Module MAU34802: The Theory of Linear Programming
Hilary Term 2021
Section 5: Duality and Convexity

D. R. Wilkins

Contents

5 Duality and Convexity

5.1 General Linear Programming Problems

Linear programming is concerned with problems seeking to maximize or minimize a linear functional of several real variables subject to a finite collection of constraints, where each constraint either fixes the values of some linear function of the variables or else requires those values to be bounded, above or below, by some fixed quantity.

The objective of such a problem involving n real variables $x_1, x_2, \ldots, x_n$ is to maximize or minimize an *objective function* of those variables that is of the form

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n,$$

subject to appropriate constraints. The coefficients $c_1, c_2, \ldots, c_n$ that determine the objective function are then fixed real numbers.

Now such an optimization problem may be presented as a minimization problem, because simply changing the signs of all the coefficients $c_1, c_2, \ldots, c_n$ converts any maximization problem into a minimization problem. We therefore suppose, without loss of generality, that the objective of the linear programming problem is to find a feasible solution satisfying appropriate constraints which minimizes the value of the objective function amongst all such feasible solutions to the problem.

Some of the constraints may simply require specific variables to be non-negative or non-positive. Now a constraint that requires a particular variable x_j to be non-positive can be reformulated as one requiring a variable to be non-negative by substituting x_j for $-x_j$ in the statement of the problem. Thus, without loss of generality, we may suppose that all constraints that simply specify the sign of a variable x_j will require that variable to be non-negative. Then all such constraints can be specified by specifying a subset J^+ of $\{1, 2, \ldots, n\}$: the constraints then require that $x_j \geq 0$ for all $j \in J^+$.

There may be further constraints in addition to those that simply specify whether one of the variables is required to be non-positive or non-negative. Suppose that there are m such additional constraints, and let them be numbered between 1 and m. Then, for each integer i between 1 and m, there exist real numbers $A_{i,1}, A_{i,2}, \ldots, A_{i,n}$ and b_i that allow the ith constraint to be expressed either as an *inequality constraint* of the form

$$A_{i,1}x_1 + A_{i,2}x_2 + \ldots + A_{i,n}x_n \geq b_i$$

or else as an *equality constraint* of the form

$$A_{i,1}x_1 + A_{i,2}x_2 + \ldots + A_{i,n}x_n = b_i.$$

It follows from the previous discussion that the statement of a general linear programming problem can be transformed, by changing the signs of some of the variables and constants in the statement of the problem, so as to ensure that the statement of the problem conforms to the following restrictions:—

- the objective function is to be minimized;
- some of the variables may be required to be non-negative;
- other constraints are either inequality constraints placing a lower bound on the value of some linear function of the variables or else equality constraints fixing the value of some linear function of the variables.

Let us describe the statement of a linear programming problem as being in *general primal form* if it conforms to the restrictions just described.

A linear programming problem is expressed in general primal form if the specification of the problem conforms to the following restrictions:—

- the objective of the problem is to find an optimal solution minimizing the objective function amongst all feasible solutions to the problem;
- any variables whose sign is prescribed are required to be non-negative, not non-positive;
- all inequality constraints are expressed by prescribing a lower bound on the value on some linear function of the variables.

A linear programming problem in general primal form can be specified by specifying the following data: an $m \times n$ matrix A with real coefficients, an m-dimensional vector $\mathbf{b}$ with real components; an n-dimensional vector $\mathbf{c}$ with real components; a subset I^+ of $\{1, 2, \ldots, m\}$; and a subset J^+ of $\{1, 2, \ldots, n\}$. The linear programming programming problem specified by this data is the following:—

> *seek* $\mathbf{x} \in \mathbb{R}^n$ *that minimizes the objective function* $\mathbf{c}^T\mathbf{x}$ *subject to the following constraints:—*
>
> - $A\mathbf{x} \geq \mathbf{b}$*;*
> - $(A\mathbf{x})_i = (\mathbf{b})_i$ *unless* $i \in I^+$*;*
> - $(\mathbf{x})_j \geq 0$ *for all* $j \in J^+$*.*

We refer to the $m \times n$ matrix A, the m-dimensional vector $\mathbf{b}$ and the n-dimensional vector $\mathbf{c}$ employed in specifying a linear programming problem in general primal form as the *constraint matrix*, *target vector* and *cost vector* respectively for the linear programming problem. Let us refer to the subset I^+ of $\{1, 2, \ldots, m\}$ specifying those constraints that are inequality constraints as the *inequality constraint specifier* for the problem, and let us refer to the subset J^+ of $\{1, 2, \ldots, n\}$ that specifies those variables that are required to be non-negative for a feasible solution as the *variable sign specifier* for the problem.

We denote by $\mathrm{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ the linear programming problem whose specification in general primal form is determined by a constraint matrix A, target vector $\mathbf{b}$, cost vector $\mathbf{c}$, inequality constraint specifier I^+ and variable sign specifier J^+.

A linear programming problem formulated in general primal form can be reformulated as a problem in Dantzig standard form, thus enabling the use of the Simplex Method to find solutions to the problem.

Indeed consider a linear programming problem $\mathrm{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ where the constraint matrix A is an $m \times n$ matrix with real coefficients, the target vector $\mathbf{b}$ and the cost vector $\mathbf{c}$ are vectors of dimension m and n respectively with real coefficients. Then the inequality constraint specifier I^+ is a subset of $\{1, 2, \ldots, m\}$ and the variable sign specifier J^+ is a subset of $\{1, 2, \ldots, n\}$. The problem is already in Dantzig standard form if and only if $I^+ = \emptyset$ and $J^+ = \{1, 2, \ldots, n\}$.

If the problem is not in Dantzig standard form, then each variable x_j for $j \notin J^+$ can be replaced by a pair of variables x_j^+ and x_j^- satisfying the constraints $x_j^+ \geq 0$ and $x_j^- \geq 0$: the difference $x_j^+ - x_j^-$ of these new variables is substituted for x_j in the objective function and the constraints. Also a *slack variable* z_i can be introduced for each $i \in I^+$, where z_i is required to satisfy the sign constraint $z_i \geq 0$, and the inequality constraint

$$A_{i,1}x_1 + A_{i,2}x_2 + \ldots + A_{i,n}x_n \geq b_i$$

is then replaced by the corresponding equality constraint

$$A_{i,1}x_1 + A_{i,2}x_2 + \ldots + A_{i,n}x_n - z_i = b_i.$$

The linear programming problem $\mathrm{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ in general primal form can therefore be reformulated as a linear programming problem in Dantzig standard form as follows:—

determine values of x_j for all $j \in J^+$, x_j^+ and x_j^- for all $j \in J^0$, where $J^0 = \{1, 2, \ldots, n\} \setminus J^+$, and z_i for all $i \in I^+$ so as to *minimize the objective function*

$$\sum_{j\in J^+} c_j x_j + \sum_{j\in J^0} c_j x_j^+ - \sum_{j\in J^0} c_j x_j^-$$

subject to the following constraints:—

(i) $\sum_{j\in J^+} A_{i,j}x_j + \sum_{j\in J^0} A_{i,j}x_j^+ - \sum_{j\in J^0} A_{i,j}x_j^- = b_i$ *for each* $i \in \{1,2,\ldots,n\}\setminus I^+$;

(ii) $\sum_{j\in J^+} A_{i,j}x_j + \sum_{j\in J^0} A_{i,j}x_j^+ - \sum_{j\in J^0} A_{i,j}x_j^- - z_i = b_i$ *for each* $i \in I^+$;

(iii) $x_j \geq 0$ *for all* $j \in J^+$;

(iv) $x_j^+ \geq 0$ *and* $x_j^- \geq 0$ *for all* $j \in J^0$;

(v) $z_i \geq 0$ *for all* $i \in I^+$.

Once the problem has been reformulated in Dantzig standard form, techniques based on the Simplex Method can be employed in the search for solutions to the problem.

5.2 Duals of Linear Programming Problems

Every linear programming problem $\mathrm{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ in general primal form determines a corresponding linear programming problem, which we shall denote by $\mathrm{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$, in *general dual form*. The second linear programming problem is referred to as the *dual* of the first, and the first linear programming problem is referred to as the *primal* of its dual.

We shall give the definition of the dual problem associated with a given linear programming problem, and investigate some important relationships between the primal linear programming problem and its dual.

Let $\mathrm{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ be a linear programming problem in general primal form specified in terms of an $m\times n$ constraint matrix A, m-dimensional target vector $\mathbf{b}$, n-dimensional cost vector $\mathbf{c}$, inequality constraint specifier I^+ and variable sign specifier J^+. The corresponding dual problem $\mathrm{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ may be specified in *general dual form* as follows:

seek $\mathbf{p} \in \mathbb{R}^m$ *that maximizes the objective function* $\mathbf{p}^T\mathbf{b}$ *subject to the following constraints:—*

- $\mathbf{p}^T A \leq \mathbf{c}^T$;

- $(\mathbf{p})_i \geq 0$ *for all* $i \in I^+$;
- $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ *unless* $j \in J^+$.

Lemma 5.1 *Let* $\text{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ *be a linear programming problem expressed in general primal form with constraint matrix* A *with* m *rows and* n *columns, target vector* $\mathbf{b}$*, cost vector* $\mathbf{c}$*, inequality constraint specifier* I^+ *and variable sign specifier* J^+*. Then the feasible and optimal solutions of the corresponding dual linear programming problem* $\text{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ *are those of the problem* $\text{Primal}(-A^T, -\mathbf{c}, -\mathbf{b}, J^+, I^+)$.

Proof An m-dimensional vector $\mathbf{p}$ satisfies the constraints of the dual linear programming problem $\text{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ if and only if $\mathbf{p}^T A \leq \mathbf{c}^T$, $(\mathbf{p})_i \geq 0$ for all $i \in I^+$ and $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ unless $j \in J^+$. On taking the transposes of the relevant matrix equations and inequalities, we see that these conditions are satisfied if and only if $-A^T\mathbf{p} \geq -\mathbf{c}$, $(\mathbf{p})_i \geq 0$ for all $i \in I^+$ and $(-A^T\mathbf{p})_j = (-\mathbf{c})_j$ unless $j \in J^+$. But these are the requirements that the vector $\mathbf{p}$ must satisfy in order to be a feasible solution of the linear programming problem $\text{Primal}(-A^T, -\mathbf{c}, -\mathbf{b}, J^+, I^+)$. Moreover $\mathbf{p}$ is an optimal solution of $\text{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ if and only if it maximizes the value of $\mathbf{p}^T\mathbf{b}$, and this is the case if and only if it minimizes the value of $-\mathbf{b}^T\mathbf{p}$. The result follows. ∎

A linear programming problem in Dantzig standard form is specified by specifying integers m and n a constraint matrix A which is an $m \times n$ matrix with real coefficients, a target vector $\mathbf{b}$ belonging to the real vector space $\mathbb{R}^m$ and a cost vector $\mathbf{c}$ belonging to the real vector space $\mathbb{R}^m$. The objective of the problem is to find a feasible solution to the problem that minimizes the quantity $\mathbf{c}^T\mathbf{x}$ amongst all n-dimensional vectors $\mathbf{x}$ for which $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.

The objective of the dual problem is then to find some feasible solution to the problem that maximizes the quantity $\mathbf{p}^T\mathbf{b}$ amongst all m-dimensional vectors $\mathbf{p}$ for which $\mathbf{p}^T A \leq \mathbf{c}$.

5.3 Complementary Slackness and the Weak Duality Theorem

Theorem 5.2 (Weak Duality Theorem for Linear Programming Problems in Dantzig Standard Form)
Let m *and* n *be integers, let* A *be an* $m \times n$ *matrix with real coefficients, let* $\mathbf{b} \in \mathbb{R}^m$ *and let* $\mathbf{c} \in \mathbb{R}^n$*. Let* $\mathbf{x} \in \mathbb{R}^n$ *satisfy the constraints* $A\mathbf{x} = \mathbf{b}$ *and* $\mathbf{x} \geq \mathbf{0}$*, and let* $\mathbf{p} \in \mathbb{R}^m$ *satisfy the constraint* $\mathbf{p}^T A \leq \mathbf{c}$*. Then* $\mathbf{p}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}$*. Moreover* $\mathbf{p}^T\mathbf{b} = \mathbf{c}^T\mathbf{x}$ *if and only if the following complementary slackness condition is satisfied:*

- $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ *for all integers* j *between* 1 *and* n *for which* $(\mathbf{x})_j > 0$.

Proof Let $x_j = (\mathbf{x})_j$ and $c_j = (\mathbf{c})_j$ for $j = 1, 2, \ldots, n$. The constraints satisfied by the vectors $\mathbf{x}$ and $\mathbf{p}$ ensure that

$$\begin{aligned} \mathbf{c}^T\mathbf{x} - \mathbf{p}^T\mathbf{b} &= (\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} + \mathbf{p}^T(A\mathbf{x} - \mathbf{b}) \\ &= (\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x}, \end{aligned}$$

because $A\mathbf{x} - \mathbf{b} = \mathbf{0}$. But also $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{c}^T - \mathbf{p}^T A \geq \mathbf{0}$, and therefore $(\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} \geq 0$. Moreover

$$(\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} = \sum_{j=1}^{n} (c_j - (\mathbf{p}^T A)_j) x_j,$$

where $c_j - (\mathbf{p}^T A)_j \geq 0$ and $x_j \geq 0$ for $j = 1, 2, \ldots, n$. It follows that $(\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} = 0$ if and only if $c_j - (\mathbf{p}^T A)_j = 0$ for all integers j between 1 and n for which $x_j > 0$. The result follows. ∎

Corollary 5.3 *Let a linear programming problem in Dantzig standard form be specified by an* $m \times n$ *constraint matrix* A*, and* m*-dimensional target vector* $\mathbf{b}$ *and an* n*-dimensional cost vector* $\mathbf{c}$*. Let* $\mathbf{x}^*$ *be a feasible solution of this primal problem, and let* $\mathbf{p}^*$ *be a solution of the dual problem. Then* $\mathbf{p}^{*T} A \leq \mathbf{c}^T$*. Suppose that the complementary slackness conditions for this primal-dual pair are satisfied, so that* $(\mathbf{p}^{*T} A)_j = (\mathbf{c})_j$ *for all integers* j *between* 1 *and* n *for which* $(\mathbf{x}^*)_j > 0$*. Then* $\mathbf{x}^*$ *is an optimal solution of the primal problem, and* $\mathbf{p}^*$ *is an optimal solution of the dual problem.*

Proof Because the complementary slackness conditions for this primal-dual pair are satisfied, it follows from the Weak Duality Theorem that $\mathbf{c}^T\mathbf{x}^* = \mathbf{p}^{*T}\mathbf{b}$ (see Theorem 5.2). But it then also follows from the Weak Duality Theorem that

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{p}^{*T}\mathbf{b} = \mathbf{c}^T\mathbf{x}^*$$

for all feasible solutions $\mathbf{x}$ of the primal problem. It follows that $\mathbf{x}^*$ is an optimal solution of the primal problem. Similarly

$$\mathbf{p}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}^* = \mathbf{p}^{*T}\mathbf{b}$$

for all feasible solutions $\mathbf{p}$ of the dual problem. It follows that $\mathbf{p}^*$ is an optimal solution of the dual problem, as required. ∎

Another special case of duality in linear programming is exemplified by a primal-dual pair of problems in *Von Neumann Symmetric Form.* In this case the primal and dual problems are specified in terms of an $m \times n$ constraint matrix A, an m-dimensional target vector $\mathbf{b}$ and an n-dimensional cost vector $\mathbf{c}$. The objective of the problem is minimize $\mathbf{c}^T\mathbf{x}$ amongst n-dimensional vectors $\mathbf{x}$ that satisfy the constraints $A\mathbf{x} \geq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$. The dual problem is to maximize $\mathbf{p}^T\mathbf{b}$ amongst m-dimensional vectors $\mathbf{p}$ that satisfy the constraints $\mathbf{p}^T A \leq \mathbf{c}^T$ and $\mathbf{p} \geq \mathbf{0}$.

Theorem 5.4 (Weak Duality Theorem for Linear Programming Problems in Von Neumann Symmetric Form)
Let m and n be integers, let A be an $m \times n$ matrix with real coefficients, let $\mathbf{b} \in \mathbb{R}^m$ and let $\mathbf{c} \in \mathbb{R}^n$. Let $\mathbf{x} \in \mathbb{R}^n$ satisfy the constraints $A\mathbf{x} \geq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$, and let $\mathbf{p} \in \mathbb{R}^m$ satisfy the constraints $\mathbf{p}^T A \leq \mathbf{c}$ and $\mathbf{p}^T \geq \mathbf{0}$. Then $\mathbf{p}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}$. Moreover $\mathbf{p}^T\mathbf{b} = \mathbf{c}^T\mathbf{x}$ if and only if the following complementary slackness conditions are satisfied:

- $(A\mathbf{x})_i = (\mathbf{b})_i$ *for all integers i between 1 and m for which* $(\mathbf{p})_i > 0$;
- $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ *for all integers j between 1 and n for which* $(\mathbf{x})_j > 0$;

Proof The constraints satisfied by the vectors $\mathbf{x}$ and $\mathbf{p}$ ensure that

$$\mathbf{c}^T\mathbf{x} - \mathbf{p}^T\mathbf{b} = (\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} + \mathbf{p}^T(A\mathbf{x} - \mathbf{b}).$$

But $\mathbf{x} \geq \mathbf{0}$, $\mathbf{p} \geq \mathbf{0}$, $A\mathbf{x} - \mathbf{b} \geq \mathbf{0}$ and $\mathbf{c}^T - \mathbf{p}^T A \geq \mathbf{0}$. It follows that $\mathbf{c}^T\mathbf{x} - \mathbf{p}^T\mathbf{b} \geq 0$. and therefore $\mathbf{c}^T\mathbf{x} \geq \mathbf{p}^T\mathbf{b}$. Moreover $\mathbf{c}^T\mathbf{x} - \mathbf{p}^T\mathbf{b} = 0$ if and only if $(\mathbf{c}^T - \mathbf{p}^T A)_j(\mathbf{x})_j = 0$ for $j = 1, 2, \ldots, n$ and $(\mathbf{p})_i(A\mathbf{x} - \mathbf{b})_i = 0$, and therefore $\mathbf{c}^T\mathbf{x} = \mathbf{p}^T\mathbf{b}$ if and only if the complementary slackness conditions are satisfied. ∎

Theorem 5.5 (Weak Duality Theorem for Linear Programming Problems in General Primal Form)
Let $\mathbf{x} \in \mathbb{R}^n$ be a feasible solution to a linear programming problem

$$\mathrm{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$$

expressed in general primal form with constraint matrix A with m rows and n columns, target vector $\mathbf{b}$, cost vector $\mathbf{c}$, inequality constraint specifier I^+ and variable sign specifier J^+, and let $\mathbf{p} \in \mathbb{R}^m$ be a feasible solution to the corresponding dual programming problem

$$\mathrm{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+).$$

Then $\mathbf{p}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}$. Moreover $\mathbf{p}^T\mathbf{b} = \mathbf{c}^T\mathbf{x}$ if and only if the following complementary slackness conditions are satisfied:—

- $(A\mathbf{x})_i = \mathbf{b}_i$ *whenever* $(\mathbf{p})_i \neq 0$*;*
- $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ *whenever* $(\mathbf{x})_j \neq 0$.

Proof The feasible solution $\mathbf{x}$ to the primal problem satisfies the following constraints:—

- $A\mathbf{x} \geq \mathbf{b}$;
- $(A\mathbf{x})_i = (\mathbf{b})_i$ unless $i \in I^+$;
- $(\mathbf{x})_j \geq 0$ for all $j \in J^+$.

The feasible solution $\mathbf{p}$ to the dual problem satisfies the following constraints:—

- $\mathbf{p}^T A \leq \mathbf{c}^T$;
- $(\mathbf{p})_i \geq 0$ for all $i \in I^+$;
- $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ unless $j \in J^+$.

Now

$$\begin{aligned}\mathbf{c}^T\mathbf{x} - \mathbf{p}^T\mathbf{b} &= (\mathbf{c}^T - \mathbf{p}^T A)\mathbf{x} + \mathbf{p}^T(A\mathbf{x} - \mathbf{b}) \\ &= \sum_{j=1}^{n} (\mathbf{c}^T - \mathbf{p}^T A)_j (\mathbf{x})_j + \sum_{i=1}^{m} (\mathbf{p})_i (A\mathbf{x} - \mathbf{b})_i.\end{aligned}$$

Let j be an integer between 1 and n. If $j \in J^+$ then $(\mathbf{x})_j \geq 0$ and $(\mathbf{c}^T - \mathbf{p}^T A)_j \geq 0$, and therefore

$$(\mathbf{c}^T - \mathbf{p}^T A)_j (\mathbf{x})_j \geq 0.$$

If $j \notin J^+$ then $(\mathbf{p}^T A)_j = (\mathbf{c})_j$, and therefore

$$(\mathbf{c}^T - \mathbf{p}^T A)_j (\mathbf{x})_j = 0,$$

irrespective of whether $(\mathbf{x})_j$ is positive, negative or zero. It follows that

$$\sum_{j=1}^{n} (\mathbf{c}^T - \mathbf{p}^T A)_j (\mathbf{x})_j \geq 0.$$

Moreover

$$\sum_{j=1}^{n} (\mathbf{c}^T - \mathbf{p}^T A)_j (\mathbf{x})_j = 0$$

if and only if $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ for all indices j for which $(\mathbf{x})_j \neq 0$.

Next let i be an index between 1 and m. If $i \in I^+$ then $(\mathbf{p})_i \geq 0$ and $(A\mathbf{x} - \mathbf{b})_i \geq 0$, and therefore $(\mathbf{p})_i(A\mathbf{x} - \mathbf{b})_i \geq 0$. If $i \notin I^+$ then $(A\mathbf{x})_i = (\mathbf{b})_i$, and therefore $(\mathbf{p})_i(A\mathbf{x} - \mathbf{b})_i = 0$, irrespective of whether $(\mathbf{p})_i$ is positive, negative or zero. It follows that

$$\sum_{i=1}^{m} (\mathbf{p})_i (A\mathbf{x} - \mathbf{p})_i \geq 0.$$

Moreover

$$\sum_{i=1}^{m} (\mathbf{p})_i (A\mathbf{x} - \mathbf{p})_i = 0.$$

if and only if $(A\mathbf{x})_i = (\mathbf{b})_i$ for all indices i for which $(\mathbf{p})_i \neq 0$. The result follows. ∎

Corollary 5.6 *Let $\mathbf{x}^* \in \mathbb{R}^n$ be a feasible solution to a linear programming problem* $\mathrm{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$ *expressed in general primal form with constraint matrix A with m rows and n columns, target vector $\mathbf{b}$, cost vector $\mathbf{c}$, inequality constraint specifier I^+ and variable sign specifier J^+, and let $\mathbf{p}^* \in \mathbb{R}^m$ be a feasible solution to the corresponding dual programming problem* $\mathrm{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$. *Suppose that the complementary slackness conditions are satisfied for this pair of problems, so that $(A\mathbf{x})_i = \mathbf{b}_i$ whenever $(\mathbf{p})_i \neq 0$, and $(\mathbf{p}^T A)_j = (\mathbf{c})_j$ whenever $(\mathbf{x})_j \neq 0$. Then $\mathbf{x}^*$ is an optimal solution for the primal problem and $\mathbf{p}^*$ is an optimal solution for the dual problem.*

Proof Because the complementary slackness conditions for this primal-dual pair are satisfied, it follows from the Weak Duality Theorem that $\mathbf{c}^T\mathbf{x}^* = \mathbf{p}^{*T}\mathbf{b}$ (see Theorem 5.5). But it then also follows from the Weak Duality Theorem that

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{p}^{*T}\mathbf{b} = \mathbf{c}^T\mathbf{x}^*$$

for all feasible solutions $\mathbf{x}$ of the primal problem. It follows that $\mathbf{x}^*$ is an optimal solution of the primal problem. Similarly

$$\mathbf{p}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}^* = \mathbf{p}^{*T}\mathbf{b}$$

for all feasible solutions $\mathbf{p}$ of the dual problem. It follows that $\mathbf{p}^*$ is an optimal solution of the dual problem, as required. ∎

Example Consider the following linear programming problem in general primal form:—

find values of x_1*,* x_2*,* x_3 *and* x_4 *so as to minimize the objective function*

$$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4$$

subject to the following constraints:—

- $a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 = b_1$*;*
- $a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + a_{2,4}x_4 = b_2$*;*
- $a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + a_{3,4}x_4 \geq b_3$*;*
- $x_1 \geq 0$ *and* $x_3 \geq 0$*.*

Here $a_{i,j}$, b_i and c_j are constants for $i = 1, 2, 3$ and $j = 1, 2, 3, 4$. The dual problem is the following:—

find values of p_1*,* p_2 *and* p_3 *so as to maximize the objective function*

$$p_1b_1 + p_2b_2 + p_3b_3$$

subject to the following constraints:—

- $p_1a_{1,1} + p_2a_{2,1} + p_3a_{3,1} \leq c_1$*;*
- $p_1a_{1,2} + p_2a_{2,2} + p_3a_{3,2} = c_2$*;*
- $p_1a_{1,3} + p_2a_{2,3} + p_3a_{3,3} \leq c_3$*;*
- $p_1a_{1,4} + p_2a_{2,4} + p_3a_{3,4} = c_4$*;*
- $p_3 \geq 0$*.*

We refer to the first and second problems as the *primal problem* and the *dual problem* respectively. Let (x_1, x_2, x_3, x_4) be a feasible solution of the primal problem, and let (p_1, p_2, p_3) be a feasible solution of the dual problem. Then

$$\begin{aligned}\sum_{j=1}^{4} c_jx_j - \sum_{i=1}^{3} p_ib_i &= \sum_{j=1}^{4}\left(c_j - \sum_{i=1}^{3} p_ia_{i,j}\right)x_j \\ &\quad + \sum_{i=1}^{3} p_i\left(\sum_{j=1}^{4} a_{i,j}x_j - b_i\right).\end{aligned}$$

Now the quantity $c_j - \sum_{i=1}^{3} p_ia_{i,j} = 0$ for $j = 2$ and $j = 4$, and $\sum_{j=1}^{4} a_{i,j}x_j - b_i = 0$ for $i = 1$ and $i = 2$. It follows that

$$\sum_{j=1}^{4} c_jx_j - \sum_{i=1}^{3} p_ib_i = \left(c_1 - \sum_{i=1}^{3} p_ia_{i,1}\right)x_1$$

$$+\left(c_3 - \sum_{i=1}^{3} p_i a_{i,3}\right) x_3$$

$$+ p_3 \left(\sum_{j=1}^{4} a_{3,j} x_j - b_3\right).$$

Now $x_1 \geq 0$, $x_3 \geq 0$ and $p_3 \geq 0$. Also

$$c_1 - \sum_{i=1}^{3} p_i a_{i,1} \geq 0, \quad c_3 - \sum_{i=1}^{3} p_i a_{i,3} \geq 0$$

and

$$\sum_{j=1}^{4} a_{3,j} x_j - b_3 \geq 0.$$

It follows that

$$\sum_{j=1}^{4} c_j x_j - \sum_{i=1}^{3} p_i b_i \geq 0.$$

and thus

$$\sum_{j=1}^{4} c_j x_j \geq \sum_{i=1}^{3} p_i b_i.$$

Now suppose that

$$\sum_{j=1}^{4} c_j x_j = \sum_{i=1}^{3} p_i b_i.$$

Then

$$\begin{aligned} \left(c_1 - \sum_{i=1}^{3} p_i a_{i,1}\right) x_1 &= 0, \\ \left(c_3 - \sum_{i=1}^{3} p_i a_{i,3}\right) x_3 &= 0, \\ p_3 \left(\sum_{j=1}^{4} a_{3,j} x_j - b_3\right) &= 0, \end{aligned}$$

because a sum of three non-negative quantities is equal to zero if and only if each of those quantities is equal to zero.

It follows that

$$\sum_{j=1}^{4} c_j x_j = \sum_{i=1}^{3} p_i b_i$$

if and only if the following three complementary slackness conditions are satisfied:—

- $\sum_{i=1}^{3} p_i a_{i,1} = c_1$ if $x_1 > 0$;
- $\sum_{i=1}^{3} p_i a_{i,3} = c_3$ if $x_3 > 0$;
- $\sum_{j=1}^{4} a_{3,j} x_j = b_3$ if $p_3 > 0$.

5.4 Open and Closed Sets in Euclidean Spaces

Let m be a positive integer. The *Euclidean norm* $|\mathbf{x}|$ of an element $\mathbf{x}$ of $\mathbb{R}^m$ is defined such that

$$|\mathbf{x}|^2 = \sum_{i=1}^{m} (\mathbf{x})_i^2.$$

The *Euclidean distance function* d on $\mathbb{R}^m$ is defined such that

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{y} - \mathbf{x}|$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$. The Euclidean distance function satisfies the Triangle Inequality, together with all the other basic properties required of a distance function on a metric space, and therefore $\mathbb{R}^m$ with the Euclidean distance function is a metric space.

A subset U of $\mathbb{R}^m$ is said to be *open* in $\mathbb{R}^m$ if, given any point $\mathbf{b}$ of U, there exists some real number ε satisfying $\varepsilon > 0$ such that

$$\{\mathbf{x} \in \mathbb{R}^m : |\mathbf{x} - \mathbf{b}| < \varepsilon\} \subset U.$$

A subset of $\mathbb{R}^m$ is *closed* in $\mathbb{R}^m$ if and only if its complement is open in $\mathbb{R}^m$.

Every union of open sets in $\mathbb{R}^m$ is open in $\mathbb{R}^m$, and every finite intersection of open sets in $\mathbb{R}^m$ is open in $\mathbb{R}^m$.

Every intersection of closed sets in $\mathbb{R}^m$ is closed in $\mathbb{R}^m$, and every finite union of closed sets in $\mathbb{R}^m$ is closed in $\mathbb{R}^m$.

Lemma 5.7 *Let m be a positive integer, let $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(m)}$ be a basis of $\mathbb{R}^m$, and let*

$$F = \left\{ \sum_{i=1}^{m} s_i \mathbf{u}^{(i)} : s_i \geq 0 \text{ for } i = 1, 2, \ldots, m \right\}.$$

Then F is a closed set in $\mathbb{R}^m$.

Proof Let $T: \mathbb{R}^m \to \mathbb{R}^m$ be defined such that

$$T(s_1, s_2, \ldots, s_m) = \sum_{i=1}^{m} s_i \mathbf{u}^{(i)}$$

for all real numbers $s_1, s_2, \ldots, s_m$. Then T is an invertible linear operator on $\mathbb{R}^m$, and $F = T(G)$, where

$$G = \{\mathbf{x} \in \mathbb{R}^m : (\mathbf{x})_i \geq 0 \text{ for } i = 1, 2, \ldots, m\}.$$

Moreover the subset G of $\mathbb{R}^m$ is closed in $\mathbb{R}^m$.

Now it is a standard result of real analysis that every linear operator on a finite-dimensional vector space is continuous. Therefore $T^{-1}: \mathbb{R}^m \to \mathbb{R}^m$ is continuous. Moreover $T(G)$ is the preimage of the closed set G under the continuous map T^{-1}, and the preimage of any closed set under a continuous map is itself closed. It follows that $T(G)$ is closed in $\mathbb{R}^m$. Thus F is closed in $\mathbb{R}^m$, as required. ∎

Lemma 5.8 *Let m be a positive integer, let F be a non-empty closed set in $\mathbb{R}^m$, and let $\mathbf{b}$ be a vector in $\mathbb{R}^m$. Then there exists an element $\mathbf{g}$ of F such that $|\mathbf{x} - \mathbf{b}| \geq |\mathbf{g} - \mathbf{b}|$ for all $\mathbf{x} \in F$.*

Proof Let R be a positive real number chosen large enough to ensure that the set F_0 is non-empty, where

$$F_0 = F \cap \{\mathbf{x} \in \mathbb{R}^m : |\mathbf{x} - \mathbf{b}| \leq R\}.$$

Then F_0 is a closed bounded subset of $\mathbb{R}^m$. Let $f: F_0 \to \mathbb{R}$ be defined such that $f(\mathbf{x}) = |\mathbf{x} - \mathbf{b}|$ for all $\mathbf{x} \in F$. Then $f: F_0 \to \mathbb{R}$ is a continuous function on F_0.

Now it is a standard result of real analysis that any continuous real-valued function on a closed bounded subset of a finite-dimensional Euclidean space attains a minimum value at some point of that set. It follows that there exists an element $\mathbf{g}$ of F_0 such that

$$|\mathbf{x} - \mathbf{b}| \geq |\mathbf{g} - \mathbf{b}|$$

for all $\mathbf{x} \in F_0$. If $\mathbf{x} \in F \setminus F_0$ then

$$|\mathbf{x} - \mathbf{b}| \geq R \geq |\mathbf{g} - \mathbf{b}|.$$

It follows that

$$|\mathbf{x} - \mathbf{b}| \geq |\mathbf{g} - \mathbf{b}|$$

for all $\mathbf{x} \in F$, as required. ∎

5.5 A Separating Hyperplane Theorem

Definition A subset K of $\mathbb{R}^m$ is said to be *convex* if $(1-\mu)\mathbf{x}+\mu\mathbf{x}' \in K$ for all elements $\mathbf{x}$ and $\mathbf{x}'$ of K and for all real numbers μ satisfying $0 \leq \mu \leq 1$.

It follows from the above definition that a subset K of $\mathbb{R}^{>}$ is a convex subset of $\mathbb{R}^m$ if and only if, given any two points of K, the line segment joining those two points is wholly contained in K.

Theorem 5.9 *Let m be a positive integer, let K be a closed convex set in $\mathbb{R}^m$, and let $\mathbf{b}$ be a vector in $\mathbb{R}^m$, where $\mathbf{b} \notin K$. Then there exists a linear functional $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ and a real number c such that $\varphi(\mathbf{x}) > c$ for all $\mathbf{x} \in K$ and $\varphi(\mathbf{b}) < c$.*

Proof It follows from Lemma 5.8 that there exists a point $\mathbf{g}$ of K such that $|\mathbf{x}-\mathbf{b}| \geq |\mathbf{g}-\mathbf{b}|$ for all $\mathbf{x} \in K$. Let $\mathbf{x} \in K$. Then $(1-\lambda)\mathbf{g}+\lambda\mathbf{x} \in K$ for all real numbers λ satisfying $0 \leq \lambda \leq 1$, because the set K is convex, and therefore

$$|(1-\lambda)\mathbf{g}+\lambda\mathbf{x}-\mathbf{b}| \geq |\mathbf{g}-\mathbf{b}|$$

for all real numbers λ satisfying $0 \leq \lambda \leq 1$. Now

$$(1-\lambda)\mathbf{g}+\lambda\mathbf{x}-\mathbf{b} = \mathbf{g}-\mathbf{b}+\lambda(\mathbf{x}-\mathbf{g}).$$

It follows by a straightforward calculation from the definition of the Euclidean norm that

$$\begin{aligned} |\mathbf{g}-\mathbf{b}|^2 &\leq |(1-\lambda)\mathbf{g}+\lambda\mathbf{x}-\mathbf{b}|^2 \\ &= |\mathbf{g}-\mathbf{b}|^2 + 2\lambda(\mathbf{g}-\mathbf{b})^T(\mathbf{x}-\mathbf{g}) \\ &\qquad + \lambda^2|\mathbf{x}-\mathbf{g}|^2 \end{aligned}$$

for all real numbers λ satisfying $0 \leq \lambda \leq 1$. In particular, this inequality holds for all sufficiently small positive values of λ, and therefore

$$(\mathbf{g}-\mathbf{b})^T(\mathbf{x}-\mathbf{g}) \geq 0$$

for all $\mathbf{x} \in K$.

Let

$$\varphi(\mathbf{x}) = (\mathbf{g}-\mathbf{b})^T\mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^m$. Then $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ is a linear functional on $\mathbb{R}^m$, and $\varphi(\mathbf{x}) \geq \varphi(\mathbf{g})$ for all $\mathbf{x} \in K$. Moreover

$$\varphi(\mathbf{g}) - \varphi(\mathbf{b}) = |\mathbf{g}-\mathbf{b}|^2 > 0,$$

and therefore $\varphi(\mathbf{g}) > \varphi(\mathbf{b})$. It follows that $\varphi(\mathbf{x}) > c$ for all $\mathbf{x} \in K$, where $c = \frac{1}{2}\varphi(\mathbf{b}) + \frac{1}{2}\varphi(\mathbf{g})$, and that $\varphi(\mathbf{b}) < c$. The result follows. ∎

5.6 Convex Cones

Definition Let m be a positive integer. A subset C of $\mathbb{R}^m$ is said to be a *convex cone* in $\mathbb{R}^m$ if $\lambda\mathbf{v} + \mu\mathbf{w} \in C$ for all $\mathbf{v}, \mathbf{w} \in C$ and for all real numbers λ and μ satisfying $\lambda \geq 0$ and $\mu \geq 0$.

Lemma 5.10 *Let m be a positive integer. Then every convex cone in $\mathbb{R}^m$ is a convex subset of $\mathbb{R}^m$.*

Proof Let C be a convex cone in $\mathbb{R}^m$ and let $\mathbf{v}, \mathbf{w} \in C$. Then $\lambda\mathbf{v} + \mu\mathbf{w} \in C$ for all non-negative real numbers λ and μ. In particular $(1-\lambda)\mathbf{w} + \lambda\mathbf{v} \in C$. whenever $0 \leq \lambda \leq 1$, and thus the convex cone C is a convex set in $\mathbb{R}^m$, as required. ∎

Lemma 5.11 *Let S be a subset of $\mathbb{R}^m$, and let C be the set of all elements of $\mathbb{R}^m$ that can be expressed as a linear combination of the form*

$$s_1\mathbf{a}^{(1)} + s_2\mathbf{a}^{(2)} + \cdots + s_n\mathbf{a}^{(n)},$$

where $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ are vectors belonging to S and $s_1, s_2, \ldots, s_n$ are non-negative real numbers. Then C is a convex cone in $\mathbb{R}^m$.

Proof Let $\mathbf{v}$ and $\mathbf{w}$ be elements of C. Then there exist finite subsets S_1 and S_2 of S such that $\mathbf{v}$ can be expressed as a linear combination of the elements of S_1 with non-negative coefficients and $\mathbf{w}$ can be expressed as a linear combination of the elements of S_2 with non-negative coefficients. Let

$$S_1 \cup S_2 = \{\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}\}.$$

Then there exist non-negative real numbers $s_1, s_2, \ldots, s_n$ and $t_1, t_2, \ldots, t_n$ such that

$$\mathbf{v} = \sum_{j=1}^{n} s_j\mathbf{a}^{(j)} \quad \text{and} \quad \mathbf{w} = \sum_{j=1}^{n} t_j\mathbf{a}^{(j)}.$$

Let λ and μ be non-negative real numbers. Then

$$\lambda\mathbf{v} + \mu\mathbf{w} = \sum_{j=1}^{n}(\lambda s_j + \mu t_j)\mathbf{a}^{(j)},$$

and $\lambda s_j + \mu t_j \geq 0$ for $j = 1, 2, \ldots, n$. It follows that $\lambda\mathbf{v} + \mu\mathbf{w} \in C$, as required. ∎

Proposition 5.12 *Let m be a positive integer, let $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)} \in \mathbb{R}^m$, and let C be the subset of $\mathbb{R}^m$ defined such that*

$$C = \left\{ \sum_{j=1}^{n} t_j \mathbf{a}^{(j)} : t_j \geq 0 \text{ for } j = 1, 2, \ldots, n \right\}.$$

Then C is a closed convex cone in $\mathbb{R}^m$.

Proof It follows from Lemma 5.11 that C is a convex cone in $\mathbb{R}^m$. We must prove that this convex cone is a closed set.

The vectors $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ span a vector subspace V of $\mathbb{R}^m$ that is isomorphic as a real vector space to $\mathbb{R}^k$ for some integer k satisfying $0 \leq k \leq m$. This vector subspace V of $\mathbb{R}^m$ is a closed subset of $\mathbb{R}^m$, and therefore any subset of V that is closed in V will also be closed in $\mathbb{R}^m$. Replacing $\mathbb{R}^m$ by $\mathbb{R}^k$, if necessary, we may assume, without loss of generality that the vectors $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ span the vector space $\mathbb{R}^m$. Thus if A is the $m \times n$ matrix defined such that $(A)_{i,j} = (\mathbf{a}^{(j)})_i$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ then the matrix A is of rank m.

Let $\mathcal{B}$ be the collection consisting of all subsets B of $\{1, 2, \ldots, n\}$ for which the members of the set $\{\mathbf{a}^{(j)} : j \in B\}$ constitute a basis of the real vector space $\mathbb{R}^m$ and, for each $B \in \mathcal{B}$, let

$$C_B = \left\{ \sum_{i=1}^{m} s_i \mathbf{a}^{(j_i)} : s_i \geq 0 \text{ for } i = 1, 2, \ldots, m \right\},$$

where $j_1, j_2, \ldots, j_m$ are distinct and are the elements of the set B. It follows from Lemma 5.7 that the set C_B is closed in $\mathbb{R}^m$ for all $B \in \mathbb{B}$.

Let $\mathbf{b} \in C$. The definition of C then ensures that there exists some $\mathbf{x} \in \mathbb{R}^n$ that satisfies $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$. Thus the problem of determining $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ has a feasible solution. It follows from Theorem 4.2 that there exists a basic feasible solution to this problem, and thus there exist distinct integers $j_1, j_2, \ldots, j_m$ between 1 and n and non-negative real numbers $s_1, s_2, \ldots, s_m$ such that $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \ldots, \mathbf{a}^{(j_m)}$ are linearly independent and

$$\mathbf{b} = \sum_{i=1}^{m} s_i \mathbf{a}^{(j_i)}.$$

Therefore $\mathbf{b} \in C_B$ where

$$B = \{j_1, j_2, \ldots, j_m\}.$$

We have thus shown that, given any element $\mathbf{b}$ of C, there exists a subset B of $\{1, 2, \ldots, n\}$ belonging to $\mathcal{B}$ for which $\mathbf{b} \in C_B$. It follows from this that the subset C of $\mathbb{R}^m$ is the union of the closed sets C_B taken over all elements B of the finite set $\mathcal{B}$. Thus C is a finite union of closed subsets of $\mathbb{R}^m$, and is thus itself a closed subset of $\mathbb{R}^m$, as required. ∎

5.7 Farkas' Lemma

Proposition 5.13 *Let C be a closed convex cone in $\mathbb{R}^m$ and let $\mathbf{b}$ be a vector in $\mathbb{R}^m$. Suppose that $\mathbf{b} \notin C$. Then there exists a linear functional $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ such that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < 0$.*

Proof Suppose that $\mathbf{b} \notin C$. The cone C is a closed convex set. It follows from Theorem 5.9 that there exists a linear functional $\varphi\colon \mathbb{R}^m \to \mathbb{R}$ and a real number c such that $\varphi(\mathbf{v}) > c$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < c$.

Now $\mathbf{0} \in C$, and $\varphi(\mathbf{0}) = 0$. It follows that $c < 0$, and therefore $\varphi(\mathbf{b}) \leq c < 0$.

Let $\mathbf{v} \in C$. Then $\lambda\mathbf{v} \in C$ for all real numbers λ satisfying $\lambda > 0$. It follows that $\lambda\varphi(\mathbf{v}) = \varphi(\lambda\mathbf{v}) > c$ and thus $\varphi(\mathbf{v}) > \dfrac{c}{\lambda}$ for all real numbers λ satisfying $\lambda > 0$, and therefore

$$\varphi(\mathbf{v}) \geq \lim_{\lambda \to +\infty} \frac{c}{\lambda} = 0.$$

We conclude that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$.

Thus $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < 0$, as required. ∎

Lemma 5.14 (Farkas' Lemma) *Let A be a $m \times n$ matrix with real coefficients, and let $\mathbf{b} \in \mathbb{R}^m$ be an m-dimensional real vector. Then exactly one of the following two statements is true:—*

(i) there exists $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$;

(ii) there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A \geq \mathbf{0}$ and $\mathbf{y}^T\mathbf{b} < 0$.

Proof Let $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(n)}$ be the vectors in $\mathbb{R}^m$ determined by the columns of the matrix A, so that $(\mathbf{a}^{(j)})_i = (A)_{i,j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, and let

$$C = \left\{ \sum_{j=1}^{n} x_j \mathbf{a}^{(j)} : x_j \geq 0 \text{ for } j = 1, 2, \ldots, n \right\}.$$

It follows from Proposition 5.12 that C is a closed convex cone in $\mathbb{R}^m$. Moreover

$$C = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{x} \geq \mathbf{0}\}.$$

Thus $\mathbf{b} \in C$ if and only if there exists $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{b} = A\mathbf{x}$ and $\mathbf{x} \geq \mathbf{0}$. Therefore statement (i) in the statement of Farkas' Lemma is true if and only if $\mathbf{b} \in C$.

If $\mathbf{b} \notin C$ then it follows from Proposition 5.13 that there exists a linear functional $\varphi: \mathbb{R}^m \to \mathbb{R}$ such that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in C$ and $\varphi(\mathbf{b}) < 0$. Then there exists $\mathbf{y} \in \mathbb{R}^m$ with the property that $\varphi(\mathbf{v}) = \mathbf{y}^T\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^m$. Now $A\mathbf{x} \in C$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\mathbf{x} \geq \mathbf{0}$. It follows that $\mathbf{y}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\mathbf{x} \geq \mathbf{0}$. In particular $(\mathbf{y}^T A)_i = \mathbf{y}^T A\mathbf{e}^{(i)} \geq 0$ for $i = 1, 2, \ldots, m$, where $\mathbf{e}^{(i)}$ is the vector in $\mathbb{R}^m$ whose ith component is equal to 1 and whose other components are zero. Thus if $\mathbf{b} \notin C$ then there exists $\mathbf{y} \in \mathbb{R}^m$ for which $\mathbf{y}^T A \geq \mathbf{0}$ and $\mathbf{y}^T\mathbf{b} < 0$.

Conversely suppose that there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A \geq 0$ and $\mathbf{y}^T\mathbf{b} < 0$. Then $\mathbf{y}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\mathbf{x} \geq \mathbf{0}$, and therefore $\mathbf{y}^T\mathbf{v} \geq 0$ for all $\mathbf{v} \in C$. But $\mathbf{y}^T\mathbf{b} < 0$. It follows that $\mathbf{b} \notin C$. Thus statement (ii) in the statement of Farkas's Lemma is true if and only if $\mathbf{b} \notin C$. The result follows. ∎

Corollary 5.15 *Let A be a $m \times n$ matrix with real coefficients, and let $\mathbf{c} \in \mathbb{R}^n$ be an n-dimensional real vector. Then exactly one of the following two statements is true:—*

(i) there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A = \mathbf{c}^T$ and $\mathbf{y} \geq \mathbf{0}$;

(ii) there exists $\mathbf{v} \in \mathbb{R}^n$ such that $A\mathbf{v} \geq \mathbf{0}$ and $\mathbf{c}^T\mathbf{v} < 0$.

Proof It follows on applying Farkas's Lemma to the transpose of the matrix A that exactly one of the following statements is true:—

(i) there exists $\mathbf{y} \in \mathbb{R}^m$ such that $A^T\mathbf{y} = \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}$;

(ii) there exists $\mathbf{v} \in \mathbb{R}^m$ such that $\mathbf{v}^T A^T \geq \mathbf{0}$ and $\mathbf{v}^T\mathbf{c} < 0$.

But $\mathbf{v}^T\mathbf{c} = \mathbf{c}^T\mathbf{v}$. Also $A^T\mathbf{y} = \mathbf{c}$ if and only if $\mathbf{y}^T A = \mathbf{c}^T$, and $\mathbf{v}^T A^T \geq \mathbf{0}$ if and only if $A\mathbf{v} \geq \mathbf{0}$. The result follows. ∎

Corollary 5.16 *Let A be a $m \times n$ matrix with real coefficients, and let $\mathbf{c} \in \mathbb{R}^n$ be an n-dimensional real vector. Suppose that $\mathbf{c}^T\mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ satisfying $A\mathbf{v} \geq \mathbf{0}$. Then there exists some there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A = \mathbf{c}^T$ and $\mathbf{y} \geq \mathbf{0}$.*

Proof Statement (ii) in the statement of Corollary 5.15 is false, by assumption, and therefore statement (i) in the statement of that corollary must be true. The result follows. ∎

Proposition 5.17 *Let n be a positive integer, let I be a non-empty finite set, let $\varphi: \mathbb{R}^n \to \mathbb{R}$ be a linear functional on $\mathbb{R}^n$, and, for each $i \in I$, let $\eta_i: \mathbb{R}^n \to \mathbb{R}$ be a linear functional on $\mathbb{R}^n$. Suppose that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ with the property that $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I$. Then there exist non-negative real numbers g_i for all $i \in I$ such that $\varphi = \sum_{i \in I} g_i \eta_i$.*

Proof We may suppose that $I = \{1, 2, \ldots, m\}$ for some positive integer m. For each $i \in I$ there exist real numbers $A_{i,1}, A_{i,2}, \ldots, A_{i,n}$ such that

$$\eta_i(v_1, v_2, \ldots, v_n) = \sum_{j=1}^{n} A_{i,j} v_j$$

for $i = 1, 2, \ldots, m$ and for all real numbers $v_1, v_2, \ldots, v_n$. Let A be the $m \times n$ matrix whose coefficient in the ith row and jth column is the real number $A_{i,j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. Then an n-dimensional vector $\mathbf{v} \in \mathbb{R}^n$ satisfies $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I$ if and only if $A\mathbf{v} \geq \mathbf{0}$.

There exists an n-dimensional vector $\mathbf{c} \in \mathbb{R}^n$ such that $\varphi(\mathbf{v}) = \mathbf{c}^T\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^n$. Then $\mathbf{c}^T\mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ satisfying $A\mathbf{v} \geq 0$. It then follows from Corollary 5.16 that there exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T A = \mathbf{c}^T$ and $\mathbf{y} \geq \mathbf{0}$. Let $g_i = (\mathbf{y})_i$ for $i = 1, 2, \ldots, m$. Then $g_i \geq 0$ for $i = 1, 2, \ldots, m$ and $\sum_{i \in I} g_i \eta_i = \varphi$, as required. ∎

Remark The result of Proposition 5.17 can also be viewed as a consequence of Proposition 5.13 applied to the convex cone in the dual space $\mathbb{R}^{n*}$ of the real vector space $\mathbb{R}^n$ generated by the linear functionals η_i for $i \in I$. Indeed let C be the subset of $\mathbb{R}^{n*}$ defined such that

$$C = \left\{ \sum_{i \in I} g_i \eta_i : g_i \geq 0 \text{ for all } i \in I \right\}.$$

It follows from Proposition 5.12 that C is a closed convex cone in the dual space $\mathbb{R}^{n*}$ of $\mathbb{R}^n$. If the linear functional φ did not belong to this cone then it would follow from Proposition 5.13 that there would exist a linear functional $V: \mathbb{R}^{n*} \to \mathbb{R}$ with the property that $V(\eta_i) \geq 0$ for all $i \in I$ and $V(\varphi) < 0$.

But given any linear functional on the dual space of a given finite-dimensional vector space, there exists some vector belonging to the given vector space such that the linear functional on the dual space evaluates elements of the dual space at that vector (see Corollary 2.7). It follows that there would exist $\mathbf{v} \in \mathbb{R}^n$ such that $V(\psi) = \psi(\mathbf{v})$ for all $\psi \in \mathbb{R}^{n*}$. But then $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I$ and $\varphi(\mathbf{v}) < 0$. This contradicts the requirement that $\varphi(\mathbf{v}) \geq 0$

for all $\mathbf{v} \in \mathbb{R}^n$ satisfying $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I$. To avoid this contradiction it must be the case that $\varphi \in C$, and therefore there must exist non-negative real numbers g_i for all $i \in I$ such that $\varphi = \sum_{i \in I} g_i \eta_i$.

Corollary 5.18 *Let n be a positive integer, let I be a non-empty finite set, let $\varphi: \mathbb{R}^n \to \mathbb{R}$ be a linear functional on $\mathbb{R}^n$, and, for each $i \in I$, let $\eta_i: \mathbb{R}^n \to \mathbb{R}$ be a linear functional on $\mathbb{R}^n$. Suppose that there exists a subset I_0 of I such that $\varphi(\mathbf{v}) \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ with the property that $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I_0$. Then there exist non-negative real numbers g_i for all $i \in I$ such that $\varphi = \sum\limits_{i \in I} g_i \eta_i$ and $g_i = 0$ when $i \notin I_0$.*

Proof It follows directly from Proposition 5.17 that there exist non-negative real numbers g_i for all $i \in I_0$ such that $\varphi = \sum\limits_{i \in I_0} g_i \eta_i$. Let $g_i = 0$ for all $i \in I \setminus I_0$. Then $\varphi = \sum\limits_{i \in I_0} g_i \eta_i$, as required. ∎

Definition A subset X if $\mathbb{R}^n$ is said to be a *convex polytope* if there exist linear functionals $\eta_1, \eta_2, \ldots, \eta_m$ on $\mathbb{R}^n$ and real numbers $s_1, s_2, \ldots, s_m$ such that

$$X = \{\mathbf{x} \in \mathbb{R}^n : \eta_i(\mathbf{x}) \geq s_i \text{ for } i = 1, 2, \ldots, m\}.$$

Let $(\eta_i : i \in I)$ be a finite collection of linear functionals on $\mathbb{R}^n$ indexed by a finite set I, let s_i be a real number for all $i \in I$, and let

$$X = \bigcap_{i \in I} \{\mathbf{x} \in \mathbb{R} : \eta_i(\mathbf{x}) \geq s_i\}.$$

Then X is a convex polytope in $\mathbb{R}^n$. A point $\mathbf{x}$ of $\mathbb{R}^n$ belongs to the convex polytope X if and only if $\eta_i(\mathbf{x}) \geq s_i$ for all $i \in I$.

Proposition 5.19 *Let n be a positive integer, let I be a non-empty finite set, and, for each $i \in I$, let $\eta_i: \mathbb{R}^n \to \mathbb{R}$ be non-zero linear functional and let s_i be a real number. Let X be the convex polytope defined such that*

$$X = \bigcap_{i \in I} \{\mathbf{x} \in \mathbb{R} : \eta_i(\mathbf{x}) \geq s_i\}.$$

(Thus a point $\mathbf{x}$ of $\mathbb{R}^n$ belongs to the convex polytope X if and only if $\eta_i(\mathbf{x}) \geq s_i$ for all $i \in I$.) Let $\varphi: \mathbb{R}^n \to \mathbb{R}$ be a non-zero linear functional on $\mathbb{R}^n$, and let $\mathbf{x}^* \in X$. Then $\varphi(\mathbf{x}^*) \leq \varphi(\mathbf{x})$ for all $\mathbf{x} \in X$ if and only if there exist non-negative real numbers g_i for all $i \in I$ such that $\varphi = \sum\limits_{i \in I} g_i \eta_i$ and $g_i = 0$ whenever $\eta_i(\mathbf{x}^*) > s_i$.

Proof Let $K = \{i \in I : \eta_i(\mathbf{x}^*) > s_i\}$. Suppose that there do not exist non-negative real numbers g_i for all $i \in I$ such that $\varphi = \sum_{i \in I} g_i \eta_i$ and $g_i = 0$ when $i \in K$. Corollary 5.18 then ensures that there must exist some $\mathbf{v} \in \mathbb{R}^n$ such that $\eta_i(\mathbf{v}) \geq 0$ for all $i \in I \setminus K$ and $\varphi(\mathbf{v}) < 0$. Then

$$\eta_i(\mathbf{x}^* + \lambda \mathbf{v}) = \eta_i(\mathbf{x}^*) + \lambda \eta_i(\mathbf{v}) \geq s_i$$

for all $i \in I \setminus K$ and for all $\lambda \geq 0$. If $i \in K$ then $\eta_i(\mathbf{x}^*) > s_i$. The set K is finite. It follows that there must exist some real number λ_0 satisfying $\lambda_0 > 0$ such that $\eta_i(\mathbf{x}^* + \lambda \mathbf{v}) \geq s_i$ for all $i \in K$ and for all real numbers λ satisfying $0 \leq \lambda \leq \lambda_0$.

Combining the results in the cases when $i \in I \setminus K$ and when $i \in K$, we find that $\eta_i(\mathbf{x}^* + \lambda \mathbf{v}) \geq s_i$ for all $i \in I$ and $\lambda \in [0, \lambda_0]$, and therefore $\mathbf{x}^* + \lambda \mathbf{v} \in X$ for all real numbers λ satisfying $0 \leq \lambda \leq \lambda_0$. But

$$\varphi(\mathbf{x}^* + \lambda \mathbf{v}) = \varphi(\mathbf{x}^*) + \lambda \varphi(\mathbf{v}) < \varphi(\mathbf{x}^*)$$

whenever $\lambda > 0$. It follows that the linear functional φ cannot attain a minimum value in X at any point $\mathbf{x}^*$ for which either $K = I$ or for which K is a proper subset of I but there exist non-negative real numbers g_i for all $i \in I \setminus K$ such that $\varphi = \sum_{i \in I \setminus K} g_i \eta_i$. The result follows. ∎

5.8 Strong Duality

Example Consider again the following linear programming problem in general primal form:—

> *find values of x_1, x_2, x_3 and x_4 so as to minimize the objective function*
>
> $$c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4$$
>
> *subject to the following constraints:—*
>
> - $a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 = b_1$;
> - $a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + a_{2,4}x_4 = b_2$;
> - $a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + a_{3,4}x_4 \geq b_3$;
> - $x_1 \geq 0$ *and* $x_3 \geq 0$.

Now the constraint

$$a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 = b_1$$

can be expressed as a pair of inequality constraints as follows:

$$\begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 & \geq & b_1 \\ -a_{1,1}x_1 - a_{1,2}x_2 - a_{1,3}x_3 - a_{1,4}x_4 & \geq & -b_1. \end{array}$$

Similarly the equality constraint involving b_2 can be expressed as a pair or inequality constraints.

Therefore the problem can be reformulated as follows:—

> *find values of* x_1*,* x_2*,* x_3 *and* x_4 *so as to minimize the objective function*
>
> $$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4$$
>
> *subject to the following constraints:—*
>
> - $a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4 \geq b_1$*;*
> - $-a_{1,1}x_1 - a_{1,2}x_2 - a_{1,3}x_3 - a_{1,4}x_4 \geq -b_1$*;*
> - $a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + a_{2,4}x_4 \geq b_2$*;*
> - $-a_{2,1}x_1 - a_{2,2}x_2 - a_{2,3}x_3 - a_{2,4}x_4 \geq -b_2$*;*
> - $a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + a_{3,4}x_4 \geq b_3$*;*
> - $x_1 \geq 0$*;*
> - $x_3 \geq 0$.

Let

$$\varphi(x_1, x_2, x_3, x_4) = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4,$$

and let

$$\begin{array}{rcl} \eta_1^+(x_1, x_2, x_3, x_4) & = & a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + a_{1,4}x_4, \\ \eta_1^-(x_1, x_2, x_3, x_4) & = & -\eta_1(x_1, x_2, x_3, x_4), \\ \eta_2^+(x_1, x_2, x_3, x_4) & = & a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + a_{2,4}x_4, \\ \eta_2^-(x_1, x_2, x_3, x_4) & = & -\eta_3(x_1, x_2, x_3, x_4), \\ \eta_3(x_1, x_2, x_3, x_4) & = & a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + a_{3,4}x_4, \\ \zeta_1(x_1, x_2, x_3, x_4) & = & x_1, \\ \zeta_3(x_1, x_2, x_3, x_4) & = & x_3, \end{array}$$

Then (x_1, x_2, x_3, x_4) is a feasible solution to the primal problem if and only if this element of $\mathbb{R}^4$ belongs to the convex polytope X, where X is the subset of $\mathbb{R}^4$ consisting of all points $\mathbf{x}$ of $\mathbb{R}^4$ that satisfy the following constraints:—

- $\eta_1^+(\mathbf{x}) \geq b_1$;
- $\eta_1^-(\mathbf{x}) \geq -b_1$;
- $\eta_2^+(\mathbf{x}) \geq b_2$;
- $\eta_2^-(\mathbf{x}) \geq -b_2$;
- $\eta_3(\mathbf{x}) \geq b_3$;
- $\zeta_1(\mathbf{x}) \geq 0$;
- $\zeta_3(\mathbf{x}) \geq 0$.

An inequality constraint is said to be *binding* for a particular feasible solution $\mathbf{x}$ if equality holds in that constraint at the feasible solution. Thus the constraints on the values of η_1^+, η_1^-, η_2^+ and η_2^- are always binding at points of the convex polytope X, but the constraints determined by η_3, ζ_1 and ζ_3 need not be binding.

Suppose that the linear functional φ attains its minimum value at a point $\mathbf{x}^*$ of X, where $\mathbf{x}^* = (x_1^*, x_2^*, x_3^*, x_4^*)$. It then follows from Proposition 5.19 that there exist non-negative real numbers p_1^+, p_1^-, p_2^+, p_2^-, p_3, q_1 and q_3 such that

$$p_1^+\eta_1^+ + p_1^-\eta_1^- + p_2^+\eta_2^+ + p_2^-\eta_2^- + p_3\eta_3 + q_1\zeta_1 + q_3\zeta_3 = \varphi.$$

Moreover $p_3 = 0$ if $\eta_3(\mathbf{x}^*) > b_3$, $q_1 = 0$ if $\zeta_1(\mathbf{x}^*) > 0$, and $q_3 = 0$ if $\zeta_3(\mathbf{x}^*) > 0$.

Now $\eta_1^- = -\eta_1^+$ and $\eta_2^- = -\eta_2^+$. It follows that

$$p_1\eta_1^+ + p_2\eta_2^+ + p_3\eta_3 + q_1\zeta_1 + q_3\zeta_3 = \varphi,$$

where $p_1 = p_1^+ - p_1^-$ and $p_2 = p_2^+ - p_2^-$. Moreover $p_3 = 0$ if $\sum_{i=1}^{4} a_{3,j}x_j^* > b_3$, $q_1 = 0$ if $x_1^* > 0$, and $q_3 = 0$ if $x_3^* > 0$.

It follows that

$$\begin{aligned}
p_1a_{1,1} + p_2a_{2,1} + p_3a_{3,1} &\leq c_1, \\
p_1a_{1,2} + p_2a_{2,2} + p_3a_{3,2} &= c_2, \\
p_1a_{1,3} + p_2a_{2,3} + p_3a_{3,3} &\leq c_3, \\
p_1a_{1,4} + p_2a_{2,4} + p_3a_{3,4} &= c_4, \\
p_3 &\geq 0.
\end{aligned}$$

Moreover $p_3 = 0$ if $\sum_{i=1}^{4} a_{3,j}x_j^* > b_3$, $\sum_{i=1}^{3} p_i a_{i,1} = c_1$ if $x_1^* > 0$, and $\sum_{i=1}^{3} p_i a_{i,3} = c_3$ if $x_3^* > 0$. It follows that (p_1, p_2, p_3) is a feasible solution of the dual problem to the feasible primal problem.

Moreover the complementary slackness conditions determined by the primal problem are satisfied. It therefore follows from the Weak Duality Theorem (Theorem 5.5) that (p_1, p_2, p_3) is an optimal solution to the dual problem.

Theorem 5.20 (Strong Duality for Linear Programming Problems with Optimal Solutions)
Let $\mathbf{x}^* \in \mathbb{R}^n$ be an optimal solution to a linear programming problem

$$\text{Primal}(A, \mathbf{b}, \mathbf{c}, I^+, J^+)$$

expressed in general primal form with constraint matrix A with m rows and n columns, target vector $\mathbf{b}$, cost vector $\mathbf{c}$, inequality constraint specifier I^+ and variable sign specifier J^+. Then there exists an optimal solution $\mathbf{p}^*$ to the corresponding dual programming problem

$$\text{Dual}(A, \mathbf{b}, \mathbf{c}, I^+, J^+),$$

and moreover $\mathbf{p}^{*T}\mathbf{b} = \mathbf{c}^T\mathbf{x}^*$.

Proof Let $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)$, and let $A_{i,j} = (A)_{i,j}$, $b_i = (\mathbf{b})_i$ and $c_j = (\mathbf{c})_j$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. Then optimal solution $\mathbf{x}^*$ minimizes $\mathbf{c}^T\mathbf{x}^*$ subject to the following constraints:—

- $A\mathbf{x}^* \geq \mathbf{b}$;
- $(A\mathbf{x}^*)_i = b_i$ unless $i \in I^+$;
- $x_j^* \geq 0$ for all $j \in J^+$.

Let $\mathbf{p}$ be a feasible solution to the dual linear programming problem, where $\mathbf{p} = (p_1, p_2, \ldots, p_m)$. Then $\mathbf{p}$ must satisfy the following constraints:—

- $\mathbf{p}^T A \leq \mathbf{c}^T$;
- $p_i \geq 0$ for all $i \in I^+$;
- $(\mathbf{p}^T A)_j = c_j$ unless $j \in J^+$.

Now the constraints of the primal problem can be expressed in inequality form as follows:—

- $(A\mathbf{x}^*)_i \geq b_i$ for all $i \in I^+$;
- $(A\mathbf{x}^*)_i \geq b_i$ for all $i \in I \setminus I^+$; $(-A\mathbf{x}^*)_i \geq -b_i$ for all $i \in I \setminus I^+$;
- $x_j^* \geq 0$ for all $j \in J^+$.

Let

$$\varphi(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{n} c_j x_j,$$

$$\eta_i(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{n} (A)_{i,j} x_j \quad (i = 1, 2, \ldots, m)$$

$$\zeta_j(x_1, x_2, \ldots, x_n) = x_j \quad (j = 1, 2, \ldots, n)$$

It follows from Proposition 5.19 that if there exists an optimal solution to the primal problem then there exist non-negative quantities p_i for all $i \in I^+$, p_i^+ and p_- for all $i \in I \setminus I^+$ and q_j for all $j \in J^+$ such that

$$\varphi = \sum_{i \in I^+} p_i \eta_i + \sum_{i \in I \setminus I^+} (p_i^+ - p_i^-) \eta_i + \sum_{j \in J^+} q_j \zeta_j.$$

Moreover $p_i = 0$ whenever $i \in I^+$ and $\eta_i(x_1^*, x_2^*, \ldots, x_n^*)_i > b_i$ and $q_j = 0$ whenever $x_j^* > 0$.

Let $\mathbf{p}^* \in \mathbb{R}^m$ be defined such that $(\mathbf{p}^*)_i = p_i$ for all $i \in I^+$ and $(\mathbf{p}^*)_i = p_i^+ - p_i^-$ for all $i \in I \setminus I^+$. Then $(\mathbf{p}^{*T} A)_j \leq c_j$ for $j = 1, 2, \ldots, n$, $(\mathbf{p}^*)_i \geq 0$ for all $i \in I^+$, and $(\mathbf{p}^{*T} A)_j = c_j$ unless $j \in J^+$. Moreover $(\mathbf{p}^*)_i = 0$ whenever $(A\mathbf{x}^*)_i > b_i$ and $q_i = 0$ whenever $x_j > 0$. It follows that $\mathbf{p}^*$ is a feasible solution of the dual problem. Moreover the relevant complementary slackness conditions are satisfied by $\mathbf{x}^*$ and $\mathbf{p}^*$. It is then a consequence of the Weak Duality Theorem that $\mathbf{c}^T \mathbf{x}^* = \mathbf{p}^{*T} \mathbf{b}$, and that therefore $\mathbf{p}^*$ is an optimal solution of the dual problem (see Corollary 5.6). The result follows. ∎

5.9 Kuhn-Tucker Theory

We consider the *General Maximum Problem* of nonlinear programming. This problem may be stated as follows:

> **(The General Maximum Problem of Nonlinear Programming)** *Let $g, f_1, f_2, \ldots, f_m$ be differentiable real-valued functions on the set*
>
> $$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\},$$
>
> *and let*
>
> $$X = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0} \text{ and } f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\}.$$
>
> *Determine $\mathbf{x}^* \in X$ such that $g(\mathbf{x}^*) \geq g(\mathbf{x})$ for all $\mathbf{x} \in X$.*

Let $g,\ f_1, f_2, \ldots, f_m$ be differentiable real-valued functions on the set

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\},$$

and let

$$X = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0} \text{ and } f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\}.$$

Let $\mathbf{x}^*$ be an element of X with the property that $g(\mathbf{x}^*) \geq g(\mathbf{x})$ for all $\mathbf{x} \in X$.

Let $\gamma\colon (-\delta, \delta) \to \mathbb{R}^m$ be a differentiable path in $\mathbb{R}^m$, defined over an open interval $(-\delta, \delta)$ centred on 0, where $\delta > 0$, with the properties that $\gamma(t) \in X$ for all real numbers t satisfying $0 \leq t < \delta$ and $\gamma(0) = \mathbf{x}^*$, and let

$$\mathbf{v} = \gamma'(0) = \left.\frac{\partial \gamma(t)}{\partial t}\right|_{t=0}.$$

Let

$$I^0 = \{i \in \mathbb{N} : 1 \leq i \leq m \text{ and } f_i(\mathbf{x}^*) = 0\}$$

and

$$J^0 = \{j \in \mathbb{N} : 1 \leq j \leq n \text{ and } (\mathbf{x}^*)_j = 0\}.$$

If $i \in I^*$ then $f_i(\gamma(0)) = 0$ and $f_i(\gamma(t)) \geq 0$ for all $t \in [0, \delta)$. It follows from the Chain Rule of multivariable differential calculus that

$$(Df_i)_{\mathbf{x}^*}(\mathbf{v}) = (\text{grad } f_i)^T_{\mathbf{x}^*}\mathbf{v} = \sum_{j=1}^{n} (\mathbf{v})_j \left.\frac{\partial f_i}{\partial x_j}\right|_{\mathbf{x}^*} = \left.\frac{df_i(\gamma(t))}{dt}\right|_{t=0} \geq 0.$$

Thus $(Df_i)_{\mathbf{x}^*}(\mathbf{v}) \geq 0$ for all $i \in I^0$.

Also if $j \in J^0$ then $(\gamma(0))_j = 0$ and $(\gamma(t))_j \geq 0$ for all $t \in [0, \delta)$ and therefore $(\mathbf{v})_j \geq 0$.

Definition Let X be a subset of $\mathbb{R}^n$, let g be a differentiable real-valued function defined throughout some open neighbourhood of X, and let $\mathbf{x}^*$ be a point of X. We say that the function g achieves a local maximum on X at the point $\mathbf{x}^*$, if the inequality $g(\mathbf{x}) \leq g(\mathbf{x}^*)$ for all points $\mathbf{x}$ of X that lie sufficiently close to the point $\mathbf{x}^*$.

Let g be a differentiable real-valued function defined throughout some open neighbourhood of the set X, and let $\mathbf{x}^*$ be a point of X. Suppose that the function g achieves a local maximum on X at the point $\mathbf{x}^*$. Let $\gamma\colon (-\delta, \delta) \to \mathbb{R}^n$ be a differentiable curve, where $\delta > 0$, $\gamma(0) = \mathbf{x}^*$, and $\gamma(t) \in X$ for all real numbers t satisfying $0 \leq t < \delta$. Then $g(\gamma(t)) \leq g(\gamma(0))$ for all real numbers t satisfying $0 \leq t < \delta$, and therefore

$$(Dg)_{\mathbf{x}^*}(\mathbf{v}) = \left.\frac{d(g(\gamma(t)))}{dt}\right|_{t=0} \leq 0,$$

where

$$\mathbf{v} = \gamma'(0) = \left.\frac{d(\gamma(t))}{dt}\right|_{t=0}.$$

We have shown that if a vector $\mathbf{v}$ is tangent to a differentiable curve $\gamma\colon (-\delta, \delta) \to \mathbb{R}^n$ for which $\gamma(0) = \mathbf{x}^*$ and $\gamma(t) \in X$ when $0 \leq t < \delta$ then $(Df_i)_{\mathbf{x}^*}(\mathbf{v}) \geq 0$ for all $i \in I^0$ and $(\mathbf{v})_j \geq 0$ for all $j \in J^0$. Those points $\mathbf{x}^*$ where these properties characterize tangent vectors to diffentiable curves entering the region X at $\mathbf{x}^*$ are said to satisfy the *constraint qualification* (CQ). This constraint qualification is thus formally defined as follows.

Definition Let $f_1, f_2, \ldots, f_m$ be differentiable real-valued functions $\mathbb{R}^n$, let

$$X = \{\mathbf{x} \in \mathbb{R}^n : f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\},$$

and let $\mathbf{x}^* \in X$. The *constraint qualification* (CQ) is said to be satisfied at $\mathbf{x}^*$ if, given any vector $\mathbf{v} \in \mathbb{R}^n$ with the properties that $(Df_i)_{\mathbf{x}^*}(\mathbf{v}) \geq 0$ for all $i \in I^0$, where the set I^0 consists of those indices i between 1 and m for which $f_i(\mathbf{x}^*) = 0$, there exists a differentiable curve $\gamma\colon (-\delta, \delta) \to \mathbb{R}^n$, where $\delta > 0$, with the property that

$$\mathbf{v} = \left.\frac{d\gamma(t)}{dt}\right|_{t=0}.$$

Theorem 5.21 (Karush-Kuhn-Tucker) *Let $f_1, f_2, \ldots, f_m$ be differentiable real-valued functions on $\mathbb{R}^n$, let*

$$X = \{\mathbf{x} \in \mathbb{R}^n : f_i(\mathbf{x}) \geq 0 \textit{ for } i = 1, 2, \ldots, m\}$$

and let $g\colon X \to \mathbb{R}$ be a real-valued function on X. Suppose that the function g achieves a local maximum at some point $\mathbf{x}^*$ of X and is differentiable there. Suppose also that $f_i(\mathbf{x}^*) = 0$ for $i = 1, 2, \ldots, m$ and that the constraint qualification (CQ) is satisfied at the point $\mathbf{x}^*$. Then there exist non-negative real numbers $\lambda_1, \lambda_2, \ldots \lambda_m$ such that

$$\left.\frac{\partial g}{\partial x_j}\right|_{\mathbf{x}^*} + \sum_{i=1}^{m} \lambda_i \left.\frac{\partial f_i}{\partial x_j}\right|_{\mathbf{x}^*} = 0$$

for $j = 1, 2, \ldots, n$.

Proof Let C be the subset of $\mathbb{R}^n$ consisting of those vectors $\mathbf{v} \in \mathbb{R}^n$ with the properties that $(Df_i)_{\mathbf{x}^*}(\mathbf{v}) \geq 0$ for $i = 1, 2, \ldots, m$. Then C is a closed convex cone in $\mathbb{R}^n$. Let $\mathbf{v} \in C$. The constraint qualification (CQ) ensures that there exists a differentiable curve $\gamma\colon (-\delta, \delta) \to \mathbb{R}^n$, where $\delta > 0$, such that $\gamma(0) = \mathbf{x}^*$, $\gamma(t) \in X$ when $0 \leq t < \delta$ and

$$\left.\frac{d\gamma(t)}{dt}\right|_{t=0} = \mathbf{v}.$$

But then

$$(Dg)_{\mathbf{x}^*}(\mathbf{v}) = \left.\frac{dg(\gamma(t))}{dt}\right|_{t=0} \leq 0.$$

Let A be the $m \times n$ matrix whose coefficient in the ith row and jth column is $\dfrac{\partial f_i}{\partial x_j}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, and let $\mathbf{c}$ be the n-dimensional vector whose jth component is $\dfrac{\partial g}{\partial x_j}$ for $j = 1, 2, \ldots, n$. Then $\mathbf{c}^T\mathbf{v} \leq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ satisfying $A\mathbf{v} \geq \mathbf{0}$. It then follows from Corollary 5.16 that there exists $\mathbf{y} \in \mathbb{R}^m$ for which $\mathbf{y}^T A = -\mathbf{c}$. Let

$$\mathbf{y}^T = (\lambda_1, \lambda_2, \ldots, \lambda_m).$$

Then

$$\sum_{i=1}^{m} \lambda_j \frac{\partial f_i}{\partial x_j} = -\frac{\partial g}{\partial x_j}.$$

The result follows. ∎

Let $f_1, f_2, \ldots, f_m$ be differentiable real-valued functions on the set $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}$, let

$$X = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0} \text{ and } f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\},$$

and let $\mathbf{x}^* \in X$. The *constraint qualification* (CQ) is said to be satisfied at $\mathbf{x}^*$ if, given any vector $\mathbf{v} \in \mathbb{R}^n$ with the properties that $(Df_i)_{\mathbf{x}^*}(\mathbf{v}) \geq 0$ for all $i \in I^0$ and $(\mathbf{v})_j \geq 0$ for all $j \in J^0$, where the set I^0 consists of those indices i between 1 and m for which $f_i(\mathbf{x}^*) = 0$ and the set J^0 consists of those indices j between 1 and n for which $(\mathbf{x}^*)_j = 0$, there exists a differentiable curve $\gamma\colon (-\delta, \delta) \to \mathbb{R}^n$ (where $\delta > 0$) with the property that

$$\mathbf{v} = \left.\frac{d\gamma(t)}{dt}\right|_{t=0}.$$

Corollary 5.22 *Let $f_1, f_2, \ldots, f_m$ be differentiable real-valued functions on the set $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}$, let*

$$X = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0} \textit{ and } f_i(\mathbf{x}) \geq 0 \textit{ for } i = 1, 2, \ldots, m\}$$

and let $g\colon X \to \mathbb{R}$ be a real-valued function on X. Suppose that the function g achieves a local maximum at some point $\mathbf{x}^*$ of X and is differentiable there. Let I^0 be the set consisting of those indices i between 1 and m for which $f_i(\mathbf{x}^*) = 0$ and let J^0 be the set consisting of those indices j between 1 and n for which $(\mathbf{x}^*)_j = 0$. Suppose that the constraint qualification (CQ) is satisfied at the point $\mathbf{x}^*$. Then there exist real numbers $\lambda_1, \lambda_2, \ldots \lambda_m$, and $\mu_1, \mu_2, \ldots, \mu_n$ for which the following properties are satisfied:—

(i) $\left.\dfrac{\partial g}{\partial x_j}\right|_{\mathbf{x}^*} + \displaystyle\sum_{i=1}^{m} \lambda_i \left.\dfrac{\partial f_i}{\partial x_j}\right|_{\mathbf{x}^*} + \mu_j = 0$ *for* $j = 1, 2, \ldots, n$;

(ii) $\lambda_i \geq 0$ *for* $i = 1, 2, \ldots, m$ *and* $\mu_j \geq 0$ *for* $j = 1, 2, \ldots, n$;

(iii) $\lambda_i = 0$ *unless* $i \in I^0$, *and* $\mu_j = 0$ *unless* $j \in J^0$.

Proof We may assume, without loss of generality, that $I^0 = \{1, 2, \ldots, m\}$ and that if j is an index between 1 and n for which $(\mathbf{x}^*)_j = 0$ then the coordinate function $\mathbf{x} \mapsto (\mathbf{x})_j$ is included amongst the functions $f_1, f_2, \ldots, f_m$. This follows from the observation that we can, without loss of generality, ignore those functions f_i for which $f_i(\mathbf{x}^*) > 0$. Also we can augment the functions f_i for $i \in I^0$ with the functions $\mathbf{x} \mapsto (\mathbf{x})_j$ for all $j \in J^0$ in order to reduce the general problem to one in which the function g is defined over a subset X of $\mathbb{R}^n$ of the form

$$X = \{\mathbf{x} \in \mathbb{R}^n : f_i(\mathbf{x}) \geq 0 \text{ for } i = 1, 2, \ldots, m\},$$

where $X \subset \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}$ and $f_i(\mathbf{x}^*) = 0$ for $i = 1, 2, \ldots, m$. The result then follows on application of Theorem 5.21. ∎

Example This example was presented by Kuhn and Tucker in 1950. Let

$$X = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0,\ \ x_2 \geq 0 \text{ and } f(x_1, x_2) \geq 0\},$$

where

$$f(x_1, x_2) = (1 - x_1)^3 - x_2.$$

and let $g\colon \mathbb{R}^2 \to \mathbb{R}$ be defined so that $g(x_1, x_2) = x_1$. Then the maximum value of the function g on X is achieved at $(1, 0)$. At this point the gradient of g is $(1, 0)$ and the gradient of f is $(0, -1)$. These gradients are not collinear. This is not a counter example to the Kuhn-Tucker conditions stated in Theorem 5.21 because the constraint qualification (CQ) is not satisfied at $(1, 0)$.