5 Introduction to concave programming: Lagrange and Kuhn-Tucker

5.1 Introduction

We have already looked in some detail at maximization problems which required us to find the largest value of $f: R^n \to R$. To find such a value we were allowed to search through *all* $\mathbf{x} \in R^n$ to find an $\mathbf{x}^*$ at which f attains its largest value. The typical maximization problem thrown up by microeconomic theory is not of this form. Usually we are only allowed to search through those $\mathbf{x} \in R^n$ which satisfy certain *constraints* and we want to find the largest value of f *subject to* those constraints. The general concave programming problem is exactly such a constrained maximization problem. If $f: R^n \to R$ and $g_i: R^n \to R$, $i = 1, \ldots, m$ are $m + 1$ concave functions this general concave programming problem is

$$\text{(GCPP)} \quad \max f(\mathbf{x}) \quad \text{subject to} \quad g_1(\mathbf{x}) \geqslant 0, \ldots, g_m(\mathbf{x}) \geqslant 0$$

Notice that the constraints are all in *weak* inequality form. The g_i functions, $i = 1, \ldots, m$ are known as *constraint functions* and f is the *objective function*. The *feasible set* is

$$K = \{\mathbf{x} \in R^n | g_i(\mathbf{x}) \geqslant 0 \quad i = 1, \ldots, m\}$$

Alternatively (GCCP) can be written

$$\max f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{x} \in K$$

A *solution* to this problem is $\mathbf{x}^* \in K$ such that $f(\mathbf{x}^*) \geqslant f(\mathbf{x})$, $\forall \mathbf{x} \in K$; $\mathbf{x}^*$ is then said to be a *global* maximum of f on K and $f(\mathbf{x}^*)$ is the *optimal value* of f on K.

The task of this chapter is to introduce an important *characterization* of solutions to (GCCP). For the earlier (unconstrained) maximization problems we found that (assuming f is C^1 and concave)

stationarity characterized the solutions. However, stationarity of f is in general neither necessary nor sufficient for solutions to the new concave programming problem. For instance consider

$$\max -x^2 \quad \text{subject to} \quad x \geqslant 1 \text{ and } x \leqslant 2$$

The second constraint is $-x \geqslant -2$ and this is a special case of (GCCP) with feasible set $[1, 2]$. Inspection of the behaviour of $-x^2$ on this interval reveals that $x^* = 1$ is the (unique) solution. However, it is not a stationary point of f; $f'(1) = -2 \neq 0$. So stationarity of f is not *necessary* for solutions to (GCCP) in general. Moreover f has a stationary point at 0 but this is not in the feasible set and so is not a solution. So mere stationarity of f is not *sufficient* to ensure solutions to (GCCP) in general either.

On the other hand if in (GCCP) f has a stationary point at $\mathbf{x}^*$ *in* K, it follows that $\mathbf{x}^* \in K$ and since f is concave, $f(\mathbf{x}^*) \geqslant f(\mathbf{x})$, $\forall \mathbf{x}$; in particular then $f(\mathbf{x}^*) \geqslant f(\mathbf{x})$, $\forall \mathbf{x} \in K$ and $\mathbf{x}^*$ is a solution of (GCCP). So stationarity of f *in* K is sufficient to ensure a solution to (GCCP). But this condition is still not necessary, as the above one-variable example still demonstrates. In other words there is not an obvious *characterization* of solutions to (GCCP) in terms of the stationarity of f.

The characterization we want to introduce in this chapter in fact evolves from an attempt to produce *a new function* whose stationary points do characterize solutions to (GCCP). This new function is called the *Lagrangean function* for (GCCP) and the characterizing conditions are *Kuhn-Tucker conditions*, involving in fact a little more than mere stationarity of the Lagrangean. We introduce these concepts shortly by looking at a simple special case of (GCCP) where $m = n = 1$; i.e. there is one variable and one constraint. The generalization to n variables is studied in the next chapter, and the case of many constraints is not needed until much later.

Before proceeding two remarks are in order. The *linear programming problem* is

$$\text{(LP)} \quad \max f(\mathbf{x}) \quad \text{subject to} \quad g_i(\mathbf{x}) \geqslant 0, i = 1, \ldots, m$$
$$\text{and} \quad x_i \geqslant 0, i = 1, \ldots, n$$

where f and all g_i are *linear functions*. This is a *special case* of (GCCP) with $m + n$ linear (and so concave) constraints and a linear (hence concave) objective. Once we have mastered (GCCP) we can apply our knowledge to this (LP) special case. Because of the linearity, however, we will be able to find new results and procedures for (LP) not available in general for (GCCP). The discussions of these points come much later.

Secondly we will also want to generalize (GCCP) to the case where all functions are merely *quasi-concave*. This will lead us into *quasi-concave programming* and naturally we defer study of this until we have covered concave programming and quasi-concave functions more fully.

5.2 Concave programming on R with one constraint

The problem is:

$$\max f(x) \quad \text{subject to} \quad g(x) \geqslant 0 \tag{5.1}$$

Since there is only one constraint we do not bother subscripting g. $f: R \to R$ and $g: R \to R$ are assumed concave and C^1, although C^1 is not needed for all our statements. The feasible set is $K = \{x \in R | g(x) \geqslant 0\}$. Assuming 0 is in the range of f, K is non-empty and $K = UC_g(0)$. However, the characterization of solutions to (5.1) which we want needs somewhat more than the non-emptiness of K. We assume that the following so-called constraint qualification (CQ) is satisfied:

$$\text{(CQ)} \quad \exists\, x \in R \quad \text{where } g(x) > 0$$

Since $K = UC_g(0)$ we know from chapter 2 that K is a convex subset of R; i.e. K is an interval of R. In fact we can be more specific about the nature of this interval. For suppose $UC_g(0)$ is of the form (a, b). Then for $\epsilon > 0$ and 'small', $g(a+\epsilon) \geqslant 0$. But g is continuous; so $\lim_{\epsilon \to 0} g(a+\epsilon) = g(a)$. Since $g(a+\epsilon) \geqslant 0$, $\forall \epsilon > 0$ and sufficiently small, the weak inequality must also remain true in the limit $\epsilon = 0$; hence $g(a) \geqslant 0$ and $a \in UC_g(0)$ as well, and $UC_g(0)$ cannot be of the form (a, b) as it must include a. Similarly it must include b. In fact this reasoning shows that $UC_g(0)$ must be an interval of one of the following forms:

$$\text{(i) } [a, b], \quad \text{(ii) } [a, +\infty), \quad \text{(iii) } (-\infty, b], \quad \text{(iv) } (-\infty, +\infty)$$

In case (iv), (5.1) reduces to: $\max f(x)$, and stationarity of f does still characterize the solutions to (5.1). To find a more general characterization which in fact covers all four possibilities we consider case (i) in detail. Assuming (CQ), fig. 5.1 illustrates the typical K in this case. Notice that $x = a$ is the unique x where $g(x) = 0$ and $g'(x) > 0$ while $x = b$ is defined by $g(x) = 0$ and $g'(x) < 0$.

Now if f has a stationary point at $x^* \in [a, b]$ this will be solution to (5.1). If f does not have a stationary point in $[a, b]$ then $f'(x) \neq 0$, $\forall x \in [a, b]$ and so either $f'(x) > 0$, $\forall x \in [a, b]$ or $f'(x) < 0$, $\forall x \in [a, b]$ since f is C^1. In the first case G_f is increasing throughout $[a, b]$ and

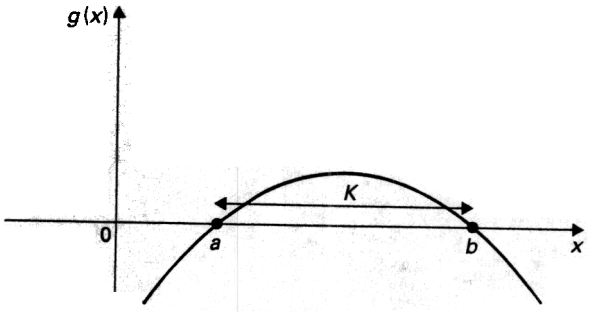

Figure 5.1

$x^* = b$ will be the solution: conversely $x^* = a$ is the solution in the second case. These are the only three possibilities: x^* solves (5.1) if and only if

either (1) $x^* \in [a, b]$ and $f'(x^*) = 0$
or (2) $x^* = b$ and $f'(x^*) > 0$
or (3) $x^* = a$ and $f'(x^*) < 0$.

Equivalently, x^* solves (7.3) if and only if

either (1) $g(x^*) \geqslant 0$ and $f'(x^*) = 0$
or (2) $g(x^*) = 0, g'(x^*) < 0$ and $f'(x^*) > 0$
or (3) $g(x^*) = 0, g'(x^*) > 0$ and $f'(x^*) < 0$.

This is a characterization of solutions to (5.1) but it does not have a useful generalization in higher dimensions. We try an alternative tack: can we find some new function whose stationary points do characterize solutions to (5.1) in case (i)?

Consider the *Lagrangean function* for (5.1) which is a function of the variables in the problem (x) and a new set of variables, one for each constraint, called *Lagrange multipliers*. Here there is only one constraint, so the one new variable is denoted λ and the Lagrangean function is

$$L : R^2 \to R \quad \text{where } L(x, \lambda) = f(x) + \lambda g(x)$$

We will give an intuitive justification for this choice of functional form later.

The introduction of λ has given us a new 'degree of freedom' in our search. Pose the question: can we find, for every possible solution to (5.1), a value of λ, λ^* say, which makes the solution a stationary point of L with respect to x? YES.

(1) If x^* is a 'type (1)' solution set $\lambda^* = 0$. Then $L(x, \lambda^*) = L(x, 0) = f(x)$ and $f'(x^*) = 0$ implies

$$\frac{\partial L}{\partial x}(x^*, 0) = 0$$

(2) If x^* is a 'type (2)' solution set

$$\lambda^* = -\frac{f'(x^*)}{g'(x^*)} > 0$$

Then

$$L(x, \lambda^*) = f(x) - \frac{f'(x^*)}{g'(x^*)} g(x)$$

$$\frac{\partial L}{\partial x}(x, \lambda^*) = f'(x) - \frac{f'(x^*)}{g'(x^*)} g'(x)$$

and so

$$\frac{\partial L}{\partial x}(x^*, \lambda^*) = 0 \text{ with } \lambda^* = -\frac{f'(x^*)}{g'(x^*)}$$

(3) If x^* is a 'type (3)' solution,

$\lambda^* = -\dfrac{f'(x^*)}{g'(x^*)} > 0$ ensures $\dfrac{\partial L}{\partial x}(x^*, \lambda^*) = 0$ as in (2).

Notice that $\lambda^* \geqslant 0$ in all cases.

So if x^* solves (5.1) in case (i) there is always a value of $\lambda^* \geqslant 0$ such that $(\partial L/\partial x)(x^*, \lambda^*) = 0$. In addition of course we must have $g(x^*) \geqslant 0$. But also either $\lambda^* = 0$ (case (1)) or $g(x^*) = 0$ (cases (2) and (3)): in all cases $\lambda^* g(x^*) = 0$.

The reader can check that our reasoning applies to feasible sets (ii), (iii) and (iv) as well. In case (iv), type (1) is the only possibility; in case (iii), (1) or (2) are the only possibilities; and (1) or (3) are the only possibilities for case (ii). In all cases, however, we can say: if x^* solves (5.1), there exists a real number λ^* such that

(a) $\dfrac{\partial L}{\partial x}(x^*, \lambda^*) = 0$

(b) $\lambda^* \geqslant 0$

(c) $\lambda^* g(x^*) = 0$

(d) $g(x^*) \geqslant 0$.

Condition (a) can be written alternatively. For $\lambda \geqslant 0$ and since f and g are concave it follows that $f(x) + \lambda g(x)$ is a concave function of x, for $x \in R$. For such functions stationarity characterizes the global maxima. Hence (a) is equivalent to, $L(x^*, \lambda^*) \geqslant L(x, \lambda^*)$, $\forall x \in R$. We have shown therefore that

If x^* solves (5.1), then there exists a real number λ^* such that

$$\left.\begin{array}{ll} \text{(A)} & L(x^*, \lambda^*) \geqslant L(x, \lambda^*) \quad \forall x \in R \\ \text{(B)} & \lambda^* \geqslant 0 \\ \text{(C)} & \lambda^* g(x^*) = 0 \\ \text{(D)} & g(x^*) \geqslant 0 \end{array}\right\} \quad \text{(K–T)}$$

The conditions indicated by (K–T) are the *Kuhn–Tucker* conditions for (5.1) and we have shown that they are *necessary conditions* at a solution to (5.1). Remarkably they are also sufficient, as follows.

From (A) $f(x^*) + \lambda^* g(x^*) \geqslant f(x) + \lambda^* g(x) \quad \forall x$

Using (C) $f(x^*) \geqslant f(x) + \lambda^* g(x) \quad \forall x$

From (B) $\lambda^* \geqslant 0$; hence, $\forall x$ such that $g(x) \geqslant 0$, $\lambda^* g(x) \geqslant 0$ and $f(x^*) \geqslant f(x)$, $\forall x$ such that $g(x) \geqslant 0$, i.e. $\forall x \in K$

From (D) $g(x^*) \geqslant 0$. So $x^* \in K$ and $f(x^*) \geqslant f(x)$, $\forall x \in K$. x^* is indeed a solution to (5.1).

(K–T) are therefore *necessary and sufficient* for, i.e. they characterize, solutions to (5.1) under our assumptions. We have proved:

Theorem 5.1 Suppose that f and g in (5.1) are concave and C^1 and suppose that (CQ) is satisfied. Then x^* solves (5.1) if and only if the Kuhn–Tucker conditions are satisfied.

A few remarks are in order. First our proof of theorem (5.1) has made use of the C^1 nature of f and g. The general proof of the next chapter does not need this: theorem 5.1 is true even if f and g are only concave. However, when f and g are C^1, (A) can be replaced by the previous Lagrangean stationarity condition (a). Secondly (CQ) is needed for the first part of our argument, the necessity of K–T. For instance suppose the graph of g is shown in fig. 5.2 (see example 3.3 for a specific such function): g is concave, K is convex but (CQ) is not satisfied. In fact $K = [0, 1]$. Suppose $f(x) = x$. Clearly $x^* = 1$ is the unique solution. But $g'(1) = 0$, $f'(1) = 1$ and there is no real number λ^* such that $1 + \lambda^*. 0 = 0$; so (CQ) is needed for necessity.

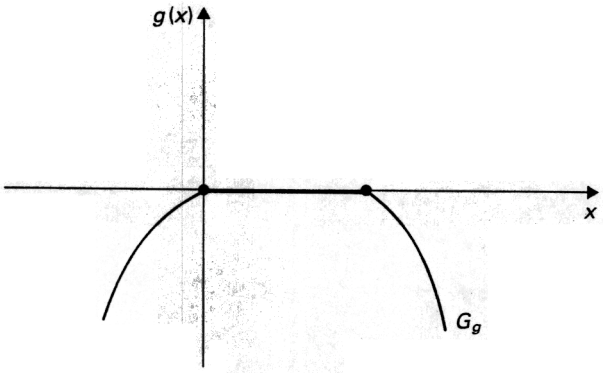

Figure 5.2

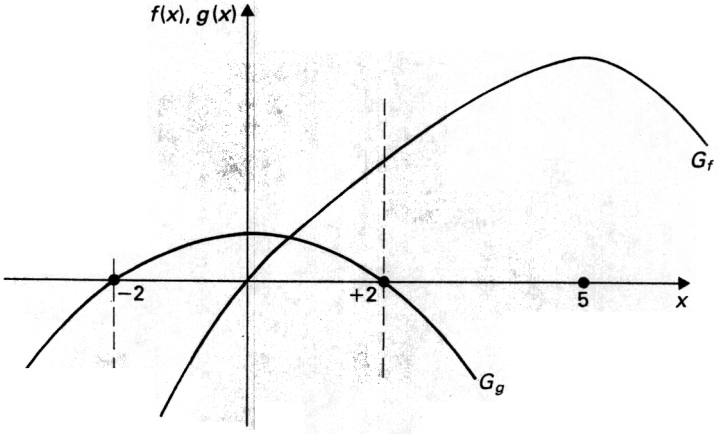

Figure 5.3

However, (CQ) is *not* needed in our sufficiency proof. Nevertheless in actual examples it is as well to check that (CQ) is satisfied so that you are sure that you are picking up all solutions to the problem.

Example 5.1 $\max 10x - x^2$ subject to $4 - x^2 \geqslant 0$. $f''(x) = -2 \leqslant 0$, $\forall x$, $g''(x) = -2 \leqslant 0$, $\forall x$ and f, g are concave; for instance, $4 - 0 > 0$ and (CQ) is satisfied. However, it is not necessary to use K–T to solve this. Figure 5.3 sketches G_g and G_f; the feasible set is $[-2, +2]$ and the solution is obviously $x^* = 2$. Nevertheless we work through the K–T derivation of this result to familiarize the reader with the mechanics. The Lagrangean is

$$L(x, \lambda) = 10x - x^2 + \lambda(4 - x^2)$$

(K–T) are:

(A) $\dfrac{\partial L}{\partial x} = 10 \quad 2x \quad 2\lambda x = 0$

(B) $\lambda \geqslant 0$

(C) $\lambda = 0$ or $4 - x^2 = 0$

(D) $4 - x^2 \geqslant 0$.

We have to find x^*, λ^* satisfying this system of equations and inequalities. There are three possibilities in (C): (a) $\lambda = 0$, (b) $x = -2$ or (c) $x = +2$. We try these in turn (even though we know it is the last one we want) to see if (A), (B) and (D) can also be satisfied.

(a) $\lambda = 0$. (A) $\Rightarrow x = 5$. (B) is satisfied, but $4 - 25 < 0$ and (D) is not.

(b) $x = -2$. (A) $\Rightarrow \lambda = -\frac{14}{4} < 0$ and (B) is not satisfied.

(c) $x = 2$. (A) $\Rightarrow \lambda = \frac{6}{4} = \frac{3}{2} > 0$ and (B) is satisfied. $4 \quad 4 = 0$ and (D) is satisfied.

So $x^* = +2$, $\lambda^* = \frac{3}{2}$ are the unique solutions to (K–T) and $x^* = +2$ *is the unique solution* to the original problem with *optimal value 16*.

Manipulation of K–T can be rather tedious even for very simple problems. Sometimes the objective function satisfies a condition which makes things easier. Suppose f has no *global unconstrained maximum*. In the C^1 case this is where f has no stationary point on R: we say f is *non-stationary*. Condition (A) can then never be satisfied with $\lambda^* = 0$. In this case (B) becomes '$\lambda^* > 0$', (C) therefore becomes '$g(x^*) = 0$' and (D) is redundant:

Corollary to theorem 5.1 Suppose in addition to the suppositions of theorem 5.1 that f is non-stationary. Then x^* solves (5.1) if and only if there exists a number λ^* such that,

(I) $L(x^*, \lambda^*) \geqslant L(x, \lambda^*) \quad \forall x \in R$

(II) $\lambda^* > 0$ and

(III) $g(x^*) = 0$.

(I)–(III) are the Kuhn–Tucker conditions for non-stationary objective functions. Of course non-stationarity of the objective is a strong assumption (not satisfied in example 5.1 for instance). Often in

economics it will be satisfied, however, and the mechanics become easier.

Example 5.2 max $7x-4$ subject to $4-x^2 \geqslant 0$
f is linear and so concave and non-stationary; g is concave and (CQ) is satisfied as in example 5.1. The Lagrangean is

$$L(x, \lambda) = 7x - 4 + \lambda(4-x^2)$$

(I) $\dfrac{\partial L}{\partial x} = 7 - 2\lambda x = 0$

(II) $\lambda > 0$

(III) $4 - x^2 = 0$

(III) $\Rightarrow x = -2$ or $x = +2$. If $x = -2$, (I) $\Rightarrow \lambda = -\frac{7}{4}$ and (II) is not satisfied. If $x = +2$, (I) $\Rightarrow \lambda = \frac{7}{4} > 0$ and (II) is satisfied. So $x^* = +2$, $\lambda^* = \frac{7}{4}$ are the unique solutions to (K–T) and $x^* = +2$ *is the unique solution* with *optimal value 10.*

5.3 Unique solutions

Concave programming problems like (5.1) may have no solution (e.g. max x subject to $x \geqslant 0$), one solution (e.g. examples 5.1 and 5.2) or more than one solution (see fig. 5.4); in fig. 5.4 any $x^* \in K$ is in fact a solution (the reader can supply functional forms for this: e.g. use example 3.3). As in the earlier unconstrained maximization discussion if there is more than one solution to a concave programming

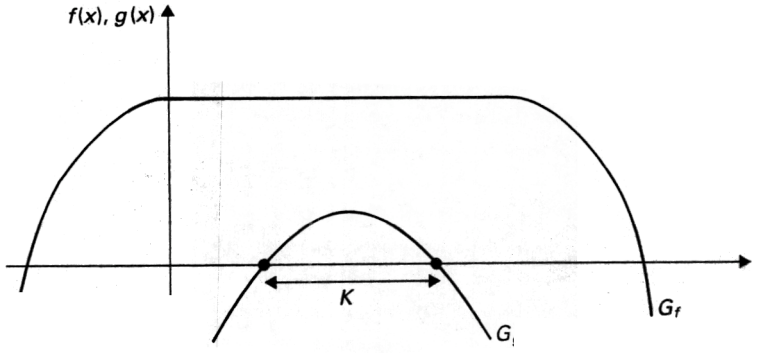

Figure 5.4

problem there are an infinite number of such solutions, and the set of solutions is convex (as is the case in fig. 5.4). To see why suppose x^1 and x^2 are two solutions to (5.1). Then $f(x^1) = f(x^2)$ and

$$g(x^1) \geqslant 0 \quad \text{and} \quad f(x^1) \geqslant f(x) \quad \forall x \in K$$

and

$$g(x^2) \geqslant 0 \quad \text{and} \quad f(x^2) \geqslant f(x) \quad \forall x \in K$$

Since g is concave,

$$g[\lambda x^1 + (1-\lambda)\, x^2] \geqslant \lambda g(x^1) + (1-\lambda)\, g(x^2) \geqslant 0$$
$$\forall \lambda \in [0, 1]$$

Since f is concave,

$$f[\lambda x^1 + (1-\lambda)\, x^2] \geqslant \lambda f(x^1) + (1-\lambda)\, f(x^2) = f(x^1) = f(x^2)$$
$$\forall \lambda \in [0, 1]$$
$$\geqslant f(x) \quad \forall x \in K, \forall \lambda \in [0, 1]$$

So, $g[\lambda x^1 + (1-\lambda)\, x^2] \geqslant 0$ and $f[\lambda x^1 + (1-\lambda)\, x^2] \geqslant f(x)$, $\forall x \in K$ and any convex combination of x^1 and x^2 is also a solution to (5.1).

As in chapter 3, strict concavity assumptions help rule out multiple solutions. There is now another case to note, however.

Theorem 5.2 Suppose that f and g in (5.1) are concave and C^1. If either (a) f is non-stationary or (b) f is strictly concave, then there is at most one solution to (5.1).

Proof (a) If f is non-stationary and C^1 either $f'(x) > 0$, $\forall x \in R$ or $f'(x) < 0$, $\forall x \in R$. In the former case the right-hand boundary of K (if there is one) must be the unique solution, while in the latter the left-hand boundary of K (if it exists) is the solution.

(b) Suppose f is strictly concave and let $x^1 \neq x^2$ be two solutions. Then

$$f[\lambda x^1 + (1-\lambda)\, x^2] > \lambda f(x^1) + (1-\lambda)\, f(x^2) \quad \forall \lambda \in (0, 1)$$
$$= f(x^1) = f(x^2)$$

since x^1, x^2 are solutions. Since g is concave, $g[\lambda x^1 + (1-\lambda)\, x^2] \geqslant \lambda g(x^1) + (1-\lambda)\, g(x^2) \geqslant 0$, $\forall \lambda \in [0, 1]$ since $x^1, x^2 \in K$. Hence, $\forall \lambda \in (0, 1)$

$$g[\lambda x^1 + (1-\lambda)\, x^2] \geqslant 0 \quad \text{and} \quad f[\lambda x^1 + (1-\lambda)\, x^2] > f(x^1)$$

and x^1 cannot be a solution. So there can be at most one solution.

Q.E.D.

Now suppose that (CQ) is satisfied in the concave programming problem (5.1) and suppose that (5.1) has a unique solution x^*. Then (K-T) will have a unique solution in x, namely x^*. Does it follow that there will be a unique value of λ^* satisfying (K-T) also? YES.

Theorem 5.3 Suppose that f and g in (5.1) are concave and C^1, and suppose that (CQ) is satisfied. If x^* is the unique solution to (5.1) there is a unique value of λ^* such that x^*, λ^* satisfies (K-T).

Proof If x^* solves (5.1) either (i) $g(x^*) > 0$ or (ii) $g(x^*) = 0$. In case (i), $\lambda^* = 0$ from (C) of (K-T) is the unique value of λ^* such that x^*, λ^* satisfy (K-T). In case (ii) suppose $\lambda_1^* \neq \lambda_2^*$ both satisfy (K-T) with x^*. From (C) of (K-T), $\lambda_1^*, \lambda_2^* > 0$. From (A)

$$f'(x^*) + \lambda_1^* g'(x^*) = f'(x^*) + \lambda_2^* g'(x^*) = 0$$

Therefore $(\lambda_1^* - \lambda_2^*) g'(x^*) = 0$ and $g'(x^*) = 0$ since $\lambda_1^* \neq \lambda_2^*$. But g is concave and so $g'(x^*) = 0$ implies x^* is a global maximum of g. Remembering $g(x^*) = 0$ in this case it follows that $g(x) \leqslant 0$, $\forall x \in R$, contradicting (CQ). So $\lambda_1^* = \lambda_2^*$ and there is a unique value of λ^* which satisfies (K-T) with x^*. *Q.E.D.*

So if either (a) or (b) of theorem 5.2 is satisfied, and if there is a solution to (5.1), there will be exactly one solution, x^* say, to (5.1) and exactly one solution x^*, λ^* to (K-T).

Both examples (5.1) and (5.2) exhibited such unique solutions, as do almost all the cases in this book. The uniqueness theorems 5.2 and 5.3 become very important in chapter 7.

5.4 Minimization problems

Consider:

$$\min f(x) \quad \text{subject to} \quad g(x) \geqslant 0 \tag{5.2}$$

For a solution, we require $x^* \in R$ with $g(x^*) \geqslant 0$ such that

$$f(x^*) \leqslant f(x), \forall x \quad \text{such that} \quad g(x) \geqslant 0$$

or

$$-f(x) \leqslant -f(x^*), \forall x \quad \text{such that} \quad g(x) \geqslant 0$$

But this last statement means exactly that x^* is a solution to the problem of type (5.1)

$$\max -f(x) \quad \text{subject to} \quad g(x) \geqslant 0 \tag{5.3}$$

So (5.3) is an equivalent maximization problem to the original minimization problem (5.2), in the sense that both have the same solution. But we can solve (5.3) now provided that $-f$ is concave and g is concave and satisfies (CQ). Hence, when presented with (5.2) we transform to the equivalent (5.3) and solve as before. Notice that $-f$ must be concave which means the original f is *convex*. To apply (K–T) to minimization problems the objective must be convex, with everything else as before.

Examples 5.3

$$\min x^2 - 4x + 1 \quad \text{subject to} \quad 1 - x^4 \geqslant 0$$

The equivalent maximization problem is

$$\max -x^2 + 4x - 1 \quad \text{subject to} \quad 1 - x^4 \geqslant 0$$

Now f is concave (quadratic with negative x^2 coefficient). $g''(x) = -12x^2 \leqslant 0$, $\forall x$ and g is concave. $g(0) = 1 > 0$, for instance and (CQ) is satisfied. f is not non-stationary so we need Kuhn–Tucker conditions (A)–(D). The Lagrangean is:

$$L(x, \lambda) = x^2 - 4x + 1 + \lambda(1 - x^4)$$

We need:

(A) $\dfrac{\partial L}{\partial x} = 2x \quad 4 \quad 4x^3\lambda = 0$

(B) $\lambda \geqslant 0$

(C) $\lambda = 0 \quad \text{or} \quad 1 \quad x^4 = 0$

(D) $1 - x^4 \geqslant 0$

Try $\lambda = 0$ first. Then, from (A) $x = 2$ and $1 - x^4 = 1 - 16 = -15 \geqslant 0$ and (D) is not satisfied. So there is no solution with $\lambda = 0$. Therefore try $1 - x^2 = 0$ or $x = \pm 1$. With $x = +1$, $\lambda = (-2/-4) = 1 > 0$ and $x = 1 = \lambda^*$ solves (A)–(D). Hence $x^* = 1$ solves our minimization problem since it solves the equivalent maximization problem. With $x = -1$, $\lambda = (-6/4) < 0$ and there is no solution. The optimal value in the min problem is -2.

Example 5.4

$$\min 3x + 1 \quad \text{subject to} \quad x^2 + x - 1 \leqslant 0$$

The constraint here is not in the form '$g(x) \geqslant 0$' which it must have for our results to apply. So we first rewrite the constraint as

$$1-x-x^2 \geqslant 0$$

and our problem is

$$\min 3x+1 \quad \text{subject to} \quad 1-x-x^2 \geqslant 0$$

The equivalent maximization problem is:

$$\max -3x-1 \quad \text{subject to} \quad 1-x-x^2 \geqslant 0$$

f is linear and hence concave and non-stationary. g is concave and $g(0)=1>0$ (for instance) so (CQ) is satisfied. The Lagrangean is:

$$L(x,\lambda) = -3x-1+\lambda(1-x-x^2)$$

and Kuhn–Tucker conditions (I)–(III) require:

(I) $\dfrac{\partial L}{\partial x} = -3-\lambda(1+2x) = 0$

(II) $\lambda > 0$

(III) $1-x-x^2 = 0$

From (III)

$$x = \frac{-1 \pm \sqrt{1+4}}{2} = -\frac{1}{2}+\frac{1}{2}\sqrt{5} \quad \text{or} \quad -\frac{1}{2}-\frac{1}{2}\sqrt{5}$$

With

$$x = -\frac{1}{2}+\frac{1}{2}\sqrt{5}, \text{ from (I)}, \lambda = \frac{-3}{\sqrt{5}} < 0$$

and there is no solution.

With

$$x = -\frac{1}{2}-\frac{1}{2}\sqrt{5}, = \frac{-3}{-\sqrt{5}} > 0$$

and we have a solution,

$$x^* = -\frac{1}{2}-\frac{1}{2}\sqrt{5}, \lambda^* = \frac{3}{5}$$

Hence

$$x^* = -\frac{1}{2}-\frac{1}{2}\sqrt{5}$$

solves our minimization problem with optimal value

$$-\frac{1}{2}-\frac{3}{2}\sqrt{5}.$$

5.5 Interpreting Lagrangeans

The introduction of the Lagrangean function earlier evolved from the desire to find a function whose stationary points or global maxima on R characterized solutions to the concave programming problem. However, we chose the Lagrangean functional form without any justification. To allay any worries the reader may have about this magical Lagrangean appearance we now tell the reader a story which provides some motivation for the choice of the Lagrangean functional form.

Think of a person and call him (or her) 'the maximizer'. We are going to present the maximizer with a (concave) function on R with values $M(x)$, say, and the maximizer is going to choose an $x \in R$. We will then give to the maximizer an amount of money equal to $M(x)$. The maximizer, being greedy and being able to find global maxima of concave functions on R, will choose the global maximum of M. Now our problem is: if x^* solves the concave programming problem (5.1) can we find a function M such that the maximizer's choice is x^*? $M = f$ suffices if x^* is a 'type 1' solution to (5.1) but not otherwise.

The problem with $M = f$ when we have a type (2) or (3) solution is, of course, that the maximizer's choice (a global max of M on R) will then violate the constraint. How can we stop the maximizer violating the constraint? We cannot keep $M = f$ and forbid choices outside the constraint, as all we are then doing is presenting the maximizer with the concave programming problem and our whole objective is that the maximizer's problem be unconstrained. So we cannot *ban* choices outside the feasible set; but maybe we could make such choices less attractive? From the basis of the original suggestion of $M = f$, a natural possibility for us to try is that we offer the maximizer $f(x)$ if the choice is x but we reduce this by an amount which is proportional to the extent of his constraint violation at x. So if the choice is x outside the feasible set, i.e. where $g(x) < 0$, the maximizer would get $f(x)$ less a 'forfeit' which would be proportional to the size of $g(x)$. With a non-negative constant of proportionality $k \geqslant 0$, this would lead us to present the maximizer with $M(x) = f(x) + kg(x)$. So if the choice is x with $g(x) < 0$, the

maximizer would get $f(x)$ plus the negative amount $kg(x)$. Incidentally, of course, we are then not only presenting the maximizer with a disincentive to choosing outside the feasible set (a forfeit of $kg(x) \leqslant 0$) but we are also simultaneously providing a positive incentive to choose inside this set (a 'bonus' of $kg(x) \geqslant 0$). Writing k as λ the Lagrangean functional form $f(x) + \lambda g(x)$ has emerged as a natural candidate to answer our original question. Section 5.2 now contains a demonstration that the Lagrangean does answer this question provided that λ is chosen as the λ^* which with x^* satisfies the Kuhn–Tucker conditions.

Hopefully this story helps to justify the appearance of the Lagrangean functional form. More importantly, and additionally, it suggests that the number λ^* may convey some useful information about solutions to concave programming problems: λ^* has emerged as a 'price' or 'cost' associated with constraint violations. In chapter 7 we give a precise description of this interpretation of Lagrange multipliers.

First we generalize to the case of n variables.

Exercise 5

1 Solve the following one variable maximization problems using (if possible) K–T:

(a) max $1-x^2$ subject to $(x-1)(x-2) \leqslant 0$
(b) max $-e^x$ subject to $1-x^2 \geqslant 0$

2 (a) Solve the following one-variable minimization problem using (if possible) K–T:

$$\min (x-1)(x-2) \quad \text{subject to} \quad 1-x^2 \geqslant 0$$

(b) Compare the answer with that to question 1(a).
(c) Show that if x^* solves max $f(x)$ subject to $g(x) \geqslant 0$, then x^* also solves max $g(x)$ subject to $f(x) \geqslant f(x^*)$.

3 Prove that the following problem has at most one solution if f is strictly convex and g is concave:

$$\min f(x) \text{ subject to } g(x) \geqslant 0.$$

6 Concave programming with one constraint

6.1 Introduction

We now generalize the Lagrangean/Kuhn–Tucker story of chapter 5 to the context of the following n-variable maximization problem:

$$\max f(\mathbf{x}) \quad \text{subject to } g(\mathbf{x}) \geqslant 0 \tag{6.1}$$

There is still only one constraint. Much later we will look at 'many-constraint' problems.

Throughout we assume that $f: R^n \to R$ and $g: R^n \to R$ are concave functions, of course, so that we do have a concave programming problem in (6.1). The discussion of chapter 5 generalizes for the most part in the obvious fashion.

6.2 The Kuhn–Tucker characterization

The feasible set for (6.1) is $K = UC_g(0) = \{\mathbf{x} \in R^n \mid g(\mathbf{x}) \geqslant 0\}$.

The constraint qualification for (6.1) is (CQ) $\exists\, \mathbf{x} \in R^n$ where $g(\mathbf{x}) > 0$.

The *Lagrangean function* for (6.1) is $L: R^{n+1} \to R$ where

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x})$$

The *Kuhn–Tucker conditions* for (6.1) are:

there exists a real number λ^* such that

(A) $L(\mathbf{x}^*, \lambda^*) \geqslant L(\mathbf{x}, \lambda^*) \quad \forall \mathbf{x}$

(B) $\lambda^* \geqslant 0$

(C) $\lambda^* g(\mathbf{x}^*) = 0$

(D) $g(\mathbf{x}^*) \geqslant 0$

Without more ado we state:

Theorem 6.1
Suppose f and g in (6.1) are concave and suppose (CQ) is satisfied. Then $\mathbf{x}^*$ solves (6.1) if and only if the Kuhn–Tucker conditions are satisfied.

Proof of 'if' Suppose $\mathbf{x}^*, \lambda^*$ satisfy (A)–(D).

From (D) $\mathbf{x}^* \in K$

From (A) $f(\mathbf{x}^*) + \lambda^* g(\mathbf{x}^*) \geqslant f(\mathbf{x}) + \lambda^* g(\mathbf{x}) \quad \forall \mathbf{x}$

From (C) $f(\mathbf{x}^*) \geqslant f(\mathbf{x}) + \lambda^* g(\mathbf{x}) \quad \forall \mathbf{x}$

From (B) $\lambda^* g(\mathbf{x}) \geqslant 0$ whenever $g(\mathbf{x}) \geqslant 0$

Hence

$$f(\mathbf{x}^*) \geqslant f(\mathbf{x}) \ \forall \mathbf{x} \quad \text{where } g(\mathbf{x}) \geqslant 0.$$

That is, $f(\mathbf{x}^*) \geqslant f(\mathbf{x})$, $\forall \mathbf{x} \in K$. Since $\mathbf{x}^* \in K$ it follows that $\mathbf{x}^*$ solves (6.1). *Q.E.D.*

Notice that (CQ) is not needed in this proof; it is, however, needed in the proof of 'only if'. The latter is a bit complicated and we relegate it to the end of the chapter.

Condition (A) is that $\mathbf{x}^*$ be a global maximum of $L(\mathbf{x}, \lambda^*)$ with respect to $\mathbf{x}$. Since f and g are concave and since $\lambda^* \geqslant 0$, L is a concave function of $\mathbf{x}$. So if f and g are in addition C^1 (then L is a C^1 function of $\mathbf{x}$), condition (A) is equivalent to the stationarity of L with respect to $\mathbf{x}$ at $\mathbf{x}^*$:

Corollary 1 to theorem 6.1 If f and g are in addition C^1 then (A in theorem 6.1 can be replaced with:

$$(1) \quad \frac{\partial L}{\partial x_i}(\mathbf{x}^*, \lambda^*) = 0 \quad i = 1, \quad ., n$$

or

$$(2) \quad f_i'(\mathbf{x}^*) + \lambda^* g_i'(\mathbf{x}^*) = 0 \quad i = \qquad n$$

We give some examples to illustrate the mechanics of (K–T) application to higher dimensions.

Example 6.1

$$\max \ \cdot x_1^2 - x_2^2 \quad \text{subject to } x_1 + x_2 \leqslant$$

We rewrite this as: max $-x_1^2 - x_2^2$ subject to $-x_1 - x_2 - 1 \geqslant 0$. f and g are concave and C^1, (CQ) is satisfied; e.g. when $x_1 = x_2 = -2$, $-x_1 - x_2 - 1 = 3 > 0$. The Lagrangean is:

$$L(\mathbf{x}, \lambda) = -x_1^2 - x_2^2 + \lambda(-x_1 - x_2 - 1)$$

(K-T) are

(A) $-2x_1 - \lambda = 0, -2x_2 \quad \lambda = 0$

(B) $\lambda \geqslant 0$

(C) $\lambda = 0$ or $-x_1 - x_2 \quad = 0$

(D) $-x_1 - x_2 - 1 \geqslant 0$

From (C) either $\lambda = 0$ or $-x_1 - x_2 - 1 = 0$. Try $\lambda = 0$ first. From (A), $\mathbf{x} = (0, 0)$ but then (D) is not satisfied. So try $-x_1 - x_2 - 1 = 0$. From (A), $x_1 = x_2$ and so $x_1 = x_2 = -\frac{1}{2}$. From (A), $\lambda = 1 \geqslant 0$ and (B) is satisfied. So $\mathbf{x}^* = (-\frac{1}{2}, -\frac{1}{2})$ is the unique solution and the optimal value is $-\frac{1}{2}$.

We can usefully depict the solution visually by sketching the feasible set and contours of the objective function (see fig. 6.1). $C_g(0)$ is the line $x_1 + x_2 = -1$ and $UC_g(0)$, i.e. the feasible set, is the set of points on or below the line, which is the shaded half-space in the diagram. The range of the objective function is $(-\infty, 0]$ and for $y \in (-\infty, 0]$, $C_f(y)$ has equation $x_1^2 + x_2^2 = -y$. Of course these contours are circles, centre the origin and radius $\sqrt{-y}$; when $y = 0$ the circle degenerates to a point, namely the origin. The global (unconstrained) maximum of f occurs at the origin where the value of f is zero and value of f increases as we move to circles closer to the origin. Our constrained maximization problem is visually to find the point in the feasible set which lies on the nearest contour to the origin. Clearly at such a point the contour through it will be tangential to the feasible set boundary, $C_g(0)$. Our calculations have found that $\mathbf{x}^* = (-\frac{1}{2}, -\frac{1}{2})$ is the unique point with this property, as indicated in fig. 6.1.

Example 6.2

max $x_1^2 \quad x_2^2$ subject to $x_1 + x_2 \geqslant$

This is the same as example 6.1 except that the sense of the constraint inequality is reversed. The new feasible set is the set of points on or *above* the line $x_1 + x_2 = -1$ in fig. 6.1, and now the global unconstrained maximum of f (viz. $\mathbf{x}^* = (0, 0)$ is feasible, and must

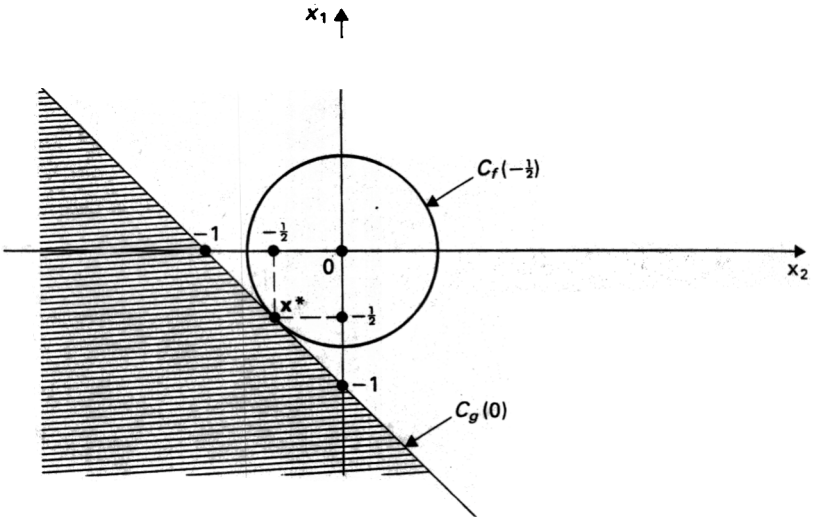

Figure 6.1

be the solution to the new problem. To confirm this, the new Lagrangean and (K–T) are:

$$L(\mathbf{x}, \lambda) = -x_1^2 - x_2^2 + \lambda(x_1 + x_2 + 1)$$

(A) $-2x_1 + \lambda = 0, -2x_2 + \lambda = 0$

(B) $\lambda \geqslant 0$

(C) $\lambda = 0$ or $x_1 + x_2 + 1 = 0$

(D) $x_1 + x_2 + 1 \geqslant 0$

With $\lambda = 0$ in (C), (B) is satisfied, $\mathbf{x} = (0, 0)$ satisfies (A) and $0 + 0 + 1 \geqslant 0$ satisfies (D). The reader can check that there is no solution with $x_1 + x_2 + 1 = 0$. So as expected $\mathbf{x}^* = (0, 0)$ is the unique solution with optimal value 0.

Analogously to chapter 5 if we assume f has no global unconstrained maximum (or in the C^1 case, f has no stationary point or is non-stationary), (A) cannot be satisfied with $\lambda^* = 0$. So (B) becomes $\lambda^* > 0$, (C) is $g(\mathbf{x}^*) = 0$ and (D) is redundant. Hence:

Corollary 2 to theorem 6.1

(a) Assuming f has no global unconstrained maximum and making the assumptions of theorem (6.1): $\mathbf{x}^*$ solves (6.1) if and only if

there exists a real number λ^* such that

(I) $L(\mathbf{x}^*, \lambda^*) \geqslant L(\mathbf{x}, \lambda^*) \quad \forall \mathbf{x}$

(II) $\lambda^* > 0$

(III) $g(\mathbf{x}^*) = 0$

(b) If in addition the assumptions of corollary 1 are satisfied then (I) can be replaced with (1) or (2) of corollary 1 to theorem 6.1.

Example 6.3

$$\max x_1 + x_2 \quad \text{subject to } x_1^2 + x_2^2 \leqslant 1$$

We rewrite as

$$\max x_1 + x_2 \quad \text{subject to } 1 - x_1^2 - x_2^2 \geqslant 0$$

f and g are concave and C^1; (CQ) is satisfied, e.g. by $\mathbf{x} = (0, 0)$. Moreover f is now non-stationary. The Lagrangean is

$$L(\mathbf{x}, \lambda) = x_1 + x_2 + \lambda(1 - x_1^2 - x_2^2)$$

We can use (K-T) in the (I)–(III) version:

(I) $1 - 2\lambda x_1 = 0,\ 1 - 2\lambda x_2 = 0$

$\lambda > 0$

$1 - x_1^2 - x_2^2 = 0$

From (I) $x_1 = x_2$. From (III) $x_1 = x_2 = \pm(1/\sqrt{2})$. From (I) $\lambda = 1/2x_1$ and (II) is satisfied only when $x_1 > 0$. Hence $\mathbf{x}^* = (1/\sqrt{2}, 1/\sqrt{2})$ is the unique solution with optimal value $= 2/\sqrt{2} = \sqrt{2}$.

Figure 6.2 sketches the feasible set (shaded) and the solution, similarly to fig. 6.1. Notice again that at the solution the contour of f through the solution is tangential to $C_g(0)$.

The tangency of the solution between $C_g(0)$ and the contour of f through the solution in examples 6.1 and 6.3 is no fluke. It will characterize all solutions to (6.1) with two variables *provided* $\lambda^* > 0$. For if $\lambda^* > 0$ then $g(\mathbf{x}^*) = 0$ and

$$\frac{\partial L}{\partial x_1} = f_1'(\mathbf{x}^*) + \lambda^* g_1'(\mathbf{x}^*) = 0$$

$$\frac{\partial L}{\partial x_2} = f_2'(\mathbf{x}^*) + \lambda^* g_2'(\mathbf{x}^*) = 0$$

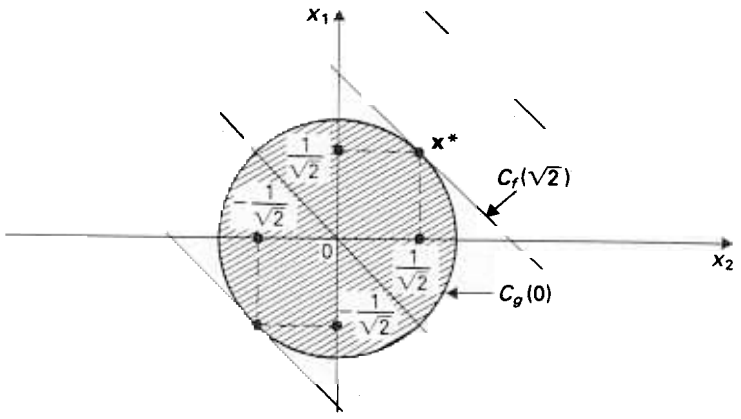

Figure 6.2

Eliminating λ^* from these two equations gives

$$-\frac{f_2'(\mathbf{x}^*)}{f_1'(\mathbf{x}^*)} = -\frac{g_2'(\mathbf{x}^*)}{g_1'(\mathbf{x}^*)}$$

which says that the contours of f and g through $\mathbf{x}^*$ have the same slope at $\mathbf{x}^*$. Since $g(\mathbf{x}^*) = 0$ the solution is indeed therefore characterized by tangency between the contour of f through the solution and $C_g(0)$.

In general the concave programming problem (6.1) may have 0, 1 or many solutions. However, if there is more than one solution then the set of solutions will be an (infinite) convex set; the reasoning in chapter 5 is straightforwardly adapted to show this. Conditions which rule out multiple solutions to (6.1) involve strict concavity and are rather different from those of theorem 5.2.

Theorem 6.2 Suppose that f and g in (6.1) are concave and C^1. If either (a) f is non-stationary *and* g is strictly concave or (b) f is strictly concave, then there is at most one solution to (6.1).

The proof of (b) follows exactly that of theorem 5.2(b)—details omitted.

The proof of (a) is given at the end of this chapter. If in addition (CQ) is satisfied then (K-T) characterize solutions to (6.1). If there is a unique solution $\mathbf{x}^*$ say to (6.1) (e.g. because of theorem 6.2) then there will be a unique solution $\mathbf{x}^*, \lambda^*$ say to (K-T). The proof is very similar to that of theorem (5.3)—details are left to the reader.

Theorem 6.3 Suppose that f and g in (6.1) are concave and C^1 and suppose (CQ) is satisfied. If $\mathbf{x}^*$ is the unique solution to (6.1) then there is a unique value of λ^* such that $\mathbf{x}^*, \lambda^*$ satisfies (K–T).

All three earlier examples exhibited unique solutions and unique Lagrange multipliers. Non-unique solutions are exhibited by the following trivial problems:

(a) $\max x_1 + x_2$ subject to $x_1 + x_2 \leqslant 0$

(b) $\max x_1 + x_2$ subject to $x_1 + x_2 \geqslant 0$

Using a sketch or by applying K–T the reader should find that (b) has no solution while (a) has an infinite (convex) set of solutions, namely $\{\mathbf{x} \in R^2 | x_1 + x_2 = 0\}$. Most cases studied in this book will enjoy uniqueness. Theorems 6.2 and 6.3 give us conditions which ensure such uniqueness of solutions and multipliers, and these are useful later.

6.3 Minimization problems

If our problem is

$$\min f(\mathbf{x}) \quad \text{subject to } g(\mathbf{x}) \geqslant 0$$

then exactly as in chapter 5 a maximizing problem with the same solution is

$$\max -f(\mathbf{x}) \quad \text{subject to } g(\mathbf{x}) \geqslant 0$$

If g is concave and satisfies (CQ) and if $-f$ is concave (i.e. f is *convex*) we can solve the maximizing problem and hence the minimizing problem.

Example 6.4

$$\min x_1 + x_2 \quad \text{subject to } x_1^2 + x_2^2 \leqslant 1$$

The equivalent maximizing problem is

$$\max -x_1 - x_2 \quad \text{subject to } 1 - x_1^2 - x_2^2 \geqslant 0$$

As in example 6.3 g is concave and C^1 and (CQ) is satisfied. f is linear and so concave C^1 and non-stationary. The Lagrangean and (K–T) are

$$L(\mathbf{x}, \lambda) = -x_1 - x_2 + \lambda(1 - x_1^2 - x_2^2)$$

(I) $-1-2\lambda x_1 = 0, -1-2\lambda x_2 = 0$

(II) $\lambda > 0$

(III) $1-x_1^2-x_2^2 = 0$

From (I) and (III) $x_1 = x_2 = \pm(1/\sqrt{2})$. From (I), $\lambda = -(1/2x_1) > 0$ only when $x_1 < 0$. Hence $\mathbf{x}^* = (-(1/\sqrt{2}), -(1/\sqrt{2}))$ is the unique solution (to the minimization problem) with optimal value $-(2/\sqrt{2}) = -\sqrt{2}$.

Visually the feasible set for this minimization problem is the same as in example 6.3, the shaded area of fig. 6.2. f is also the same as in example 6.3 but now we want to *minimize* f. We have to find the *lowest* valued contour of f attainable in the feasible set and clearly this is the 'lowest' contour of f shown in fig. 6.2, the solution occurring at $(-(1/\sqrt{2}), -(1/\sqrt{2}))$.

6.4 Proof of 'only if' in theorem 6.1

Suppose $\mathbf{x}^*$ solves (6.1). We have to show that (K–T) are then satisfied. (D) follows immediately. Let

$$A = \{\mathbf{x} \in R^n \mid f(\mathbf{x}) > f(\mathbf{x}^*)\}$$

and

$$B = \{\mathbf{x} \in R^n \mid g(\mathbf{x}) > 0\}$$

B is non-empty from (CQ). Since $\mathbf{x}^*$ solves (6.1) it must be that $A \cap B = \emptyset$. Let

$$C = \{(\gamma, \delta) \in R^2 \mid \exists\, \mathbf{x} \in R^n \text{ with } f(\mathbf{x}) - f(\mathbf{x}^*) > \gamma \text{ and } g(\mathbf{x}) > \delta\}$$

Now if $\gamma, \delta \geqslant 0$ there *cannot* exist $\mathbf{x} \in R^n$ with $f(\mathbf{x}) - f(\mathbf{x}^*) > \gamma$ and $g(\mathbf{x}) > \delta$ since otherwise there would be $\mathbf{x} \in R^n$ with $f(\mathbf{x}) - f(\mathbf{x}^*) > 0$ and $g(\mathbf{x}) > 0$ contradicting the disjointness of A and B. So $(\gamma, \delta) \in R^2_+$ cannot belong to C, or

$$C \cap R^2_+ = \emptyset$$

R^2_+ is non-empty and convex. C is non-empty since $R^2_{--} \subset C$ for if $\gamma < 0$ and $\delta < 0$, $f(\mathbf{x}^*) - f(\mathbf{x}^*) = 0 > \gamma$ and $g(\mathbf{x}^*) \geqslant 0 > \delta$ and $x^* \in R^n$ satisfies the requirements for γ, δ to be in C. C is also convex, as follows.

Suppose (γ_1, δ_1) and $(\gamma_2, \delta_2) \in C$. Then $\exists\, \mathbf{x}^1, \mathbf{x}^2 \in R^n$ with

(1) $f(\mathbf{x}^1) - f(\mathbf{x}^*) > \gamma_1$

(2) $g(\mathbf{x}^1) > \delta_1$

(3) $f(\mathbf{x}^2) - f(\mathbf{x}^*) > \gamma_2$

(4) $g(\mathbf{x}^2) > \delta_2$

We have to show that $(\lambda\gamma_1 + (1-\lambda)\gamma_2, \lambda\delta_1 + (1-\lambda)\delta^2) \in C$, for all $\lambda \in (0, 1)$; $\lambda = 0$ or 1 is covered by the supposition. So we have to show that $\forall \lambda \in (0, 1)$, $\exists\ \mathbf{x} \in R^n$ such that

(5) $f(\mathbf{x}) - f(\mathbf{x}^*) > \lambda\gamma_1 + (1-\lambda)\gamma_2$

(6) $g(\mathbf{x}) > \lambda\delta_1 + (1-\lambda)\,\delta_2$

Let $\lambda \in (0, 1)$ and let $\mathbf{x} = \lambda\mathbf{x}^1 + (1 \quad \lambda)\,\mathbf{x}^2$. Since f and g are concave

7) $f(\mathbf{x}) \geqslant \lambda f(\mathbf{x}^1) + (1-\lambda) f(\mathbf{x}^2)$

(8) $g(\mathbf{x}) \geqslant \lambda g(\mathbf{x}^1) + (1-\lambda)\, g(\mathbf{x}^2)$

Using (1) and (3) in (7) we get, since $\lambda \in (0, 1)$,

$$f(\mathbf{x}) > \lambda\gamma_1 + \lambda f(\mathbf{x}^*) + (1-\lambda)\,\gamma_2 + (1-\lambda)\, f(\mathbf{x}^*)$$

and so

$$f(\mathbf{x}) - f(\mathbf{x}^*) > \gamma_1 + (1-\lambda)\,\gamma_2$$

which establishes (5). Similarly (2), (4) and (8) lead to (6). So C is convex.

Hence C and R^2_+ are two disjoint, non-empty convex sets in R^2. By the separating hyperplane theorem there is a hyperplane in R^2 which separates them; i.e. $\exists\ \alpha_1, \alpha_2$ and β where α_1 or α_2 is non-zero such that

(9) $\alpha_1\gamma + \alpha_2\delta \leqslant \beta \quad \forall(\gamma, \delta) \in C$

and

(10) $\alpha_1\gamma + \alpha_2\delta \geqslant \beta \quad \forall(\gamma, \delta) \in R^2_+$

Now (10) implies $\beta \leqslant 0$ (since $(0, 0) \in R^2_+$). In addition we must have $\alpha_1 \geqslant 0$ and $\alpha_2 \geqslant 0$, since $\delta = 0$ and γ very large would violate (10) if $\alpha_1 < 0$; similarly for α_2. From (9) then

(11) $\alpha_1\gamma + \alpha_2\delta \leqslant 0 \quad \forall(\gamma, \delta) \in C$

Now for any $\mathbf{x} \in R^n$ and $\epsilon > 0$, let

$$y_1 = f(\mathbf{x}) - f(\mathbf{x}^*) - \epsilon$$

and

$$y_2 = g(\mathbf{x}) - \epsilon$$

Then $f(\mathbf{x}) - f(\mathbf{x}^*) = y_1 + \epsilon > y_1$ and $g(\mathbf{x}) = y_2 + \epsilon > y_2$. So $(y_1, y_2) \in C$. From (11)

$$\alpha_1 y_1 + \alpha_2 y_2 \leqslant 0$$

Therefore, for any $\mathbf{x} \in R^n$ and $\epsilon > 0$

$$\alpha_1[f(\mathbf{x}) - f(\mathbf{x}^*) - \epsilon] + \alpha_2[g(\mathbf{x}) \quad \epsilon] \leqslant 0$$

or

$$\alpha_1[f(\mathbf{x}) - f(\mathbf{x}^*)] + \alpha_2 g(\mathbf{x}) \leqslant \epsilon(\alpha_1 + \alpha_2)$$

Since this is true for *any* $\epsilon > 0$ we must in fact have

$$\alpha_1[f(\mathbf{x}) - f(\mathbf{x}^*)] + \alpha_2 g(\mathbf{x}) \leqslant 0$$

or

$$\alpha_1 f(\mathbf{x}) + \alpha_2 g(\mathbf{x}) \leqslant \alpha_1 f(x^*) \quad \forall \mathbf{x} \in R^n$$

Suppose $\alpha_1 = 0$: then $\alpha_2 > 0$ and $g(\mathbf{x}) \leqslant 0$, $\forall \mathbf{x} \in R^n$ contradicting (CQ). So $\alpha_1 > 0$ and writing $\lambda^* = (\alpha_2/\alpha_1) \geqslant 0$, we have

$$f(\mathbf{x}^*) \geqslant f(\mathbf{x}) + \lambda^* g(\mathbf{x}) \quad \forall \mathbf{x} \in R^n$$

When $\mathbf{x} = \mathbf{x}^*$ we get $\lambda^* g(\mathbf{x}^*) \leqslant 0$. On the other hand we know $\lambda^* \geqslant 0$ and $g(\mathbf{x}^*) \geqslant 0$. So $\lambda^* g(x^*) \geqslant 0$. It must be that $\lambda^* g(x^*) = 0$ completing the proof.

6.5 Proof of theorem 6.2(a)

Suppose $\mathbf{x}^1 \neq \mathbf{x}^2$ are two solutions to (6.1) where f is non-stationary and g is strictly concave. We derive a contradiction. From the supposition $f(\mathbf{x}^1) = f(\mathbf{x}^2)$; also

$$g(\mathbf{x}^1) \geqslant 0 \text{ and } f(\mathbf{x}^1) \geqslant f(\mathbf{x}) \quad \forall \mathbf{x} \in K$$

and

$$g(\mathbf{x}^2) \geqslant 0 \text{ and } f(\mathbf{x}^2) \geqslant f(\mathbf{x}) \quad \forall \mathbf{x} \in K$$

Let $\mathbf{x} = \lambda \mathbf{x}^1 + (1-\lambda)\,\mathbf{x}^2$ for some $\lambda \in (0, 1)$. Then since g is strictly concave:

$$g(\mathbf{x}) > \lambda g(\mathbf{x}^1) + (1-\lambda)\, g(\mathbf{x}^2)$$

and so

$$g(\mathbf{x}) > 0$$

Since f is non-stationary, $f_i'(\mathbf{x}) \neq 0$, some i. Without loss of generality suppose $f_1'(\mathbf{x}) > 0$. Let $\mathbf{y} = (x_1 + \epsilon, x_2, x_3, \ldots, x_n)$; for $\epsilon > 0$ and sufficiently small we have $g(\mathbf{y}) > 0$ since $g(\mathbf{x}) > 0$ and g is continuous.

Since $f_1'(\mathbf{x}) > 0$, for $\epsilon > 0$ and sufficiently small it also follows that $f(\mathbf{y}) > f(\mathbf{x})$. But from the concavity of f

$$f(\mathbf{x}) \geqslant \lambda f(\mathbf{x}^1) + (1-\lambda) f(\mathbf{x}^2) = f(\mathbf{x}^1) = f(\mathbf{x}^2)$$

Hence $f(\mathbf{y}) > f(\mathbf{x}^1)$ in particular. Therefore

$$g(\mathbf{y}) > 0$$

and

$$f(\mathbf{y}) > f(\mathbf{x}^1)$$

and $\mathbf{x}^1$ cannot be a solution to (6.1), a contradiction. So there is at most one solution to (6.1).

Exercise 6

Convex programming. Consider the problem:

$$\min f(\mathbf{x}) \quad \text{subject to } g(\mathbf{x}) \leqslant 0$$

where $\exists\, \mathbf{x}$ with $g(\mathbf{x}) < 0$ and $f: R^n \to R$ and $g: R^n \to R$ are convex functions. Show that if $\mathbf{x}^*$, λ^* satisfy the following four conditions then x^* solves this minimization problem:

(a) $L(\mathbf{x}^*, \lambda^*) \leqslant L(\mathbf{x}, \lambda^*) \quad \forall \mathbf{x}$

(b) $\lambda^* \geqslant 0$

(c) $\lambda^* g(\mathbf{x}^*) = 0$

(d) $g(\mathbf{x}^*) \leqslant 0$

where $L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x})$.

2. Solve the following two variable problems

(a) $\max x_2 - x_1 \quad \text{subject to } x_1 \geqslant e^{x_2}$

(b) $\min x_1^2 + (x_2 - 1)^2 \quad \text{subject to } x_1 \geqslant e^{x_2}$

3. For the problems in question 2 sketch in R^2 the feasible set and selected contours of the objective function indicating the solution.

4. Prove theorem 6.3.