Convex Programming

for The New Palgrave Dictionary of Economics, 2nd ed.

Lawrence E. Blume

Headwords: CONVEXITY, DUALITY, LAGRANGE MULTIPLIERS, QUASI-CONCAVITY

1   Introduction

Firms maximize profits and consumers maximize preferences. This is the core of microeconomics, and under conventional assumptions about decreasing returns it is an application of convex programming. The paradigm of convex optimization, however, runs even deeper through economic analysis. The idea that competitive markets perform well, which dates back at least to Adam Smith, has been interpreted since the neoclassical revolution as a variety of conjugate duality for the primal optimization problem of finding Pareto-optimal allocations. The purpose of this entry and the companion entry on DUALITY is (in part) to explain this sentence. This entry surveys without proof the basic mathematics of convex sets and convex optimization with an eye towards their application to microeconomic and general equilibrium theory, some of which can be found under DUALITY.

Unfortunately there is no accessible discussion of concave and convex optimization outside textbooks and monographs of convex analysis such as Rockafellar (1970,1974). Rather than just listing theorems, then, this entry attempts to provide a sketch of the main ideas. It is certainly no substitute for the sources. This entry covers only convex optimization in finite-dimensional vector spaces. While many of these ideas carry over to infinite-dimensional vector spaces and to important applications in infinite horizon economies and economies with non-trivial uncertainty, the mathematical subtleties of infinite-dimensional topological vector spaces raise issues which cannot reasonably

be treated here. The reader looking only for a statement of the Kuhn-Tucker Theorem is advised to read backwards from the end, to find the Theorem and notation.

A word of warning: This entry is written from the perspective of constrained maximization of concave functions because this is the canonical problem in microeconomics. Mathematics texts typically discuss the constrained minimization of convex functions, so textbook treatments will look slightly different.

2   Convex Sets

A subset C of a Euclidean space is convex if it contains the line segment connecting any two of its members. That is, if x and y are vectors in C and t is a number between 0 and 1, the vector $tx + (1 - t)y$ is also in C. A linear combination with non-negative weights which sum to 1 is a *convex combination* of elements of C; a set C is convex if it contains all convex combinations of its elements.

The key fact about convex sets is the famous *separation theorem*. A linear function p from V to $\mathbf{R}$, and a real number a define a *hyperplane*, the solutions to the equation $p \cdot x = a$. Every hyperplane divides V into two *half-spaces*; the upper (closed) half-space, containing those vectors x for which $p \cdot x \geq a$, and the lower (closed) half-space, containing those vectors x for which $p \cdot x \leq a$. The separation theorem uses linear functionals to describe closed convex sets. If a given vector is not in a closed convex set, then there is a hyperplane such that the set lies strictly inside the upper half-space while the vector lies strictly inside the lower half-space:

Separation Theorem: If C is a closed convex set and x is not in C, then there is a linear functional $p \in V^*$ and a real number a such that $p \cdot y > a$ for all $y \in C$, and $p \cdot x < a$.

This theorem implies that every closed convex set is the intersection of the half-spaces containing it. This half-space description is a dual description of closed convex sets, since it describes them with elements of the dual space. From the separation theorem the existence of a *supporting hyperplane* can also be deduced. If x is on the boundary of a closed convex set C, then there is a (non-zero) linear functional p such that $p \cdot y \geq p \cdot x$ for all $y \in C$; p is the hyperplane that supports C at x.

The origin of the term 'duality' lies in the mathematical construct of the dual to a vector space. The *dual space* of a vector space V is the collection of all *linear functionals*, that is, real-valued linear functions, defined on V. The distinction between vector spaces and their duals is obscured in finite dimensional spaces because each such space is its own dual. If an n-dimensional Euclidean vector space is represented by column vectors of length n, the linear functionals are $1 \times n$ matrices; that

is the dual to $\mathbf{R^n}$ is $\mathbf{R^n}$. (This justifies the notation used above.) Self-duality (called reflexivity in the literature) is not generally true in infinite-dimensional spaces, which is reason enough to avoid discussing them here. Nonetheless, although V will be $\mathbf{R^n}$ throughout this entry, the usual notation V^* will be used to refer to the dual space of V simply because it is important to know when we are discussing a vector in V and when we are discussing a member of its dual, a linear functional on V.

If the weights in a linear combination sum to 1 but are not constrained to be non-negative, then the linear combination is called an *affine combination.* Just as a convex set is a set which contains all convex combinations of its elements, an affine set in a vector space V is a set which contains all affine combinations of its elements. The set containing all affine combinations of elements in a given set C is an affine set, $A(C)$. The purpose of all this is to define the *relative interior* of a convex set C, $\operatorname{ri} C$. The relative interior of a convex set C is the interior of C relative to $A(C)$. A line segment in $\mathbf{R^2}$ has no interior, but it's relative interior is everything on the segment but its endpoints.

3   Concave Functions

The neoclassical assumptions of producer theory imply that production functions are concave and cost functions are convex. The quasi-concave functions which arise in consumer theory share much in common with concave functions, and quasi-concave programming has a rich duality theory.

In convex programming it is convenient to allow concave functions to take on the value $-\infty$ and convex functions to take on the value $+\infty$. A function f defined on some $\mathbf{R^n}$ with range $[-\infty, \infty)$ is concave if the set $\{(x, a) : a \in \mathbf{R}, \ a \leq f(x)\}$ is convex. This set, a subset of $\mathbf{R^{n+1}}$, is called the *hypograph* of f and is denoted $\operatorname{hypo} f$. Geometrically, it is the set of points in $\mathbf{R^{n+1}}$ that lie on or below the graph of f. Similarly, the *epigraph* of f is the set of points in $\mathbf{R^{n+1}}$ that lie on or above the graph of f: $\operatorname{epi} f = \{(x, a) : a \in \mathbf{R}, \ a \geq f(x)\}$. A function f with range $(-\infty, \infty]$ is convex $-f$ is concave, and convexity of f is equivalent to convexity of the set $\operatorname{epi} f$. Finally, the *effective domain* of a concave function is the set $\operatorname{dom} f = \{x \in \mathbf{R^n} : f(x) > -\infty\}$, and similarly for a convex function. Those familiar with the literature will note that attention here is restricted to *proper* concave and convex functions. Functions that are everywhere $+\infty$ will also be considered concave, and those everywhere $-\infty$ will be assumed convex when Lagrangeans are discussed below.

Convex optimization does not require that functions be differentiable or even continuous. Our main tool is the separation theorem, and for that closed convex sets are needed. A concave function f is *upper semi-continuous* (usc) if its hypograph is closed; a convex function is *lower semi-continuous* (lsc) if its epigraph is closed. Upper and lower semi-continuity apply to any functions, but these concepts interact nicely conveniently with convex and concave functions. In particular, usc concave

and lsc convex functions are continuous on the relative interiors of their domain. A famous example of a usc concave function that fails to be continuous is $f(x,y) = -y^2/2x$ for $x > 0$, 0 at the origin and $-\infty$ otherwise. Along the curve $y = \sqrt{\alpha x}$, $y \to 0$ as $x \to 0$, but f is constant at $-\alpha/2$, so f is not continuous at $(0,0)$, but it is usc because the supremum of the limits at the origin is 0.

It is useful to know that if f is concave and usc, then $f(x) = \inf q(x) = a \cdot x + b$ where the infimum is taken over all affine functions that are everywhere at least as big as f. This is another way of saying that, since $\mathrm{hypo}\, f$ is closed, it is the intersection of all halfspaces containing it.

4   The Fenchel Transform

The concave Fenchel transform associates with each usc function on a Euclidean space V, not necessarily concave, a usc concave function on its dual space V^* (which, we recall, happens to be V since its dimension is finite). The adjective concave is applied because a similar transform is defined slightly differently for convex functions. Henceforth we shall transform only concave functions, and misplace the adjective. The concave Fenchel transform of f is

$$f^*(p) = \inf_{x \in V}\{p \cdot x - f(x)\}.$$

which is often called the *conjugate* of f. (From here on out we will drop the braces.) The conjugate f^* of f is concave because, for fixed x, $p \cdot x - f(x)$ is linear, hence concave, in p, and the pointwise infimum of concave functions is concave. The textbooks all prove that if $\mathrm{hypo}\, f$ is closed, so is $\mathrm{hypo}\, f^*$, that is, upper semi-continuity is preserved by conjugation. So what is this transformation doing, and why is it interesting?

The conjugate f^* of a concave function f describes all the non-vertical half-spaces containing $\mathrm{hypo}\, f$. This should be checked. A half-space in $\mathbf{R}^{n+1}$ can be represented by the inequality $(p,q) \cdot (x,y) \geq a$ where q is a real number (as is a) and $p \in V^*$. The half-space is non-vertical if $p^* \neq 0$. In $\mathbf{R}^2$ this means geometrically that the line defining the boundary of the half-space is not vertical. So choose a linear functional $p \neq 0$ in V^*. For any $(x,z) \in \mathrm{hypo}\, f$, and any $p \in V^*$,

$$\begin{aligned}
p \cdot x - z &\geq p \cdot x - f(x) \\
&\geq \inf_{x \in V} p \cdot x - f(x) \\
&= f^*(p).
\end{aligned}$$

In other words, the upper half-space $(p,-1) \cdot (x,z) \geq f^*(p)$ contains $\mathrm{hypo}\, f$. It actually supports $\mathrm{hypo}\, f$ because of the infimum operation: If $a > f^*(p)$, there is an $(x,z) \in \mathrm{hypo}\, f$ such that $p \cdot x - z < a$, so the upper half-space fails to contain $\mathrm{hypo}\, f$.

Before seeing what the Fenchel transform is good for, we must answer an obvious question. If it is good to transform once, why not do it again? Define

$$f^{**}(x) = \inf_{p \in V^*} p \cdot x - f^*(p),$$

the *double dual* of f. The fundamental fact about the Fenchel transform is the following theorem, which is the function version of the dual descriptions of closed convex sets.

Conjugate Duality Theorem: If f is usc and concave, then $f^{**} = f$.

This is important enough to explain. Notice that just as p is a linear functional acting on x, so x is a linear functional acting on p. Suppose that f is concave and usc. For all x and p, $p \cdot x - f(x) \geq f^*(p)$, and so $p \cdot x - f^*(p) \geq f(x)$. Taking the infimum on the left, $f^{**}(x) \geq f(x)$.

On the other hand, take a $p \in V^*$ and a real number b such that the half-space $(p, -1) \cdot (x, z) \geq b$ in $\mathbf{R}^{n+1}$ contains the hypograph of f. This is true if and only if $p \cdot x - b \geq f(x)$ for all x and because f is usc, $f(x)$ is the infimum of $p \cdot x - b$ over all such p and b. Since $p \cdot x - f(x) \geq b$ for all x, take the infimum on the left to conclude that $f^*(p) \geq b$. Thus $p \cdot x - b \geq p \cdot x - f^*(p)$, and taking the infimum now on the right, $p \cdot x - b \geq f^{**}(x)$. Taking the infimum on the left over all the p and b such that the half-space contains $\text{hypo}\, f$, $f(x) \geq f^{**}(x)$.

It is worthwhile to compute an example to get the feel of the concave Fenchel transform. If C is a closed convex set, the *concave indicator function* of C is $\phi(x)$ which is 0 for x in C, and $-\infty$ otherwise. This is a good example to see the value of allowing infinite values. The Fenchel transform of C is $\phi^*(p) = \inf_{x \in \mathbf{R}^n} p \cdot x - \phi(x)$. Clearly the infimum cannot be reached at any $x \notin C$, for the value of ϕ at such an x is $-\infty$, and so the value of $p \cdot x - \phi(x)$ is $+\infty$. Consequently $\phi^*(p) = \inf_{x \in C} p \cdot x$. This function has the enticing property of positive homogeneity: If t is a positive scalar, then $\phi^*(tp) = t\phi^*(p)$.

Compute the double dual, first for $x \notin C$. The separating hyperplane theorem claims the existence of some p in V^* and a real number a such that $p \cdot x < a \leq p \cdot y$ for all $y \in C$. Take the infimum on the right to conclude that $p \cdot x < \phi^*(p)$, which is to say that $p \cdot x - \phi^*(p) < 0$. Then, multiply both sides by an arbitrary positive scalar t to conclude that $tp \cdot x - \phi^*(tp)$ can be made arbitrarily negative. Hence $\phi^{**}(x) = -\infty$ if $x \notin C$. And if x is in C? Then $p \cdot x - 0 \geq \phi^*(p)$ for all p (recall $\phi(x) = 0$). So $p \cdot x - \phi^*(p) \geq 0$. But $\phi^*(0) = 0$, so $\phi^{**}(x)$, the infimum of the left hand side over all possible p functionals, is 0, and so the Fenchel transform of ϕ^* recovers ϕ.

A particularly interesting version of this problem is to suppose that C is an 'at least as good as' set for level u of some upper semi-continuous and quasi-concave utility function (or, more generally,

a convex preference relation with closed weak upper contour sets). Then $\phi^*(p)$ is just the minimum expenditure necessary to achieve utility u at price p. See DUALITY for more discussion. Another interesting exercise is to apply the Fenchel transform to concave functions which are not usc, and to non-concave functions. These constructions have important applications in optimization theory which we will not pursue.

The theory of convex functions is exactly the same if, rather than the *concave Fenchel Transform*, the *convex Fenchel transform* is employed: $\sup_{x \in \mathbf{R^n}} p \cdot x - f(x)$. This transform maps convex lsc functions on V into convex lsc functions on V^*. Both the concave and convex Fenchel transforms will be important in what follows.

5   The Subdifferential

The separation theorem applied to $\operatorname{hypo} f$ implies that usc concave functions have tangent lines: For every $x \in \operatorname{ri dom} f$ there is a linear functional p_x such that $f(y) \leq f(x) + p_x \cdot (y - x)$. This inequality is called the *subgradient inequality,* and p_x is a *subgradient* of f; p_x defines a tangent line for the graph of f, and the graph lies on or underneath it. The set of subgradients of f at $x \in \operatorname{dom} f$ is denoted $\partial f(x)$, and is called the *subdifferential* of f at x. Subdifferentials share many of the derivative's properties. For instance, if $0 \in \partial f(x)$, then x is a global maximum of f. In fact, if $\partial f(x)$ contains only one subgradient p_x, then f is differentiable at x and $Df(x) = p_x$. The set $\partial f(x)$ need not be single-value, however, because f may have kinks. The graph of the function f defined on the real line such that $f(x) = -\infty$ for $x < 0$ and $f(x) = \sqrt{x}$ for $x \geq 0$ illustrates why the subdifferential may be empty at the boundary of the effective domain. At 0, a subgradient would be infinitely steep.

There is a corresponding *subgradient inequality* for convex f: $f(y) \geq f(x) + p_x \cdot (y - x)$. With these definitions, $\partial(-f)(x) = -\partial f(x)$. Note that some texts refer to superdifferentials for concave functions and subdifferentials for convex functions. Others do not multiply the required terminology, and we follow them.

The multivalued map $x \mapsto \partial f(x)$, is called the *subdifferential correspondence* of f. An important property of subdifferential correspondences is monotonicity. From the subgradient inequality, if $p \in \partial f(x)$ and $q \in \partial f(y)$, then $f(y) \leq f(x) + p \cdot (y - x)$ and $f(x) \leq f(y) + q \cdot (x - y)$, and it follows that $(p - q) \cdot (x - y) \leq 0$. For convex f the inequality is reversed.

The Fenchel transforms establish a clear relationship between the subdifferential correspondences of concave functions and their duals. If f is concave, then the subdifferential inequality says that $p \in \partial f(x)$ if and only if for all $z \in X$, $p \cdot x - f(x) \leq p \cdot z - f(z)$. The the map $z \mapsto p \cdot z - f(z)$

is minimized at $z = x$, and so p is in $\partial f(x)$ if and only if $f^*(p) = p \cdot x - f(x)$. If f is usc, then $f^{**} \equiv f$, and so $f^{**}(x) = f(x) = p \cdot x - f^*(x)$. That is, $p \in \partial f(x)$ if and only if $x \in \partial f^*(p)$.

6 Optimization and Duality

Economics most often presents us with constrained maximization problems. Within the class of problems with concave objective functions, there is no formal difference between constrained and unconstrained maximization. The constrained problem of maximizing concave and usc f on a closed convex set C is the same as the unconstrained problem of maximizing $f(x) + I_C(x)$ on $\mathbf{R^n}$, where $I_C(x)$ is the concave indicator function of C.

The general idea of duality schemes in optimization theory is to represent maximization (or minimization) problems as half of a minimax problem which has a saddle value. There are several reasons why such a seemingly odd construction can be useful. In economics it often turns out that the other half of the minimax problem, the dual problem, sheds additional light on properties and interpretations of the primal problem. This is the source of the 'shadow price' concept: The shadow price is the value of relaxing a constraint. Perhaps the most famous example of this is the Second Theorem of welfare economics.

6.1 Lagrangeans

The *primal problem* (problem P) is to maximize $f(x)$ on a Euclidean space V. Suppose there is a function $L : V \times V^* \to \mathbf{R}$ such that $f(x) = \inf_{p \in V^*} L(x, p)$. Define $g(p) = \sup_{x \in V} L(x, p)$, and consider the problems of maximizing $f(x)$ on V and minimizing $g(p)$ on V^*. The first problem is the primal problem, and the second is the *dual problem*. For all x and p it is clear that $f(x) \le L(x, p) \le g(p)$, and thus that

$$\sup_x \inf_p L(x, p) = \sup_x f(x) \le \inf_p g(p) = \inf_p \sup_x L(x, p).$$

If the inequality is tight, that is, it holds with equality, then the common value is called a *saddle value* of L. In particular, a saddle value exists if there is a *saddle point* of L, a pair (x^*, p^*) such that for all $x \in V$ and $p \in V^*$, $L(x, p^*) \le L(x^*, p^*) \le L(x^*, p)$. A pair (x^*, p^*) is a saddle point if and only if x^* solves the primal, p^* solves the dual, and a saddle value exists. The function L is the *Lagrangean*, which is familiar from the analysis of smooth constrained optimization problems. Here it receives a different foundation.

The art of duality schemes is to identify an interesting L, and here is where the Fenchel transforms come in. Interesting Lagrangeans can be generated by embedding the problem $\max f$ in a parametric class of concave maximization problems. Suppose that there is a (Euclidean) parameter space P, and a usc and concave function $F : V \times Y \to \mathbf{R}$ such that $f(x) = F(x, 0)$, and consider all the problems $\max_{x \in V} F(x, y)$. A particularly interesting object of study is the *value function* $\phi(y) = \sup_x F(x, y)$, which is the indirect utility function in consumer theory, and the cost function in the theory of the firm (with concave replaced by convex and max by min). The map $y \mapsto -F(x, y)$ is closed and convex for each x, so define on $V \times V^*$

$$L(x, p) = \sup_y p \cdot y + F(x, y),$$

its (convex) Fenchel transform. The map $p \mapsto L(x, p)$ is closed and convex on V^*. Transform again to see that $F(x, y) = \inf_{p \in V^*} L(x, p) - p \cdot y$. In particular, $f(x) = \inf_p L(x, p)$.

An example of this scheme is provided by the usual concave optimization problem given by a concave objective function f, K concave constraints $g_k(x) \geq 0$, and an implicit constraint $x \in C$: $\max_x f(x)$ subject to the constraints $g_k(x) \geq 0$ for all k and $x \in C$. Introduce parameters y so that $g_k(x) \geq y_k$, and define $F(x, y)$ to be $f(x)$ if all the constraints are satisfied and $-\infty$ otherwise. The supremum defining the Lagrangean cannot be realized for y such that x is infeasible, and so

$$L(x, p) = \begin{cases} f(x) + \sum_k p_k g_k(x) & \text{if } x \in C \text{ and } p \in \mathbf{R}_+^K, \\ +\infty & \text{if } x \in C \text{ and } p \notin \mathbf{R}_+^K, \\ -\infty & \text{if } x \notin C. \end{cases} \tag{1}$$

If there are no feasible x, then $F(x, y)$ is everywhere $-\infty$, and so $L(x, p) \equiv -\infty$.

Here, in summary, are the properties of the Lagrangean for the problems discussed here:

Lagrangean Theorem: If $F(x, y)$ is lsc and concave then (1) the Lagrangean L is lsc and convex in p for each $x \in V$, (2) L is concave in x for each $p \in V^*$, and (3) $f(x) = \inf_p L(x, p)$.

Following the original scheme, the objective for the dual problem is $g(p) = \sup_x L(x, p)$, and the *dual problem* (problem D) is to maximize g on V^*. Perhaps the central fact of this dual scheme is the relationship between the dual objective function g and the value function ϕ. The function ϕ is easily seen to be concave, and simply by writing out the definitions, one sees that $g(p) = \sup_y \cdot (p \cdot y + \phi(y))$, the convex Fenchel transform of the convex function $-\phi$. So $g(p)$ is lsc and convex, $g(p) = (-\phi)^*(p)$ and whenever ϕ is usc, $\inf_p g(p) = \phi(0)$.

To make the duality scheme complete, the min problem should be embedded in a parametric class of problems in a complementary way. Take $G(p,q) = \sup_{x \in V} L(x,p) - q \cdot x$, so that $g(p) = G(p,0)$. With this definition, $-G(p,q) = \inf_{x \in V} q \cdot x - L(x,p)$, the concave Fenchel transform of $x \mapsto L(x,p)$. The value function for the parametric class of minimization problems is $\gamma(q) = \inf_p G(p,q)$. The relationship between F and G is computed by combining the definitions:

$$G(p,q) = \sup_{x,y} F(x,y) - q \cdot x + p \cdot y = -\inf_{x,y} q \cdot x - p \cdot y - F(x,y) = -F^*(q,-y)$$

and so

$$F(x,y) = \inf_{p,q} G(p,q) + q \cdot x - p \cdot y$$

where the F^* is the concave Fenchel transform of the map $(x,y) \mapsto F(x,y)$. Computing from the definitions, $f(x) = \inf_{p,q} q \cdot x + G(p,q) = \inf_q q \cdot x + \gamma(q)$, so $f = (-\gamma)^*$, and whenever γ is lsc, $\sup_x f(x) = \gamma(0)$.

In summary, if $F(x,y)$ is concave in its arguments, and usc, then we have constructed a Lagrangean and a dual problem of minimizing a concave and lsc $G(p,q)$ over p. If the value functions $\phi(y)$ and $\gamma(q)$ are usc and lsc, respectively, then $\sup_x F(x,0) = \gamma(0)$ and $\inf_p G(p,0) = \phi(0)$, so a saddle value exists. Upper and lower semicontinuity of the value functions can be an issue. The hypograph of ϕ is the set of all pairs (y,a) such that $\sup_x F(x,y) \geq a$, and this is the projection onto y and a of the set of all triples (x,y,a) such that $F(x,y) \geq a$, that is, $\operatorname{hypo} F$. Unfortunately, even if $\operatorname{hypo} F$ is closed, its projection may not be, so upper semi-continuity of ϕ does not follow from the upper semi-continuity of F.

In the constrained optimization problem with Lagrangean (1), the parametric class of dual minimization problems is to minimize $G(p,q) = \sup_{x \in C} f(x) + \sum_k p_k g_k(x) - q \cdot x$ if $y \in \mathbf{R}_+^K$ and $+\infty$ otherwise. Specialize this still further by considering linear programming. The canonical linear program is to $\max a \cdot x$ subject to the explicit constraints $b_k \cdot x \leq c_k$ and the implicit constraint $x \geq 0$. Rewrite the constraints as $-b_k \cdot x + c_k \geq 0$ to be consistent with the formulation of (1). Then

$$G(p,q) = \sup_{x \geq 0} a \cdot x - \sum_k p_k(b_k \cdot x - c_k) - q \cdot x = \sum_k c_k p_k + \sup_{x \geq 0}\left(a - \sum_k p_k b_k - q\right) \cdot x$$

for $p \in \mathbf{R}_+^K$ and $+\infty$ otherwise. The dual problem is to minimize this over p. The sup term in G will be $+\infty$ unless the vector in parentheses is non-positive, in which case the sup will be 0. So the dual problem, taking $q = 0$, is to minimize $\sum_k c_k p_k$ over p subject to the constraints that $\sum_k p_k b_k \geq a$ and $p \in \mathbf{R}_+^K$. If the primal constraints are infeasible, $\phi(0) = -\infty$. If the dual is infeasible, $\gamma(0) = +\infty$,

and this serves as an example of how the dual scheme can fail over lack of continuity. For linear programs there is no problem with the hypographs of ϕ and γ, because these are *polyhedral convex sets*, the intersection of a finite number of closed half-spaces, and projections of closed polyhedral convex sets are closed.

6.2 Solutions

Subdifferentials act like partial derivatives, particularly with respect to identifying maxima and minima: x^* in V solves problem P if and only if $0 \in \partial f(x^*)$. When f is identically $-\infty$, there are no solutions which satisfy the constraints. Thus $\text{dom}\, f$ is the set of *feasible solutions* to the primal problem P. Similarly, $p^* \in V^*$ solves the dual problem D if and only if $\partial g(p^*) = 0$, and here $\text{dom}\, g$ is the set of *dual-feasible solutions*. Saddle points of the Lagrangean also have a subdifferential characterization. Adapting the obvious partial differential notion and notation, (x^*, p^*) is a saddle point for L if and only if $0 \in \partial_x L(x^*, p^*)$ and $0 \in \partial_p L(x^*, p^*)$ (these are different 0's since they live in different spaces), which we write $(0,0) \in \partial L(x^*, p^*)$. This condition is often called the *Kuhn-Tucker* condition. The discussion so far can be summarized in the following theorem, which is less general than can be found in the sources:

Kuhn-Tucker Theorem: Suppose that $F(x, y)$ is concave and usc. Then the following are equivalent:

1. $\sup f = \inf g$,

2. ϕ is usc and concave,

3. the saddle value of the Lagrangean L exists,

4. γ is lsc and convex.

In addition, the following are equivalent:

5. x^* solves P, p^* solves D, and the saddle value of the Lagrangean exists.

6. (x^*, p^*) satisfy the Kuhn-Tucker condition.

For economists, the most interesting feature of the dual is that it often describes how the value of the primal problem will vary with parameters. This follows from properties of the subdifferential and the key relation between the primal value function and the dual objective function, $g = (-\phi)^*$: $-\partial\phi(0) = \partial(-\phi)(0)$, and this equals the set $\{p : (-\phi)^*(p) = p \cdot 0 - (-\phi)(0)\}$, and this is precisely the set $\{p : g(p) = \sup_x f(x)\}$. In words, if p is a solution to the dual problem D, then $-p$ is in the subgradient of the primal value function. When $\partial\phi(0)$ is a singleton, there is a unique

solution to the dual, and it is the derivative of the value function with respect to the parameters. More generally, from the subdifferential of a convex function one can construct directional derivatives for particular changes in parameter values. Similarly, $-\partial\gamma(0) = \{x : f(x) = \inf_p g(p)\}$, with an identical interpretation. In summary, add to the Kuhn Tucker theorem the following equivalence:

7. $-p^* \in \partial\phi(0)$ and $-x^* \in \partial\gamma(0)$.

The remaining question is, when is any one of these conditions satisfied? A condition guaranteeing that the subdifferentials are non-empty is that $0 \in \mathrm{ri\,dom}\,\phi$, since concave functions always have subdifferentials on the relative interior of their effective domain. In the constrained optimization problem whose Lagrangean is described in equation (1), an old condition guaranteeing the existence of saddlepoints is the *Slater condition*, that there is an $x \in \mathrm{ri}\,C$ such that for all k, $g_k(x) > 0$. This condition implies that $0 \in \mathrm{ri\,dom}\,\phi$, because there is an open neighborhood around 0 such that for y in the neighborhood and for all k, $g_k(x) > y_k$. Thus $\phi(y) \geq F(x, y) > -\infty$ for all y in the neighborhood. Conditions like this are called *constraint qualifications*. In the standard calculus approach to constrained optimization, they give conditions under which derivatives sufficiently characterize the constraint set for calculus approximations to work. See Arrow, Hurwicz and Uzawa (1961).

Finally, it is worth noting that infinite dimensional constrained optimization problems, such as those arising in dynamic economic models and the study of uncertainty, can be addressed with extensions of the methods discussed here. The main difficulty is that most infinite dimensional vector spaces are not like $\mathbf{R^n}$. There is no 'natural' vector space topology, and which topology one chooses has implications for demonstrating the existence of optima. The existence of separating hyperplanes is also a difficulty in infinite dimensional spaces. These and other problems are discussed in Mas-Colell and Zame (1991). Nonetheless, much of the preceding development does go through. See Rockafellar (1974).

References

ARROW, K., HURWICZ, L. AND UZAWA, H. (1961): "Constraint qualifications in maximization problems." *Naval Logistics Research Quarterly*, 8, 175–191.

MAS-COLELL, A. AND ZAME, W. (1991): "Equilibrium theory in infinite dimensional spaces," In *Handbook of Mathematical Economics* vol. IV, W. Hildenbrand and H. Sonnenschein, eds., North-Holland, Amsterdam.

ROCKAFELLAR, R.T. (1970): *Convex Analysis*. Princeton University Press, Princeton, NJ.

Rockafellar, R.T. (1974): *Conjugate Duality and Opttimization*. CBMS Regional Conference Series No. 16, Society for Industrial and Applied Mathematics, Philadelphia PA.