not change the fact that metal stamping and engine assembly capacity constraints are binding.

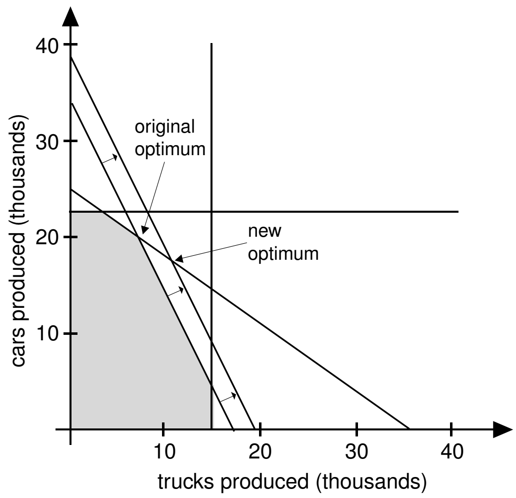

Figure 4.2: Changes in the optimal solution brought about by a small increase in capacity for engine assembly.

This observation does not hold when we consider large changes. As illustrated in Figure 4.3, sufficiently large changes can result in a different set of binding constraints. The figure shows how after a large increase in engine assembly capacity, the associated constraint is no longer binding. Instead, the truck assembly capacity constraint becomes binding.

The *sensitivity* y_i of profit to quantity of the ith resource is the rate at which $z(\Delta)$ increases as Δ_i increases, starting from $\Delta_i = 0$. It is clear that small changes in non binding capacities do not influence profit. Hence, $y_1 = y_2 = 0$. From the preceding discussion, we have $z(\Delta) - z(0) = c^T \overline{A}^{-1} \overline{\Delta}$, and therefore

$$\begin{bmatrix} y_3 & y_4 \end{bmatrix} = c^T \overline{A}^{-1} = \begin{bmatrix} 3 & 2.5 \end{bmatrix} \begin{bmatrix} 0.389 & -0.185 \\ -0.195 & 0.259 \end{bmatrix} = \begin{bmatrix} 0.681 & 0.092 \end{bmatrix}.$$

In other words, the sensitivity is about \$0.681 million per percentage of metal stamping capacity and \$0.092 million per percentage of engine assembly capacity. If a 1% increase in metal stamping capacity requires the same investment as a 1% increase in engine assembly, we should invest in metal stamping.

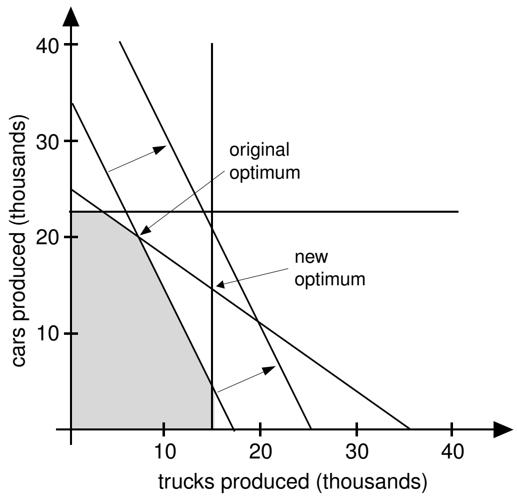

Figure 4.3: Changes in the optimal solution brought about by a large increase in capacity for engine assembly.

4.1.2 Shadow Prices and Valuation of the Firm

The sensitivities of profit to resource quantities are commonly called *shadow prices*. Each ith resource has a shadow price y_i. In our example of building cars and trucks, shadow prices for car and truck assembly capacity are zero. Shadow prices of engine assembly and metal stamping capacity, on the other hand, are \$0.092 and \$0.681 million per percent. Based on the discussion in the previous section, if the metal stamping and engine assembly capacity constraints remain binding when resource quantities are set at $b + \Delta$, the optimal profit is given by $z(\Delta) = z(0) + y^T \Delta$.

A shadow price represents the maximal price at which we should be willing to buy additional units of a resource. It also represents the minimal price at which we should be willing to sell units of the resource. A shadow price might therefore be thought of as the value per unit of a resource. Remarkably, if we compute the value of our entire stock of resources based on shadow prices, we get our optimal profit! For instance, in our example of building cars and trucks, we have

$$0.092 \times 100 + 0.681 \times 100 = 77.3.$$

As we will now explain, this is not just a coincidence but reflects a fundamental property of shadow prices.

From the discussion above we know that as long as the metal stamping and engine assembly constraints are binding, that $z(\Delta) = z(0) + y^T \Delta$. If we let $\Delta = -b$, then the resulting linear program has 0 capacity at each plant, so the optimal solution is 0, with associated profit of 0. Moreover, both the metal stamping and engine assembly constraints are still binding. This means that $0 = z(-b) = z(0) + y^T(-b)$. Rearranging this gives that $z(0) = y^T b$. This is a remarkable fundamental result: the net value of our current resources, valued at their shadow prices, is equal to the maximal profit that we can obtain through operation of the firm – i.e., the value of the firm.

4.1.3 The Dual Linear Program

Shadow prices solve another linear program, called the *dual*. In order to distinguish it from the dual, the original linear program of interest – in this case, the one involving decisions on quantities of cars and trucks to build in order to maximize profit – is called the *primal*. We now formulate the dual.

To understand the dual, consider a situation where we are managing the firm but do not know linear programming. Therefore, we do not know exactly what the optimal decisions or optimal profit are. Company X approaches us and expresses a desire to purchase capacity at our factories. We enter into a negotiation over the prices $y \in \Re^4$ that we should charge per percentage of capacity at each of our four factories.

To have any chance of interesting us, the prices must be nonnegative: $y \geq 0$. We also argue that there are fixed bundles of capacity that we can use to manufacture profitable products, and the prices y must be such that selling such a bundle would generate at least as much money as manufacturing the product. In other words, we impose requirements that

$$4.44y_1 + 4y_3 + 3y_4 \geq 3 \quad \text{and} \quad 6.67y_2 + 2.86y_3 + 6y_4 \geq 2.5.$$

The first constraint ensures that selling a bundle of capacity that could be used to produce a car is at least as profitable as producing the car. The second constraint is the analog associated with production of trucks.

Given our requirements, Company X solves a linear program to determine prices that minimize the amount it would have to pay to purchase all of our capacity:

$$\begin{array}{lll}
\text{minimize} & 100y_1 + 100y_2 + 100y_3 + 100y_4 & \text{(cost of capacity)} \\
\text{subject to} & 4.44y_1 + 4y_3 + 3y_4 \geq 3 & \text{(car production)} \\
 & 6.67y_2 + 2.86y_3 + 6y_4 \geq 2.5 & \text{(truck production)} \\
 & y \geq 0 & \text{(nonnegative prices).}
\end{array}$$

In matrix notation, we have

$$\begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \geq c \\ & y \geq 0. \end{array}$$

The optimal solution to this linear program is

$$y = \begin{bmatrix} 0 \\ 0 \\ 0.092 \\ 0.681 \end{bmatrix},$$

and the minimal value of the objective function is 77.3. Remarkably, we have recovered the shadow prices and the optimal profit!

It is not a coincidence that the minimal cost in the dual equals the optimal profit in the primal and that the optimal solution of the dual is the vector of shadow prices – these are fundamental relations between the primal and the dual. We offer an intuitive explanation now and a more in-depth analysis in the next section.

The constraints ensure that we receive at least as much money from selling as we would from manufacturing. Therefore, it seems clear that the minimal cost in the dual is at least as large as the maximal profit in the primal. This fact is known as *weak duality*. Another result, referred to as *strong duality*, asserts that the minimal cost in the dual *equals* the maximal profit in the primal. This is not obvious. It is motivated to some extent, though, by the fact that Company X is trying to get the best deal it can. It is natural to think that if Company X negotiates effectively, it should be able to acquire all our resources for an amount of money equal that we would obtain as profit from manufacturing. This would imply strong duality.

Why, now, should an optimal solution to the dual provide shadow prices? To see this, consider changing the resource quantities by a small amount $\Delta \in \Re^4$. Then, the primal and dual become

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b + \Delta \\ & x \geq 0 \end{array} \qquad \text{and} \qquad \begin{array}{ll} \text{minimize} & (b+\Delta)^T y \\ \text{subject to} & A^T y \geq c \\ & y \geq 0. \end{array}$$

The maximal profit in the primal and the minimal cost in the dual are both equal to $z(\Delta)$. Suppose that the optimal solution to the dual is unique – as is the case in our example of building cars and trucks. Then, for sufficiently small Δ, the optimal solution to the dual should not change, and therefore the optimal profit should change by $z(\Delta) - z(0) = (b+\Delta)^T y - b^T y = \Delta^T y$. It follows that the optimal solution to the dual is the vector of shadow prices.

4.2 Duality Theory

In this section, we develop weak and strong duality in general mathematical terms. This development involves intriguing geometric arguments. Developing intuition about the geometry of duality is often helpful in generating useful insights about optimization problem.

Duality theory applies to general linear programs, that can involve greater-than, less-than, and equality constraints. However, to keep things simple, we will only study in this section linear programs that are in *symmetric form.* Such linear programs take the form:

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \geq 0. \end{array}$$

for some matrix $A \in \Re^{M \times N}$ and vectors $b \in \Re^M$ and $c \in \Re^N$. The decision variables – called the *primal variables* – make up a vector $x \in \Re^N$. As we will discuss later in the chapter, general linear programs can be converted to symmetric form, so our development of duality theory in this context also applies to general linear programs.

The dual of a symmetric form linear program takes the form

$$\begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \geq c \\ & y \geq 0. \end{array}$$

The decision variables – called the *dual variables* – form a vector $y \in \Re^M$.

Note that each decision variable in the primal problem corresponds to a constraint in the dual problem, and each constraint in the primal problem corresponds to a variable in the dual problem.

4.2.1 Weak Duality

Suppose that x is a feasible solution of the primal and y is a feasible solution of the dual. Then, $Ax \leq b$, $y^T A \geq c^T$, $x \geq 0$, and $y \geq 0$. It follows that $y^T Ax \geq c^T x$ and $y^T Ax \leq y^T b$. Hence, $c^T x \leq b^T y$. This is the weak duality theorem, which we state below:

Theorem 4.2.1. (Weak Duality) *For any feasible solutions x and y to primal and dual linear programs, $c^T x \leq b^T y$.*

The following theorem states one immediate implication of weak duality.

Theorem 4.2.2. (Certificate of Optimality) *If x and y are feasible solutions of the primal and dual and $c^T x = b^T y$, then x and y must be optimal solutions to the primal and dual.*

There is another interesting consequence of weak duality that relates infiniteness of profit/cost in the primal/dual with feasibility of the dual/primal, as we now explain. Let y be a feasible solution of the dual. By weak duality, we have $c^T x \leq b^T y$ for all feasible x. If the optimal profit in the primal is ∞, then $\infty \leq b^T y$. This is not possible, so the dual cannot have a feasible solution. The following theorem captures this fact, together with its converse, which can be established via a symmetric argument.

Theorem 4.2.3. (Infiniteness and Feasibility in Duality) *If the optimal profit in the primal is ∞, then the dual must be infeasible. If the optimal cost in the dual is $-\infty$, then the primal must be infeasible.*

4.2.2 Strong Duality

Theorem 4.2.2 asserts that if x and y are feasible solutions of the primal and dual and $c^T x = b^T y$, then x and y must be optimal solutions of the primal and dual. This does not imply that there are feasible solutions x and y such that $c^T x = b^T y$. However the strong duality guarantees this.

Theorem 4.2.4. (Strong Duality) *The dual has an optimal solution if and only if the primal does. If x^* and y^* are optimal solutions to the primal and dual, then $c^T x^* = b^T y^*$.*

Note that here, and throughout the book, when we refer to an optimal solution, it is implicitly assumed to be finite. If the optimal value can get arbitrarily large, we say the objective is unbounded. There are two slightly different sorts of unboundedness we discuss, subtly different. In the first, the feasible region is unbounded, and in this situation we say the *problem* is unbounded. In the second, the objective function can get arbitrarily large, and we say that the *objective value* is unbounded. Note that the second sort of unboundedness implies the first.

In order to prove the Strong Duality Theorem, we first have an aside, and discuss optimality of slightly more general functions.

4.2.3 First Order Necessary Conditions

It is possible to establish the Strong Duality Theorem directly, but the KKT conditions (given later in this section) are useful in their own right, and strong duality is an immediate consequence.

Before given the KKT conditions, we digress still more, and talk about convex sets and hyperplanes. Given two sets, U and V, we say that a hyperplane H *separates* U and V if all of U is on one side of the hyperplane, and all of V is on the other side. In other words, if H is given by $\{x|a^Tx = b\}$, then H separates $\Re^N$ into two sets, $H^+ = \{x|a^Tx \geq b\}$ and H^- $\{x|a^Tx \leq b\}$. H separates U and V if U is contained in H^+ and V is contained in H^- or vice versa.

Theorem 4.2.5. (Separating Hyperplane) *Let U and V be two disjoint convex sets. Then, there exists a hyperplane separating U and V.*

"Picture Proof:" Let $\delta = \inf_{u\in U, v\in V} \|u - v\|$ We will demonstrate the result only for the case where $\delta > 0$ and there is a $u \in U$ and $v \in V$ with $\|u - v\| = \delta$. This case is all that will be needed to cover all the applications we will use of the theorem, and the full result is beyond the scope of this book.

Take $u \in U$ and $v \in V$ with $\|u - v\| = \delta$. Let H be the hyperplane through v that is perpendicular to $u - v$. We claim that H is a separating hyperplane for U and V.

Suppose this were not the case. Then, without loss of generality, we can assume that $v = 0$ (we can translate every point by $-v$). The means that H will be given by $\{x|u^Tx = 0\}$. Note that $0 \in H^-$. Suppose not all of U is in H^+. Then there would be some $v \in U$ with $u^Tv < 0$. If $d = v - u$ and $\alpha = -\frac{u^Td}{d^Td} \in (0, 1)$, and let $w = u + \alpha d = \alpha v + (1 - \alpha)u$. w must be in U because it is a convex combination of things in U, and the length of w is

$$\begin{aligned} w^Tw &= (u + \alpha d)^T(u + \alpha d) \\ &= u^Tu + 2\alpha u^Td + \alpha^2 d^Td \\ &= u^Tu + \alpha d^Td(2\frac{u^Td}{d^Td} + \alpha) \\ &< u^Tu \end{aligned}$$

because $u^Td < 0$. This contradicts the fact that u was the point in U closest to the origin. Thus, all of U is in H^+. A similar argument shows that each point of V must lie in H^- or else the convexity of V would generate a point closer to u than 0, and so H is a separating hyperplane for U and V. □

As discussed in the above proof, we will only be needing a restricted form of the separating hyperplane here. In particular, the following result which follows from the fact that polyhedra are closed (they are the intersection of closed half spaces), and for any point x and any closed set P, there is a point in P that is closest to x.[1]

Corollary 4.2.1. ***Separating Hyperplane Theorem*** *If P is a polyhedron, and x is a point distinct from P, then there is a vector s such that $s^T x < s^T p$ for all $p \in P$.*

A corollary of separating hyperplanes is Farkas' lemma.

Lemma 4.2.1. ***Farkas' Lemma*** *For any $A \in \Re^{M \times N}$ and $b \in \Re^M$, exactly one of the following two alternatives holds:*
(a) *There exists a vector $x \in \Re^N$ such that $x \geq 0$, $Ax = b$.*
(b) *There exists a vector $y \in \Re^M$ such that $b^T y < 0$ and $A^T y \geq 0$.*

Proof:
If (a) and (b) are both true, then $0 > b^T y = y^T b = y^T Ax = x^T A^T y \geq 0$, which is a contradiction. This means that (a) being true makes (b) false.

Suppose that (a) is false, then b is not in the polyhedron $P = \{Ax, x \geq 0\}$[2]. Let y be the vector guaranteed by the separating hyperplane theorem. This means that $y^T b < y^T p$ for all $p \in P$. 0 is in P, so this means that $b^T y = y^T b < 0$. Suppose $y^T A_{j*} < 0$ for some j. Then for some α we must have $y^T(\alpha A_{j*}) < y^T b$ violating the fact that αA_{j*} is in P. Thus $y^T A_{j*} \geq 0$ for each j, or $A^T y \geq 0$ so that (b) is true. □

We can now give the KKT conditions. Note that these conditions are necessary, but not sufficient for a maximizer. There is a slight technical condition on the maximizing point, it needs to be *regular*. A regular point is one where if all active constraints are linearized (that is, replaced with tangent hyperplanes), the set of feasible directions remains the same.[3]. This rules out some extremely rare coincidences of constraints, and note that in linear systems every point is regular.

That caveat aside, here are the KKT conditions.

Theorem 4.2.6. *Suppose that $f, g^1, g^2, \ldots g^M$ are differentiable functions from $\Re^N$ into $\Re$. Let x be the point that maximizes f subject to $g^i(x) \leq 0$ for each i, and assume the first k constraints are active and x is regular. Then there exists $y_1, \ldots, y_k \geq 0$ such that $\nabla f(x) = y_1 \nabla g^1(x) + \ldots + y_k \nabla g^k(x)$.*

[1] Despite being obvious if drawn, this result is typically established using the fact that distance is a continuous function and applying Weierstrass' theorem.

[2] To see that P is a polyhedron, it can also be written as $\{y | y = Ax, x \geq 0\}$.

[3] If there is a feasible curve from x whose initial direction is d, then d is a feasible direction.

Proof: Let $c = \nabla f(x)$. The directional derivative of f in the direction of unit vector z is $z^T c$. Thus, every feasible direction z must have $z^T c \leq 0$. Let $A \in \Re^{k \times N}$ be the matrix whose ith row is $(\nabla g^i(x))^T$. Then a direction z is feasible if $Az \leq 0$. Thus, there is no z with $Az \leq 0$ and $c^T z > 0$. From Farkas' lemma, we can now say that there is a y such that $y \geq 0$ and $A^T y = c$, which is the statement of the theorem. $\square$

Note that the conditions describe ∇f as a combination of active constraint gradients only. Another way of stating the conditions is to say that $\nabla f = y_1 \nabla g^1(x) + \ldots + y_n \nabla g^n(x)$ where $y^i g^i(x) = 0$. Now the sum is over all constraints (not just the active ones), but the second condition says that the coefficients of non-active constraints must be 0. The condition $y^i g^i(x) = 0$ is called a complementarity condition, and is another certificate of optimality.

The Strong Duality Theorem is an application of the KKT conditions to the particular case where each of the functions being considered is linear.

Proof of strong duality: The primal problem is given by

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \geq 0. \end{array}$$

Letting $f(x) = c^T x$, $g^1(x) = (Ax - b)_1, \ldots, g^M(x) = (Ax - b)_M$, $g^{M+1}(x) = -x_1, \ldots, g^{M+N}(x) = -x_N$, we see that the primal is equivalent to

$$\begin{array}{ll} \text{maximize} & f(x) \\ \text{subject to} & g(x) \leq 0. \end{array}$$

Note that $\nabla g(x^*) = c$, $\nabla g^k(x) = A_{k*}$ for $k = 1, \ldots, M$, and $\nabla g^k(x) = -e_k$ for $k = M+1, \ldots, M+N$.

Suppose x^* is an optimal solution. The KKT conditions ensure existence $y \in \Re^M$ and $z \in \Re^N$ such that $y \geq 0$, $z \geq 0$, $c = A^T y - z$, and $(Ax^* - b)^T y - (x^*)^T z = 0$. It follows that $A^T y \geq c$ and $b^T y = (A^T y - z)^T x^* = c^T x$. The result follows. $\square$

4.2.4 Complementary Slackness

Recall that if x^* and y^* are optimal solutions to primal and dual linear programs, each dual variable y_i^* can be viewed as the sensitivity of the objective value to the value of b_i. If the constraint $A_{i*}^T x \leq b_i$ is not binding, the objective value should not be sensitive to the value of b_i, and therefore, y_i^* should be equal to zero. The fact that this is true for every i can be expressed concisely in terms of an equation: $(b - Ax^*)^T y^* = 0$; since all components of both $b - Ax^*$ and y^* are nonnegative, the only way the inner product can be equal to 0 is if, for each ith component, either $A_{i*}^T x = b_i$ or $y_i = 0$. Similarly, since the primal is the dual of the dual, each x_j^* represents sensitivity of the objective value to c_j, and we have $(A^T y^* - c)^T x^* = 0$.

The preceding discussion suggests that, for any optimal solutions x^* and y^* to the primal and dual, $(b - Ax^*)^T y^* = 0$ and $(A^T y^* - c)^T x^* = 0$. Interestingly, in addition to this statement, the converse is true: for any feasible solutions x and y to the primal and dual, if $(b - Ax)^T y = 0$ and $(A^T y - c)^T x = 0$ then x and y are optimal solutions to the primal and dual. These facts are immediate consequences of duality theory. They are captured by the complementary slackness theorem, which we now state and prove.

Theorem 4.2.7. Complementary Slackness) *Let x and y be feasible solutions to symmetric form primal and dual linear programs. Then, x and y are optimal solutions to the primal and dual if and only if $(b - Ax)^T y = 0$ and $(A^T y - c)^T x = 0$.*

Proof: Feasibility implies that $(b - Ax)^T y \geq 0$ and $(A^T y - c)^T x \geq 0$. Further, if x and y are optimal,

$$(A^T y - c)^T x + (b - Ax)^T y = b^T y - c^T x = 0,$$

by strong duality (Theorem 4.2.4). Hence, if x and y are optimal, $(b - Ax)^T y = 0$ and $(A^T y - c)^T x = 0$.

For the converse, suppose that x and y are feasible and that $(b - Ax)^T y = 0$ and $(A^T y - c)^T x = 0$. Then,

$$0 = (A^T y - c)^T x + (b - Ax)^T y = b^T y - c^T x,$$

which provides a certificate of optimality (Theorem 4.2.2). □

There are many interesting consequences to complementary slackness. We will consider in Section 4.5 one application involving allocation of a labor force among interdependent production processes.

4.3 Duals of General Linear Programs

For a linear program that isn't in the symmetric form we can still construct the dual problem. To do this, you can transform the linear program to symmetric form, and then construct the dual from that. Alternatively, you can apply the KKT conditions directly. Either approach results in an equivalent problem.

For example, suppose the linear program is minimize $c^T x$ subject to $Ax \leq b, x \geq 0$. Then because minimizing $c^T x$ is the same as maximizing $-c^T x =$

$(-c)^T x$, the linear program is the same as maximize $(-c)^T x$ subject to $Ax \leq b, x \geq 0$.

The modifications needed are summarized below.

- If the objective function, $c^T x$ is minimized rather than maximized, then replace c by $-c$.
- If a constraint is a greater than constraint, $a^T x \geq \beta$, then take the negative of the constraint to get $(-a)^T x \leq -\beta$.
- If a constraint is an equality constraint, $a^T x = \beta$, then treat it as a greater than constraint and a less than constraint to get $a^T x \leq \beta$ and $(-a)^T x \leq -\beta$.
- If x is not constrained to be positive, then replace x by x^+ and $x-$, there x^+ represents the positive part of x, and $x-$ represents the negative part of x, just like the arbitrage example of the previous chapter.

As an example, suppose the linear program is

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & A^1 x \leq b^1 \\ & A^2 x \geq b^2 \\ & A^3 x = b^3 \end{array}$$

Then rearranging into symmetric form would give

$$\begin{array}{ll} \text{maximize} & (-c)^T x^+ + c^T x^- \\ \text{subject to} & A^1 x^+ - A^1 x^- \leq b^1 \\ & (-A^2) x^+ + A^2 x^- \leq -b^2 \\ & A^3 x^+ - A^3 x^- \leq b^3 \\ & (-A^3) x^+ + A^3 x^- \leq -b^3 \\ x^+, x^- \geq 0 & \end{array}$$

Note that if x^+, x^- is a solution, then so is $x^+ + y, x^- + y$ for any $y \geq 0$. However, if $y \neq 0$, then this will not represent a vertex.

Taking the dual of the above linear program gives

$$\begin{array}{ll} \text{minimize} & (b^1)^T y^1 - (b^2)^T y^2 + (b^3)^T y^3 - (b^3)^T y^4 \\ \text{subject to} & (A^1)^T y^1 - (A^2)^T y^2 + (A^3)^T y^3 - (A^3)^T y^4 \geq -c \\ & (-A^1)^T y^1 + (A^2)^T y^2 - (A^3)^T y^3 + (A^3)^T y^4 \geq c \\ & y^1, y^2, y^3, y^4 \geq 0 \end{array}$$

Notice that y^3 and y^4 are exactly what they would have been if one had replaced an unconstrained y with $y^3 = y^+$ and $y^4 = y^-$. Thus, writing $y = y^3 - y^4$, we can rewrite the dual as

$$\begin{array}{ll} \text{minimize} & (b^1)^T y^1 - (b^2)^T y^2 + (b^3)^T y \\ \text{subject to} & (A^1)^T y^1 - (A^2)^T y^2 + (A^3)^T y \geq -c \\ & (-A^1)^T y^1 + (A^2)^T y^2 - (A^3)^T y \geq c \\ & y^1, y^2 \geq 0 \end{array}$$

The fact that equality constraints in the primal correspond to unconstrained variables in the dual is but one aspect that can be observed by looking at the above dual. Other features are summarized in the table below, which describes how to take the dual of a general linear program.

PRIMAL	maximize	minimize	DUAL
	$\leq b_i$	≥ 0	
constraints	$\geq b_i$	≤ 0	*variables*
	$= b_i$	*unconstrained*	
	≥ 0	$\geq c_j$	
variables	≤ 0	$\leq c_j$	*constraints*
	unconstrained	$= c_j$	

Note, using the rules in the above table, the dual of

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & A^1 x \leq b^1 \\ & A^2 x \geq b^2 \\ & A^3 x = b^3 \end{array}$$

becomes

$$\begin{array}{ll} \text{maximize} & (b^1)^T y^1 + (b^2)^T y^2 + (b^3)^T y^3 \\ \text{subject to} & (A^1)^T y^1 + (A^2)^T y^2 + (A^3)^T y^3 = c \\ & y^1 \leq 0 \\ & y^2 \geq 0 \end{array}$$

Since the dual of the dual is the primal, reorganizing the above table yields an alternative procedure for converting primals that involve minimization to their duals.

PRIMAL	minimize	maximize	DUAL
	$\leq b_i$	≤ 0	
constraints	$\geq b_i$	≥ 0	*variables*
	$= b_i$	*unconstrained*	
	≥ 0	$\leq c_j$	
variables	≤ 0	$\geq c_j$	*constraints*
	unconstrained	$= c_j$	

4.4 Two-Player Zero-Sum Games

In this section, we consider games in which each of two opponents selects a strategy and receives a payoff contingent on both his own and his opponent's selection. We restrict attention here to zero-sum games – those in which a payoff to one player is a loss to his opponent. Let us recall our example from Chapter **??** that illustrates the nature of such problems.

Example 4.4.1. (drug running) *A South American drug lord is trying to get as many of his shipments across the border as possible. He has a fleet of boats available to him, and each time he sends a boat, he can choose one of three ports at which to unload. He could choose to unload in San Diego, Los Angeles, or San Francisco.*

The USA Coastguard is trying to intercept as many of the drug shipments as possible but only has sufficient resources to cover one port at a time. Moreover, the chance of intercepting a drug shipment differs from port to port. A boat arriving at a port closer to South America will have more fuel with which to evade capture than one arriving farther away. The probabilities of interception are given by the following table:

Port	Probability of interception
San Diego	*1/3*
Los Angeles	*1/2*
San Francisco	*3/4*

The drug lord considers sending each boat to San Diego, but the coastguard realizing this would always choose to cover San Diego, and only $2/3$ *of his boats would get through. A better strategy would be to pick a port at random (each one picked with* $1/3$ *probability). Then, the coastguard should cover port 3, since this would maximize the number of shipments captured. In this scenario,* $3/4$ *of the shipments would get through, which is better than* $2/3$. *But is this the best strategy?*

Clearly, the drug lord should consider randomized strategies. But what should he optimize? We consider as an objective maximizing the probability that a ship gets through, assuming that the Coastguard knows the drug lord's choice of randomized strategy. We now formalize this solution concept for general two-person zero-sum games, of which our example is a special case.

Consider a game with two players: player 1 and player 2. Suppose there are N alternative decisions available to player 1 and M available to player 2. If player 1 selects decision $j \in \{1, \ldots, N\}$ and player 2 selects decision $i \in \{1, \ldots, M\}$, there is an expected payoff of P_{ij} to be awarded to player 1 at the expense of player 2. Player 1 wishes to maximize expected payoff, whereas player 2 wishes to minimize it. We represent expected payoffs for all possible decision pairs as a matrix $P \in \Re^{M \times N}$.

A randomized strategy is a vector of probabilities, each associated with a particular decision. Hence, a randomized strategy for player 1 is a vector $x \in \Re^N$ with $e^T x = 1$ and $x \geq 0$, while a randomized strategy for player 2 is a vector $y \in \Re^M$ with $e^T y = 1$ and $y \geq 0$. Each x_j is the probability that player 1 selects decision j, and each y_i is the probability that player 2 selects decision i. Hence, if the players apply randomized strategies x and y, the probability of payoff P_{ij} is $y_i x_j$ and the expected payoff is $\sum_{i=1}^M \sum_{j=1}^N y_i x_j P_{ij} = y^T P x$.

How should player 1 select a randomized policy? As a solution concept, we consider selection of a strategy that maximizes expected payoff, assuming that player 2 knows the strategy selected by player 1. One way to write this is as

$$\max_{\{x \in \Re^N | e^T x = 1, x \geq 0\}} \min_{\{y \in \Re^M | e^T y = 1, y \geq 0\}} y^T P x.$$

Here, y is chosen with knowledge of x, and x is chosen to maximize the worst-case payoff. We will now show how this optimization problem can be solved as a linear program.

First, consider the problem of optimizing y given x. This amounts to a linear program:

$$\begin{array}{ll} \text{minimize} & (Px)^T y \\ \text{subject to} & e^T y = 1 \\ & y \geq 0. \end{array}$$

It is easy to see that the basic feasible solutions of this linear program are given by $e^1, \ldots, e^M$, where each e^i is the vector with all components equal to 0 except for the ith, which is equal to 1. It follows that

$$\min_{\{y \in \Re^M | e^T y = 1, y \geq 0\}} y^T P x = \min_{i \in \{1, \ldots, M\}} (Px)_i.$$

This minimal value can also be expressed as the solution to a linear program:

$$\begin{array}{ll} \text{maximize} & v \\ \text{subject to} & ve \leq Px, \end{array}$$

where $v \in \Re$ is the only decision variable and x is fixed. In particular, the optimal value v^* resulting from this linear program satisfies

$$v^* = \min_{\{y \in \Re^M | e^T y = 1, y \geq 0\}} y^T P x.$$

To determine an optimal strategy for player 1, we find the value of x that maximizes v^*. In particular, an optimal strategy is delivered by the following linear program:

$$\begin{array}{ll} \text{maximize} & v \\ \text{subject to} & ve \leq Px \\ & e^T x = 1 \\ & x \geq 0, \end{array}$$

where $v \in \Re$ and $x \in \Re^N$ are decision variables. An optimal solution to this linear program provides a stochastic strategy x that maximizes the payoff v, assuming that player 2 knows the randomized strategy of player 1 and selects a payoff-minimizing counter-strategy. We illustrate application of this linear program through a continuation of Example 4.4.2.

Example 4.4.2. (linear programming for drug running) *To determine an optimal drug running strategy, we formulate the problem in the terms we have introduced. The drug lord's strategy is represented as a vector $x \in \Re^3$ of three probabilities. The first, second, and third components represent the probabilities that a ship is sent to San Diego, Los Angeles, or San Francisco, respectively. The payoff is 1 if a ship gets through, and 0 otherwise. Hence, the expected payoff P_{ij} is the probability that a ship gets through if player 1 selects decision j and player 2 selects decision i. The payoff matrix is then*

$$P = \left[\begin{array}{ccc} 2/3 & 1 & 1 \\ 1 & 1/2 & 1 \\ 1 & 1 & 1/4 \end{array} \right].$$

The optimal strategy for the drug lord is given by a linear program:

$$\begin{array}{ll} \text{maximize} & v \\ \text{subject to} & ve \leq Px \\ & e^T x = 1 \\ & x \geq 0. \end{array}$$

Suppose that the drug lord computes an optimal randomized strategy x^* by solving the linear program. Over time, as this strategy is used to guide shipments, the drug lord can estimate the Coastguard's strategy y. Given y, he may consider adjusting his own strategy in response to y, if that will increase expected payoff. But should it be possible for the drug lord to improve his expected payoff after learning the Coastguard's strategy? Remarkably, if the coastguard selects a randomized strategy through an approach analogous to that we have described for the drug lord, neither the drug lord nor the Coastguard should ever need to adjust their strategies. We formalize this idea in the context of general two-player zero-sum games.

Recall from our earlier discussion that player 1 selects a randomized strategy x^* that attains the maximum in

$$\max_{\{x \in \Re^N | e^T x = 1, x \geq 0\}} \min_{\{y \in \Re^M | e^T y = 1, y \geq 0\}} y^T P x,$$

and that this can be done by solving a linear program

$$\begin{array}{ll} \text{maximize} & v \\ \text{subject to} & ve \leq Px \\ & e^T x = 1 \\ & x \geq 0. \end{array}$$

Consider determining a randomized strategy for player 2 through an analogous process. An optimal strategy will then be a vector y^* that attains the minimum in

$$\min_{\{y \in \Re^M | e^T y = 1, y \geq 0\}} \max_{\{x \in \Re^N | e^T x = 1, x \geq 0\}} y^T P x.$$

Similarly with the case of finding a strategy for player 1, this new problem can be converted to a linear program:

$$\begin{array}{ll} \text{minimize} & u \\ \text{subject to} & ue \geq P^T y \\ & e^T y = 1 \\ & y \geq 0. \end{array}$$

A remarkable fact is that – if player 1 uses x^* and player 2 uses y^* – neither player should have any reason to change his strategy after learning the strategy being used by the other player. Such a situation is referred to as an *equilibrium*. This fact is an immediate consequence of the minimax theorem:

Theorem 4.4.1. (Minimax) *For any matrix $P \in \Re^{M \times N}$,*

$$\max_{\{x \in \Re^N | e^T x = 1, x \geq 0\}} \min_{\{y \in \Re^M | e^T y = 1, y \geq 0\}} y^T P x = \min_{\{y \in \Re^M | e^T y = 1, y \geq 0\}} \max_{\{x \in \Re^N | e^T x = 1, x \geq 0\}} y^T P x.$$

The minimax theorem is a simple corollary of strong duality. In particular, it is easy to show that the linear programs solved by players 1 and 2 are duals of one another. Hence, their optimal objective values are equal, which is exactly what the minimax theorem states.

Suppose now that the linear program solved by player 1 yields an optimal solution x^*, while that solved by player 2 yields an optimal solution y^*. Then, the minimax theorem implies that

$$(y^*)^T P x \leq (y^*)^T P x^* \leq y^T P x^*,$$

for all $x \in \Re^N$ with $e^T x = 1$ and $x \geq 0$ and $y \in \Re^M$ with $e^T y = 1$ and $y \geq 0$. In other words, the pair of strategies (x^*, y^*) yield an equilibrium.

4.5 Allocation of a Labor Force

Our economy presents a network of interdependent industries. Each both produces and consumes goods. For example, the steel industry consumes coal to manufacture steel. Reciprocally, the coal industry requires steel to support its own production processes. Further, each industry may be served by multiple manufacturing technologies, each of which requires different resources per unit production. For example, one technology for producing steel starts with iron ore while another makes use of scrap metal. In this section, we consider a hypothetical economy where labor is the only limiting resource. We will develop a model to guide how the labor force should be allocated among industries and technologies.

In our model, each industry produces a single good and may consume others. There are M goods, indexed $i = 1, \ldots, M$. Each can be produced by one or more technologies. There are a total of $N \geq M$ technologies, indexed $j = 1, \ldots, N$. Each jth technology produces $A_{ij} > 0$ units of some ith good per unit of labor. For each $k \neq i$, this jth industry may consume some amount of good k per unit labor, denoted by $A_{kj} \leq 0$. Note that this quantity A_{kj} is nonpositive; if it is a negative number, it represents the quantity of good k consumed per unit labor allocated to technology j. The productivity and resource requirements of all technologies are therefore captured by a matrix $A \in \Re^{M \times N}$ in which each column has exactly one positive entry and each row has at least one positive entry. We will call this matrix A the *production matrix*. Without loss of generality, we will assume that A has linearly independent rows.

Suppose we have a total of one unit of labor to allocate over the next year. Let us denote by $x \in \Re^N$ our allocation among the N technologies.

Hence, $x \geq 0$ and $e^T x \leq 1$. Further, the quantity of each of the M goods produced is given by a vector Ax.

Now how should we allocate labor? One objective might be to optimize social welfare. Suppose that the amount society values each unit of each ith good is $c_i > 0$, regardless of the quantity produced. Then, we might define the social welfare generated by production activities to be $c^T Ax$. Optimizing this objective leads to a linear program:

$$\begin{array}{ll} \text{maximize} & c^T Ax \\ \text{subject to} & Ax \geq 0 \\ & e^T x \leq 1 \\ & x \geq 0. \end{array} \tag{4.2}$$

A production matrix A is said to be *productive* if there exists a labor allocation x (with $x \geq 0$ and $e^T x = 1$) such that $Ax > 0$. In other words, productivity means that some allocation results in positive quantities of every good. It turns out that – when the production matrix is productive – only M technologies are beneficial, and the choice of M technologies is independent of societal values. This remarkable result is known as the substitution theorem:

Theorem 4.5.1. (substitution) *If a production matrix A is productive, there is a set of M technologies such that for any vector c of societal values, social welfare can be maximized by an allocation of labor among only these M technologies.*

In the remainder of this section, we will leverage linear algebra and duality theory to prove the substitution theorem.

4.5.1 Labor Minimization

Consider a related problem with an objective of minimizing the labor required to generate a particular "bill of goods" $b \in \Re^M$. Here, each b_i is nonnegative and represents the quantity of good i demanded by society. This problem is captured by the following linear program:

$$\begin{array}{ll} \text{minimize} & e^T x \\ \text{subject to} & Ax \geq b \\ & x \geq 0. \end{array} \tag{4.3}$$

As before, each x_j is the amount of labor allocated to the jth technology. The requirement is that we produce at least b_i units of each ith good, and we wish to minimize the quantity $e^T x$ of labor used to accomplish this. The following lemma relates solutions of (4.3) to (4.2).

Lemma 4.5.1. *Let x^* be an optimal solution to (4.2) and let $b = Ax^*$. Then, the set of optimal solutions to (4.3) is the same as the set of optimal solutions to (4.2).*

Proof: Note that $e^T x^* = 1$; if this were not true, $x^*/e^T x^*$ would be another feasible solution to (4.2) with objective value $c^T A x^*/e^T x^* > c^T x^*$, which would contradict the fact that x^* is an optimal solution.

We now show that x^* (and therefore any optimal solution to (4.2)) is an optimal solution to (4.3). Let $\overline{x}$ be an optimal solution to (4.3). Assume for contradiction that x^* is not an optimal solutoin to (4.3). Then, $e^T \overline{x} < e^T x^*$ and

$$c^T A\overline{x}/e^T\overline{x} = c^T b/e^T\overline{x} > c^T b/e^T x^* = c^T A x^*.$$

Since $\overline{x}/e^T\overline{x}$ is a feasible solution to (4.2), this implies that x^* is not an optimal solution to (4.2), which is a contradiction. The conclusion is that x^* is an optimal solution to (4.3).

Since the fact that x^* is an optimal solution to (4.3) implies that $e^T\overline{x} = e^T x^* = 1$. It follows that $\overline{x}$ is a feasible solutoin to (4.2). Since $c^T A\overline{x} \geq c^T b = c^T A x^*$, $\overline{x}$ is also an optimal solution to (4.2). □

4.5.2 Productivity Implies Flexibility

Consider M technologies, each of which produces one of the M goods. Together they can be described by an $M \times M$ production matrix A. Interestingly, if A is productive then any bill of goods can be met exactly by appropriate application of these technologies. This represents a sort of flexibility – any demands for goods can be met without any excess supply. This fact is captured by the following lemma.

Lemma 4.5.2. *If a square production matrix $A \in \Re^{M\times M}$ is productive then for any $b \geq 0$, the equation $Ax = b$ has a unique solution $x \in \Re^M$, which satisfies $x \geq 0$.*

Proof: Since A is productive, there exists a vector $\overline{x} \geq 0$ such that $A\overline{x} > 0$. The fact that only one element of each column of A is positive implies that $\overline{x} > 0$.

Since the rows of A are linearly independent, $Ax = b$ has a unique solution $x \in \Re^N$. Assume for contradiction that there is some $\hat{b} \geq 0$ and $\hat{x}$ with at least one negative component such that $A\hat{x} = \hat{b}$. Let

$$\alpha = \min\{\alpha \geq 0 | \alpha\overline{x} + \hat{x} \geq 0\},$$

and note that $\alpha > 0$ because some component of $\hat{x}$ is negative. Since only one element of each column of A is positive, this implies that at least one component of $A(\alpha\overline{x} + \hat{x})$ is nonpositive.

Recall that $A\overline{x} > 0$ and $A\hat{x} \geq 0$, and therefore

$$A\hat{x} < \alpha A\overline{x} + A\hat{x} = A(\alpha\overline{x} + \hat{x}),$$

contradicting the fact that at least one component of $A(\alpha\overline{x}+\hat{x})$ is nonpositive.
□

4.5.3 Proof of the Substitution Theorem

Since A is productive, there exists a vector $\overline{x} \geq 0$ such that $A\overline{x} > 0$. Let $b^1 = A\overline{x}$ and consider the labor minimization problem:

$$\begin{array}{ll} \text{minimize} & e^T x \\ \text{subject to} & Ax \geq b^1 \\ & x \geq 0. \end{array}$$

Let x^1 be an optimal basic feasible solution and note that $Ax^1 \geq b^1 > 0$. Since x^1 is a basic feasible solution, at least $N - M$ components must be equal to zero, and therefore, at most M components can be positive. Hence, the allocation x^1 makes use of at most M technologies. Lemma 4.5.2 implies that these M technologies could be used to fill any bill of goods.

Given an arbitrary bill of goods $b^2 \geq 0$, we now know that there is a vector $x^2 \geq 0$, with $x^2_{j_k} = 0$ for $k = 1, \ldots, N - M$, such that $Ax^2 = b^2$. Note that $x^2 \geq 0$ and satisfies N linearly independent constraints of and is therefore a basic feasible solution of the associated labor minimization problem:

$$\begin{array}{ll} \text{minimize} & e^T x \\ \text{subject to} & Ax \geq b^2 \\ & x \geq 0. \end{array}$$

Let y^1 be an optimal solution to the dual of the labor minimization problem with bill of goods b^1. By complementary slackness, we have $(e - A^T y^1)^T x^1 = 0$. Since $x^2_j = 0$ if $x^1_j = 0$, we also have $(e - A^T y^1)^T x^2 = 0$. Further, since $Ax^2 = b^2$, we have $(Ax^2 - b^2)^T y^1 = 0$. Along with the fact that x^2 is a feasible solution, this gives us the complementary slackness conditions required to ensure that x^2 and y^1 are optimal primal and dual solutions to the labor minimization problem with bill of goods b^2.

We conclude that there is a set of M technologies that is sufficient to attain the optimum in the labor minimization problem for any bill of goods. It follows from Lemma 4.5.1 that the same set of M technologies is sufficient to attain the optimum in the social welfare maximization problem for any societal values. □

4.6 Exercises

Question 1

Consider the following linear program (LP).

$$\begin{array}{rrcl} & \max\ x_1 - x_2 & & \\ s.t. & -2x_1 - 3x_2 & \leq & -4 \\ & -x_1 + x_2 & \leq & -1 \\ & x_1, x_2 & >= & 0 \end{array}$$

(a) Plot the feasible region of the primal and show that the primal objective value goes to infinity.

(b) Formulate the dual, plot the feasible region of the dual, and show that it is empty.

Question 2

Convert the following optimization problem into a symmetric form linear program, and then find the dual.

$$\begin{array}{rrcl} \max & -x_1 - 2x_2 - x_3 & & \\ s.t. & x_1 + x_2 + x_3 & = & 1 \\ & |x_1| & \leq & 4 \\ & x_1, x_2, x_3 & \geq & 0 \end{array}$$

Note: $|x|$ denotes the absolute value of x.

Question 3

consider the LP

$$\begin{array}{rrcl} \max & -x_1 - x_2 & & \\ s.t. & -x_1 - 2x_2 & \leq & -3 \\ & -x_1 + 2x_2 & \leq & 4 \\ & x_1 + 7x_2 & \leq & 6 \\ & x_1, x_2 & \geq & 0 \end{array}$$

(a) solve this problem in Excel using solver. After solver finds an optimal solution, ask it to generate the sensitivity report.

(b) Find the dual of the above LP. Read the shadow prices from the sensitivity report, and verify that it satisfies the dual and gives the same dual objective value as the primal.

Question 4

Consider the LP

$$
\begin{array}{rrcl}
\min & 2x_1 + x_2 & & \\
s.t. & x_1 + x_2 & \leq & 6 \\
 & x_1 + 3x_2 & \geq & 3 \\
 & x_1, x_2, x_3 & \geq & 0
\end{array}
$$

Note: there is an x_3.
(a) plot the feasible region and solve the problem graphically.
(b) Rearrange into symmetric form and find the dual. Solve the dual graphically.
(c) Verify that primal and dual optimal solutions satisfy the Strong Duality Theorem.

Question 5

(a) Consider the problem of feeding an army presented in Section 3.1.2. Provide a dual linear program whose solution gives sensitivities of the cost of an optimal diet to nutritional requirements. Check to see whether the sensitivities computed by solving this dual linear program match the sensitivities given by Solver when solving the primal.

(b) Suppose a pharmaceutical company wishes to win a contract with the army to sell digestible capsules containing pure nutrients. They sell three types of capsule, with 1 grams of carbohydrates, 1 gram of fiber, and 1 grams of protein, respectively. The army requires that the company provide a price for each capsule independently, and that substituting the nutritional value of any food item with capsules should be no more expensive than buying the food item. Explain the relation between the situation faced by the pharmaceutical company and the dual linear program.

Question 6

Consider a symmetric form primal linear program:

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \geq 0. \end{array}$$

(a) Find the dual of this problem. Noting that $\max f(x) = \min -f(x)$ rearrange the dual into symmetric form. Take the dual of you answer, and rearrange again to get into symmetric form.

(b) Explain why the sensitivities of the optimal dual objective to the dual constraints should be optimal primal variables.

Question 7

Recall Question 8, Homework 4, the diet Problem for the pigs. We found an optimal solution for that problem (see Solution Homework 4). Now, suppose that Dwight doesn't have a good estimate of the price of Feed type A because of some turbulence in the market. Therefore, he would like to know how sensitive is his original optimal solution with respect to changes in the price of Feed Type A. In particular, in what range around \$0.4 can the price of Feed Type A change, without changing the original optimal solution? For prices out of this range, what are the new optimal solutions? Now suppose Dwight doesn't have a good estimate of the requirements of vitamins. In what range around 700 does the requirement of vitamins can change without changing the original optimal solution? For values out of this range, what are the new optimal solutions? Your arguments should be geometric (and not based on Excel), that is, you should draw the problem in R^2 and see what is going on.

Question 8

Consider the linear program

$$\begin{array}{ll} \text{maximize} & -q^T z \\ \text{subject to} & Mz \leq q \\ & z \geq 0. \end{array}$$

where

$$M = \begin{bmatrix} 0 & A \\ -A^T & 0 \end{bmatrix}, \quad q = \begin{bmatrix} c \\ -b \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}.$$

a) Derive the dual
b) Show that the optimal solution to the dual is given by that of the primal and vice-versa.
c) Show that the primal problem has an optimal solution if and only if it is feasible.

Question 9

Consider the following LP.

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \geq 0. \end{array}$$

Show that if A, b and c are all positive, then both the primal and dual feasible regions are non-empty.

Question 10

Why is it that, if the primal has unique optimal solution x^*, there is a sufficiently small amount by which c can be altered without changing the optimal solution?

Question 11 - Games and Duals

Show that the linear programs given in the notes to determine the strategies of player 1 and player 2 are indeed duals of one another.

Question 12 - Drug Runners

Consider the drug running example from the notes. Imagine the drug lord has a 4th alternative which involves transporting his drugs overland.

a) Suppose that the DEA can reassign its agents from coastguard duty to guarding the border, and if it does so any shipment of drugs transported overland will be caught. Will the drug lord ever choose to send shipments overland? If so, with what probability? If not, why not.

b) Suppose that guarding the border does not require the coastguard to reassign its agents, so that it can still guard a port (for instance, suppose the customs agents at the border are sufficiently equipped to detect most drug shipments). Suppose that an overland shipment will get through 85% of the time. Will the drug lord ever choose to go overland? With what probability?

c) Suppose the percentage in part b) were changed to %80. What would be the probability now.

d) Suppose the percentage in part b) were changed to %90. What would be the probability now.

e) Suppose that if the DEA reassigns its agents to the border, they will certainly catch any overland drug shipments, but if the agents are not reassigned, then the customs agents will catch %80 of all drug shipments. Should the DEA ever assign their agents to the border? With what probability?

Question 13 - Elementary Asset Pricing and Arbitrage

You are examining a market to see if you can find arbitrage opportunities. For the same of simplicity, imagine that there are M states that the market could be in next year, and you are only considering buying a portfolio now, and selling it in a years time. There are also N assets in the market, with price vector $\rho \in \Re^N$. The payoff matrix is P. So, the price of asset i is ρ_i and the payoff of asset i in state j is P_{ji}.

You have observed some relationships amongst the prices of the assets, in particular you have found that the price vector is in the *row space* of P.

a) Suppose that $\rho^T = q^T P$ for some $q \in \Re^M$. An *elementary asset* is another name for an Arrow-Debreu security. That is, it is an asset that pays \$1 in a particular state of the market, and \$0 in any other state. So, there are M elementary assets, and the ith elementary asset pays \$1 if the market is in state i, and \$0 otherwise. Suppose that the ith elementary asset can be constructed from a portfolio $x \in \Re^N$. What is the price of x in terms of q and ρ? (Assume that there are no arbitrage opportunities.)

b) Suppose that however I write ρ as a linear combination of rows of P, the coefficient of each row is non-negative. Is it possible for there to be an arbitrage opportunity? If not, why not. If so, give an example.

c) Suppose that ρ can be written as a linear combination of rows of P, where the coefficient of each row is non-negative. Is it possible for their to be an arbitrage opportunity? If not, why not. If so, give an example. Note, this is a weaker condition than (b) because there may be many ways of writing ρ as a linear combination of rows of P.

Hint: Note that ρ being in the row space of P means that there is a vector q such that $\rho^T = q^T P$. Consider a portfolio x.