Nonlinear programming
(theory and methods)

Dinh the Luc

Publicaciones del Departamento de Matemáticas
Centro de Investigación y de Estudios Avanzados del I.P.N.

Lectures on

Nonlinear programming

(Theory and methods)

by

Dinh The Luc

Contents

CHAPTER 1

CONVEX ANALYSIS

1.1 Convex Sets

Let E denote a real topological vector space. Given a set $A \subseteq E$, int A and cl A stand for the interior and the closure of A in E.

Definition 1.1.

A set $A \subseteq E$ is said to be convex if for any two points $x_1, x_2 \in A$, the line segment joining them $[x_1, x_2] = \{x \in E: x = t_1 + (1-t)x_2, 0 \leq t \leq 1\}$ belongs to A.

Example 1.2.

1) Denote E' the topological dual space of E, i.e. E' is the space of continuous linear functionals on E. Let $\xi \in E'$. Then the following sets are convex:

$$H = \{x \in E: \xi(x) = t\},$$
$$H_{\geq} = \{x \in E: \xi(x) \geq t\},$$
$$H_{>} = \{x \in E: \xi(x) > t\},$$

where t is a fixed number. H is called a hyperplane generated by ξ, while $H_{>}$ and $H_{\geq}$ are called the closed and open half spaces generated by ξ.

2) If E is a normed space, for a number $t > 0$ and a point $a \in E$, the set.

$$B(a,t) = \{ x \in E: \|x-a\| \leq t \}$$

is convex. It is called a ball with a center at a amd radius t.

Proposition 1.3.

We have the following:

i) The intersection of an arbitrary collection of convex sets is convex;

ii) The Cartesian product of convex sets is convex;

iii) The image and inverse image of a convex set under a linear maps are convex. In particular, for any two convex two convex sets $A,B \subseteq E$ and a number t, the sets tA, $A + B$ are convex.

Proof. This is immediate from definition.

Proposition 1.4.

Given a convex set $A \subseteq E$. Then

i) int A and cl A are convex;

ii) for each $y \in A$, $x \in$ int A, the set $[x,y) = \{tx + (1-t)y: 0 < t \leq 1\}$ belongs to int A;

iii) If int A is empty,
$\operatorname{cl} A = \operatorname{cl}(\operatorname{int} A)$,
$\operatorname{int}(\operatorname{cl} A) = \operatorname{int} A$.

Proof. i) and iii) follow from ii). To see ii) let U be a neighborhood of x in A.

Then $tU + (1-t)y \subseteq A$ and it is a neighborhood of $tx + (1-t)y$ whenever $0 < t \leq 1$. Hence the conclusion. ■

Definition 1.5.

Given n points $x_1, \ldots, x_n \in E$. The point $x = \sum_{i=1}^{n} t_i x_i$ with $t_i \geq 0$, $\sum_{i=1}^{n} t_i = 1$ is called a convex combination of $x_1, \ldots, x_n$.

The convex hull (resp. convex closure) of A is the intersection of all convex (resp. convex closed) sets containing A. These sets are denoted by conv A and cl conv A respectively.

Proposition 1.6

The following assertions are true:

i) the convex hull of a set A coincides with the set of all convex combinations of points of A;

ii) the closure of the convex hull of a set A coincides with its convex closure.

Proof. For the first assertion, denote by B the set of all convex combinations of points from A. Then B is convex and it contains A. Hence $\operatorname{conv} A \subseteq B$. Moreover, if C is a convex set with $A \subseteq C$, then C contains every convex combination from A. This implies that $B \subseteq C$. Thus, $B = \operatorname{conv} A$.

For the second assertion, observe that by the first part, any closed convex set which contains A contains also conv A, hence $\operatorname{cl}(\operatorname{conv} A)$.

On the other hand, in view of Proposition 1.4, $cl(\text{conv } A)$ is a closed convex set which contains A. Therefore, $cl \text{ conv } A = cl(\text{conv } A)$. ■

Definition 1.7

A subset K of E is said to be a cone if $tx \in K$, for any $x \in K$ and $t \geq 0$. A cone is called pointed if it contains no nontrivial linear subspace.

Example 1.8

1) Let $E = R^n$ (the n dimensional Euclidean space). The non-negative orthant $R^n_+ = \{x = (x^1,\ldots,x^n) \in R^n : x^i \geq 0\}$ is a convex pointed cone.

2) Let Ω be the space of sequences $\{x_n\}$ such that $x_n = 0$ for all but a finite number of choices for n. A norm in Ω can be given by

$$\|\{x_n\}\| = \max\{x_n : n = 1,2,\ldots\} .$$

The set Ω_+ consisting of zero and sequences whose last nonzero term is positive, is a convex pointed cone. This cone has the property that its interior is empty but $\text{lin } \Omega_+ = \Omega$. It is called a ubiquitous cone.

Given a set $A \subseteq E$. The cone generated by A is

$$\text{cone } A = \{ta : a \in A, t \geq 0\} .$$

Proposition 1.9

A subset $K \subseteq E$ is a convex cone if and only if

$$K + K \subseteq K$$
$$tK \subseteq K \text{ , for all } t \geq 0 .$$

Proof. The proof is straightforward.■

1.2 Separation of Convex Sets

Let us recall the Hahn-Banach theorem from Functional Analysis: if A is an open convex set and L is a linear subspace in E (a topological vector space) with the property that $A \cap L = \phi$, then there exists a continuous linear functional $\xi \in E'$ such that

$$\xi(x) > \xi(y) = 0 \text{ , for all } x \in A \text{ , } y \in L .$$

Definition 2.1

We say that a linear functional $\xi \in E'$ separates (resp., strongly separates) two sets A and B in E if

$$\xi(x) \leq \xi(y) \text{ , for all } x \in A \text{ , } y \in B$$

(resp., $\xi(x) \leq \xi(y) - \varepsilon$, for all $x \in A$, $y \in B$, some $\varepsilon > 0$) .

The following result is a geometric version of the Hahn-Banach theorem.

Theorem 2.2

Let A and B be disjoint convex sets in E. If int A or int B is nonempty, then there exists a nonzero functional $\xi \in E'$ which separates A and B.

Proof. Consider the set $C = A - B$. In view of Proposition 1.3, this set is convex. Moreover it has a nonempty interior which is also convex by Proposition 1.4. It is clear that $0 \notin \text{int } C$. Apply the Hahn-Banach theorem to obtain a nonzero functional $\xi \in E'$ such that

$$\xi(x) > \xi(0) = 0 \text{ , for all } z \in \text{int } C .$$

Since ξ is continuous and $C = A - B$, one has

$$\xi(x) \geq \xi(y) \text{ , for all } x \in A \text{ , } y \in B . \blacksquare$$

Theorem 2.3

Let A be a closed convex set, B a compact convex set in E, where E is assumed to be a separated locally convex space. If A and B are disjoint, then there exists a functional $\xi \in E'$ which strongly separates A and B.

Proof. Since $E \backslash A$ is open and contains B and since B is compact, there exists a neighborhood U of zero in E which may be considered to be convex such that $B + U \subseteq E \backslash A$. Two convex sets $B + U$ and A are disjoint and $\text{int}(B + U) \neq \phi$. By Theorem 2.2, there is $\xi \in E'$, $\xi \neq 0$ such that

$$\xi(x) \leq \xi(y+u) \text{ , for all } x \in A \text{ , } y \in B \text{ , } u \in U \text{ .}$$

Taking $\varepsilon = -\inf\{\xi(u) : u \in U\}$ which is positive, we see that

$$\xi(x) \leq \xi(y) - \varepsilon \text{ , for all } x \in A \text{ , } y \in B \text{ . } \blacksquare$$

Corollary 2.4

In a separated locally convex space a closed convex set is the intersection of all half spaces containing it. Consequently, a convex set is closed if and only if it is closed in the weak topology.

Proof. The first part is immediate from Theorem 2.3. The second part follows from the first part and from the fact that every closed half space is weakly closed. ■

1.3 Convex Functions

Let f be a function from E to the extended real line $\bar{R} = R \cup \{\pm\infty\}$. The effective domain of f is

$$\text{dom } f = \{x \in E: f(x) < +\infty\},$$

and the epigraph of f is

$$\text{epi } f = \{(x,t) \in E \times R : t \geq f(x)\} \text{ .}$$

Definition 3.1

A function f is said to be convex if epi f is a convex set in the product space $E \times R$. It is called proper if dom $f \neq \phi$ and $f(x) > -\infty$ everywhere.

Obviously, if f is proper, then it is convex if and only if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \text{ for all } x,y \in E, t \in [0,1].$$

Example 3.2

All the functions given below are convex.

1) Affine function: $f(x) = \xi(x) + t$, some $\xi \in E'$, $t \in R$.

2) The indicator function of a convex set $A \subseteq E$:

$$\delta(x|A) = \begin{cases} 0, & \text{if } x \in A \\ +\infty & \text{else.} \end{cases}$$

3) The support function of a set $\Delta \subseteq E'$:

$$s(x|\Delta) = \sup\{\xi(x): \xi \in \Delta\}.$$

4) The distance function of a nonempty convex set A in a normed space:

$$d(x|A) = \inf\{ \|x-a\| : a \in A \}.$$

5) The gauge (or Minkowski) function of a set $A \subseteq E$:

$$\rho(x|A) = \inf\{ t > 0 : x \in tA \}.$$

We turn now to the operations with convex functions.

Proposition 3.3

Let $f_1,\dots,f_m$ be proper convex functions. Then the following functions are convex:

i) the sum: $(f_1 + \dots + f_m)(x) = \sum_{i=1}^{m} f_i(x)$;

ii) the infimal convolution:

$$(f_1 \square \dots \square f_m)(x) = \inf\{ \sum_{i=1}^{m} f_i(x_i) : \sum_{i=1}^{m} x_i = x \} .$$

Proof. This is straightforward from the definition. ■

Proposition 3.4

Let $\{f_\alpha : \alpha \in I\}$ be a collection of convex functions. Then the following functions are convex:

i) the pointwise supremum: $f(x) = \sup\{f_\alpha(x) : \alpha \in I\}$;

ii) the convex hull:

$$(\operatorname{conv}_{\alpha \in I} f_\alpha)(x) = \inf\{ \sum_{\alpha \in I} t_\alpha f_\alpha(x_\alpha) : t_\alpha \geq 0 , \sum t_\alpha = 1$$

and $\sum t_\alpha x_\alpha = x$, where only finite number of t_α are nonzero} .

Proof. This is straightforward from the definition.■

Proposition 3.5

Let L be a linear map from a topological vector space E_1 to another topological vector space E_2. Let g be a convex function on E_1 and h a convex function on E_2. Then the following functions are convex:

i) the image of g under L :

$$(Lg)(y) = \inf \{g(x) : x \in E_1 , L(x) = y\} ;$$

ii) the inverse image of h under L :

$$(hL)(x) = h(Lx) .$$

Proof. The proof is straightforward.■

Given a function f on E. The closure of f is the function $\operatorname{cl} f$ whose epigraph is the closure of the epigraph of f, i.e.

$$\operatorname{epi}(\operatorname{cl} f) = \operatorname{cl}(\operatorname{epi} f) .$$

The convex closure of f is the function $\operatorname{cl} \operatorname{conv} f$ with

$$\operatorname{epi}(\operatorname{cl} \operatorname{conv} f) = \operatorname{cl} \operatorname{conv}(\operatorname{epi} f) .$$

We recall that f is lower semicontinuous (resp., upper semicontinuous) at $x \in E$ if

$$f(x) = \liminf_{y \to x} f(y)$$

(resp. $f(x) = \limsup_{y \to x} f(y)$) .

A convex function f is said to be closed if $\operatorname{cl} f = f$. It is clear that a proper convex function is closed if and only if it is lower semicontinuous. Next, we study continuity properties of convex functions.

Theorem 3.6

Let f be a proper convex function on E . The following statements are equivalent

i) f is bounded above on a neighborhood of some point x_0 ;

ii) f is continuous at some point x_0 ;

iii) $\operatorname{int}(\operatorname{epi} f) \neq \phi$;

iv) $\operatorname{int}(\operatorname{dom} f) \neq \phi$ and f is continuous on $\operatorname{int}(\operatorname{dom} f)$.

Proof. The schema of the proof is:

$$\text{i)} \iff \text{ii)} \Longleftarrow \text{iv)} \Longleftarrow \text{iii)} \Longleftarrow \text{i)} .$$

The implication i) $\Longleftarrow$ ii) is obvious. For i) $\Longrightarrow$ ii) let $f(x) \leq c < \infty$

for $x \in U$, a neighborhood of x_0 where c is a positive fixed number. Without loss of generality we may assume that $x_0 = 0$ and $f(x_0) = 0$. Choose a positive ε with $\varepsilon < c$ and set

$$V_\varepsilon = \frac{\varepsilon}{c} U \cap (-\frac{\varepsilon}{c})U .$$

Then V_ε is a neighborhood of zero. We wish to prove that

$$|f(x)| < \varepsilon \quad , \text{ for all } \quad x \in V_\varepsilon , \tag{3.1}$$

which shows ii). For this purpose, let $x \in V_\varepsilon$. Then $\frac{c}{\varepsilon} x \in U$ and by convexity of f one has

$$f(x) < \frac{\varepsilon}{c} f(\frac{c}{\varepsilon} x) + (1-\frac{\varepsilon}{c})f(0) \leq \varepsilon \quad . \tag{3.2}$$

On the other hand, $-\frac{c}{\varepsilon} x \in U$. Hence using the expression

$$0 = \frac{1}{1+\varepsilon/c} x + \frac{\varepsilon/c}{1+\varepsilon/c} (-\frac{c}{\varepsilon} x) ,$$

One has

$$0 = f(0) \leq \frac{1}{1+\varepsilon/c} f(x) + \frac{\varepsilon/c}{1+\varepsilon/c} f(-\frac{c}{\varepsilon} x) ,$$

which implies $-\varepsilon \leq f(x)$. This and (3.2) prove (3.1).

The implication iv) ⟹ ii) is trivial.

For the implication i) ⟹ iii) note that if $c \geq f(x)$ for all $x \in U$, then

$$\{ (x,\alpha) \in E \times R : x \in U , \alpha > c \} \subseteq \text{epi } f .$$

The set in the right hand side is open, hence $\text{int}(\text{epi } f) \neq \phi$. The last task is to show iii) ⟹ iv). Let $(x,\alpha) \in \text{int}(\text{epi } f)$. Then f is bounded above on a neighborhood of x. By the equivalence between i) and ii) we conclude that f is continuous at this point. Moreover, observe that

$$\text{int}(\text{dom } f) = \{x \in E : \text{ there is } \alpha \in R \text{ with } (x,\alpha) \in \text{int}(\text{epi } f)\} .$$

Hence iv) follows. ∎

Definition 3.7

Assume that E is a normed space. We say that f satisfies a Lipschitz condition on $X \subseteq E$ if there is a nonnegative number ℓ such that

$$|f(x) - f(x')| \leq \ell \, \|x-x'\| , \text{ for all } x, x' \in X .$$

We say that f is Lipschitz near $x \in E$ if for some $\varepsilon > 0$, it satisfies a Lipschitz condition on the ball $B(x;\varepsilon)$.

Trivially, if f is Lipschitz near x, then it is continuous on a small neighborhood of x. The converse is of course not true.

Proposition 3.8

Let f be a convex function which is bounded above on a neighborhood of a point x. Then f is Lipschitz near x.

Proof. In view of Theorem 3.6, f is continuous in a neighborhood U of x. Hence for a small positive ε, there is a constant ℓ_0 such that

$$|f(x)| \leq \ell_0 \text{ , for all } x \in B(x,2\varepsilon) .$$

For every $y,y' \in B(x,\varepsilon)$ with $y \neq y'$, we set

$$z = y + (y - y') \frac{\varepsilon}{\|y - y'\|} .$$

Then

$$y = ty' + (1-t)z \text{ , with } t = \frac{\varepsilon}{\|y - y'\| + \varepsilon} .$$

Since f is convex, one gets

$$f(y) \leq t\, f(y') + (1-t)f(z) ,$$

which implies

$$f(y) - f(y') \leq (1-t)(f(z) - f(y')) \leq \frac{\|y-y'\|}{\varepsilon} |f(z) - f(y')| \leq \frac{2\ell_0}{\varepsilon} \|y-y'\| .$$

Interchange the roles of y and y' to complete the proof. ■

1.4 Conjugate Functions

Definition 4.1

Let f be a function on E. The conjugate function of f is the function on E' which is defined by the rule:

$$f^*(\xi) = \sup_{x \in E} \{\xi(x) - f(x)\} , \quad \xi \in E' .$$

Example 4.2

1) Let f be an affine function, i.e.

$$f(x) = \xi_0(x) + t , \text{ some } \xi_0 \in E' , t \in R .$$

Then

$$f^*(\xi) = \begin{cases} -t , \text{ if } \xi = \xi_0 \\ \infty , \text{ otherwise.} \end{cases}$$

2) For the indicator function of the set A,

$$f(x) = \delta(x|A) ,$$

one has

$$f^*(\xi) = \sigma(\xi|A) ,$$

which is the support function of the set A.

3) For the gauge function

$$f(x) = \rho(x|A) ,$$

the conjugate function is

$$f^*(\xi) = \delta(\xi|A^\circ) ,$$

where $A^\circ = \{\xi \in E' : \sigma(\xi|A) \leq 1\}$, the polar of A.

Proposition 4.3

For a given function f, one has

i) $f \geq f^{**}$

ii) f^* is convex and weak*-closed

iii) f^* is proper if f is proper closed convex.

Proof. For the first assertion, by definition one has

$$f^*(\xi) \geq \xi(x) - f(x) .$$

Hence $F(x) \geq \sup_{\xi \in E'} \{\xi(x) - f^*(\xi)\} = f^{**}(x)$.

As to the second assertion, for every fixed $x \in E$, the function $\xi(x) - f(x)$ is affine on E', which is continuous in the weak*-topology (i.e. the topology on E' in which for every fixed $x \in E$, the map $\xi \to \xi(x)$ is continuous). The epigraph of f^* is then the intersection of a family of convex weak*-closed sets. Hence f^* is convex weak*-closed.

For the last assertion, observe first that if $x_0 \in \operatorname{dom} f$, then $f^*(\xi) \geq \xi(x_0) - f(x_0) > -\infty$. It remains to show that $\operatorname{dom} f^* \neq \phi$. It is clear that $(x_0, f(x_0) - 1) \notin \operatorname{epi} f$. By the separation theorem there exists a pair $(\xi, \beta) \in E' \times R$ such that

$$\xi(x_0) + \beta(f(x_0) - 1) > \sup_{(x,\alpha) \in \operatorname{epi} f} \{\xi(x) + \beta\alpha\}.$$

Obviously, $\beta \neq 0$. On the other hand, β cannot be positive because α can take its value as large as we want. Hence

$$\xi(x_0)/\beta + f(x_0) - 1 < \sup_{(x,\alpha) \in \operatorname{epi} f} \{\xi(x)/\beta + \alpha\} ,$$

or equivalently

$$\infty > 1 - f(x_0) - \xi(x_0)/\beta > \sup_{x \in \operatorname{dom} f} \{-\xi(x)/\beta - f(x)\} .$$

In particular, $\xi/\beta \in \operatorname{dom} f^*$. ■

Theorem 4.4 (Fenchel-Moreau)

Suppose that $f(x) > -\infty$ for every x . Then $f = f^{**}$ if and only if f is convex and closed.

Proof. If $f = f^{**}$, then in view of Proposition 4.3, f is convex closed. Furthermore, if $f \equiv +\infty$, then $f = f^{**}$. In this way by Proposition 4.3, to complete the proof it suffices to show that $f \leq f^{**}$ for f being proper closed convex.

Suppose to the contrary that there exists a point $x_0 \in \operatorname{dom} f^{**}$ such that

$$f^{**}(x_0) < f(x_0) .$$

Then we may strictly separate the point $(x_0, f^{**}(x_0))$ and epi f . There is a pair $(\xi,\beta) \in E' \times R$ with

$$\xi(x_0) + \beta f^{**}(x_0) > \sup_{(x,\alpha) \in \operatorname{epi} f} \{\xi(x) + \beta\, \alpha\}. \qquad (4.1)$$

We wish to show $\beta < 0$. It is evident that $\beta \leq 0$. If $\beta = 0$, (4.1) implies

$$\xi(x_0) > \sup_{x \in \operatorname{dom} f} \xi(x) . \tag{4.2}$$

We know from Proposition 4.3 that $\operatorname{dom} f^* \neq \phi$.

Pick any point ξ_0 from $\operatorname{dom} f^*$. Then for every $t \in R$,

$$f^*(\xi_0 + t\xi) = \sup_{x \in \operatorname{dom} f} \{(\xi_0 + t\xi)(x) - f(x)\}$$

$$\leq f^*(\xi_0) + t \sup_{x \in \operatorname{dom} f} \xi(x) .$$

This and (4.2) show that

$$f^{**}(x_0) \geq (\xi_0 + t\xi)(x_0) - f^*(\xi_0 + t\xi) \geq$$

$$\geq \xi_0(x_0) - f^*(\xi_0) + t[\xi(x_0) - \sup_{x \in \operatorname{dom} f} \xi(x)] ,$$

which must be $+\infty$ because t can run to ∞. This contradicts the fact that $x_0 \in \operatorname{dom} f^{**}$ and indeed $\beta < 0$. Dividing (4.1) by $|\beta|$ and setting $\xi_* = \xi/|\beta|$ one has

$$\xi_*(x_0) - f^{**}(x_0) > \sup_{(x,\alpha) \in \operatorname{epi} f} \{\xi_*(x) - \alpha\} = f^*(\xi_*) ,$$

which contradicts the inequality obtained from the definition of f^{**}. ∎

Theorem 4.5

Let f_1, f_2 be two functions on E. Then

i) $(f_1 \square f_2)^* = f_1^* + f_2^*$

ii) $(f_1 + f_2)^* \leq f_1^* \square f_2^*$,

the equality occurs if f_1 and f_2 are proper convex and if there is at least a point where one of them is continuous and the other is finite;

iii) $(\mathrm{conv}(f_1, f_2))^* = \sup(f_1^*, f_2^*)$

iv) $(\sup(f_1, f_2))^* \leq \mathrm{conv}(f_1^*, f_2^*)$,

the equality occurs if f_1 and f_2 are convex finite on E and if at least one of them is continuous.

Proof. We prove the first two assertions. Other parts can be done in a similar way.

The fist formula is calculated directly by definition:

$$(f_1 \square f_2)^*(\xi) = \sup_x\{\xi(x) - \inf_z(f_1(x-z) + f_2(z))\}$$

$$= \sup_{z,y}\{\xi(y) - f_1(y) + \xi(z) - f_2(z)\}$$

$$= f_1^*(\xi) + f_2^*(\xi) .$$

For ii), by definition one has

$$f_1^*(\xi_1) + f_2^*(\xi_2) \geq (\xi_1 + \xi_2)(x) - f_1(x) - f_2(x) ,$$

for all $\xi_1 , \xi_2 \in E'$, $x \in E$. Therefore,

$$f_1^*(\xi_1) + f_2^*(\xi_2) \geq (f_1 + f_2)^*(\xi_1 + \xi_2) .$$

In particular when $\xi_1 + \xi_2 = \xi$, one has

$$f_1^* \square f_2^* \geq (f_1 + f_2)^* .$$

This implies also that in the case $\operatorname{dom}(f_1 + f_2)^* = \phi$, equality holds trivially. So we may proceed to the nontrivial case: there is $\xi \in E'$ with

$$(f_1 + f_2)^*(\xi) = \alpha_0 < \infty ,$$

and assume that f_1 is continuous at a point of $\operatorname{dom} f_2$. Then $\operatorname{dom}(f_1 + f_2) \neq \phi$. Hence $(f_1 + f_2)^* > -\infty$, in particular $\alpha_0 > -\infty$.

Let us consider the set

$$A = \{(x,\alpha) \in E \times R: \quad \alpha \leq \xi(x) - f_2(x) - \alpha_0 \} .$$

It is clear that A is convex and

$$A \cap \operatorname{int}(\operatorname{epi} f_1) = \phi .$$

In fact, if the intersection is nonempty, say it contains a point (x,α) ,

then

$$f_1(x) < \alpha \leq \xi(x) - f_2(x) - \alpha_0 .$$

and we arrive at the contradiction:

$$\alpha_0 < \xi(x) - f_1(x) - f_2(x) \leq (f_1 + f_2)^*(\xi) = \alpha_0 .$$

Apply theorem 2.2. to separate A and epi f_1. There is a pair $(\xi_0, \beta) \in E' \times R$ such that

$$\sup_{(x,\alpha) \in \text{epi } f_1} \{\xi_0(x) + \beta\alpha\} \leq \inf_{(x,\alpha) \in A} \{\xi_0(x) + \beta\alpha\} . \qquad (4.3)$$

It is evident that $\beta \leq 0$. Actually $\beta < 0$ because if $\beta = 0$, (4.3) shows that dom f_1 and dom f_2 are separated, and f_1 cannot be continuous at a point of dom f_2. Dividing (4.3) by $|\beta|$ and set $\xi_1 = \xi_0 / |\beta|$, we obtain

$$f_1^*(\xi_1) = \sup_{x \in E}\{\xi_1(x) - f_1(x)\} = \sup_{(x,\alpha) \in \text{epi } f_1} \{\xi_1(x) - \alpha\}$$

$$\leq \inf_{(x,\alpha) \in A} \{\xi_1(x) - \alpha\} = \inf_{x \in \text{dom } f_2} \{(\xi_1 - \xi)(x) + f_2(x)\} + \alpha_0$$

$$= - f_2^*(\xi - \xi_1) + \alpha_0 .$$

The equality then follows. ■

Theorem 4.6

Let L be a continuous linear map from E_1 to E_2, let g and f be functions on E_1 and E_2 respectively. Then

i) $(Lg)^* = g^* L^*$

ii) $(fL)^* \leq L^* f^*$,

the equality occurs if f is convex, continuous at a point of the image of L.

Proof. For i), using the definition one has

$$(Lg)^*(\xi) = \sup_y\{\xi(y) - \inf_{x:L(x)=y} g(x)\}$$

$$= \sup_{x:L(x)=y} \{\xi(y) - g(x)\}$$

$$= \sup_x\{L^*(\xi)(x) - g(x)\}$$

$$= g^* L^*(\xi) .$$

Similarly for ii):

$$L^* f^*(\xi) = \inf_{\eta:L^*(\eta)=\xi} f^*(\eta)$$

$$= \inf_{\eta:L^*(\eta)=\xi} \sup_y\{\eta(y) - f(y)\}$$

$$\geq \inf_{\eta:L^*(\eta)=\xi} \sup_x\{\eta(Lx) - f(Lx)\}$$

$$= \sup_x\{\xi(x) - (fL)(x)\}$$

$$= (fL)^*(\xi) .$$

It remains to prove the inequality

$$(L^*f^*)(\xi) \leq (fL)^*(\xi)$$

under the assumptions that f is convex continuous at a point of the image of L and that $\xi \in \mathrm{dom}(fL)^*$.

For this purpose, set

$$\alpha_0 = (fL)^*(\xi) .$$

Since $\mathrm{dom}\, f \cap \mathrm{Im}\, L \neq \phi$, the function fL has finite values. In particular α_0 is a finite number. Consider the set

$$M = \{(y,\alpha) \in E_2 \times R : \text{there is } x \in E_1 \text{ with } y = Lx \text{ and } \alpha = \xi(x) - \alpha_0\}.$$

It is a convex set which does not meet int epi f.
In fact if it intersects int epi f at a point x, then

$$f(Lx) < \xi(x) - \alpha_0 .$$

which gives a contradiction:

$$\alpha_0 < \xi(x) - f(Lx) \leq (fL)^*(\xi) = \alpha_0 .$$

One can separate M and epi f by a nonzero vector $(\eta,\beta) \in E_2' \times R$:

$$\sup_{(y,\alpha)\in \text{epi } f} \{\eta(y) + \beta\alpha\} \leq \inf_{(y,\alpha)\in M} \{\eta(y) + \alpha\} . \qquad (4.4)$$

It is obvious that $\beta \leq 0$. If $\beta = 0$, then $\eta \neq 0$ and it separates dom f and Im L . This is impossible. Thus, $\beta < 0$. Dividing (4.4) by $|\beta|$ and setting $\eta_0 = \eta / |\beta|$ we obtain

$$f^*(\eta_0) \leq \inf_{(y,\alpha)\in M} \{\eta_0(y) - \alpha\}$$

$$= \inf_{x\in E_1} \{\eta_0(Lx) - \xi(x) + \alpha_0\} .$$

This implies that $\xi = L^*(\eta_0)$ and one has

$$L^*f^*(\xi) \leq f^*(\eta_0) \leq \alpha_0 = (fL)^*(\xi) .$$

The proof is complete. ■

1.5 Subdifferentials

Let f be a convex function on E . We recall that directional derivative of f in direction $v \in E$ at $x \in E$ is the following limit if such exists

$$f'(x;v) = \lim_{t\downarrow 0} \frac{f(x+tv) - f(x)}{t} .$$

Proposition 5.1

Let f be a proper convex function on E. Then f has a directional derivative at any point of dom f and

$$f'(x;v) = \inf_{\lambda>0} \frac{f(x+\lambda v) - f(x)}{\lambda} .$$

Moreover, $f'(x;v)$ is convex homogeneous in v.

Proof. For every fixed $x \in \text{dom } f$, $v \in E$ we consider the function

$$\varphi(t) = f(x + tv) .$$

It is a proper convex function on R with $0 \in \text{dom } \varphi$. For every three numbers $t_1 < t_2 < t_3$ one has

$$\varphi(t_2) \leq \frac{t_3 - t_2}{t_3 - t_1} \varphi(t_1) + \frac{t_2 - t_1}{t_3 - t_1} \varphi(t_2) ,$$

which implies

$$\varphi \frac{(t_2) - \varphi(t_1)}{t_2 - t_1} \leq \frac{\varphi(t_3) - \varphi(t_1)}{t_3 - t_1} \leq \frac{\varphi(t_3) - \varphi(t_2)}{t_3 - t_2} .$$

In other words, $\frac{\varphi(t+\lambda) - \varphi(t)}{\lambda}$ does not increase as λ decreases towards 0. Therefore

$$\varphi'_+(t) = \varphi'(t;1) = \lim_{\lambda \downarrow 0} \frac{\varphi(t+\lambda) - \varphi(t)}{\lambda}$$

exists for any $t \in \operatorname{dom} \varphi$. In particular, $\varphi'_+(0) = f'(x;v)$ exists and equals $\inf_{\lambda>0} \frac{f(x+\lambda v) - f(x)}{\lambda}$.

The second part of the proposition is straighforward. ∎

Proposition 5.2

Let f be a proper convex function on E which is continuous at every point of a set $U \subseteq E$. If for some $v \in E$ with $x + v \in U$ the derivative $f'(x;v)$ is finite, then $f'(x;\cdot)$ is continuous on $U - x$. Moreover, if f is continuous at x, then $f'(x;\cdot)$ is finite, continuous on E...

Proof. We show that $f'(x;\cdot)$ is a proper function. Observe first that $|f'(x;v)| < \infty$, hence $x \in \operatorname{dom} f$. By Proposition 5.1,

$$f'(x;u) \leq f(x+u) - f(x), \text{ for all } u \in E.$$

If for some u_1, $f'(x;u_1) = -\infty$, then since $x + v \in \operatorname{int} \operatorname{dom} f$, for ε small enough one has

$$x + (v + \varepsilon(v-u_1)) \in \operatorname{dom} f.$$

Denote $v_1 = v + \varepsilon(v-u_1)$. By the convexity of f, for every $\lambda > 0$,

$$f(x+\lambda v) \leq \frac{1}{1+\varepsilon} f(x+\lambda v_1) + \frac{\varepsilon}{1+\varepsilon} f(x+\lambda u_1) .$$

We arrive at the following contradiction:

$$f'(x;v) \leq \frac{1}{1+\varepsilon} f'(x;v_1) + \frac{\varepsilon}{1+\varepsilon} f'(x;u_1) = -\infty .$$

Now, let $u \in U - x$. Then f is bounded by a constant c on a small neighborhood V of $x + u$. Hence for every $y \in V-x$, one has

$$f'(x;y) \leq f(x+y) - f(x) \leq c - f(x) .$$

Thus, $f'(x;\cdot)$ is finite and bounded on $V-x$, which implies its continuity at u.

The second part of the proposition is obvious. ■

Definition 5.3.

The subdifferential of f at x is the set

$$\partial f(x) = \{\xi \in E' : f(y) - f(x) \geq \xi(y-x), \text{ for all } y \in E\} .$$

If $\partial f(x) \neq \phi$, we say that f is subdifferentiable at x.

Example 5.4

1) Let f be the norm function: $f(x) = \|x\|$. Then

$$\partial f(x) = \begin{cases} \{\xi \in E' : \|\xi\| \leq 1\}, & \text{if } x = 0 \\ \{\xi \in E' : \|\xi\| = 1 \ , \ \xi(x) = \|x\| \}, & \text{otherwise .} \end{cases}$$

2) Let f be the indicator function of a set A: $f(x) = \delta(x|A)$. Then

$$\partial\delta(x|A) = \{\xi \in E' : \xi(z-x) \leq 0 \ , \text{ for all } z \in A\} \ .$$

It is a cone called the normal cone to A at x and is denoted by $N(x|A)$. In particular, if $A = L$ is a subspace, then for each $x \in L$, $N(x|L) = L^{\perp}$. If A is a cone, $N(0|A) = A^{\circ}$ – the polar cone.

Proposition 5.5

For a convex function f , the following conditions are equivalent:

i) $\xi \in \partial f(x)$

ii) $f(x) + f^*(\xi) = \xi(x)$

iii) $f'(x;v) \geq \xi(v)$, for all $v \in E$.

Proof. For the implication i) $\Longrightarrow$ ii) , let $\xi \in \partial f(x)$.

Then $\xi(x) - f(x) \geq \xi(y) - f(y)$, for all $y \in E$.

Hence $\xi(x) - f(x) = \sup_y\{\xi(y) - f(y)\} = f^*(\xi)$.

For the implication ii) ==>iii) observe that

$$f(x+\lambda v) \geq \xi(x+\lambda v) - f^*(\xi),$$

which implies

$$\frac{f(x+\lambda v) - f(x)}{\lambda} \geq \frac{\xi(\lambda v)}{\lambda} = \xi(v) \text{ , for all } \lambda > 0 .$$

Hence $f'(x;v) \geq \xi(v)$, for all $v \in E$.

As to the last implication iii) ==> i) using Proposition 5.1 , one has

$$f(x+ (y-x)) - f(x) \geq f'(x;y-x) .$$

This and iii) show that

$$f(y) - f(x) \geq \xi(y-x) \text{ , for all } y \in E .$$

Hence $\xi \in \partial f(x)$. ■

Corollary 5.6.

If f is convex, then $\partial f(x) = \mathrm{dom}[f'(x;\cdot)]^*$.

Proof. Let $\xi \in \partial f(x)$. By Proposition 5.5,

$$f'(x;v) \geq \xi(v) \text{ , for each } v \in E \text{ .}$$

Hence $[f'(x;\cdot)]^*(\xi) \leq 0$, and $\xi \in \text{dom } [f'(x;\cdot)]^*$.

Conversely, let $\xi \in \text{dom } [f'(x;\cdot)]^*$. By definition,

$$[f'(x;\cdot)]^*(\xi) = \sup_v \{\xi(v) - f'(x;v)\} < \infty \text{ .}$$

Since $f'(x;v)$ is positively homogeneous, the above inequality is possible only in the case

$$\xi(v) - f'(x;v) \leq 0 \text{ , for every } v \in E \text{ .}$$

In view of Proposition 5.5 , $\xi \in \partial f(x)$. ■

Proposition 5.7

Let f be a proper convex function. Then

i) f is subdifferentiable at $x \in \text{dom } f$ if and only if its directional derivative at this point is lower semicontinuous at 0;

ii) if f is continuous at x_0 , then $\partial f(x_0)$ is a nonempty weak*-bounded set.

Proof. For the first assertion suppose that f is subdifferentiable at $x \in \operatorname{dom} f$. By Proposition 5.5,

$$f'(x;v) \geq \xi(v), \text{ for all } v \in E.$$

This implies

$$\liminf_{v \to 0} f'(x;v) = 0,$$

and $f'(x;\cdot)$ is lower semicontinuous at 0.

Conversely, if $f'(x;\cdot)$ is lower semicontinuous at 0, then $\operatorname{cl} f'(x;0) = 0$. But $f'(x;\cdot)$ is homogeneous, one has

$$f'(x;v) > -\infty, \text{ for all } v \in E.$$

In view of Proposition 4.3, $[f'(x;\cdot)]^*$ is a proper function. Consequently, by Corollary 5.6, $\partial f(x) \neq \phi$.

Finally, if f is continuous at x_0, in virtue of Proposition 5.2, $f'(x_0;\cdot)$ is continuous on E. Hence, by the first part, $\partial f(x_0) \neq \phi$. Moreover, for every $v \in E$,

$$\sup_{\xi \in \partial f(x_0)} \xi(v) = f'(x_0;v) < \infty,$$

which means that $\partial f(x_0)$ is bounded in the weak* topology of the space E'. ■

Theorem 5.8

Let f_1 , f_2 be proper convex functions on E . Then

$$\partial f_1(x) + \partial f_2(x) \subseteq \partial(f_1+f_2)(x),$$

for every $x \in E$. If at a point $x_0 \in \text{dom} f_1 \cap \text{dom } f_2$ one of these functions is continuous, then equality occurs at any point $x \in E$.

Proof. The first part is immediate from the definition. Let now f_1 be continuous at $x_0 \in \text{dom} f_2$. If $\partial(f_1+f_2)(x) = \phi$, equality holds trivially. We consider the case $\partial(f_1+f_2)(x) \neq \phi$. Let ξ be an element of this set. Then

$$x \in \text{dom } (f_1+f_2) = \text{dom } f_1 \cap \text{dom } f_2 .$$

In view of Proposition 5.2 , $f_1'(x;\cdot)$ is continuous at $x_0 - x$. On the other hand

$$f_2'(x;x_0-x) \leq f_2(x_0) - f_2(x) < \infty ,$$

which implies $x_0 - x \in \text{dom } f_2'(x;\cdot)$.

If $f_2'(x;\cdot)$ is a proper function, then by Theorem 4.5 and corollary 5.6,

$$\partial f_1(x) + \partial f_2(x) = \mathrm{dom}[f_1'(x;\cdot)]^* + \mathrm{dom}[f_2'(x;\cdot)]^*$$

$$= \mathrm{dom}[f_1+f_2)'(x;\cdot)]^*$$

$$= \partial(f_1+f_2)(x) \ ,$$

which completes the proof. In this way our final debt is to show that $f_2'(x;\cdot)$ is proper. Suppose to the contrary that there is $z \in E$ with

$$f_2'(x;z-x) = -\infty \ .$$

Then for small $\lambda > 0$, $y = x + \lambda(z-x) \in \mathrm{dom}\ f_2$ and $f_2'(x;\ y-x) = -\infty$. Set

$$x(\alpha) = \alpha x_0 + (1-\alpha)y \ , \text{ for } \alpha \in [0,1] \ .$$

Then $x(\alpha) \in \mathrm{dom}\ f_2$. We also have $x(\alpha) \in \mathrm{dom}\ f_1$ whenever α is sufficiently close to 1 because $x_0 \in \mathrm{int\ dom}\ f_1$. Hence

$$\Omega$$

$$-\infty < \xi(x(\alpha) - x)$$

$$\leq (f_1+f_2)'(x;x(\alpha) - x)$$

$$= f_1'(x;x(\alpha) - x) + f_2'(x;x(\alpha) - x)$$

$$\leq f_1'(x;x(\alpha) - x) + \alpha f_2'(x;x_0-x) + (1-\alpha)f_2'(x;y-x)$$

$$= -\infty \ .$$

This contradiction shows that $f_2'(x;\cdot)$ is proper and the theorem is proven. ■

Theorem 5.9

Let L be a continuous linear operator from E_1 to E_2. If f is a convex function on E_2, then

$$L^*\partial f(Lx) \subseteq \partial(fL)(x) ,$$

for every $x \in E$. If in addition f is continuous at a point of the set Im L (the image of L), then equality holds for all $x \in E_1$.

Proof. The inclusion of the theorem is immediate from the definition. Therefore equality holds trivially if $\partial(fL)(x)$ is empty. Now, assume that f is continuous at Lx_0 and $\partial(fL)(x) \neq \phi$.

It is obvious that

$$(fL)'(x;v) = f'(Lx;Lv) .$$

In view of Proposition 5.2, $f'(Lx;\cdot)$ is continuous at $L(x_0-x) \in$ Im L. Using Theorem 4.6 and Corollary 5.6, one has

$$\begin{aligned} \partial(fL)(x) &= \operatorname{dom}[(fL)'(x;\cdot)]^* \\ &= \operatorname{dom} L^*[f'(Lx;\cdot)]^* \\ &= L^*\partial f(Lx) , \end{aligned}$$

which completes the proof. ■

Theorem 5.10

Let T be a nonempty parameter set and

$$f(x) = \max_{t \in T} f_t(x) ,$$

$$T(x) = \{t \in T : f(x) = f_t(x)\} ,$$

and let $P(x)$ denote the collection of probability Radon measures μ with support in $T(x)$ for which

$$\int_T f_t(x)\mu(dt) > -\infty, \quad x \in M .$$

Assume that the following conditions hold:

i) T is a compact space

ii) $f_t(\cdot)$ is finite convex on a set $M \subseteq E$ with $\operatorname{int} M \neq \phi$ for each fixed $t \in T$, and the function $t \to f_t(x)$ is upper semi-continuous on T for every fixed $x \in M$.

Then for every $x_0 \in M$ one has

$$\partial f(x_0) = \bigcup_{\mu \in P(x_0)} \{\partial[\int_T f_t(\cdot)\mu(dt)](x_0)\} .$$

Proof. Denote the set in the right hand side by A. We show first that

$$A \subseteq \partial f(x_0) .$$

In fact, let $\xi \in A$, i.e. there is $\mu \in P(x_0)$ with

$$\xi \in \partial[\int_T f_t(\cdot)\mu(dt)](x_0) .$$

By definition,

$$\int_T f_t(x)\mu(dt) - \int_T f_t(x_0)\mu(dt) \geq \xi(x-x_0) , \text{ for all } x \in E .$$

But $f(x) \geq f_t(x)$ for all $t \in T$ and $f(x_0) = f_t(x_0)$ for all $t \in T(x_0)$, the above inequality becomes

$$f(x) - f(x_0) \geq \int_T f_t(x)\mu(dt) - \int_T f_t(x_0)\mu(dt) \geq \xi(x-x_0) ,$$

which shows that $\xi \in \partial f(x_0)$.

It follows that equality holds trivially if $\partial f(x_0) = \phi$. Now, let $\xi \in \partial f(x_0)$. We want to show first that

$$\xi \in \partial f^\circ(x_0) , \tag{5.1}$$

where $f^\circ(x) = \max_{t \in T(x_0)} f_t(x)$.

For this purpose, for every fixed positive ε, put

$$T_\varepsilon = \{t \in T : f_t(x_0) \geq f(x_0) - \varepsilon\} .$$

Condition ii) implies that the above set is closed. Set

$$f^\varepsilon(x) = \max_{t \in T_\varepsilon} f_t(x) .$$

We want to establish the relation

$$\xi \in \partial f^\varepsilon(x_0) , \text{ for every } \varepsilon > 0 . \tag{5.2}$$

To do this, for a fixed $x \in M$, choose $\delta_0 > 0$ so that

$$|f((1-\delta)x_0 + \delta x) - f(x_0)| \leq \varepsilon/3 , \text{ for } \delta: \quad 0 < \delta \leq \delta_0 ,$$

and set

$$x_1 = (1 - \frac{\delta_0}{2}) x_0 + \frac{\delta_0}{2} x$$

$$= \frac{1}{2} x_0 + \frac{1}{2} ((1 - \delta_0)x_0 + \delta_0 x) . \tag{5.3}$$

Then for $t \notin T_\varepsilon$ one has

$$\begin{aligned} f_t(x_1) &\le \frac{1}{2} f_t(x_0) + \frac{1}{2} f_t((1-\delta_0)x_0 + \delta_0 x) \\ &< \frac{1}{2}[f(x_0) - \varepsilon] + \frac{1}{2} f((1-\delta_0)x_0 + \delta_0 x) \\ &\le \frac{1}{2}[f(x_0) - \varepsilon] + \frac{1}{2}[f(x_0) + \frac{1}{3}\varepsilon] \\ &\le f(x_1) + (\frac{1}{3} + \frac{1}{6} - \frac{1}{2})\varepsilon \\ &= f(x_1) . \end{aligned}$$

Hence $f^\varepsilon(x_1) = f(x_1)$. Furthermore, since $f^\varepsilon(x_0) = f(x_0)$ and $\xi \in \partial f(x_0)$, one has

$$f^\varepsilon(x_0) - \xi(x_0) \le f^\varepsilon(x_1) - \xi(x_1) . \tag{5.4}$$

From (5.3) and (5.4) and from the convexity of f^ε , one has

$$f^\varepsilon(x_0) - \xi(x_0) \le f^\varepsilon(x) - \xi(x) . \tag{5.5}$$

This proves (5.2). With this we are going to establish (5.1). Observe first that by definition, $f^\varepsilon(x)$ is nondecreasing in $\varepsilon > 0$, and

$$f^\varepsilon(x) \ge f^0(x) , \text{ for all } x \in M .$$

Hence the following limit exists and

$$\lim_{\varepsilon \downarrow 0} f^{\varepsilon}(x) \geq f^{\circ}(x) , \quad x \in M . \tag{5.6}$$

We fix an $x \in M$, and for each $\varepsilon > 0$, choose $t_{\varepsilon} \in T_{\varepsilon}$ such that

$$f^{\varepsilon}(x) = f_{t_{\varepsilon}}(x) .$$

Consider the sequence $\{t_{\varepsilon_n}\}$ when ε_n tends monotonically to zero. Without loss of generality it can be assumed that $\lim t_{\varepsilon_n} = t_0 \in T$. It is obvious that

$$T(x_0) = \bigcap_{\varepsilon > 0} T_{\varepsilon} .$$

Hence $t_0 \in T(x_0)$. By the upper semicontinuity assumption,

$$\lim_{\varepsilon \downarrow 0} f^{\varepsilon}(x) = \lim_{\varepsilon \downarrow 0} f_{t_{\varepsilon}}(x) = \lim_{n} f_{t_{\varepsilon_n}}(x)$$

$$\leq f_{t_0}(x) \leq f^{\circ}(x) .$$

This and (5.6) imply

$$\lim_{\varepsilon \downarrow 0} f^{\varepsilon}(x) = f^{\circ}(x) , \text{ for all } x \in M .$$

The latter equality and (5.5) give us the relation

$$f^{\circ}(x_0) - \xi(x_0) \leq f^{\circ}(x) - \xi(x) , \quad x \in M ,$$

which establishes (5.1).

Now in the space of continuous functions on $T(x_0)$, $C(T(x_0))$, consider the set

$$N = \{ \varphi : \text{there is } x \in M \text{ with } \varphi(t) \geq f_t(x) - f_t(x_0) - \xi(x-x_0) \text{ for all } t \in T(x_0) \} .$$

It is clear that N is a convex set. Moreover it does not meet the negative cone

$$C(T(x_0))_- = \{\psi: \psi(t) < 0 , \text{ for all } t \in T(x_0)\} .$$

We separate them by a continuous linear functional ν on $C(T(x_0))$:

$$\nu(\psi) < 0 \leq \nu(\varphi) , \text{ for all } \psi \in C(T(x_0))_- , \ \varphi \in N . \qquad (5.7)$$

In particular ν is nonnegative and nonzero. Hence it is defined by a positive Radon measure on $T(x_0)$. Put

$$\mu = \nu / \nu [T(x_0)] .$$

Let $x \in M$. The function φ° from $T(x_0)$ to R defined by the rule

$$\varphi^\circ(t) = f_t(x) - f_t(x_0) - \xi(x-x_0) ,$$

is upper semicontinuous. Then it can be expressed as the lower enveloping function of the set

$$N(\varphi^\circ) = \{ \varphi \in C[T(x_0)] : \varphi(t) \geq \varphi^\circ(t) , t \in T_0 \}.$$

One has then

$$\int_{T(x_0)} \varphi^\circ(t)\mu(dt) = \inf_{\varphi \in N(\varphi^\circ)} \int_{T(x_0)} \varphi(t)\mu(dt) .$$

Since $N(\varphi^\circ) \subseteq N$, (5.7) implies

$$\int_{T(x_0)} \varphi^\circ(t)\mu(dt) \geq 0 .$$

This means that

$$\int_{T(x_0)} (f_t(x) - f_t(x_0) - \xi(x-x_0))\mu(dt) \geq 0 .$$

Since this is true for all $x \in M$, one concludes that $\mu \in P(x_0)$ and

$$\xi \in \partial[\int_T f_t \, \mu(dt)](x_0) .$$

The theorem is proven. ■

1.6 Problems

1. Let L be a continuous linear map from a topological vector space E_1 to another topological vector space E_2, and let A be a set in E_1. Is it true that

   i) $L(\text{conv } A) = \text{conv}(LA)$

   ii) $L(\text{cl conv } A) = \text{cl conv}(LA)$.

2. For a set $A \subseteq E$. Find conditions under which $\text{cl}(\text{cone } A) = \text{cone}(\text{cl } A)$.

3. Suppose that $A \subseteq R^n$ is a convex set. The relative interior $\text{ri } A$ of A is the set consisting of points $a \in A$ such that there is a positive ε with

$$B(a,\varepsilon) \cap (\text{aff } A) \subseteq A .$$

   Prove that $a \in \text{ri } A$ if and only if for each $x \in A$ there exists $t > 1$ such that

$$(1-t)x + ta \in A .$$

4. Let L be a linear transformation from R^n to R^m and A a convex set in R^n, B a convex set in R^m. Prove

i) $ri(LA) = L(ri\ A)$

$cl(LA) \supseteq L(cl\ A)$

ii) if $L^{-1}(ri\ B) \neq \phi$, then

$ri(L^{-1}B) = L^{-1}(ri\ B)$,

$cl(L^{-1}B) = L^{-1}(cl\ B)$.

5. Let E be a Banach space. Prove that for every point x of the weak closure of a set $A \subseteq X$, there exists a sequence of convex combinations of elements of A which converges to x in the norm:

6. Let $f(x) = \|x\|$ and $g(x) = \delta(x|C)$, where C is a convex set. Prove that

$$(f \square g)(x) = d(x|C),$$

where $d(x|C)$ is the distance from x to C.

7. Let $f_i(x) = \delta(x|a_i) + \alpha_i$, $i = 1,\ldots,m$, where a_i' are fixed points in E and α_i' are fixed numbers. Prove that

$$f(x) = \text{conv}\{f_1(x),\ldots,f_m(x)\} = \inf\{\Sigma \lambda_i \alpha_i : \Sigma \lambda_i a_i = x\}.$$

It is the greatest convex function satisfying

$$f(a_i) \leq \alpha_i \ , \ i = 1,\dots,m \ .$$

8. Prove that

$$(f_1 \square f_2)(x) = \inf\{\mu\colon (x,\mu) \in (\mathrm{epi}_1 + \mathrm{epi}\ f_2)\}$$

$$\mathrm{conv}\{f_1,\dots,f_m\} = \inf\{\lambda_1 f_1 \square \dots \square \lambda_m f_m\colon \lambda_i \geq 0 \ , \ \Sigma \lambda_i = 1\} \ .$$

9. In Propositions 3.3, 3.4, 3.5 the operations do not preserve the properness of functions. Give examples to show this.

10. Let $K_1,\dots,K_m$ be nonempty convex cones in R^n. Prove that

i) $(K_1+\dots+K_m)^\circ = K_1^\circ \cap \dots \cap K_m^\circ$;

ii) $(\mathrm{cl}\ K_1 \cap \dots \cap \mathrm{cl}\ K_m)^\circ = \mathrm{cl}(K_1^\circ+\dots+K_m^\circ)$

where K° denotes the polar cone of K.

11. Suppose that $E = R^n$ and $x \to f_t(x)$ is convex continuous for every fixed $t \in T$, where T is a compact space, while $t \to f_t(x)$ is upper semicontinuous on T for every fixed $x \in R^n$. Let $f(x) = \max\limits_{t \in T} f_t(x)$.

Prove that every $\xi \in \partial f(x_0)$ can be represented in the form

$$\xi = \sum_{i=1}^{r} \alpha_i \xi_i ,$$

where $r \leq n + 1$, $\sum_{i=1}^{r} \alpha_i = 1$, $\alpha_i > 0$ and

$$\xi_i \in \partial f_{t_i}(x_0) , \text{ some } t_i \in T(x_0) , i = 1,\ldots,r ,$$

where $T(x_0) = \{t \in T: f(x_0) = f_t(x_0)\}$.

12. Let $A_0, A_1,\ldots,A_n$ be convex sets in E with

$$A_0 \cap \text{int } A_1 \cap \ldots \cap \text{int } A_n \neq \phi .$$

Let $A = A_0 \cap \ldots \cap A_n$. Prove that

$$N(x/A) = N(x/A_0) + \ldots + N(x/A_n) ,$$

for every $x \in A$, where $N(x/A)$ is the normal cone to A at x .

13. Given a proper convex function f which is continuous at $x_0 \in E$. Assume that there is $x_1 \in E$ with

$$f(x_1) < f(x_0) = \alpha_0 .$$

Prove that $N(x_0 | \text{lev}_f(\alpha_0)) = \text{cone}(\partial f(x_0))$ where $\text{lev}_f(\alpha_0) = \{x: f(x) \leq \alpha_0\}$.

14. Prove that in Proposition 5.7, the set $\partial f(x_0)$ is weak*-compact.

CHAPTER 2

NONSMOOTH ANALYSIS

2.1 Classical Derivatives

Let X and Y be Banach spaces and f a map from X to Y.

The directional derivative of f at $x \in X$ in the direction $v \in X$ is defined by

$$f'(x;v) = \lim_{t \downarrow o} \frac{f(x+tv)-f(x)}{t}$$

if the limit exists.

The function f is said to be Gateaux differentiable at $x \in X$ if there is a continuous linear map $Df(x)$ from X to Y such that

$$Df(x)\,(v) = f'(x;v)\,, \text{ for all } v \in X\,.$$

The map $Df(x)$ is called the Gateaux derivative of f at x.

If $(f(x+tv) - f(x))/t$ converges to $Df(x)(v)$ uniformly with respect to v in compact sets (respectively, in bounded sets) we obtain the Hadamard derivative (resp., the Frechet derivative).

Now assume that f is Frechet differentiable on a neighborhood of a point x. Then $x \to Df(x)$ is a map from this neighborhood to $L(X,Y)$ (the space of continuous liner maps X to Y). If this map is Frechet differentiable at x, then its derivative $D^2f(x)$ is called the second derivative of f at x.

An important property of a function f whose second derivative is uniformly continuous on a neighborhood of a point x, is that

$$f(x+v) = f(x) + Df(x)\,v + \frac{1}{2} D^2f(x)(v,v) + r(v),$$

where $||r(v)||/||v||$ tends to 0 as $||v||$ tends to 0.

Definition 1.1

f is said to admit a strict derivative at x if there is a continuous linear map $D_sf(x)$ from X to Y such that

$$D_sf(x)(v) = \lim_{\substack{x' \to x \\ t \to 0}} \frac{f(x'+tv)-(f(x')}{t},$$

and the convergence is uniform for v in compact sets.

Proposition 1.2

Let f map a neighborhood of x to Y, and $L \in L(X,Y)$.

The following are equivalent

i) f is strictly differentiable at x and $D_sf(x)=L$;

ii) f is Lipschitz near x and for each $v \in X$

$$\lim_{\substack{x' \to x \\ t \downarrow 0}} \frac{f(x'+tv) - f(x')}{t} = L(v) .$$

Proof. Assume i) . The equality in ii) holds by assumption, so to prove ii) we need only show that f is Lipschitz near x . If that is not the case, there are $\{x_n\}$ and $\{x_n'\}$ converging to x such that $x_n , x_n' \in x + \frac{1}{n} B(0;1)$ and

$$\|f(x_n') - f(x_n)\| > n \|x_n' - x_n\| .$$

Let us define $t_n > 0$ and v_n via

$$x_n' = x_n + t_n v_n$$

$$\|v_n\| = 1 / \sqrt{n} .$$

Then t_n tends to 0 . Let V consist of $\{v_n\}$ and zero. Then V is compact, so that by definition of $D_s f(x)$, for any $\varepsilon > 0$ there exists n_ε such that for $n \geq n_\varepsilon$ and for $v \in V$ one has

$$\left| \frac{f(x_n+t_n v) - f(x_n)}{t_n} - D_s f(x)(v) \right| < \varepsilon .$$

But this is impossible, since for $v = v_n$

$$\left| \frac{f(x_n + t_n v) - f(x_n)}{t_n} \right| > \sqrt{n} .$$

Conversely, posit ii). Let V be any compact subset of X and ε any positive. In view of ii) there exists for each $v \in V$ a number $\delta(v) > 0$ such that

$$\left| \frac{f(x'+tv) - f(x')}{t} - L(v) \right| < \varepsilon ,$$

for all $x' \in x + \delta B(0,1)$, $t \in (0,\delta)$.
Since the function is Lipschitz, it follows from the above inequality that for a suitable redefinition of $\delta(v)$ one has

$$\left| \frac{f(x'+tv') - f(x')}{t} - L(v') \right| < 2\varepsilon ,$$

for all $x' \in x + \delta B(0;1)$, $v' \in v + \delta B(0,1)$, $t \in (0,\delta)$.
A finite number of open sets $\{ v + \delta(v)B(0;1) : v \in V \}$ will cover V , say those that correspond to $v_1, \ldots, v_n$ Set $\delta_0 = \min \delta(v_i)$. The above inequality shows that L is the strict derivative of f at x . ■

2.2 Generalized Directional Derivative

Assume that $f : E \to R$, where E is a Banach space. Definition 2.1

Let f be Lipschitz near x . The generalized (Clarke) directional derivative of f at x in direction v denoted $f^{\circ}(x;v)$ is

$$f^{\circ}(x;v) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{f(y+tv)-f(y)}{t} .$$

Proposition 2.2

If f is Lipschitz of rank ℓ near x, then

i) $f^\circ(x;\cdot)$ is a finite positively homogeneous subadditive function on E with $|f^\circ(x;v)| \leq \ell \|v\|$.

ii) $f^\circ(x;v)$ is upper semicontinuous in (x,v) and Lipschitz of rank ℓ in the variable v on E.

iii) $f^\circ(x;-v) = (-f)^\circ(x;v)$.

Proof. For the first statement, by definition one has

$$|f^\circ(x;v)| < \limsup_{\substack{y\to x \\ t\downarrow 0}} \left| \frac{f(y+tv) - f(y)}{t} \right| < \ell \|v\| .$$

The positive homogeneity property is obvious. We turn now to the subadditivity

$$\begin{aligned} f^\circ(x;v+w) &= \limsup \frac{f(y+tv+tw) - f(y)}{t} \\ &\leq \limsup \frac{f(y+tv+tw) - f(y+tw)}{t} + \limsup \frac{f(y+tw) - f(y)}{t} \\ &\leq f^\circ(x;v) + f^\circ(x;w) . \end{aligned}$$

For ii), let $\{x_n\}$, $\{v_n\}$ be arbitrary sequences converging to x and v respectively. For each n, by definition of upper limit, there exists $y_n \in E$, $t_n > 0$ such that

$$\|y_n - x_n\| t_n < \frac{1}{n} .$$

$$f^\circ(x_n; v_n) - \frac{1}{n} \leq \frac{f(y_n + t_n v_n) - f(y_n)}{t_n}$$

$$= \frac{f(y_n + t_n v) - f(y_n)}{t_n} + \frac{f(y_n + t_n v_n) - f(y_n + t_n v)}{t_n} .$$

Since

$$\left| \frac{f(y_n + t_n v_n) - f(y_n + t_n v)}{t_n} \right| \leq \ell \, \|v_n - v\| ,$$

one deduces from the relation above that

$$\limsup_n f^\circ(x_n; v_n) \leq f^\circ(x; v) ,$$

which shows that $f^\circ(x;v)$ is upper semicontinuous. To prove that it is Lipschitz, let $v , w \in E$. Then for y near x and t near 0, one has

$$f(y+tv) - f(y) \leq f(y+tw) - f(y) + \ell t \, \|v-w\| .$$

Dividing by t we obtain

$$f^\circ(x; v) \leq f^\circ(x; w) + \ell \, \|v-w\| .$$

Since this also holds with v and w switsched, ii) follows.

For iii) we calculate $f^\circ(x;-v)$ using the change $u = x' - tv$:

$$f^\circ(x;-v) = \limsup_{\substack{x' \to x \\ t \downarrow 0}} \frac{f(x'-tv) - f(x')}{t}$$

$$= \limsup_{\substack{u \to x \\ t \downarrow 0}} \frac{(-f)(u+tv) - (-f)(u)}{t}$$

$$= (-f)^\circ(x;v) .$$

The proposition is proven. ■

2.3 Generalized Gradient

Definition 3.1

The generalized gradient of f at x, denoted $\partial f(x)$, is the set

$$\partial f(x) = \{\xi \in E' : f^\circ(x;v) \geq \xi(v) \text{ , for all } v \in E\} .$$

Proposition 3.2

Let f be Lipschitz of rank ℓ near x. Then

i) $\partial f(x)$ is nonempty, convex, weak*-compact in E' and

$|\xi| \leq \ell$ for all $\xi \in \partial f(x)$.

ii) For every $v \in E$, one has

$$f^\circ(x;v) = \max_{\xi \in \partial f(x)} \xi(v) .$$

Proof. For i) observe that if f is Lipschitz near x, then by Proposition 2.2 and the Hahn-Banach theorem, there is $\xi \in E'$ which is majorized by the finite positively homogeneous subadditive function $f^\circ(x;\cdot)$:

$$f^\circ(x;v) \geq \xi(v) , \text{ for all } v \in E ,$$

which implies that $\partial f(x) \neq \phi$. The convexity of this set follows from the subadditivity of $f^\circ(x;\cdot)$. The weak*-compactness is derived from Alaoglu's theorem which says that the polar of every neighborhood of zero is weak*-compact.

For ii), by definition one has

$$f^\circ(x;v) \geq \max_{\xi \in \partial f(x)} \xi(v).$$

If there is some v_0 with

$$f^\circ(x;v_0) > \max_{\xi \in \partial f(x)} \xi(v_0) ,$$

then by a separation theorem, one can find $\xi_0 \in E'$ such that $f^\circ(x;v_0) = \xi_0(v_0)$ and

$$f^\circ(x;v) \geq \xi_0(v) \text{ , for all } v \in E \text{ .}$$

Hence $\xi_0 \in \partial f(x)$ and we arrive at a contradiction

$$\max_{\xi \in \partial f(x)} \xi(v_0) \geq \xi_0(v_0) = f^\circ(x;v_0) > \max_{\xi \in \partial f(x)} \xi(v_0) \text{ . } \blacksquare$$

For further investigations of generalized gradient we need some definitions concerning set-valued maps.

Definition 3.3

A multifunction (or set-valued map) F from X to Y is a map from X to the subsets of Y .

F is said to be closed at x if any sequence $\{(x_n,y_n)\}$ with $y_n \in F(x_n)$ converges to some (x,y) , one has $y \in F(x)$. F is said to be closed if it is closed at every point $x \in X$, or equivalently, if its graph $gr\ F = \{(x,y) \in X \times Y\colon y \in F(x) , x \in X\}$ is a closed set in the product space $X \times Y$.

F is said to be upper continuous at x if for every $\epsilon > 0$, there is $\delta > 0$ such that

$$F(x') \leq F(x) + \epsilon B_y(0;1) \text{ , for all } x' \in x + \delta B_x(0;1) \text{ .}$$

F is said to be lower continuous at x if for every $\epsilon > 0$, $y \in F(x)$, there exists $\delta > 0$ such that

$$F(x') \cap (y + \varepsilon B_Y(0;1)) \neq \phi \text{ , for all } x' \in x + \delta B_X(0;1) \text{ .}$$

Proposition 3.4

Suppose that f is Lipschitz near x. Then

i) ∂f is weak*-closed at x.

ii) $\partial f(x) = \bigcap_{\delta>0} \bigcup_{x' \in x+\delta B} \partial f(x')$.

iii) If E is finite dimensional, then ∂f is upper continuous at x.

Proof. For i), let $\{x_n\}$ be a sequence in E converging to x, $\{\xi_n\}$ a sequence in E' weakly* converging to ξ and $\xi_n \in \partial f(x_n)$. We prove $\xi \in \partial f(x)$. Since for every fixed $v \in E$, $\lim \xi_n(v) = \xi(v)$. But $f^\circ(x_n;v) \geq \xi_n(v)$, hence in view of Proposition 2.2, $f^\circ(x;v) \geq \xi(v)$, which shows $\xi \in \partial f(x)$.

The second statement is immediate from the first one.

We turn to iii). If ∂f is not upper continuous at x, then there are $x_n \in E$, $\xi_n \in E'$ with $\xi_n \in \partial f(x_n)$, $\lim x_n = x$ and a neighborhood V of $\partial f(x)$ such that $\xi_n \notin V$. Since E is finite dimensional, one may assume that $\{\xi_n\}$ converges to some $\xi \in E'$. It is obvious that $\xi \notin \partial f(x)$. This contradicts i). ■

Proposition 3.5

If f is Lipschitz near x and admits a Gateaux derivative $Df(x)$, Then $Df(x) \in \partial f(x)$.

Moreover, if f is strictly differentiable at x, then f is Lipschitz near x and $\partial f(x) = \{D_s f(x)\}$. Conversely, if f is Lipschitz near x and $\partial f(x)$ reduces to a singleton ξ, then f is strictly differentiable at x with $D_s f(x) = \xi$.

Proof. By definition,

$$f'(x;v) = Df(x)(v), \quad v \in E.$$

It is clear that $f' \leq f^\circ$ from the definition of the latter, so one has

$$f^\circ(x;v) \geq Df(x)(v), \quad v \in E,$$

which implies $Df(x) \in \partial f(x)$.

For the second part, suppose first that $D_s f(x)$ exists, so that f is Lipschitz near x by Proposition 1.2. Then by definition of f°, one has

$$f^\circ(x;v) = D_s f(x)(v), \quad v \in E.$$

Hence $\partial f(x) = \{D_s f(x)\}$.

Conversely, in view of Proposition 1.2, it suffices to show that

$$\lim_{\substack{x' \to x \\ t \downarrow 0}} \frac{f(x'+tv) - f(x')}{t} = \xi(v) .$$

We begin by showing that $f^\circ(x;v) = \xi(v)$, for every v. It is know from Proposition 3.2 that

$$f^\circ(x;v) \geq \xi(v) .$$

Using a separation theorem one can find $\xi' \in E'$ for a fixed $v_0 \in E$ such that

$$f'(x;v_0) = \xi'(v_0) .$$

$$f'(x;v) \geq \xi'(v) , \quad v \in E .$$

If $f^\circ(x;v) > \xi(v)$, then ξ' and ξ would be distinct elements of $\partial f(x)$, contrary to the hypothesis. Thus $f^\circ(x;v_0) = \xi(v_0)$, where v_0 is an arbitrary vector of E. Using this fact we calculate the limit:

$$\begin{aligned}
\liminf \frac{f(x'+ty)-f(x')}{t} &= -\limsup \frac{f(x') - f(x'+tv)}{t} \\
&= -\limsup \frac{f(x'+tv-tv) - f(x'+tv)}{t} \\
&= - f^\circ(x;-v) \\
&= \xi(v) \\
&= f^\circ(x;v) \\
&= \limsup \frac{f(x'+tv) - f(x')}{t} ,
\end{aligned}$$

which completes the proof. ■

Corollary 3.6

If f is Lipschitz near x and E is finite dimensional, then $\partial f(x')$ reduces to a singleton for every $x' \in x + \varepsilon B$ if and only if f is continuously differentiable on $x + \varepsilon B$.

Proof. Note that a point-valued map is continuous if and only if it is upper continuous in the sense of set-valued maps. This and Propositions 1.2, 3.4 prove the result. ■

Proposition 3.7

When f is convex on U and Lipschitz near x, then $\partial f(x)$ coincides with the subdifferential at x in the sense of convex analysis and $f^\circ(x;v)$ coincides with the directional derivative $f'(x;v)$, for all $v \in E$.

Proof. It is known from the previous chapter that $f'(x;v)$ exists for every $v \in E$ and that $f'(x;\cdot)$ is the support function of the subdifferential at x. It suffices therefore to show that

$$f^\circ(x;v) = f'(x;v), \text{ for all } v \in E.$$

By definition of f°, one can express it as

$$f^\circ(x;v) = \limsup_{\varepsilon \downarrow 0} \sup_{\|x-x'\| \le \varepsilon\delta} \sup_{0<t<\varepsilon} \frac{f(x'+tv) - f(x')}{t},$$

where δ is any fixed positive number. We know that $((fx'+tv) - f(x'))/t$ is nondecreasing as t increases, whence

$$f^\circ(x;v) = \limsup_{\varepsilon \downarrow 0} \sup_{\|x'-x\| < \varepsilon\delta} \frac{f(x'+\varepsilon v) - f(x')}{t}.$$

Now by the Lipschitz condition, for every $x' \in x + \varepsilon\delta B$ one has

$$\left| \frac{f(x'+\varepsilon v) - f(x')}{\varepsilon} - \frac{f(x+\varepsilon v) - f(x)}{\varepsilon} \right| \le 2\delta\ell,$$

so that

$$f^\circ(x;v) \le \lim_{\varepsilon \downarrow 0} \frac{f(x+\varepsilon v) - f(x)}{\varepsilon} + 2\delta\ell$$

$$= f'(x;v) + 2\delta\ell.$$

Since δ is arbitrary, we deduce

$$f^\circ(x;v) \le f'(x;v).$$

The converse inequality is trivial. ■

2.4 Calculus Rules

Throughout this section all functions are presumed to be Lipschitz near the point of our interest in the space E.

Definition 4.1

f is said to be regular at x if the usual directional derivative $f'(x;v)$ exists and

$$f'(x;v) = f^\circ(x;v) \text{ , for all } v \in E .$$

Proposition 4.2

We have the following

i) $\partial(tf)(x) = t\,\partial f(x)$, for every $t \in R$.

ii) $\partial(f_1+\ldots+f_n)(x) \subseteq \partial f_1(x)+\ldots+\partial f_n(x)$ and equality holds if all but at most one of these functions f_i are strictly differentiable at x , or all of them are regular at x .

Proof. For i) note that tf is also Lipschitz near x . When $t \geq 0$, $(tf)^\circ = tf^\circ$. Hence

$$\partial(tf)(x) = t\,\partial f(x) .$$

For $t = -1$, $\xi \in \partial(-f)(x)$ if and only if $(-f)^\circ(x;v) \geq \xi(v)$ for all $v \in E$. In view of Proposition 2.2, $f^\circ(x;-v) \geq \xi(v)$, which implies

$-\xi \in \partial f(x)$. Thus $\xi \in \partial(-f)(x)$ if and only if $\xi \in -\partial f(x)$.

To prove ii) it suffices to take the case where $n=2$. Note first that for every $v \in E$,

$$f_1^\circ(x;v) + f_2^\circ(x;v) \geq (f_1+f_2)^\circ(x;v) .$$

Hence in view of proposition 3.2, one has

$$\max_{\xi \in \partial f_1(x)+\partial f_2(x)} \xi(v) \geq \max_{\xi \in \partial(f_1+f_2)(x)} \xi(v) . \tag{4.1}$$

Since the gradients are convex weak*-compact (Proposition 3.2), the above inequality implies

$$\partial(f_1+f_2)(x) \subseteq \partial f_1(x) + \partial f_2(x) . \tag{4.2}$$

Now, assume that f_1 is strictly differentiable, then

$$\begin{aligned}(f_1+f_2)^\circ(x;v) &= f_1'(x;v) + f_2^\circ(x;v) \\ &= f_1^\circ(x;v) + f_2^\circ(x;v). \end{aligned} \tag{4.3}$$

Hence in (4.1) equality holds which implies equality in (4.2). In the case both of these functions are regular, (4.3) is true and again equality holds in (4.2). ■

Note that if f is Lipschitz near x and convex, then it is regular. Hence equality in Proposition 4.2 also holds if f_i are convex Lipschitz.

Theorem 4.3 (Mean-value theorem of Lebourg)

Let $[x,y]$ be an interval in an open set where f is Lipschitz. Then there exists $u \in (x,y)$ such that

$$f(y) - f(x) \in \{\xi(y-x): \ \xi \in \partial f(u)\} .$$

Proof. Observe first that if a Lipschitz function $\varphi(t)$ attains a local extremum at t then $\varphi^\circ(t;v) \geq 0$ for all v. Consequently, $0 \in \partial(t)$. Now, let us consider the function

$$g(t) = f(x + t(y-x)).$$

We calculate the gradient of g via directional derivative :

$$g^\circ(t;v) = \limsup_{\substack{t' \to t \\ \lambda \downarrow 0}} \frac{g(t'+\lambda v) - g(t')}{\lambda}$$

$$= \limsup \frac{f(x+(t'+\lambda v)(y-x)) - f(x+t'(y-x))}{\lambda}$$

$$\leq \limsup_{\substack{x' \to x+t(y-x) \\ \lambda \downarrow 0}} \frac{f(x' + \lambda v(y-x)) - f(x')}{\lambda}$$

$$= f^\circ(x+t(y-x); v(y-x)).$$

This implies $\partial q(t) \subseteq \partial f(x+t(y-x))\ (y-x)$. (4.4)

Consider another function:

$$\varphi(t) = f(x+t(y-x)) + t(f(x) - f(y)).$$

Since $\varphi(0) = \varphi(1) = f(x)$, there is a point $t \in (0,1)$ at which φ attains a local extremum. By the observation at the beginning of the proof and the formula (4.4) one has

$$0 \in \partial\varphi(t) = f(x) - f(y) + \partial f(x+t(y-x))\ (y-x),$$

Which implies the desired result.■

Theorem 4.4 (Chain Rules)

Let $h: E \to R^n$, $g: R^n \to R$ be Lipschitz functions.

Then

$$\partial(g.h)(x) \subseteq \overset{*}{W} \text{-el conv } \{\Sigma\alpha_i\xi_i : \xi_i \in \partial h_i(x),\ \alpha \in \partial g(h(x))\}$$

and equality holds in the following cases

i) h_i are regular at x, g is regular at $h(x)$ and every element $\alpha \in \partial g(h(x))$ has nonnegative components

ii) g is strictly differentiable at $h(x)$ and $n=1$;

iii) g is regular at $h(x)$ and h is strictly differentiable at x.

Proof. As before B stands for the unit ball with the center at 0 in the spaces of our consideration.

Denote

$$q_\varepsilon(v) = \max\{\Sigma\alpha_i\xi_i(v):\ \xi_i\in\partial h_i(x_i)\ ,\ \alpha\in\partial g(u)\ ,$$
$$x_i\in x+\varepsilon B\ ,\ u\in h(x) + \varepsilon\ B\ \}\ .$$

We wish to show that for each $\varepsilon > 0$,

$$q_\varepsilon(v) \geq (g\circ h)^\circ(x;v) - \varepsilon. \tag{4.5}$$

If this is true and if

$$q_\varepsilon \text{ decreases to } q_0 \text{ as } \varepsilon \text{decreases to } 0\ , \tag{4.6}$$

then

$$q_0(v) \geq (g\circ h)^\circ(x;v)\ ,\ \text{for all } v\ .$$

The latter inequality and the fact that the support function of the set on the right hand side in the theorem is given by q_0, together with Proposition 3.2 imply the set inclusion of the theorem.

Therefore our aim at this moment is to stablish (4.5) and (4.6). For (4.5), set $f = g\circ h$. By definition of generalized directional derivative,

there exists $x' \in x + \varepsilon B$ and t near to 0 such that

$$f^\circ(x;v) \leq \frac{f(x'+tv) - f(x')}{t} + \varepsilon . \tag{4.7}$$

We may choose x' and t such that

$$x' + tv \in x + \varepsilon B ,$$

$$h(x') \in h(x) + \varepsilon B ,$$

$$h(x'+tv) \in h(x) + \varepsilon B .$$

By Theorem 4.3, one has

$$f(x'+tv) - f(x') = g(h(x'+tv)) - g(h(x'))$$

$$= \Sigma\, \alpha_i [h_i(x'+tv) - h_i(x')] ,$$

where $\alpha \in \partial g(u)$, $u \in [h(x'+tv) , h(x')]$.

Again, using Theorem 4.3 for h_i one has

$$h_i(x'+tv) - h_i(x') = \xi_i(tv) ,$$

where $\xi_i \in \partial h_i(x_i)$, $x_i \in [x'+tv, x']$.

Hence,

$$f(x'+tv) - f(x') = \Sigma\, \alpha_i\, \xi_i(tv) ,$$

which combines with (4.7) to give (4.5):

$$f^\circ(x;v) \leq \Sigma \alpha_i \xi_i(v) + \varepsilon \leq q_\varepsilon(v) + \varepsilon .$$

For (4.6), observe that since $q_\varepsilon(v) \geq 0$, it suffices to show that for every $\delta > 0$, one has

$$q_\varepsilon(v) \leq q_0(v) + n\delta(1+\ell|v|) ,$$

where ℓ is a Lipschitz constant, for ε small enough. In view of Proposition 2.2, one can choose ε small so that h_i is Lipschitz of rank ℓ on $x + \varepsilon B$, and so that for all i, all $x_i \in x + \varepsilon B$:

$$h_i^\circ(x_i;\pm v) \leq h_i^\circ(x;\pm v) + \delta/\ell .$$

Since for each $\xi_i \in \partial h_i(x_i)$, $\|\xi_i\| \leq \ell$ (Proposition 3.2), we obtain

$$h_i^\circ(x_i;\xi_i v) \leq h_i^\circ(x;\xi_i v) + \delta.$$

Also by Proposition 3.4, ∂g is upper continuous, which means that when ε is small enough, one has the inclusion:

$$\partial g(h(x) + \varepsilon B) \subseteq \partial g(h(x)) + \delta B .$$

Now we are able to estimate $q_\varepsilon(v)$:

$$q_\varepsilon(v) \le \max\{ \sum_i \max[\alpha_i \xi_i(v): \xi_i \in \partial h_i(x_i)\ ,\ x_i \in x + \varepsilon B]:$$

$$\alpha \in \partial g(x) + \delta B \}$$

$$\le \max\{ \sum_i (h_i^\circ(x;\alpha_i v) + \delta): \alpha \in \partial g(x) + \delta B \}$$

$$\le \max\{ \sum_i \max[\alpha_i \xi_i(v): \xi_i \in \partial h_i(x)]: \alpha \in \partial g(x) + \delta B \}$$

$$+ n\delta$$

$$\le q_0(\varepsilon) + n\delta \ell\, |v| + n\delta\ .$$

This shows (4.6) (actually we have proven that $q_0(v) \ge \lim_{\varepsilon \downarrow 0} q_\varepsilon(v)$), and the inclusion of the theorem is established.

Now suppose additionally that i) is satisfied. Since $\alpha \ge 0$ we can calculate $q_0(v)$ as follows

$$q_0(v) = \max\{ \sum \alpha_i \max[\xi_i(v): \xi_i \in \partial h_i(x)]: \alpha \in \partial g(h(x))\}$$

$$= \max\{ \sum \alpha_i h_i'(x;v): \alpha \in \partial g(h(x)))\}$$

$$= g'(h(x)\ ;\ h'(x;v))$$

$$= \lim_{t \downarrow 0} \frac{g(h(x) + th'(x;v)) - g(h(x))}{t}$$

$$= \lim_{t \downarrow 0} \{ \frac{g(h(x+tv)) - g(h(x))}{t} + \frac{g(h(x) + th'(x;v)) - g(h(x+tv))}{t} \}$$

$$= \lim_{t \downarrow 0} \frac{g(h(x+tv)) - g(h(x))}{t}$$

$$= f'(x;v)\ . \qquad (4.8)$$

In the above calculation we have used the fact that g is Lipschitz near $h(x)$, hence the amount

$$h'(x;v) - \frac{h(x+tv) - h(x)}{t} \to 0$$

as t tends to zero. Moreover, by (4.5) and (4.6) one has on one hand

$$f^\circ(x;v) \leq q_0(v) .$$

On the other hand it follows from the definition,

$$f'(x;v) \leq f^\circ(x;v) .$$

These inequalities and (4.8) yield

$$q_0(0) = f'(x;v) = f^\circ(x;v) .$$

In particular, f is regular, and by the property of support functions, equality in the theorem holds.

Further, suppose that ii) is satisfied. Then $D_s g(h(x)) = \alpha$ is a scalar. We may assume $\alpha \geq 0$ and $q_0(v)$ can be calculated as follows

$$q_0(v) = \alpha h^\circ(x;v)$$

$$= \limsup_{\substack{x' \to x \\ t \downarrow 0}} \frac{\alpha[h(x'+tv) - h(x')]}{t}$$

$$= \limsup \frac{g(h(x'+tv)) - g(h(x'))}{t}$$

$$= f^\circ(x;v) .$$

We have exploited the fact that g is strictly differentiable at $h(x)$. Again using property of support functions one obtains the desired equality.

The case iii) is proven by a way similar to that of i). □

Theorem 4.5 (Pointwise Maxima)

Let T, $f(x)$, $T(x)$, $P(x)$ be as in Theorem 5.10 (Chapter 1). Assume that

i) T is a sequentially compact space;

ii) For every fixed y of a neighborhood U of x, $t \to f_t(y)$ is upper semicontinuous on T;

iii) For every fixed $t \in T$, $f_t(\cdot)$ is Lipschitz of rank ℓ on U, and $\{f_t(x): t \in T\}$ is bounded.

iv) E is separable, or T is metrizable.

Then

$$\partial f(x) \subseteq \bigcup_{\mu \in P(c)} \int_T \bar{\partial} f_t(x) \mu(dt) ,$$

where

$$\bar{\partial} f_t(x) = w^*\text{-cl conv } \{\xi \in E' : \text{there exist } \xi_i \in \partial f_{t_i}(x_i) \text{ with } x_i \to x , t_i \to t , \xi_i \overset{w^*}{\to} \xi\} .$$

Moreover, if $\bar{\partial} f_t(x) = \partial f_t(x)$ and f_t is regular at x for every $t \in T(x)$, then f is regular at x and equality holds.

Proof. Let us consider the function g on $U \times E$ defined by the formula

$$g(x;v) = \max\{\xi(v) : \xi \in \bar{\partial} f_t(x) , t \in T(x)\} .$$

This function is well-defined because by the Lipschitz condition, the set $\bar{\partial} f_t(x)$ is weak*-compact and $T(x)$ is closed, hence compact by i).

We wish establish the relation

$$f^\circ(x;v) \leq g(x;v) , \text{ for all } v \in E . \tag{4.9}$$

By definition of generalized directional derivative, there are $y_i \to x$, $\lambda_i \downarrow 0$ such that

$$f^{\circ}(x;v) = \lim \left(\Delta_i = \frac{f(y_i + \lambda_i v) - f(y_i)}{\lambda_i}\right).$$

Pick any $t_i \in T(y_i + \lambda_i v)$ to have

$$\Delta_i \leq \frac{f_{t_i}(y_i+\lambda_i v) - f_{t_i}(y_i)}{\lambda_i} = \delta_i$$

By the mean-value theorem one can find $\xi_i \in \partial f_{t_i}(y_i^*)$ with $y_i^* \in [y_i, y_i+\lambda_i v]$ such that $\delta_i = \xi_i(v)$.

Without loss of generality it can be assumed that $t_i \to t \in T$, $\xi_i \overset{w^*}{\to} \xi$. It is obvious that $y_i^* \to x$, hence $\xi \in \bar{\partial} f_t(x)$. Consequently, $\delta_i \to \xi(v)$ and (4.9) follows if we can show that $t \in T(x)$.

This can be seen by noticing that

$$f_{t_i}(y_i + \lambda_i v) \geq f_\tau(y_i + \lambda_i v), \text{ for } \tau \in T,$$

which together with ii), iii) implies

$$f_t(c) \geq f_\tau(x), \text{ for all } \tau \in T.$$

Now, to prove the inclusion of the theorem, take any $\xi \in \partial f(x)$. In view of (4.9),

$$g(x;v) \geq \xi(v).$$

Observe that $\xi \in \partial g(x;0)$ because $g(x,0) = 0$. Since $g(x;.)$ is a convex function of the variable v, one can apply Theorem 5.10 (Chapter 1) to see that there exists $\mu \in P(x)$ such that

$$\xi = \int_T \xi_t(.)\, \mu(dt),$$

$$\xi_t \in \bar{\partial} f_t(x) \qquad \mu\text{-almost everywhere.}$$

This establishes the set inclusion in the theorem.

Finally, let $\alpha = \liminf_{\lambda \downarrow 0} \frac{f(x+\lambda v) - f(x)}{\lambda}$. To prove the regularity of f at x, it suffices to show that

$$f^\circ(x;v) \leq \alpha \ .$$

For this purpose note that

$$\alpha \geq f_t'(x;v) = f_t^\circ(x;v) \text{ , for all } t \in T(x) \ ,$$

which implies

$$\alpha \geq \max_{t \in T(x)} f^\circ(x;v) \ . \tag{4.10}$$

By the assumption $\bar{\partial} f_t(x) = \partial f_t(x)$ one has

$$g(x;v) = \max\{\xi(v): \xi \in \partial f_t(x) \ , \ t \in T(x)\} \ .$$

Observe that $\xi \in \partial g(x;0)$ because $g(x;0) = 0$. Since $g(x;.)$ is a convex function of the variable v, one can apply Theorem 5.10 (Chapter 1) to see that there exists $\mu \in P(x)$ such that

$$\xi = \int_T \xi_t(.)\, \mu(dt),$$

$$\xi_t \in \partial f_t(x) \quad \mu\text{-almost everywhere.}$$

This establishes the set inclusion in the theorem.

Finally, let $\alpha = \liminf \frac{f(x+\lambda v) - f(x)}{2}$.

To prove the regularity of f at x, it suffices to show that

$$f^o(x;v) \leq \alpha.$$

For this purpose note that

$$\alpha \geq f_t'(x;v) = f_t^o(x;v), \text{ for all } t \in T(x),$$

which implies

$$\alpha \geq \max_{t \in T(x)} f_t^o(x;v).$$

By the assumption $\overline{\partial} f_t(x) = \partial f_t(x)$ one has

$$g(x;v) = \max \{\xi(v): \xi \in \partial f_t(x),\ t \in T(x)\}.$$

Combine the latter equality with (4.10) and (4.9) to see that $f^\circ(x;v) \leq \alpha$, and indeed f is regular at x. To show that equality holds in the theorem, let

$$\xi = \int_T \xi_t(\cdot)\mu(dt) ,$$

be any element of the set in the right hand side of the inclusion in the theorem. Then for every $v \in E$ one has

$$\begin{aligned}
\xi(v) &= \int_T \xi_t(v)\mu(dt) \\
&\leq \int_T f_t^\circ(x;v)\mu(dt) \\
&= \int_T f_t(x;v)\mu(dt) \\
&= \lim_{\lambda \downarrow 0} \frac{\int_T f_t(x+\lambda v)\mu(dt) - \int_T f_t(x)\mu(dt)}{\lambda} \\
&\leq \limsup_{\lambda \downarrow 0} \frac{\int_T f(x+\lambda v)\mu(dt) - \int_T f(x)\mu(dt)}{\lambda} \\
&= \limsup_{\lambda \downarrow 0} \frac{f(x+\lambda v) - f(x)}{\lambda} \\
&= f'(x;v) = f^\circ(x;v) .
\end{aligned}$$

Thus, $\xi(v) \leq f^\circ(x;v)$ and $\xi \in \partial f(x)$. ■

2.5 Geometric Illustrations

Let C be a nonempty subset of E. We recall that the distance function is:

$$d_C(x) = \inf\{ \|x-c\| : c \in C \}.$$

Proposition 5.1

The distance function is Lipschitz of rank 1 on E.

Proof. For every positive ε, there is $c \in C$ such that

$$d_C(x') \geq \|x'-c\| - \varepsilon .$$

Hence

$$d_C(x) \leq \|x-c\| \leq \|x-x'\| + \|x'c\| \leq \|x-x'\| + d_C(x') + \varepsilon.$$

Since ε is arbitrary, we conclude that

$$|d_C(x) - d_C(x')| \leq \|x-x'\| . \blacksquare$$

Definition 5.2

Let $x \in C$. The Clarke tangent cone of C at x is the convex cone

$$T(C;x) = \{v \in E: d_C^\circ(x;v) = 0 \}.$$

The normal cone to C at x is the polar cone of $T(C;x)$:

$$N(C;x) = \{\xi \in E' : \xi(v) \leq 0 \text{ , for all } v \in T(C;x)\} \text{ .}$$

Proposition 5.3

We have the following:

i) $N(c;x) = w^*\text{-cl cone } \partial d_C(x)$;

ii) $T(C;x) = \{v \in E$: for $x_i \in C$, $t_i > 0$ with $x_i \to x$, $t_i \downarrow 0$,

there are $v_i \in E$ with $v_i \to v$ and $x_i + t_i v_i \in C\}$.

Proof. For the first assertion we use definitions: $v \in T(C;x)$ if and only if $d_C^\circ(x;v) = 0$, which means that $v(\xi) \leq 0$ for all $\xi \in \partial d_C(x)$.

Hence $(T(C;v))^\circ = w^*\text{-cl cone } \partial d_C(x)$.

For the second part, suppose first that $v \in T(C;v)$ and $x_i \to x$, $t_i \downarrow 0$ are given. Since $d_C^\circ(x;v) = 0$, one has

$$\lim \frac{d_C(x_i+t_iv) - d_C(x_i)}{t_i} = \lim \frac{d_C(x_i+t_iv)}{t_i} = 0 \text{ .} \tag{5.1}$$

Let $c_i \in C$ satisfy:

$$\|x_i + t_iv - c_i\| \leq d_C(x_i+t_iv) + t_i/i \text{ .} \tag{5.2}$$

Set

$$v_i = (c_i - x_i)/t_i \ .$$

Then (5.1) and (5.2) imply that $\|v - v_i\| \to 0$. Moreover, $x_i + t_i v_i = c_i \in C$.

Conversely, let v be a vector as stated in ii). Choose $x_i \in E$, $x_i \to x$ and $t_i > 0$, $t_i \downarrow 0$ such that

$$d_C^\circ(x;v) = \lim \frac{d_C(x_i+t_i v) - d_C(x_i)}{t_i} \ . \tag{5.3}$$

In order to show $v \in T(C;x)$ we have to prove

$$d_C^\circ(x;v) = 0 \ .$$

It suffices to prove that $d_C^\circ(x;v) \leq 0$ because

$$\begin{aligned} d_C^\circ(x;v) &= \limsup_{\substack{x' \to x \\ t \downarrow 0}} \frac{d(x'+tv) - d(x')}{t} \\ &\geq \limsup_{t \downarrow 0} \frac{d(x+tv) - d(x)}{t} \\ &\geq 0 \ . \end{aligned}$$

Let $c_i \in C$ with the property:

$$\|c_i - x_i\| \leq d_C(x_i) + t_i/i \ . \tag{5.4}$$

Then $c_i \to x$. By assumption, there exist v_i with $v_i \to v$ and $c_i + t_i v_i \in C$.

Using Proposition 5.1 and (5.4) we obtain the relations:

$$d_C(x_i+t_i v) \leq d_C(c_i+t_i v_i) + \|x_i - c_i\| + t_i \ \|v - v_i\|$$

$$\leq d_C(x_i) + t_i(\|v-v_i\| + \tfrac{1}{i}) \ .$$

This makes the limit in (5.3) nonpositive and $v \in T(C;x)$. ∎

Definition 5.4

The contingent cone or Bouligand's cone of C at x is the cone (not necessarily convex)

$$K(C;x) = \{v \in E: \text{ there are } v_i \to v \ , \ t_i \downarrow 0$$

$$\text{such that } c_i+t_i v_i \in C\} \ .$$

Note that in the definition above it is necessary that $x \in \text{cl } C$. It is clear that $T(C;x) \subseteq K(C;x)$.

Definition 5.5

The set C is said to be regular at x provided

$$T(C;x) = K(C;x) .$$

Theorem 5.6

Assume that f is Lipschitz near x, and $0 \notin \partial f(x)$. Let C be the level set of f at $f(x)$. Then

$$\{ v \in E : f^\circ(x;v) \leq 0 \} \subseteq T(C;x) .$$

Moreover, if f is regular at x, then equality holds and C is regular at x.

Proof. Note first that since $0 \notin \partial f(x)$, there exists $v_0 \in E$ such that

$$f^\circ(x;v_0) = \sup\{\xi(v_0): \xi \in \partial f(x)\} < 0 .$$

Hence, for every v with $f^\circ(x;v) \leq 0$ and ε small enough one has

$$f^\circ(x;v + \varepsilon v_0) < 0 .$$

In this way, it suffices to show that

$$f^\circ(x;v) < 0 \quad \text{implies} \quad v \in T(C;x) .$$

Indeed, by definition of $f^\circ(x;v)$, there exist positive numbers ε and δ such that

$$x' \in x + \varepsilon B \quad \text{and} \quad t < \varepsilon \quad \text{imply} \quad f(x'+tv) - f(x') \leq -\delta t .$$

For any $x_i \to x$ and $t_i \downarrow 0$ one has

$$f(x_i) \leq f(x) ,$$

and when i is large,

$$f(x_i + t_i v) \leq f(x_i) - \delta t_i \leq f(x) - \delta t_i$$

(because $\limsup_{\substack{x' \to x \\ t \downarrow 0}} \dfrac{f(x'+tv) - f(x')}{t} < 0$) .

Consequently, $x_i + t_i v \in C$ if i is large, and in view of Proposition 5.3, $v \in T(C;x)$.

Now if in addition f is regular, we prove that

$$v \in K(C;x) \quad \text{implies} \quad f^\circ(x;v) \leq 0 ,$$

which shows that C is regular and equality holds. One has by definition:

$$\liminf_{t \downarrow 0} \frac{d_C(x+tv)}{t} = 0 .$$

In particular, for each positive ε , there exists $t_i \downarrow 0$ with

$$d_C(x + t_i v) \leq \varepsilon t_i .$$

Hence there exist $x_i \in C$ such that

$$\|x + t_i v - x_i\| \leq 2\varepsilon t_i ,$$

and of course

$$f(x_i) \leq f(x) .$$

One deduces then

$$\frac{f(x+t_i v) - f(x)}{t_i} \leq \frac{f(x_i) + 2\varepsilon \ell t_i - f(x)}{t_i} \leq 2\varepsilon \ell ,$$

where ℓ is a Lipschitz constant of f near x . This follows that

$$f^\circ(x;v) = f'(x;v) \leq 0 . \blacksquare$$

Corollary 5.7

Let C be as in Theorem 5.6 . Then

$$N(x|C) \subseteq \text{cone } \partial f(x) .$$

Moreover, equality holds if f is regular at x.

Proof. By Theorem 5.6, one has the inclusion:

$$\{v: f^\circ(x;v) \leq 0\} \subseteq T(C;x) ,$$

which implies

$$(T(C;x))^\circ \subseteq \{v: f^\circ(x;v) \leq 0\}^\circ .$$

This is the same as

$$N(x|C) \subseteq w^*\text{-cl cone } \partial f(x) .$$

To conclude the result it suffices to note that cone $\partial f(x)$ is closed because $\partial f(x)$ is weakly compact and not containing zero. ■

Corollary 5.8

Let $C = \{x \in X: f_i(x) \leq 0 , i = 1,\ldots,n\}$ and $x_0 \in C$ with $f_i(x_0) = 0$, $i = 1,\ldots,n$. Then, if f_i are strictly differentiable at x_0 and $D_S f_i(x_0)$ are positively linearly independent, it follows that C is regular at x and

$$N(x|C) = \{ \sum_{i=1}^{n} \lambda_i D_S f_i(x_0) : \lambda_i \geq 0 , i = 1,\ldots,n\} .$$

Proof. Define

$$f(x) = \max\{f_i(x) : i = 1,\dots,n\} .$$

Then f is Lipschitz near x_0 and regular at x_0. The set

$$C = \{x: f(x) \leq 0\}$$

is regular at x_0 where $f(x_0) = 0$. In view of Corollary 5.7 we have equality

$$N(x_0|C) = \text{cone } \partial f(x_0) .$$

By Theorem 4.5, $\partial f(x_0)$ is the convex hull of $D_S f_i(x_0)$. This implies the result of the corollary. ■

2.6 Ekeland's Variational Principle

Let X be a complete metric space with metric $d(x,y)$, and let $f: X \to R \cup \{+\infty\}$ be lower semicontinuous bounded below. We shall study the points which almost minimize f. For this purpose we need the following partial order in the product space $X \times R$: for a fixed $\varepsilon > 0$, $(x,\alpha) > (x',\alpha')$ if $\alpha - \alpha' + \varepsilon d(x,x') \leq 0$. It is obvious that the order $>$ above is reflexive, antisymmetric transitive. Moreover, for each fixed $(x_0,\alpha_0) \in X \times R$, the set

$$\{(x,\alpha) \subset X \times R : (x,\alpha) > (x_0,\alpha_0)\}$$

is closed.

Lemma 6.1

Let S be a closed set in $X \times R$ with the property that there is a number m such that $\alpha \geq m$ whenever $(x,\alpha) \in S$. Then for every $(x_0,\alpha_0) \subset S$, there is a maximal element (x_*,α_*) in S with

$$(x_*,\alpha_*) > (x_0,\alpha_0) .$$

Proof. Define a sequence $\{(x_n,\alpha_n)\}$ in S by induction starting with (x_0,α_0) :

$$S_n = \{(x,\alpha) \in S : (x,\alpha) > (x_n,\alpha_n)\}$$

$$m_n = \inf\{\alpha : (x,\alpha) \in S_n , \text{ some } x \in X\} .$$

It is clear that $m_n \geq m$. Take $(x_{n+1},\alpha_{n+1}) \in S$ with the property:

$$\alpha_n - \alpha_{n+1} \geq \frac{1}{2}(\alpha_n - m_n) . \tag{6.1}$$

Then S_n are closed and $S_{n+1} \subseteq S_n$.

Moreover, by (6.1) one has

$$|\alpha_{n+1} - m_{n+1}| \leq \frac{1}{2}|\alpha_n - m_n| \leq |\alpha_0 - m|/2^n .$$

Hence, for every $(x,\alpha) \in S_{n+1}$,

$$|\alpha_{n+1} - \alpha| \leq |\alpha_{n+1} - m_{n+1}| \leq |\alpha_0 - m|/2^n ,$$

$$d(x_{n+1},x) \leq |\alpha_0 - m|/2^n\alpha .$$

In other words, the diameter of S_n tends to zero.

Since X and R are complete, one has

$$\bigcap_{n\geq 1} S_n = \{(x_*,\alpha_*)\} .$$

By definition, $(x_*,\alpha_*) > (x_n,\alpha_n)$ for all n. Now we show that this element is maximal in S. In fact, if $(x',\alpha') > (x_*,\alpha_*)$, then by transitivity, $(x',\alpha') > (x_n,\alpha_n)$ for all n, whence $(x',\alpha') \in \bigcap_n S_n$, which implies that

$$(x', ') = (x_*,\alpha_*) .$$

The proof is complete. ■

Theorem 6.2

If x_0 is a point in X satisfying

$$f(x_0) \leq \inf f + \varepsilon ,$$

for some $\varepsilon > 0$, then for every $\lambda > 0$, there exists $x_\lambda \in X$ such that

i) $f(x_\lambda) \leq f(x_0)$

ii) $d(x_\lambda, x_0) \leq \lambda$

iii) for every $x \neq x_\lambda$, one has

$$f(x) + \frac{\varepsilon}{\lambda} d(x, x_\lambda) > f(x_\lambda) .$$

Proof. Set $S = \text{epi } f$, $\alpha = \frac{\varepsilon}{\lambda}$, $(x_0, \alpha_0) = (x_0, f(x_0))$. According to Lemma 6.1, there exists a minimal element $(x_\lambda, \alpha) \subset S$ with the property

$$(x_\lambda, \alpha) > (x_0, f(x_0)) . \tag{6.2}$$

Since $(x_\lambda, \alpha) \in S$ one has

$$(x_\lambda, f(x_\lambda)) > (x_\lambda, \alpha) ,$$

which implies $\alpha = f(x_\lambda)$ by maximality.

It follows from (6.2) that

$$f(x_\lambda) - f(x_0) + \alpha d(x_\lambda, x_0) \leq 0 \,, \tag{6.3}$$

which implies i) of the theorem. The maximality of $(x_\lambda, f(x_\lambda))$ in S shows that for $x \in X$, $x \neq x_\lambda$, the relation

$$(x, f(x)) > (x_\lambda, f(x_\lambda)) \,,$$

is impossible. This implies iii) of the theorem. Finally, since

$$f(x_0) \leq \inf f + \varepsilon,$$

one has

$$f(x_\lambda) \geq f(x_0) - \varepsilon \,.$$

Combine the latter inequality and (6.3) to see ii). ∎

2.7 Problems

1. Prove that if f is continuously Gateaux differentiable at x, then it is strictly differentiable at this point, and hence Lipschitz near x.

2. Calculate $f^\circ(x;\cdot)$ of the function

$$f(x_1,\ldots,x_n) = \max_{i=1,\ldots,n} x_i$$

3. Calculate $\partial f(0)$ of

$$f(x) = x^2 \sin(1/x) ,$$

to show that $\partial f(x) \neq Df(x)$.

4. Let $\xi \in L^\infty[0,1]$. Show that for

$$f(x) = \int_0^x \xi(t)dt ,$$

$\partial f(x) = [\xi^-(x), \xi^+(x)]$, where ξ^- and ξ^+ are the essential supremum and infimum of ξ at x .

5. Let f be Lipschitz on an open set. Then f is differentiable almost everywhere on that set. Let Ω_f be the set where ∇f does not exist. Then for each set S with Lebesgue measure 0, one has

$$\partial f(x) = \text{conv}\{\lim \nabla f(x_i): x_i \to x , x_i \notin S \cup \Omega_f\} .$$

Prove this fact.

6. Let f_1, f_2 be Lipschitz near x. Show that $f_1 f_2$ is Lipschitz near x $(x \in R^1)$ and

$$\partial(f_1 f_2)(x) \subseteq f_2(x)\partial f_1(x) + f_1(x)\partial f_2(x) .$$

7. Let f_1, f_2 be Lipschitz near $x \in R^1$, $f_2(x) \neq 0$. Show that f_1/f_2 is Lipschitz near x and

$$\partial(f_1/f_2)(x) \subseteq \frac{f_2(x)\partial f_1(x) - f_1(x)\partial f_2(x)}{(f_2(x))^2} .$$

8. Prove that if C is convex, then $N(x|C)$ coincides with the normal cone in the sense of convex analysis.

9. Prove that $0 \in \partial f(x) + N(x|C)$ if f is Lipschitz near x and attains a minimum over C at x.

10. Let $f: X \to X$ be a directional contraction, where Y is a complete metric space. This means that f is continuous and there exists $\varepsilon \in (0,1)$ such that for every $x \in X$ with $f(x) \neq x$, there is $y \in (x, f(x))$ with $d(f(x), f(y)) \leq \varepsilon d(x,y)$. (it is supposed that

$$d(x, f(x)) = d(x,y) + d(y, f(x)) .$$

Prove that f admits a fixed point, i.e. there is $x_* \in X$ with $f(x_*) = x_*$.

CHAPTER 3

OPTIMALITY CONDITIONS

3.1 Existence of Extrema

Let X be a nonempty set of a topological space, f a map from X to R.

Definition 1.1

A point $x^\circ \in X$ is called a minimum point (resp., a strict minimum point) of f on X if the inequality

$$f(x) \geq f(x_0) \quad (\text{resp., } f(x) > f(x_0)) \tag{1.1}$$

holds for all $x \in X$, $x \neq x_0$. It is called a local minimum point of f if there exists a neighborhood U of x_0 in X such that (1.1) holds for all $x \in U$.

Maximum and local maximum points are defined in a similar way. We shall also use the notion of infimum of f on X. It is the greatest number (denoted inf f) such that

$$f(x) \geq \inf f \text{ , for all } x \in X \text{ .}$$

the notion sup f is defined similary.

Remark: A number $t \in R$ is the infimum of f if and only if $f(x) \geq t$ for all $x \in X$ and there exists a sequence $\{x_n\}$ in X with $\lim f(x_n) = t$. A point $x_0 \in X$ is minimum if and only if $f(x_0) = \inf f$.

Lemma 1.2

A minimum point of f on X exists if and only if the set $f(X)_+ = \{t \in R : t \geq f(x) \text{ , some } x \in R\}$ is closed and has a finite lower bound.

Proof. If x_0 is a minimum point of f on X , then $f(x_0) = \inf f$ and $f(X)_+ = \{t : t \geq f(x_0)\}$, hence it is closed and has a lower bound, for instance $f(x_0)$.

Conversely, if $f(X)_+$ has a lower bound, then its infimum say t_0 is finite. Moreover, since the set is closed, this infimum belongs to it. In other words, $t_0 \geq f(x_0)$, some $x_0 \in X$. But $f(x_0) \in f(X)_+$, one obtains actually $t_0 = f(x_0)$ which shows that x_0 is a minimum point of f on X . ■

Theorem 1.3

Given f on a nonempty set X . It attains its minimum on X in the following cases

i) X is compact, f is lower semicontinuous.

ii) There is $t \in R$ such that $\mathrm{lev}_f(t)$ is compact nonempty and f is lower semicontinuous on this set.

Proof. Suppose that i) is satisfied. In view of Lemma 1.3 it suffices to show that $f(X)_+$ is closed and it has a lower bound. In fact, if it has no lower bounds, there exists a sequence $\{x_n\}$ in X such that $\lim f(x_n)=-\infty$. Since X is compact, we may assume that $\lim x_n=x_0 \in X$.

The value $f(x_0)$ is finite, neverthless, $\lim f(x_n) = -\infty$ which shows that f cannot be lower semicontinuous at x_0. Thus, $f(X)_+$ is bounded from below. Let t be the infimum of this set. Then by definition, t is also the infimum of f. Hence there is a sequence $\{x_n\}$ in X such that $\lim f(x_n) = t$. Again, since X is compact one can assume that $\lim x_n = x_0 \in X$. Due to the lower semicontinuity, $t = \lim f(x_n) \geq f(x_0)$. Actually, we have $t = f(x_0)$ and x_0 is a minimum point of f on X.

For the second case it suffices to observe that if x_0 is a minimum point of f on the set $\mathrm{lev}_f(t)$, then it is also a minimum point of f on the whole set X. ■

3.2 Optimization Problems

Let X and f be as in the preceding section. The minimization problem corresponding to X and f can be given in the form:

(P) $$\min f(x)$$
$$\text{s.t. } x \in X .$$

Furthermore, let $h_1,\ldots,h_k$, $g_1,\ldots,g_m$ be functions from X to R. The problem with constraints is given in the form:

(CP) $$\min f(x)$$
$$\text{s.t. } x \in X$$
$$g_i(x) \leq 0 , \quad i=1,\ldots,m \qquad (2.1)$$
$$h_j(x) = 0 , \quad j=1,\ldots,k \qquad (2.2)$$

Constraints (2.1) are inequality constraints, and constraints (2.2) are equality constraints. It is useful to observe that inequality constraints can be transformed to equality constraints and vice versa. For instance, given (2.1) , introduce slack variables $y_1,\ldots, y_m \in R$ and express (2.1) as

$$g_i(x) + y_i^2 = 0 , \quad i=1,\ldots,m.$$

Conversely, given (2.2), one can write it as

$$h_j(x) \leq 0 ,$$
$$-h_j(x) \leq 0 , \quad j=1,\ldots,k.$$

One calls (P) an unconstrained problem and (CP) a constrained problem. Sometimes it is supposed that X is the whole space when

examining unconstrained problems. The set of feasible solutions of (CP) is defined as

$$X_0=\{x \in X : g_i(x)=0 , h_j(x)=0 , i=1,\dots,m, j=1,\dots,k\}.$$

Definition 2.1

A point $x_0 \in X$ is said to be an optimal solution of (CP) if it is a minimum point of f on the set of feasible solutions X_0 ; in this case $f(x_0)$ in the optimal value of (CP).

Theorem 2.2

Suppose that X is compact, f , $g_1,\dots,g_m$ are lower semicontinuous, while $h_1,\dots,h_k$ are continuous on X . Then (CP) has optimal solutions whenever the feasible solution set is nonvoid.

Proof. It follows from the lower semicontinuity of g_i and continuity of h_j that X_0 is closed, hence compact. Now the result is deduced from Theorem 1.3. ■

3.3 General Optimality Conditions

Let us consider a constrained minimization problem:

(P) $$\min f(x)$$

$$\text{s.t. } x \in X ,$$

$$g_i(x) \leq 0 , \quad i=1,\ldots,m$$

$$h_j(x) = 0 , \quad j=1,\ldots,k,$$

where f, g_i, h_j are functions from a Banach space E to R, $X \subseteq E$. We shall assume that X is closed, f, g_i, h_j are Lipschitz near any point of X with a common Lipschitz constant ℓ_0. We shall write

$$g = (g_1,\ldots,g_m) , \quad h= (h_1,\ldots, h_k) .$$

Lemma 3.1

Suppose that f attains a minimum over $C \subseteq X$ at $x \in C$. Then for every $\ell \geq \ell_0$, the function

$$f(y) + \ell d_C(y)$$

attains a minimum over X at x. Moreover, if $\ell > \ell_0$ and C is closed, any other point minimizing the above function over X must also lie in C.

Proof. Suppose to the contrary that there exists $y \in C$ and $\xi > 0$ such that

$$f(y) + \ell d_C(y) < f(x) - \ell\varepsilon .$$

Let c be a point in C such that

$$\|y - c\| \leq d_C(y) + \varepsilon .$$

Then one has

$$\begin{aligned} f(c) &\leq f(y) + \ell \, \|y - c\| \\ &\leq f(y) + \ell(d_C(y) + \varepsilon) \\ &< f(x) , \end{aligned}$$

which contradicts the fact that x minimizes f over X.

Now, let $\ell > \ell_0$, and let y minimize $f(\cdot) + \ell d_C(\cdot)$ over X. Then

$$\begin{aligned} f(y) + \ell d_C(y) &= f(x) \\ &\leq f(y) + (\ell + \ell_0) d_C(y)/2 . \end{aligned}$$

Hence $(\ell - \ell_0) d_C(y) \leq 0$, and $d_C(y) = 0$. ∎

Lemma 3.2

Suppose that f attains a local minimum at x. Then $0 \in \partial f(x)$.

Proof. It is clear that $f^\circ(x;v) \geq 0$. By Proposition 3.2, $0 \in \partial f(x)$. ■

Definition 3.3

The Lagrangian of (P) is the function

$$L(x,\lambda,\xi,\eta,\ell): E \times R \times R^m \times R^k \times R \to R$$

$$L(x,\lambda,\xi,\eta,\ell) = \lambda f(x) + \xi g(x) + \eta h(x) + \ell|(\lambda,\xi,\eta)|\, d_X(x) .$$

Theorem 3.4

Let x be an optimal solution of (P) . Then for every $\ell > \ell_0$, there exists $\lambda \geq 0$, $\xi \geq 0$ and η not all zero such that

$$0 \in \partial_x L(x,\lambda,\xi,\eta,\ell)$$

$$\xi g(x) = 0 .$$

Proof. Let us consider the following parameter set:

$$T = \{t = (\lambda,\xi,\eta): \lambda \geq 0 , \xi \geq 0 , |(\lambda,\xi,\eta)| = 1\} .$$

This set is compact. For every $\varepsilon > 0$, consider the map $F: E \to R$ defined by

$$F(y) = \max\{ t(f(y) - f(x) + \varepsilon, g(y) , h(y)) : t \in T \}.$$

It is obvious that $F(x) = \varepsilon$. Moreover,

$$F(y) > 0 \text{ , for all } y \in X \text{ .} \tag{3.1}$$

In fact, if $F(y) \leq 0$, then y is a feasible solution and $f(y) \leq f(x) - \varepsilon$, which is a contradiction. In this way,

$$F(x) \leq \inf F + \varepsilon \text{ .}$$

By Theorem 6.2 (Chapter 2), there exists $u \in (x + \sqrt{\varepsilon}\, B(0,1)) \cap X$ such that for each $y \in X$ one has

$$F(y) + \sqrt{\varepsilon}\ \|y - u\| \leq F(u) \text{ .}$$

It is easy to see that given $\ell > \ell_0$, if ε is small enough, ℓ is a Lipschitz constant of the function $F(y) + \sqrt{\varepsilon}\ \|y - u\|$ near u . In view of Lemma 3.1, u is a local minimum of the function

$$y \mapsto F(y) + \sqrt{\varepsilon}\ \|y - u\| + \ell d_X(y) =$$

$$= \max_{(\lambda,\xi,\eta)\in T} \{L(y,\lambda,\xi,\eta,\ell) - \lambda f(x) + \varepsilon\lambda\} + \sqrt{\varepsilon}\ \|y - u\| \text{ .}$$

By Lemma 3.2, one has

$$0 \in \partial G(u) + \sqrt{\varepsilon}\ B_*(0,1) \text{ ,} \tag{3.2}$$

where $G(y) = \max_{T}\{L(y,\lambda,\xi,\eta,\ell) - \lambda f(x) + \varepsilon\lambda\}$

$$B_*(0,1) = \{\gamma \in E' : \|\gamma\| \leq 1\} .$$

This is because:

$$\partial\|y - u\| = B_*(0,1) \quad \text{at} \quad y = u .$$

Now we want to calculate $\partial G(u)$ by using Theorem 4.5 of Chapter 2 . We show first that the map

$$(t,y) \mapsto \partial_x L(y,t,\ell) \tag{3.3}$$

is closed in the sense that

$$\bar{\partial}_x L(y,t,\ell) = \partial_x L(y,t,\ell) ,$$

where

$$\bar{\partial}_x L(y,t,\ell) = w^*\text{-cl conv}\{\gamma \in E' : \text{there exist } \gamma_i \in \partial_x L(y_i,t_i,\ell)$$
$$\text{with } y_i \to y,\ t_i \to t,\ \gamma_i \overset{w^*}{\to} \gamma\} .$$

In fact, for any $t_1, t_2 \in T$, the function

$$y \to L(y,t_1,\ell) - L(y,t_2,\ell) = (t_1 - t_2)\ (f,g,h)(y)$$

is Lipschitz of rank $\ell|t_1 - t_2|$ near x ; thus, by Propositions 3.2 and 4.2 of Chapter 2,

$$\partial_x L(y,t_1,\ell) \subseteq \partial_x L(y,t_2,\ell) + \ell|t_1 - t_2|\, B_*(0,1) .$$

This and the closure property of the generalized gradient (Proposition 3.4, Chapter 2) imply that the map (3.3) is closed.

Moreover, since $F(u) > 0$ (because $u \in X$ and (3.1), there is a unique $t_u \in T$ at which the maximum defining F , hence G is attained. In fact,

$$\max_{t\in T} ta > 0 ,$$

if $t_1 a = t_2 a = \max_{t\in T} ta$, $t_1, t_2 \in T$ and $t_1 \neq t_2$, then taking $t_0 = (t_1 + t_2)/2$ we have that $\|t_0\| < 1$. Take $t_* = t_0/\|t_0\|$ to see that $t_* \in T$ and $t_* a = t_0 a/\|t_0\| > t_1 a$, which is a contradiction. Now applying Theorem 4.5 to (3.2) one has

$$0 \in \partial_x L(u,t_u,\ell) + \sqrt{\epsilon}\, B_*(0,1) . \tag{3.4}$$

It is clear that if $g_i(u) < 0$, then the corresponding ξ_i in t_u equals zero. (This is because

$$t_u a = \max_{t\in T} ta = \max_{t\in \operatorname{conv} T} ta ,$$

and if $\xi_i > 0$, taking $\xi_i' \to 0$ one arrives at a contradiction

$$\max_{t \in \operatorname{conv} T} ta > \max_{t \in T} ta = t_u a \ ) \ .$$

Now, with $\varepsilon_i \to 0$ one sees that $u_i \to x$ and t_{u_i} converges to some $t \in T$. This and relations (3.3), (3.4) imply

$$0 \in \partial_x L(x,t,\ell) \ .$$

The proof is complete. ■

3.4 Optimality without Constraints

Let us consider the problem

$$\min f(x)$$

$$\text{s.t. } x \in X \ ,$$

where X is a nonempty set of a Banach space E, $f: E \to R$. Suppose that the first derivative $Df(x)$ and the second derivative $D^2 f(x)$ exist and are uniformly continuous on a neighborhood of $x_0 \in X$.

Theorem 4.1 (First-Order Condition)

If x_0 is a local minimum point of f on X, then

$$Df(x_0)(v) \geq 0 \ , \text{ for all } v \in K(X;x_0) \ .$$

Conversely, if at $x_0 \in X$ where X is in a finite dimensional subspace of E,

$$Df(x_0)(v) > 0 \text{ , for all } v \in K(X;x_0) \text{ ,}$$

then x_0 is a strict local minimum point of f on X.

Proof. Let $v \in K(X;x_0)$. Then exist $t_i \in R$, $v_i \in E$ with $v_i \to v$, $t_i \downarrow 0$ such that

$$x_i = x_0 + t_i v_i \in X \text{ .}$$

By the mean-value theorem,

$$f(x_i) - f(x_0) = Df(y_i)(t_i v_i) \text{ , for some } y_i \in [x_0, x_i] \text{ .}$$

Since Df is uniformly continuous,

$$\lim Df(y_i)(v_i) = Df(x_0)(v) \text{ .}$$

Now, since x_0 is local minimum one has

$$Df(x_0)(v) = \lim Df(y_i)(v_i) = \lim \frac{f(x_i) - f(x_0)}{t_i} \geq 0 \text{ .}$$

Conversely, if x_0 is not a strict local minimum point, there exists $x_i \in X$, $x_i \neq x_0$, $\lim x_i = x_0$ such that

$$f(x_i) \leq f(x_0) .$$

By the finite dimension assumption, one may assume that $v_i = (x_i - x_0)/\|x_i - x_0\|$ converges to a vector $v \in E$. Then $v \in K(X;x_0)$. On the other hand

$$Df(x_0)(v) = \lim Df(y_i)(v_i) = \lim \frac{f(x_i) - f(x_0)}{\|x_i - x_0\|} \geq 0 ,$$

which contradicts the assumption (the points $y_i \in (x_0, x_i)$ exist by the mean-value theorem).■

Theorem 4.2 (Second-Order Condition)

Suppose that x_0 is a local minimum point of f on X where $Df(x_0) = 0$. Then for each $v \in K(X;x_0)$ one has

$$D^2f(x_0)(v,v) \geq 0 .$$

Conversely, if X is contained in a finite dimensional subspace of E and at a point $x_0 \in X$, $Df(x_0) = 0$ and for every $v \in K(X;x_0)$, $v \neq 0$ one has

$$D^2f(x_0)(v,v) > 0 ,$$

then x_0 is a strict local minimum point of f on X.

Proof. Let $v \in K(X;x_0)$. There exist $t_i \to 0$, $v_i \to v$ such that $x_i = x + t_i v_i \in X$. Since $D^2 f(x)$ is uniformly continuous, one has

$$f(x_i) - f(x_0) = Df(x_0)(x_i - x_0) + \frac{1}{2} D^2 f(x_0)(x_i - x_0, x_i - x_0) + r(x_i - x_0),$$

where $\|r(x_i - x_0)\| / \|x_i - x_0\|^2 \to 0$ as $x_i \to x_0$.

In other words,

$$D^2 f(x_0)(v,v) = 2 \lim \frac{f(x_i) - f(x_0) - Df(x_0)(x_i - x_0)}{t_i^2}. \qquad (4.1)$$

Since $Df(x_0) = 0$ and $f(x_i) \geq f(x_0)$ whenever i is large enough, (4.1) implies

$$D^2 f(x_0)(v,v) \geq 0.$$

For the converse part, suppose to the contrary that there are $x_i \in X$, $x_i \neq x_0$, $\lim x_i = x_0$ such that

$$f(x_i) \leq f(x_0).$$

We may assume that $v_i = (x_i - x_0)/\|x_i - x_0\|$ converges to some $v \in E$. Then $v \in K(X;x_0)$. Moreover, by (4.1),

$$D^2 f(x_0)(v,v) \leq 0 ,$$

which is a contradiction to the assumption. ■

3.5 Optimality with Constraints

Let us consider the following constrained problem:

(CP)
$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0 , \ i = 1,\ldots,m \\ & h_j(x) = 0 , \ j = 1,\ldots,k . \end{aligned}$$

The set of feasible solutions is denoted by $X \subseteq E$. Let $x_0 \in X$. We say that inequality constraint $g_i(x) \leq 0$ is active at x_0, in this case i is called an active index, if $g_i(x_0) = 0$. The set of active indexes is denoted by $I = \{i_1,\ldots,i_e\}$.

Denote

$$S(x_0) = \{v \in E: Dh_j(x_0)(v) = 0 , \ j = 1,\ldots,k$$
$$Dg_i(x_0)(v) \leq 0 , \ i \in I\} .$$

Lemma 5.1

For every $x_0 \in X$, one has $K(X;x_0) \subseteq S(x_0)$.

Proof. Let $v \in K(X,x_0)$. There are $x_n \in X$, $x_n \to x_0$, $t_n \downarrow 0$ such that

$$\lim \frac{x_n - x_0}{t_n} = v .$$

Since x_n are feasible solutions,

$$h_j(x_n) = h_j(x_0) = 0 , \quad j = 1,\ldots,k$$

$$g_i(x_n) \leq 0$$

$$g_i(x_0) = 0 , \quad i \in I .$$

In view of the mean-value theorem one has

$$Dh_j(x_0)(v) = \lim \frac{h_j(x_n) - h_j(x_0)}{t_n} = 0 , \quad j = 1,\ldots,k$$

$$Dg_i(x_0)(v) = \lim \frac{g_i(x_n) - g_i(x_0)}{t_n} \leq 0 , \quad i \in I . \blacksquare$$

Theorem 5.2 (Kuhn-Tucker Condition)

Suppose that x_0 is an optimal solution of (CP) and it is regular in the sense that $K(X,x_0) = S(x_0)$. Then there exists a nonzero vector $(\lambda,\mu) \in R^m \times R^k$ with $\lambda \geq 0$ called Langrange multipliers such that

$$Df(x_0) + \sum_{i=1}^{m} \lambda_i \, Dg_i(x_0) + \sum_{j=1}^{k} \mu_j \, Dh_j(x_0) = 0 \qquad (5.1)$$

$$\lambda_i g_i(x_0) = 0 \ , \ i=1,\ldots,m \ . \qquad (5.2)$$

Conversely, suppose that E is finite dimensional and a feasible solution x_0 satisfies (5.1), (5.2), and $Df(x_0)(v) \neq 0$ whenever $v \in S(x_0)$, $v \neq 0$. Then x_0 is a strict local minimum point of f on X.

Proof. According to Theorem 4.1,

$$Df(x_0)(v) \geq 0 \ , \text{ for all } \ v \in K(X, x_0) \ .$$

Since $K(X; x_0) = S(x_0)$, the above relation show that $-Df(x_0)$ belongs to the polar cone of the cone $S(x_0)$. The polar cone is the convex hull of the cone generated by vectors

$$Dg_i(x_0) \ , \ i \in I \ \text{ and } \ Dh_j(x_0) \ , \ -Dh_j(x_0) \ , \ j = 1,\ldots,k \ .$$

Hence there exists $(\lambda, \mu^1, \mu^2) \geq 0$ such that

$$-Df(x_0) = \sum_{i \in I} \lambda_i Dg_i(x_0) + \sum_{j=1}^{k} \mu_j^1 Dh_j(x_0) + \sum_{j=1}^{k} \mu_j^2 (-Dh_j(x_0)) \ ,$$

which implies (5.1) and (5.2) by taking $\lambda_i = 0$ if $i \notin I$ and $\mu_j = \mu_j^1 - \mu_j^2$.

Conversely, in view of Lemma 5.1, $K(X,x_0) \subseteq S(x_0)$, therefore one has by (5.1),

$$Df(x_0)(v)=(-\sum_{i=1}^{m} \lambda_i Dg_i(x_0)-\sum_{j-1}^{k} \mu_j Dh_j(x_0))(v) \geq 0 ,$$

for all $v \in K(X,x_0)$. This and the condition $Df(x_0)(v) \neq 0$ whenever $v \in S(x_0)$, $v \neq 0$, show that

$$Df(x_0)(v) > 0 \quad , \text{ for all } \quad v \in K(X,x_0) , \; v \neq 0 . \; \blacksquare$$

It remains to apply Theorem 4.1 to complete the proof.

Before presenting second order conditions let us make some notations. For a fixed vector $(\lambda,\mu) \in R^m \times R^k$, denote

$$L(x) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{k} \mu_j h_j(x)$$

$$S_0(x_0) = \Big\{ v \in E : \quad Dh_j(x_0) = 0 \quad \text{for } j = 1,\ldots,k$$
$$Dg_i(x_0) = 0 \quad \text{if } i \in I , \; \lambda_i > 0,$$
$$Dg_i(x_0) \leq 0 \quad \text{if } i \in I , \; \lambda_i = 0 \Big\}.$$

Theorem 5.3. (Second Order Condition)

Suppose that x_0 is a regular optimal solution of (CP). Then there exists a vector (λ,μ) as stated in Theorem 5.2 such that for each $v \in S_0(x_0)$,

$$D^2L(x_0)(v,v) \geq 0 .$$

Conversely, if E is finite dimensional and x_0 is a feasible solution where the following conditions hold:

i) there exists (λ,μ) satisfying (5.1) (5.2)

ii) $D^2 L(x_0)(v,v) > 0$ for $v \in S(x_0)$ with $Df(x_0)(v)=0$, $v \neq 0$,

then x_0 is a strict local minimum point of f over X .

Proof. For the first part, (λ,μ) exists according to Theorem 5.2. Moreover, if denote

$$X_0=\{x \in X : \; g_i(x) \leq 0 \quad \text{if} \quad \lambda_i = 0 ,$$
$$g_i(x) = 0 \quad \text{if} \quad \lambda_i > 0 \; \},$$

then x_0 minimizes L on X_0 . Observe that $DL(x_0) = 0$, hence in view of Theorem 4.2,

$$D^2L(x_0)(v,v) \geq 0 \quad \text{for} \quad v \in K(X_0,x_0) = S_0(x_0) .$$

As to the converse part, suppose to the contrary that there are $x_n \in X$, $x_n \to x_0$ with

$$f(x_n) \leq f(x_0) . \tag{5.3}$$

One can assume that

$$\lim \frac{x_n - x_0}{\|x_n - x_0\|} = v \in E .$$

It is evident that $v \in K(X, x_0)$ and by (5.1),

$$Df(x_0)(v) \geq 0 .$$

On the other hand, by the mean-value theorem and by (5.3)

$$Df(x_0)(v) \leq 0 .$$

Hence $Df(x_0)(v) = 0$. Again applying the mean-value theorem and remembering that $\lambda_i = 0$ if $i \notin I$ one has

$$D^2L(x_0)(v,v) \leq 0 ,$$

which contradicts ii). ■

3.6 Convex Problems

Let us consider a problem

(P) $$\min f(x)$$
$$\text{s.t. } g_i(x) \leq 0 , \ i = 1,\ldots,m ,$$

where $f, g_1, \ldots, g_m$ are convex functions from E to R.

We suppose also that there functions are differentiable.

Proposition 6.1

If x_0 is a local optimal solution of (P), then it is also a global optimal solution.

Proof. Suppose to the contrary that there is some feasible solution x with

$$f(x) < f(x_0) .$$

Then for every $t \in (0,1)$, $tx + (1-t)x_0$ is also feasible. Moreover, since f is convex, one has

$$f(tx + (1-t)x_0) < f(x_0) .$$

When t approaching 0, $tx + (1-t)x_0$ tends to x_0 and the above inequality contradicts the local optimality of x_0. ■

Theorem 6.2

If at a feasible solution x_0 of (P) there exists $\lambda \geq 0$ such that

$$Df(x_0) + \sum_{i=1}^{m} \lambda_i Dg_i(x_0) = 0 ,$$

$$\lambda_i g_i(x_0) = 0 , \quad i = 1, \ldots, m ,$$

then x_0 is an optimal solution of (P).

Proof. Suppose to the contrary that x_0 is not optimal, i.e. there is a feasible solution x with

$$f(x) < f(x_0) .$$

Consider the vector $v = x - x_0$. Since f is convex, one has

$$Df(x_0)(v) = \lim_{t \downarrow 0} \frac{f(x_0+tv) - f(x_0)}{t} . \tag{6.1}$$

Furthermore, by condition $\lambda_i g_i(x_0) = 0$, one has $\lambda_i = 0$ whenever i is an unactive index. Hence

$$\lambda_i Dg_i(x_0)(v) = \lim_{t \downarrow 0} \lambda_i \frac{g(x_0+tv) - g_i(x_0)}{t} \leq 0 , \tag{6.2}$$

for $i = 1,\ldots,m$. Combine (6.1) and (6.2) to see the contradiction

$$Df(x_0)(v) + \Sigma \lambda_i Dg_i(x_0)(v) < 0 .$$

The proof is complete. ■

3.7 Problems

1. Let X be the unit ball in a Banach space. Is it true that any continuous function attains its minimum on X?

2. Let X be a closed convex set in a separated topological vector space E. The recession cone of X is the cone

$$\text{Rec}(X)=\{v \in E: X + tv \subseteq X \text{ for all } t \geq 0\} .$$

Prove that if f is a convex continuous function on X with the property that $g(t) = f(x+tv)$ is an increasing function on $t \geq 0$ for any fixed $x \in X$, $v \in \text{Rec}(X)$ with $\lim_{t\to\infty} g(t) = \infty$, then f attains its infimum on X.

3. Prove that $K(X,x_0) = S(x_0)$ in the following cases

i) the constraints are linear

ii) $h_1,\ldots,h_k$ are linear, g_i, $i \in I$ are convex

iii) the constraints are convex and $\text{int } X \neq \phi$

iv) $Dh_1(x_0),\ldots,Dh_k(x_0)$, $Dg_i(x_0)$, $i \in I$ are linearly independent.

4. Give an example to show that for a convex problem in Section 6, it is not necessary for $\lambda \geq 0$ exist which yields

$$Df(x_0) + \sum_{i=1}^{m} \lambda_i Dg_i(x_0) = 0 ,$$

where x_0 is an optimal solution.

5. Prove that a point $x_0 \in \text{int } X$ is a minimum point of a convex function f if and only if

$$0 \in \partial f(x_0) .$$

CHAPTER 4

DUALITY THEORY

4.1 Duality via Conjugate Functions

Consider a minimization problem:

$$(P) \qquad \min \quad f(x)$$
$$\text{s.t.} \quad x \in X ,$$

where X is a separated topological vector space, f is a function from X to the extended real line $\bar{R} = R \cup \{\pm\infty\}$.

Let φ be a function from $X \times Y$ to $\bar{R}$, where Y is another topological vector space.

Suppose that $\varphi(x,0) = f(x)$. We consider the minimization problem

$$(P_y) \qquad \min \varphi(x,y)$$
$$\text{s.t.} \quad x \in X .$$

This problem is called a perturbed problem of (P) .

The conjugate function of φ is defined as

$$\varphi^*(\xi,\eta) = \sup\{\xi(x) + \eta(y) - (x,y): (x,y) \in X \times Y\} ,$$

for every $(\xi,\eta) \in X' \times Y'$.

One defines a maximization problem called the dual problem of (P) :

$$(P^*) \qquad \max \; -\varphi^*(0,\eta)$$
$$\text{s.t.} \quad \eta \in Y' .$$

Theorem 1.1 (Weak Duality)

For every $x \in X$, $\eta \in Y'$ one has

$$f(x) \geq -\varphi^*(0,\eta) .$$

Moreover, if there are $x_0 \in X$, $\eta_0 \in Y'$ such that

$$f(x_0) = -\varphi^*(0,\eta_0) ,$$

then x_0 is an optimal solution of (P) , η_0 is an optimal solution of (P*) and the optimal values of these problems are equal.

Proof. It follows from the definition of conjugate function that

$$\varphi^*(0,\eta) \geq \eta(y) - \varphi(x,y) \text{ , for } (x,y) \in X \times Y \text{ , } \eta \in Y' \text{ .}$$

In particular when $y = 0$ one has

$$f(x) = \varphi(x,0) \geq -\varphi^*(0,\eta) \text{ .}$$

In the case where x_0 and η_0 exist as stated in the theorem, it is immediate that they are optimal solutions of (P) and (P*) respectively. ■

Following the scheme described above one can constructs the dual (P**) for (P*) and so forth. It is obvious that

$$(P^*) = (P^{***}) \text{ .}$$

It can also be seen that if φ is a convex closed function which is not identically ∞ or $-\infty$, then

$$(P) = (P^{**}) \text{ .}$$

Set

$$h(y) = \inf_{x \in X} \varphi(x,y) \quad \text{(the marginal function) .}$$

Lemma 1.2

Assume that $\varphi(x,y)$ is a convex closed function not identical to $+\infty$ or $-\infty$. Then

i) $h(y)$ is a convex function on Y

ii) $h^*(\eta) = \varphi^*(0,\eta)$, for all $\eta \in Y'$

iii) $H^{**}(0) = \sup_{\eta \in Y'} - h^*(\eta)$.

Proof. For the convexity of h, let $y_1, y_2 \in Y$, $t \in [0,1]$. If $h(y_1)$ or $h(y_2) = +\infty$, the convexity is obvious.

Assume that both of them are smaller than $+\infty$. By definition of inf, for each $a > h(y_1)$, $b > h(y_2)$ one can find $x_1, x_2 \in X$ such that

$$h(y_1) \leq \varphi(x_1,y_1) \leq a$$

$$h(y_2) \leq \varphi(x_2,y_2) \leq b .$$

This and convexity of φ yield

$$\begin{aligned} h(ty_1 + y_2(1-t)) &= \inf_{x \in X} \varphi(x, ty_1 + (1-t)y_2) \\ &\leq \varphi(tx_1 + (1-t)x_2 , ty_1 + (1-t)y_2) \\ &\leq t\varphi(x_1,y_1) + (1-t)\varphi(x_2,y_2) \\ &\leq ta + (1-t)b . \end{aligned}$$

Let a run to $h(y_1)$, b to $h(y_2)$. We obtain

$$h(ty_1 + (1-t)y_2) \leq th(y_1) + (1-t)h(y_2) ,$$

which shows that h is convex.

For ii) let us calculate $h^*(\eta)$:

$$h^*(\eta) = \sup_{y \in Y} \{\eta(y) - h(y)\}$$

$$= \sup_y \{\eta(y) - \inf_x \varphi(x,y)\}$$

$$= \sup_y \sup_x \{\eta(y) - \varphi(x,y)\}$$

$$= \varphi^*(0,\eta) .$$

The last equality of the lemma is derived at once from ii). ■

Definition 1.3

(P) is to be normal if $h(0)$ is finite and h is lower semicontinuous at 0 .

Theorem 1.4

Assume that $\varphi(x,y)$ is convex closed not identical to $+\infty$ or $-\infty$. Then the following conditions are equivalent

i) (P) is normal

ii) (P*) is normal

iii) $\inf_{x \in X} f(x) = \sup_{\eta \in Y'} - h^*(\eta)$ and this value is finite.

Proof. We shall prove i) $\Longleftrightarrow$ iii). The equivalence between ii) and iii) follows directly from the fact that $(P^{**}) = (P)$.

Assuming (P) to be normal, we have

$$h^{**} \leq \text{cl } h \leq h \tag{1.1}$$

Since (P) is normal, $\text{cl } h(0) = h(0) \in R$. Since cl h is convex lower semicontinuous and it admits a finite value at 0, it is proper. Apply Theorem 4.4 of Chapter 1 to have the relation

$$(\text{cl } h)^{**} = \text{cl } h .$$

The inequalities of (1.1) yield

$$h^* = h^{***} \geq (\text{cl } h)^* \geq h^* .$$

Consequently, $h^* = (\text{cl } h)^*$ and $h^{**} = (\text{cl } h)^{**} = \text{cl } h$, whence $\text{cl } h(0) = h(0) = h^{**}(0)$. In view of Lemma 1.2 this implies iii).

Conversely, it follows from iii) that $h(0) = h^{**}(0) \in R$. We have then from (1.1) $h(0) = \text{cl } h(0)$, which shows that (P) is normal. ■

Definition 1.5

Problem (P) is said to be stable if $h(0)$ is finite and h is subdifferentiable at 0.

Lemma 1.6

The set of optimal solutions to (P*) is identical to $\partial h^{**}(0)$.

Proof. Let η_0 be an optimal solution of (P*). Then

$$-\varphi^*(0,\eta_0) \geq -\varphi^*(0,\eta), \text{ for all } \eta \in Y'.$$

This implies that

$$-h^*(\eta_0) = h^{**}(0),$$

which is equivalent to $\eta_0 \in \partial h^{**}(0)$. ■

Theorem 1.7

The following two conditions are equivalent to each other:

i) (P) is stable

ii) (P) is normal and (P*) has optimal solutions.

Proof. For the implication i) ⟹ ii), suppose that (P) is stable. Then $h(0)$ is finite and $\partial h(0) \neq \phi$. Hence $h(0) = h^{**}(0) \in R$. This means that (P) is normal and in addition

$$\partial h^{**}(0) = \partial h(0) \neq \phi \ .$$

In virtue of Lemma 1.6, (P*) has optimal solutions.

For the implication ii) ==> i) , if (P) is normal, $h(0) = h^{**}(0) \in R$, and if (P*) has optimal solutions, the set $\partial h^{**}(0) = \partial h(0)$ is nonempty. Hence (P) is stable. ■

4.2 Lagrangians and Saddlepoints

As in the previous section, $\varphi(x,y) : X \times Y \to \bar{R}$ is a perturbation of (P) .

Definition 2.1

The function $L: X \times Y' \to \bar{R}$ defined by

$$L(x,\eta) = \inf_{y \in Y} \{ \varphi(x,y) - \eta(y) \}$$

is called the Lagrangian (classical) of (P) relative to the given perturbations.

Proposition 2.2

$L(x,\eta)$ is a concave upper semicontinuous function in the variable η for every fixed x ; and it is convex in x for every fixed η whenever φ is a convex function.

Proof. It is clear that for every fixed $x \in X$, the function $g(y) = \varphi(x,y)$ has its conjugate

$$g^*(\eta) = -L(x,\eta) ,$$

which is convex lower semicontinuous. Hence $L(x,\eta)$ is concave upper semicontinuous in η.

Now, suppose that φ is convex. Let $x_1, x_2 \in X$, $t \in (0,1)$. The inequality

$$L(tx_1 + (1-t)x_2,\eta) \leq tL(x_1,\eta) + (1-t)L(x_2,\eta) \tag{2.1}$$

is obvious if $L(x_1,\eta)$ or $L(x_2,\eta) = +\infty$. Hence we may assume that both of them are not $+\infty$. Let $a > L(x_1,\eta)$, $b > L(x_2,\eta)$. There are $y_1, y_2 \in Y$ such that

$$L(x_1,\eta) \leq \varphi(x_1,y_1) - \eta(y_1) \leq a$$

$$L(x_2,\eta) \leq \varphi(x_2,y_2) - \eta(y_2) \leq b .$$

These inequalities and the convexity of φ show that

$$L(tx_1 + (1-t)x_2,\eta) \leq t[\varphi(x_1,y_1) - \eta(y_1)] + (1-t)[\varphi(x_2,y_2) - \eta(y_2)]$$

$$\leq ta + (1-t)b .$$

when a runs to $L(x_1,\eta)$ and b runs to $L(x_2,\eta)$, the above relation implies (2.1). ■

We shall express problems (P) and (P*) in terms of L. Without assuming anything about φ one has

$$\varphi^*(\xi,\eta) = \sup_{x,y}\{\xi(x) + \eta(y) - \varphi(x,y)\}$$

$$= \sup_x\{\xi(x) + \sup_y\{\eta(y) - \varphi(x,y)\}\}$$

$$= \sup_x\{\xi(x) - L(x,\eta)\}$$

whence the problem (P*) defined in the previous section can be written in the form

$$(P^*) \qquad \max_{\eta\in Y'} \inf_{x\in X} L(x,\eta) .$$

Similarly, if φ is assumed to be convex closed, then

$$\varphi(x,y) = g^{**}(y)$$

$$= \sup_\eta\{\eta(y) - g^*(\eta)\}$$

$$= \sup_\eta \eta(y) + L(x,\eta)\} ,$$

where $g(y) = \varphi(x,y)$ for a fixed $x \in X$.

Hence problem (P) can be written as

$$(P) \qquad \min_{x \in X} \sup_{\eta \in Y'} L(x,\eta) .$$

Remark 2.3

Given a function $L(a,b) : A \times B \to \bar{R}$. One can consider the minimax problem

$$\sup_b \inf_a L(a,b) = \inf_a \sup_b L(a,b) . \qquad (2.2)$$

In general it is true that

$$\sup_b \inf_a L(a,b) \leq \inf_a \sup_b L(a,b) .$$

Hence Theorem 1.1 can be derived from this trivial general relation. Theory of minimax problems studies conditions under which (2.2) is true.

Definition 2.4

A point $(x_0,\eta_0) \in X \times Y'$ is called a saddlepoint of L if

$$L(x_0,\eta) \leq L(x_0,\eta_0) \leq L(x,\eta_0) ,$$

for all $x \in X$, $\eta \in Y'$.

Theorem 2.5

Assume that φ is a convex closed function. Then the following two conditions are equivalent to each other:

i) (x_0,η_0) is a saddlepoint of L

ii) x_0 is an optimal solution of (P), η_0 is an optimal solution of (P*) and the optimal values of these problems are equal.

Proof. First we show the implication i) ⟹ ii). By definition one has

$$L(x_0,\eta_0) = \min_x L(x,\eta_0) = -\varphi^*(0,\eta_0)$$

$$L(x_0,\eta_0) = \max_\eta L(x_0,\eta) = \varphi(x_0,0) .$$

These equalities and Theorem 1.1 show ii).

Now suppose that ii) holds. Then

$$\varphi(x_0,0) = -\varphi^*(0,\eta_0) .$$

But one has by definition

$$\varphi(x_0,0) = \sup_h L(x_0,\eta)$$

$$-\varphi^*(0,\eta_0) = \inf_x L(x,\eta_0) .$$

Hence $\sup_{\eta} L(x_0,\eta) = L(x_0,\eta_0) = \inf_{x} L(x,\eta_0)$ which means that (x_0,y_0) is a saddlepoint.■

Proposition 2.6

Assume that φ is a convex closed function, and that (P) is stable. Then x_0 is an optimal solution of (P) if and only if there exists $\eta_0 \in Y'$ such that (x_0,η_0) is a saddlepoint of L.

Proof. The if part follows from Theorem 2.5. Now, if x_0 is an optimal solution of (P) and (P*) is stable, then (P*) has at least one optimal solution, say η_0 (Theorem 1.7) and the optimal values of these problems are equal. Apply Theorem 2.5 to see that (x_0,η_0) is a saddlepoint of L.■

Now we turn to the conditions for the existence of saddlepoints. Let us consider the general minimax problem as pointed in Remark 2.3:

$$L(a,b) : A \times B \to R ,$$

where it is assumed that

i) A, B are convex closed nonempty subsets in reflexive Banach spaces

ii) $L(a,\cdot)$ is concave upper semicontinuous in b,
$L(\cdot,a)$ is convex lower semicontinuous in a.

Proposition 2.7

Assume, in addition to the conditions i), ii) above, that for every fixed b, $L(\cdot,b)$ is Gateaux differentiable on A while for every fixed a, $L(a,\cdot)$ is Gateaux differentiable on B. Then $(a_0,b_0) \in A \times B$ is a saddlepoint of L if and only if

$$\frac{\partial L(a_0,b_0)}{\partial a}(a-a_0) \geq 0 \text{ , for all } a \in A \tag{2.3}$$

$$\frac{\partial L(a_0,b_0)}{\partial b}(b-b_0) \leq 0 \text{ , for all } b \in B \tag{2.4}$$

Proof. Suppose first that (a_0,b_0) is a saddlepoint. Then

$$\frac{\partial L(a_0,b_0)}{\partial b}(a-a_0) = \lim_{t \downarrow 0} \frac{L(a_0+t(a-a_0),b_0) - L(a_0,b_0)}{t} .$$

But $L(a_0 + t(a-a_0),b_0) \geq L(a_0,b_0)$ by definition of saddlepoints. Hence (2.3) follows. Relation (2.4) is proven in a similar way.

Conversely, if (2.3) holds, then by the convexity assumption, for each $a \in A$ one has (Proposition 5.1 of Chapter 1)

$$\begin{aligned} L(a,b_0) - L(a_0,b_0) &= L(a_0 + (a-a_0),b_0) - L(a_0,b_0) \\ &\geq \frac{L(a_0+t(a-a_0),b_0) - L(a_0,b_0)}{t} \end{aligned}$$

whenever $t \in (0,1)$. It is know that

$$\frac{\partial L(a_0,b_0)}{\partial a}(a-a_0) = \inf_{t>0} \frac{L(a_0+t(a-a_0),b_0) - L(a_0,b_0)}{t} .$$

Hence (2.3) implies $L(a,b_0) - L(a_0,b_0) \geq 0$.

Similarly, $L(a_0,b_0) - L(a_0,b) \geq 0$, which completes the proof.■

Theorem 2.8

Assume in addition to conditions i) ii), that A and B are bounded. Then L possesses at least one saddlepoint (a_0,b_0) on $A \times B$ and

$$L(a_0,b_0) = \min_a \max_b L(a,b) = \max_b \min_a L(a,b) .$$

Proof. Since the spaces are reflexive, A and B are compact in the weak topologies. Moreover, since $L(a,b)$ is convex in a , concave in b , the properties of semicontinuities are true in weak topologies.

We first prove the case where $L(\cdot,b)$ is strictly convex for each fixed b . It follows that $L(\cdot,b)$ attains its minimum over A at a unique point, say $e(b) \in A$:

$$f(b) = \min_a L(a,b) = L(e(b),b) .$$

The function $f(b)$ is concave and weakly upper semicontinuous on B. Therefore it attains its maximum over B at a point b_0:

$$f(b_0) = \max_{b \in B} f(b) = \max_{b \in B} \min_{a \in A} L(a,b) ,$$

and obviously,

$$f(b_0) \leq L(a,b_0) , \text{ for all } a \in A . \tag{2.5}$$

By the concavity assumption, for $a \in A$, $b \in B$, $\lambda \in (0,1)$ one has

$$L(a,(1-\lambda)b_0 + \lambda b) \geq (1-\lambda)L(a,b_0) + \lambda L(a,b) .$$

In particular when $a = e_\lambda = e((1-\lambda)b_0 + \lambda b)$ one obtains

$$\begin{aligned} f(b_0) &\geq f((1-\lambda)b_0 + \lambda b) \\ &= L(e_\lambda, (1-\lambda)b_0 + \lambda b) \\ &\geq (1-\lambda)L(e_\lambda,b_0) + \lambda L(e_\lambda,b) , \\ &\geq (1-\lambda)f(b_0) + \lambda L(e_\lambda,b) , \end{aligned}$$

whence

$$f(b_0) \geq L(e_\lambda,b) , \text{ for all } b \in B . \tag{2.6}$$

We may assume that e_λ converges weakly to $a_0 \in A$ when λ runs to 0. We prove that a_0 is a minimum point of $L(a,b_0)$, i.e. $a_0 = e(b_0)$. In fact,

$$L(e_\lambda,(1-\lambda)b_0 + \lambda b) \leq L(a,(1-\lambda)b_0 + \lambda b) \text{ , for } a \in A ,$$

and by the concavity one has for every $a \in A$:

$$(1-\lambda)L(e_\lambda,b_0) + \lambda L(e_\lambda,b) \leq L(a,(1-\lambda)b_0 + \lambda b) .$$

Since $L(e_\lambda,b)$ is bounded below by $f(b)$ and the function is lower semicontinuous in the first variable, one obtains

$$\begin{aligned} L(a_0,b_0) &\leq \liminf_{\lambda \downarrow 0} L(e_\lambda,b_0) \\ &= \liminf_{\lambda \downarrow 0} (1-\lambda)L(e_\lambda,b_0) + \lambda L(e_\lambda,b) \\ &\leq \limsup_{\lambda} L(a,(1-\lambda)b_0 + \lambda b) \\ &\leq L(a,b_0) \end{aligned} \tag{2.7}$$

It follows from (2.6) that

$$f(b_0) \geq \liminf_{\lambda} L(e_\lambda,b) \geq L(a_0,b)$$

for every $b \in B$. By definition of $f(b_0)$ and by the latter inequality together with (2.5) one can conclude

$$f(b_0) = \max_b \min_a L(a,b) = L(a_0,b_0) .$$

This fact and the general inequality

$$\max_b \min_a L(a,b) \leq \min_a \max_b L(a,b)$$

show that

$$\begin{aligned} L(a_0,b_0) &= \max_b \min_a L(a,b) \\ &= \min_a L(a,b_0) \\ &= \min_a \max_b L(a,b) . \end{aligned}$$

In the second equality we have used relation (2.7).

Now for the case where L is not strict convex, one considers

$$L_\varepsilon(a,b) = L(a,b) + \varepsilon\|a\| , \quad \varepsilon > 0 .$$

Then L_ε is strict convex in a and it satisfies all requirements, which yield the existence of a saddlepoint $(a_\varepsilon,b_\varepsilon)$:

$$L(a_\varepsilon,b) + \varepsilon \|a_\varepsilon\| \leq L(a_\varepsilon,b_\varepsilon) + \varepsilon \|a_\varepsilon\| \leq L(a,b_\varepsilon) + \varepsilon \|a_\varepsilon\| , \quad (2.8)$$

for all $a \in A$, $b \in B$.

By the weak compactness one can assume that

$$a_\varepsilon \overset{w}{\to} a_0 , \; b_\varepsilon \overset{w}{\to} b_0 \quad \text{when} \quad \varepsilon \to 0 .$$

Remembering that L is lower semicontinuous in the first variable and upper semicontinuous in the second one, one deduces from (2.8) that

$$L(x_0,b) \leq L(a_0,b_0) \leq L(a,b_0) , \text{ for } a \in A , b \in B .$$

This means that (a_0,b_0) is a saddlepoint.■

Theorem 2.9

Assume in addition to i), ii) that there exist $a_* \in A$ and $b_* \in B$ such that

iii) $\lim_{\|a\|\to\infty} L(a,b_*) = \infty$

iv) $\lim_{\|b\|\to\infty} L(a_*,b) = -\infty$

Then L possesses at least a saddlepoint and

$$L(a_0,b_0) = \min_a \sup_b L(a,b) = \max_b \inf_a L(a,b) .$$

Proof. For a fixed positive μ, let

$$A_\mu = \{a \in A: \quad \|a\| \leq \mu\}.$$

$$B_\mu = \{b \in B: \quad \|b\| \leq \mu\}.$$

One can choose μ large so that $x_* \subset A_\mu$, $b_* \subset B_\mu$. These sets are obviously convex closed bounded. In view of Theorem 2.8, there exists (a_μ, b_μ) a saddlepoint of L on $A_\mu \times B_\mu$, which means that

$$L(a_\mu, b) \leq L(a_\mu, b_\mu) \leq L(a, b_\mu), \text{ for } (a,b) \in A_\mu \times B_\mu \qquad (2.9)$$

Since $L(\cdot, b_*)$ is convex lower semicontinuous, condition iii) implies that it is bounded below:

$$-\infty < \alpha \leq L(a, b_*), \text{ for all } a \in A.$$

Similarly,

$$+\infty > \beta \geq L(a_*, b), \text{ for all } b \in B.$$

In particular, using (2.9) and two inequalities above one has

$$L(a_\mu, b_*) \leq L(a_*, b_\mu) \leq \beta$$

$$L(a_*, b_\mu) \geq L(a_\mu, b_*) \geq \alpha.$$

for all μ . It follows from conditions iii) and iv) that $\{a_\mu\}$, $\{b_\mu\}$ are bounded. One can assume that they converge weakly to a_0 and b_0 respectively. Relation (2.9) yields

$$L(a_0,b) \leq L(a,b_0) \text{ , for all } a \in A \text{ , } b \in B \text{ .}$$

Hence (a_0,b_0) is a saddlepoint of L on $A \times B$. ■

4.3 Special Cases

Case 1. Let us consider the problem

$$(P) \qquad \min \quad f(x,Ax)$$
$$\text{s.t. } x \in X \text{ ,}$$

where A is a linear continuous operator from X to Y and $f: X \times Y \to R$.

We consider the perturbations

$$\varphi(x,y) = f(x, Ax-y) \text{ .}$$

It is easy to see that the dual problem is of the form

$$(P^*) \qquad \max \quad -\varphi^*(0,\eta) = -f^*(A^*\eta,-\eta)$$
$$\text{s.t. } \eta \in Y' \text{ .}$$

It is also clear that φ is convex if so is f, and φ is convex closed not identical to $+\infty$ or $-\infty$ if so is f.

Theorem 3.1

Assume that f is convex, $\inf_x f(x,Ax)$ is finite and that there exists $x_0 \in X$ with $f(x_0,Ax_0) < +\infty$ such that the function $f(x_0,y)$ is continuous in y at Ax_0.

Then (P) is stable:

$$\inf_x f(x,Ax) = \sup_\eta - f^*(A^*\eta, -\eta)$$

and (P*) has at least one optimal solution η.

Proof. It is evident that the function

$$h(y) = \inf_x f(x,Ax-y)$$

is convex with $h(0)$ finite. Moreover, $f(x_0,y)$ is continuous at Ax_0, hence it is bounded on some small neighborhood of Ax_0. Consequently, $h(y)$ is bounded on a neighborhood of zero, which implies that it is continuous at zero. By Proposition 5.7 of Chapter 1, $\partial h(0)$ is nonempty. This means that (P*) is stable. Theorem 1.7 shows that (P*) has an optimal solution, say η. With this η, the equality stated in the theorem holds.■

Case 2. Consider (P) with $f(x,Ax) = f(x) + g(Ax)$. It is easy to verify that

$$f^*(\xi,\eta) = f^*(\xi) + g^*(\eta) .$$

Hence the dual problem can be written as

$$(P^*) \qquad \max - f^*(A^*\eta) - g^*(-\eta)$$
$$\text{s.t.} \quad \eta \in Y' .$$

Observe that if f and g are convex closed not identical to $+\infty$, or $-\infty$, then so is φ.

Corollary 3.2

Assume that f and g are convex, $\inf\{f(x) + g(Ax)\}$ is finite and that there exists a point $x_0 \in X$ with $f(x_0) < \infty$, $g(Ax_0) < \infty$ such that g is continuous at Ax_0. Then (P) is stable and (P*) has at least one optimal solution.

Proof. Invoke this to Theorem 3.1. ■

Corollary 3.3

The two following conditions are equivalent:

i) x_0 solves (P), η_0 solves (P*) and optimal values of these problems are equal

ii) $A^*\eta_0 \in \partial f(x_0)$, $-\eta_0 \in \partial g(Ax_0)$.

Proof. It follows from i) that

$$0 = f(x_0, Ax_0) + f^*(A^*\eta_0, -\eta_0)$$

$$= f(x_0) + f^*(A^*\eta_0) + g(Ax_0) + g^*(-\eta_0)$$

$$= \{f(x_0) + f^*(A^*\eta_0) - A^*\eta_0(x_0)\} + \{g(Ax_0) + g^*(-\eta_0) - \eta_0(Ax_0)\}.$$

Observe that each of the two terms of the last line is nonegative, hence it must be zero. In view of Proposition 5.5 of Chapter 1, one concludes

$$A^*\eta_0 \in \partial f(x_0) \quad \text{and} \quad -\eta_0 \in \partial g(Ax_0) .$$

Conversely, if ii) holds then

$$0 = f(x_0, Ax_0) + f^*(A^*\eta_0, -\eta_0) ,$$

which implies i) .■

Case 3. Let Y be ordered by a pointed convex cone C, i.e. $y_1 \geq y_2$ if and only if $y_1 - y_2 \in C$.

Denote

$$C^* = \{\eta \in Y' : \eta(y) \geq 0 \text{ for all } y \in C\}.$$

Suppose that X_0 is a nonempty convex closed subset of X and f is a convex lower semicontinuous function on X_0, g is a map from X_0 to Y which is convex with respect to the order in Y.

Let us consider the problem

$$(P) \qquad \min \ f(x)$$

$$\text{S.t} \ \ x \in X_0,$$

$$g(x) \leq 0.$$

Define a perturbation φ as follws

$$\varphi(x,y) = \begin{cases} f(x) & \text{if } x \in X_0,\ g(x) \leq y \\ +\infty & \text{otherwise.} \end{cases}$$

It is obvious that φ is a proper convex function. We compute $\varphi^*(0,\eta)$:

$$\begin{aligned} \varphi^*(0,\eta) &= \sup_{x,y} \{\eta(y) - \varphi(x,y)\} \\ &= \sup_{x\in X_0, y\in Y, g(x)\leq y} \{\eta(y) - \varphi(x,y)\} \\ &= \sup_{x\in X_0} \ \sup_{z\in Y, z\geq 0} \{\eta(g(x)) + \eta(z) - f(x)\} \\ &= \begin{cases} \inf_{x\in X_0} \{-\eta(g(x)) + f(x)\} & \text{if } \eta \leq 0 \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

The dual problem can be written as

$$(P^*) \qquad \sup_{\eta<0} \inf_{x\in A} \{-\eta(g(x) + f(x)\}$$

Proposition 3.4.

Assume that $\inf_{x\in X_0, g(x)\leq 0} f(x)$ is finite and there exists $x_0 \in X_0$ with $-g(x_0) < 0$. Then (P) is stable.

Proof. Under the above hypothesis, the function $y \to \varphi(x_0, y)$ is finite and continuous at 0. Hence $h(\cdot)$ is finite and continuous at 0. The argument of the proof of Theorem 3.1 is applicable. ■

4.4. Problems

1. Assume that in the first section, φ is convex, $\inf f(x)$ is finite and there is a point $x_0 \in X$ such that $y \to (x_0, y)$ is finite and continuous at $0 \in Y$. Prove that (P) is stable.

2. Is Theorem 2.8 true if L is continuous in both variables, quasiconvex in a, quasiconcave in b (recall that a function is quasiconvex if its level sets are convex, and it is quasiconcave if its minus is quasiconvex)?

3. Find the dual problem of a linear problem

$$\begin{aligned} &\min\ cx \\ &\text{s.t. } Ax \geq b \\ &\qquad x \geq 0\ , \end{aligned}$$

where A is a matrix, b and c are fixed vectors, $x \in R^n$. Study the normality and stability of it.

CHAPTER 5

UNCONSTRAINED OPTIMIZATION TECHNIQUES

5.1 Descent Algorithms and Convergence

Suppose that we have to solve an optimization problem, say

$$(P) \qquad \min f(x) \quad \text{s.t.} \quad x \in X ,$$

Where f is a function from a topological space E to R, X is a nonempty subset of E.

Definition 1.1

An algorithm on E is a set-valued map A from E to E.

Given an algorithm A on E and an initial point $x_0 \in X$, one can obtain a sequence of points through iteration by the rule

$$x_{k+1} \in A(x_k) . \tag{1.1}$$

Definition 1.2

A subset $S \subseteq X$ is called a generalized solution set of (P) if the points of S satisfy certain necessary optimality conditions.

It is frecuent to take in the role of S the set of optimal solutions of (P). Sometimes one considers a larger set, for instance in the case f is differentiable and X is open ,

$$S = \{x \in X : \nabla f(x) = 0 \};$$

or in the case f is Lipschitz ,

$$S = \{x \in X : \quad 0 \in \partial f(x) \}.$$

Definition 1.3

Let S be a generalized solution set of (P) , A an algorithm on E. A continuous function $\varphi : E \to R$ is said to be a descent function for S and A if

i) $\varphi(x) > (y)$ for all $y \in A(x)$, $x \notin S$

ii) $\varphi(x) \geq (y)$ for all $y \in A(x)$, $x \in S$.

We recall that A is said to be closed at x_0 , if for any sequence $\{x_k\}$ converging to x_0 , and $y_k \in A(x_k)$, $\{y_k\}$ converging to y_0 , one has $y_0 \in A(x_0)$.

Theorem 1.4

Let A be an algorithm on E, S a generalized solution set of (P). Let $\{x_k\}$ be a sequence generated by (1.1) with an initial point x_0. Assume that

i) all x_k are contained in a compact subset of E

ii) there exists a descent function φ for S and A

iii) A is closed at any point outside S.

Then the limit of any convergent subsequence of $\{x_k\}$ belongs to S.

Proof. Assume that a subsequence $\{x_{k_i}\}$ of $\{x_k\}$ converges to x_*. Then by the continuity of φ,

$$\lim \varphi(x_{k_i}) = \varphi(x_*).$$

Since $\{\varphi(x_k)\}$ is nonincreasing, one has

$$\lim \varphi(x_k) = \varphi(x_*) . \tag{1.2}$$

We show that $x_* \in S$. Suppose to the contrary that $x_* \in S$. Consider the subsequence $\{x_{k_i+1}\}$. Without loss of generality one can assume that it converges to some point y. Since $x_{k_i+1} \in A(x_{k_i})$ by the construction (1.1), and since $x_0 \in S$ where A is closed, one has

$$y \in A(x_*) . \tag{1.3}$$

Apply φ to the points x_* and y to deduce from (1.3) the inequality

$$\varphi(y) < \varphi(x_*) ,$$

which contradicts (1.2). ■

Most of descent algorithms can be described as follows. Starting from a point x_k one chooses a direction d_k and minimizes f on the line x_k+td_k , $t \geq 0$. A minimum is taken as x_{k+1} and the procedure is repeated until some optimality criteria are satisfied. The process of determining the minimum point x_{k+1} is called line search. One assumes of course. that the space E is linear.

Definition 1.5

The line search algorithm A is defined by

$$A(x,d) = \{y \in E : y = x + \alpha d \quad \text{with} \quad f(y) = \min_{\alpha \geq 0} f(x+\alpha d)\} .$$

Theorem 1.6

If f is continuous, the line search algorithm is closed at (x,d) in the case $d \neq 0$.

Proof. Assume that $\{x_k\} \to x$, $\{d_k\} \to d \neq 0$. Let $y_k \in A(x_k, d_k)$ and $y_k \to y$. We have to show that $y \in A(x,d)$. By definition,

$$y_k = x_k + \alpha_k d_k .$$

Hence

$$\alpha_k = \frac{|y_k - x_k|}{|d_k|} .$$

It is obvious that

$$\lim \alpha_k = \frac{|y - x|}{|d|} .$$

Hence $y = x + \alpha_0 d$, where $\alpha_0 = \lim \alpha_k$.

Since

$$f(y_k) \leq f(x_k + \alpha d_k) , \text{ for all } \alpha \geq 0 , \; k = 1,2,\ldots,$$

the continuity of f implies that

$$f(y) \leq f(x+\alpha d) , \text{ for all } \alpha \geq 0 ,$$

which means that $y \in A(x,d)$.■

Definition 1.7

Let $\{\lambda_k\}$ be a sequence of real numbers converging to λ_0. The order of convergence of $\{\lambda_k\}$ is the supremum of the number p such that

$$0 \leq \limsup_{k\to\infty} \frac{|\lambda_{k+1} - \lambda_0|}{|\lambda_k - \lambda_0|^p} < \infty .$$

If

$$\lim_{k\to\infty} \frac{|\lambda_{k+1} - \lambda_0|}{|\lambda_k - \lambda_0|^p} = \beta < 1 ,$$

we say that the sequence converges linearly to λ_0 with convergence ratio β. The case $\beta = 0$ is referred to as superlinear convergence.

Definition 1.8

Let $\{\lambda_k\}$ converge to λ_0. The average order of convergence is the infimum of the number $p > 1$ such that

$$\limsup_k |\lambda_k - \lambda_0|^{1/p^k} = 1 .$$

In the case $p = 1$, the amount

$$\limsup_k |\lambda_k - \lambda_0|^{1/k}$$

is called the average convergence ratio.

Definition 1.9

Let $\{x_k\}$ be a sequence in E converging to x_0. Any continuous function φ on E used to measure convergence of $\{x_k\}$ is called the error function.

Proposition 1.10

Let φ and ψ be two error functions with $\varphi(x_0) = \psi(x_0) = 0$. If for each $x \in X$,

$$0 \leq \alpha_1 \varphi(x) \leq \psi(x) \leq \alpha_2 \varphi(x), \text{ some } \alpha_1, \alpha_2 > 0. \tag{1.4}$$

Then a sequence $\{x_k\}$ converges linearly to x_0 with average ratio β with respect to one of these functions, it also does so with respect to the other.

Proof. Assume that $\{x_k\}$ converges linearly to x_0 with average ratio β with respect to ψ, then by (1.4),

$$\begin{aligned}
\limsup[\varphi(x_k)]^{1/k} &= \limsup[\alpha_1 \varphi(x_k)]^{1/k} \\
&\leq \limsup[\psi(x_k)]^{1/k} \\
&= \beta \\
&\leq \limsup[\alpha_2 \varphi(x_k)]^{1/k} \\
&= \limsup[\varphi(x_k)]^{1/k}.
\end{aligned}$$

The converse part is proven in a similar way because (1.4) yields also

$$0 \leq \frac{1}{\alpha_2}\psi(x) \leq \varphi(x) \leq \frac{1}{\alpha_1}\psi(x) \ . \ \blacksquare$$

A useful example of φ and ψ is given by the functions:

$$\psi(x) = (x - x_0)Q(x - x_0)$$

$$\varphi(x) = |x - x_0|^2 \ ,$$

where Q is a positive definite symmetrix matrix.

The numbers α_1 , α_2 can be taken as the smallest and largest eigenvalues of Q .

5.2 One-dimensional Search Techniques

Let $f: R \to R$ be a unimodal function, which means that it has a minimum point x_* and $f(x)$ decreases to its minimum as x monotonically tends to x_* .

Lemma 2.1

Suppose that $x_* \in [a,b]$ and $a < x_1 < x_2 < b$. Then one has

$$x_* \in [a,x_2] \quad \text{if} \quad f(x_2) > f(x_1)$$

$$x_* \in [x_1,b] \quad \text{otherwise} \ .$$

Proof. If $f(x_2) > f(x_1)$ then x_* cannot lie in $[x_2,b]$ because otherwise f decreases from x_1 to x_* and one would have $f(x_2) \leq f(x_1)$. The other case is proven similarly. ■

Definition 2.2

An interval $[a,b]$ containing x_* is called an interval of uncertainty.

Four methods that we are going to describe next are used to reduce intervals of uncertainty to as small as possible.

Search with Fixed Step Size

1) Choose an initial point $x_0 \in R$ and a step length s.

2) Set $x_1 = x_0 + s$

3) If $f(x_1) < f(x_0)$, in view of Lemma 1.1, $x_* \geq x_0$. Search $x_i = x_{i-1} + s$ can be continued until $f(x_i) > f(x_{i-1})$, which shows $x_* \in [x_{i-2},x_i]$.

4) If $f(x_1) > f(x_0)$, in view of Lemma 1.1, $x_* \leq x_0$. Search $x_{-i} = x_0 - (i-1)s$ must be carried.

5) If $f(x_1) = f(x_0)$, one has $x_* \in [x_0,x_1]$.

6) If $f(x_1)$ and $f(x_{-2}) > f(x_0)$, one has $x_* \in [x_{-2},x_1]$.

By this procedure we are able to define an interval of uncertainty of length no bigger than $2s$.

Dichotomous Search

Suppose that $x_* \in [a_n, b_n]$. Let c_n be the midpoint.

1) Choose a small positive ε and set

$$x_n = c_n - \varepsilon/2$$

$$y_n = c_n + \varepsilon/2 .$$

2) If $f(y_n) > f(x_n)$, by Lemma 1.1, $x_* \in [a_n, y_n]$, otherwise, $x_* \in [x_n, b_n]$.

The length of the interval of uncertainty is reduced by a factor near to $1/2$.

Fibonacci Search

This method uses Fibonacci numbers to reduce the length of the interval of uncertainty Fibonacci was an italian scientist of the 13th century who created a sequence of numbers F_n with the property:

$$F_0 = F_1 = 1 , \; F_n = F_{n-1} + F_{n-2} . \tag{2.1}$$

To understand how these numbers occur let us assume that it takes one month for rabbits to mature to fertitity and another month to produce a litter of two. Letting F_n to be the number of pairs of rabbits alive after n months and starting with one pair of rabbits, one sees that F_n is obtained exactly by (2.1).

In Fibonacci search it is assumed that the initial interval of uncertainty is given, say $[a_1,b_1]$ and the total number of experiments to be done is also given, say n .

1) At iteration k , the interval of uncertainty is $[a_k,b_k]$.
Set

$$x_k = a_k + \frac{F_{n-k}}{F_{n+2-k}} (b_k - a_k)$$

$$y_k = b_k - \frac{F_{n-k}}{F_{n+2-k}} (b_k - a_k)$$

2) Compute $f(x_k)$, $f(y_k)$.
If $f(x_k) < f(y_k)$, then $x_* \in [a_k,y_k]$, otherwise $x_* \in [x_k,b_k]$.

3) Denote the new interval of uncertainty by $[a_{k+1},b_{k+1}]$ and continue this procedure until $k = n-1$.

4) At the last iteration, $x_{n-1} = y_{n-1}$ is the midpoint of $[a_{n-1},b_{n-1}]$. To reduce the interval of uncertainty to about one-half its length, take ε small and set

$$x_{n-1} = a_{n-1} + \varepsilon$$

$$y_{n-1} = b_{n-1} + \varepsilon .$$

After n iteratives, the length of the interval of uncertainty is

$$b_n - a_n = \frac{b_1 - a_1}{F_{n+1}} + \varepsilon .$$

Golden Section Search

It is supposed that the number of experiments is large $(n \to \infty)$ in the Fibonacci search.

Denote

$$\frac{1}{\alpha} = \lim_{n\to\infty} \frac{F_{n-1}}{F_n} \approx 0.618 .$$

In the Fibonacci search one changes in 1):

$$x_k = a_k + \frac{\alpha - 1}{\alpha} (b_k - a_k)$$

$$y_k = a_k + \frac{1}{\alpha} (b_k - a_k) .$$

With this, the length of the interval of uncertainty after k iterations is reduced by a factor of $(\frac{1}{\alpha})^{k-1}$. The name "golden section" comes from the belief of ancient greek architects that a building with the sides a,b $(a > b)$ yielding the relation $\frac{a+b}{a} = \frac{a}{b} = \alpha$ will have the most pleasant properties.

The four methods described above have advantage that they are easy to performance and there is always convergence. Disadvantage is that they require a large number of function evaluations in order to achieve reasonable accuracy in the location of x_*.

Interpolation Methods

The basis for these methods is to approximate f by a polynomial, the minimum of which can be determined analytically. We present here only a quadratic approximation.

Assume that $x_* \geq x_0$ (say $x_0 = 0$). We are going to determine a quadratic function

$$q(x) = a + bx + cx^2 .$$

To do this we need to know the values at 3 points x_1, x_2, x_3. It is important to choose x_1, x_2, x_3 so that by solving equations $q(x_i) = f(x_i)$ one obtains the coefficient $c > 0$ which guarantees the existence of minimum of $q(x)$.

The procedure is as follows:

1) Fix $t > x_0$ and compute $f(t)$.

2) If $f(t) > f(x_0)$, compute $f(t/k)$, $k = 2,4,\ldots$ until

$$f(t/k) < f(x_0) .$$

Take $x_1 = x_0$, $x_2 = t/k$, $x_3 = 2t/k$, and determine a, b and c. The condition $f(t) > f(x_0) > f(t/k)$ assures the existence of the minimum of $q(x)$.

3) if $f(t) \leq f(x_0)$, compute $f(2^k t)$, $k = 1,2,\ldots$ until $f(2^k t) > f(2^{k-1} t)$ at the first time.

Take $x_1 = x_0$, $x_2 = 2^{k-1} t$, $x_3 = 2^k t$ and determine a,b,c .

The minimum of $q(x)$ exists because

$$f(x_0) > f(2^{k-1} t) > f(2^k t) .$$

It is useful to observe that the minimum of $q(x)$ can be obtained easily by solving linear equation $dq(x)/dx = 0$.

5.3 The Method of Steepest Descent

Assume that f is a function of class C^1 on R^n . The gradient of f at $x \in R^n$ is denoted by $\nabla f(x)$. The steepest descent algorithm is defined by

$$A(x) = x - \alpha \nabla f(x) , \tag{3.1}$$

where α is a number minimizing the function $f(x-\alpha\nabla f(x))$ over $\alpha \geq 0$.

The theoretical basis for the name of the algorithm is as follows. Suppose that $\nabla f(x) \neq 0$. Let v be a unit norm vector in R^n . The rate of change of f with respect to the step length dt along v is

$$\left. \frac{df(x + tv)}{dt} \right|_{t=0} = < \nabla f(x), v > ,$$

where $< \cdot, \cdot >$ denotes the inner product. We want to choose v so that this quantity is minimal, i.e. we solve the problem

$$\min < \nabla f(x), v >$$

$$\text{s.t. } v \in R^n , \ \|v\|^2 = 1 .$$

Using Lagrange function and optimality condition (section 3.5) one can obtain the optimal solution

$$v = - \nabla f(x) \ / \ \| \nabla f(x) \| .$$

Proposition 3.1

The steepest descent algorithm described by (3.1) is closed if $\nabla f(x) \neq 0$. Moreover, if we take

$$S = \{x \in X : \nabla f(x) = 0\}$$

as the generalized solution set, then $\varphi(x) = f(x)$ is a descent function for A and S . Consequently if starting from x_0 one generates the sequence $\{x_k\}$ by

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) , \tag{3.2}$$

and this sequence is bounded (α_k is a number minimizing $f(x_k - \alpha \nabla f(x_k))$ over $\alpha \geq 0$) , then any cluster point of the sequence belongs to S .

Proof. The closedness of A is immediate from Theorem 1.6. Furthermore, under the condition $\nabla f(x) \neq 0$, one has

$$\min_{\alpha \geq 0} f(x-\alpha \nabla f(x)) < f(x) ,$$

which shows that in fact the function $\varphi = f$ is a descent function for A and S. The last part of the proposition is derived from Theorem 1.4.■

Let us now take up a particular case where f is a quadratic function of the form

$$f(x) = \frac{1}{2} xQx - bx ,$$

where Q is a positive definite symmetric $(n \times n)$-matrix, $b \in R^n$. It is known that Q has n positive eigenvalues $0 < \lambda_1 \leq \ldots \leq \lambda_n$. The function f is strictly convex and it has the unique minimum x_* which can found by solving the equation :

$$0 = \nabla f(x_*) = Qx_* - b .$$

The algorithm (3.2) takes the form

$$x_{k+1} = x_k - \frac{< \nabla f(x_k), \nabla f(x_k) >}{\nabla f(x_k)\, Q\, \nabla f(x_k)} \nabla f(x_k) , \qquad (3.3)$$

where $\nabla f(x_k) = Qx_k - b$.

Denote

$$f_0(x) = f(x) + \frac{1}{2} x_* Q x_* .$$

Lemma 3.2

The iterative process (3.3) satisfies the relation

$$f_0(x_{k+1}) = \left[1 - \frac{< \nabla f(x_k), \nabla f(x_k) >^2}{(\nabla f(x_k) Q \nabla f(x_k))(\nabla f(x_k) Q^{-1} \nabla f(x_k))}\right] f_0(x_k) .$$

Proof. Let us calculate

$$\frac{f_0(x_k) - f_0(x_{k+1})}{f_0(x_k)} = \frac{2\alpha_k \nabla f(x_k) Q(x_k - x_*) - \alpha_k^2 \nabla f(x_k) Q \nabla f(x_k)}{(x_k - x_*) Q (x_k - x_*)}$$

$$= \frac{< \nabla f(x_k), \nabla f(x_k) >^2}{(\nabla f(x_k) Q \nabla f(x_k))(\nabla f(x_k) Q^{-1} \nabla f(x_k))} . \blacksquare$$

Lemma 3.3 (Kantorovich Inequality)

For every vector $x \in R^n$ one has

$$\frac{< x, x >^2}{(xQx)(xQ^{-1}x)} \geq \frac{4\lambda_1 \lambda_n}{(\lambda_1 + \lambda_n)^2} .$$

Proof. By an appropriate change of coordinates, the matrix Q becames diagonal with diagonal $(\lambda_1,\ldots,\lambda_n)$. One has then

$$\frac{< x,x >^2}{(xQx)(xQ^{-1}x)} = \frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} \lambda_i x_i^2 \quad \sum_{i=1}^{n} x_i^2/\lambda_i}$$

$$= \frac{1}{\sum_{i=1}^{n} t_i\lambda_i \quad \sum_{i=1}^{n} t_i/\lambda_i}, \tag{3.4}$$

where

$$t_i = \frac{x_i^2}{\sum_{i=1}^{n} x_i^2}.$$

Denote the amount of (3.4) by $\alpha(y)$ where $y = (t_1,\ldots,t_n)$. Observe that

$$\lambda = \sum_{i=1}^{n} t_i\lambda_i$$

is a convex combination of λ_i, and

$$h(y) = \sum_{i=1}^{n} t_i/\lambda_i$$

is a convex combination of $1/\lambda_i$.

We prove that

$$h(y) \leq \frac{\lambda_1 + \lambda_n - \lambda}{\lambda_1 \lambda_n} . \tag{3.5}$$

Let us do it by induction on n. The case $n=1$ is trivial. Supposing (3.5) is true for $n-1$, we consider it for n.

$$\sum_{i=1}^{n} \frac{t_i}{\lambda_i} = \frac{t_i}{\lambda_1} + \ldots + \frac{t_{n-2}}{\lambda_{n-2}} + \frac{t_{n-1}+t_n}{\lambda_n} + t_{n-1}\left(\frac{1}{\lambda_{n-1}} - \frac{1}{\lambda_n} \right)$$

$$\leq \frac{\lambda_1+\lambda_n - (\lambda_1 t_1 + \ldots + \lambda_{n-2} t_{n-2} + \lambda_n (t_{n-1}+t_n))}{\lambda_1 \lambda_n} +$$

$$+ t_{n-1}\left(\frac{1}{\lambda_{n-1}} - \frac{1}{\lambda_n} \right)$$

$$= \frac{\lambda_1+\lambda_n - \Sigma_{i=1}^{n} \lambda_i t_i}{\lambda_1 \lambda_n} + \frac{t_{n-1}(\lambda_{n-1} - \lambda_n)}{\lambda_1 \lambda_n} + t_{n-1}\left(\frac{1}{\lambda_{n-1}} - \frac{1}{\lambda_n}\right)$$

$$= \frac{\lambda_1+\lambda_n - \Sigma_{i=1}^{n} \lambda_i t_i}{\lambda_1 \lambda_n} + \frac{t_{n-1}}{\lambda_1 \lambda_n \lambda_{n-1}} \left[\lambda_{n-1}^2 - \lambda_n \lambda_{n-1} + \lambda_1 \lambda_n - \lambda_1 \lambda_{n-1}\right]$$

$$\leq \frac{\lambda_1+\lambda_n - \Sigma_{i=1}^{n} \lambda_i t_i}{\lambda_1 \lambda_n}$$

$$= \frac{\lambda_1+\lambda_n - \lambda}{\lambda_1 \lambda_n} .$$

In this way ,

$$\alpha(y) \geq \min_{\lambda} \frac{\lambda_i \lambda_n}{\lambda(\lambda_1+\lambda_n-\lambda)} .$$

The minimum is attained at

$$\lambda = \frac{\lambda_1 + \lambda_n}{2} ,$$

which implies the required inequality.■

Theorem 3.4

For any $x_0 \in R^n$, the method of steepest descent (3.3) converges to the unique minimum point x_* of f .

Furthermore with $f_0(x)$ defined before Lemma 3.2 , at each step k one has

$$f_0(x_{k+1}) \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 f_0(x_k) . \tag{3.6}$$

Proof. The relation (3.6) is derived at once from Lemmas 3.2 and 3.3. This implies that $f_0(x_k)$ converges to 0 . Since Q is positive definite , $\{x_k\}$ converges to the minimum point x_* .■

Note that by Definition 1.7, with respect to the error function f_0 , the method of steepest descent converges linearly with a ratio no bigger than $[\ (\lambda_n - \lambda_1) / (\lambda_n + \lambda_1)]^2$. This means that if Q has eigenvalues

with the difference $\lambda_n - \lambda_1$ small then the convergence is quick .

In general, the algorithm (3.1) behaves well at first iterations . After some iterations the convergence is slow because it uses only linear approximations at each step and the directions obtained at the last iterations are generally not effective.

5.4 Conjugate Gradient Methods

Originally these methods have been creating for solving quadratic problems. Then they are extended to more general problems and nowdays they belong to the best general purpose methods of nonlinear optimization. In this section Q denotes a symmetric positive definite $n \times n$ - matrix .

Definition 4.1

A family of vectors $v_1, \ldots, v_k \in R^n$ are said to be Q-orthogonal, or conjugate with respect to Q if

$$v_i Q v_j = 0 \text{ , for all } i,j, i \neq j. \tag{4.1}$$

Proposition 4.2

If $v_1, \ldots, v_k$ are nonzero, conjugate, then they are linearly independent .

Proof. Suppose that there are some scalars x_i with $\sum_{i=i}^{k} \alpha_i v_i = 0$. Multiplying this by Qv_i yields

$$\alpha_i \, v_i \, Q v_i = 0 \text{ .}$$

Since $v_i \neq 0$ and Q is positive definite, we conclude $\alpha_i = 0$. ■

Theorem 4.3

Let $v_1, \ldots, v_n$ be nonzero conjugate. Starting from $x_1 \in R^n$ and following the algorithm

$$x_{k+1} = x_k + \alpha_k v_k , \tag{4.2}$$

where $\alpha_k = - v_k \nabla f(x_k) / v_k Q v_k$, we can reach the unique minimum x_* of the function

$$f(x) = \frac{1}{2} x Q x - bx$$

after $k \leq n$ steps.

Proof. By Proposition 4.2, $\{v_1, \ldots, v_n\}$ forms a basis set of R^n, hence there are $\alpha_1, \ldots, \alpha_n$:

$$x_* - x_1 = \sum_{i=1}^{n} \alpha_i v_i .$$

Multiplying this by Qv_i we obtain

$$\alpha_i = \frac{v_i Q(x_* - x_1)}{v_i Q\, v_i}$$

$$= \frac{v_i(b - Qx_1)}{v_i Q v_i}$$

$$= -\frac{v_i \nabla f(x_1)}{v_i Q v_i}$$

$$= -\frac{v_i \nabla f(x_i)}{v_i Q v_i} \quad .$$

We have used the fact that $Qx_* = b$ and $x_i = x_1 + \alpha_1 v_1 + \ldots + \alpha_{i-1} v_{i-1}$ according to (4.2) .■

Let us denote by $\mathrm{Lin}\{v_1,\ldots,v_k\}$ the subspace spanned by vectors $v_1,\ldots,v_k$.

Theorem 4.4

The points x_{k+1} obtained in Theorem 4.3 minimizes the function $f(x) = \frac{1}{2}\, xQx - bx$ on the linear manifold

$$x_1 + \mathrm{Lin}\{v_1,\ldots,v_k\}.$$

Consequently, $\nabla f(x_{k+1})\, v_i = 0$ for all $i \leq k$.

Proof. Observe that f is strictly convex, therefore the result follows if it can be shown that $\nabla f(x_{k+1})$ is orthogonal to $\mathrm{Lin}\{v_1,\ldots,v_k\}$. We

prove this by induction. For $k=0$, the hypothesis is trivial. Assuming $\nabla f(x_k)$ is orthogonal to $\text{Lin}\{v_1,\ldots,v_{k-1}\}$, we show that $\nabla f(x_{k+1})$ is orthogonal to $\text{Lin}\{v_1,\ldots,v_k\}$. By definition,

$$\nabla f(x_{k+1}) = Q(x_{k+1}) - b$$

$$= \nabla f(x_k) + \alpha_k Q v_k .$$

It follows from the definition of α_k that

$$v_k \nabla f(x_{k+1}) = v_k \nabla f(x_k) + \alpha_k v_k Q v_k = 0 .$$

For $i < k$, the hypothesis of induction and the conjugacy imply

$$v_i \nabla f(x_{k+1}) = v_i \nabla f(x_k) + \alpha_k v_i Q v_k = 0 .$$

Hence $\nabla f(x_{k+1})$ is orthogonal to $\text{Lin}\{v_1,\ldots,v_k\}$. ∎

The basic idea of the conjugate gradient method is to use conjugate directions which are obtained with the aid of gradients in the role of search directions.

The algorithm is as follows:

1) Starting at any point $x_1 \in R^n$, $v_1 = -\nabla f(x_1) = Qx_1 - b$; $k = 1$

2) Define

$$x_{k+1} = x_k + \alpha_k v_k \text{ , where } \alpha_k = - \frac{v_k \nabla f(x_k)}{v_k Q v_k} , \tag{4.3}$$

$$v_{k+1} = - \nabla f(x_{k+1}) + \beta_k v_k \text{ , where } \beta_k = \frac{\nabla f(x_{k+1}) Q v_k}{v_k Q v_k} . \tag{4.4}$$

3) Terminate when $\nabla f(x_k) = 0$.

heorem 4.5

The conjugate gradient algorithm possesses the following properties:

i) $\text{Lin}\{\nabla f(x_1),\ldots,\nabla f(x_k)\} = \text{Lin}\{v_1,\ldots,v_k\} = \text{Lin}\{\nabla f(x_1), Q\nabla f(x_1),\ldots, Q^{k-1}\nabla f(x_1)\}$.

ii) $v_1,\ldots,v_k$ are conjugate

iii) α_k and β_k can be calculated by

$$\alpha_k = \frac{\|\nabla f(x_k)\|^2}{v_k Q v_k}$$

$$\beta_k = \frac{\|\nabla f(x_{k+1})\|^2}{\|\nabla f(x_k)\|^2}$$

roof. We prove i), ii) simultaneously by induction. For k=1 , the hy-
othesis is trivial. We show it for k+1 , supposing that it is true for
.

By definition of x_{k+1} one has

$$\nabla f(x_{k+1}) = \nabla f(x_k) + \alpha_k \, Qv_k \ .$$

By induction both $\nabla f(x_k)$ and v_k belong to the subspace $\mathrm{Lin}\{\nabla f(x_1) \ , \ Q\nabla f(x_1) \ , \ldots \ , \ Q^{k-1}\nabla f(x_1)\}$. Hence $\nabla f(x_{k+1})$ belongs to $\mathrm{Lin}\{\nabla f(x_1), \ Q\nabla f(x_1) \ , \ldots \ , \ Q^{k}\nabla f(x_1)\}$. Moreover, by induction of ii) , $v_1,\ldots,v_k$ are conjugate, hence in view of Theorem 4.4, $\nabla f(x_{k+1})$ is orthogonal to $\mathrm{Lin}\{v_1,\ldots,v_k\} = \mathrm{Lin}\{\nabla f(x_1) \ , \ldots \ , \ Q^{k-1}\nabla f(x_1)\}$. This fact implies that $\nabla f(x_{k+1}) \notin \mathrm{Lin}\{\nabla f(x_1) \ , \ldots \ , \ Q^{k-1}\nabla f(x_1)\}$. Hence one can conclude

$$\mathrm{Lin}\{\nabla f(x_1),\ldots,\nabla f(x_{k+1})\} = \mathrm{Lin}\{\nabla f(x_1),\ldots,Q^{k}\nabla f(x_1)\} \ .$$

Use the above part and the fact that $v_{k+1} = - \nabla f(x_{k+1}) + \beta_k v_k$ to use that

$$\mathrm{Lin}\{v_1,\ldots,v_{k+1}\} = \mathrm{Lin}\{\nabla f(x_1), \ Q\nabla f(x_1) \ , \ldots \ , \ Q^{k}\nabla f(x_1)\} \ .$$

To show that $v_1,\ldots,v_{k+1}$ are conjugate, remember that

$$v_{k+1} \, Qv_i = - \nabla f(x_{k+1})Qv_i + \beta_k \, v_k Qv_i \ .$$

If $i=k$, the right hand side is zero by definition of β_k . If $i < k$, the second term of the right hand side is zero by induction. The first term is also zero because by the first part

$$Qv_i \in \mathrm{Lin}\{v_1,\ldots,v_{i+1}\} \subseteq \mathrm{Lin}\{v_1,\ldots,v_k\}$$

and by induction, $v_1,\ldots,v_k$ are conjugate, hence in view of Theorem 4.4, $\nabla f(x_{k+1})$ is orthogonal to $\mathrm{Lin}\{v_1,\ldots,v_k\}$. In this way $v_1,\ldots,v_{k+1}$ are conjugate.

To calculate α_k, observe that

$$-v_k \nabla f(x_k) = \nabla f(x_k)\nabla f(x_k) - \beta_{k-1} v_{k-1} \nabla f(x_k) . \qquad (4.5)$$

By Theorem 4.4 and ii), the second term of the right hand side is zero. Replacing (4.5) to (4.3) we obtain the formula for α_k.

As to β_k observe first that

$$\nabla f(x_{k+1})\nabla f(x_k) = 0 ,$$

because $\nabla f(x_k) \in \mathrm{Lin}\{v_1,\ldots,v_k\}$ and $\nabla f(x_{k+1})$ is orthogonal to $\mathrm{Lin}\{v_1,\ldots,v_k\}$. Moreover,

$$\nabla f(x_{k+1}) = Q(x_{k+1}) - b = \nabla f(x_k) + \alpha_k Q v_k ,$$

hence

$$Qv_k = \frac{\nabla f(x_{k+1}) - \nabla f(x_k)}{\alpha_k} .$$

These observations and (4.4) show the formula for β_k in the theorem. ∎

Now we extend the algorithm for f of class C^1 :

1) Start with $x_1 \in X$

2) Set $v_1 = - \nabla f(x_1)$, $k=1$

3) Find x_{k+1} - the minimum point of f along $x_k + \alpha v_k$, $\alpha \geq 0$.

4) Set $v_{k+1} = - \nabla f(x_{k+1}) + \beta_k v_k$ where

$$\beta_k = \frac{\|\nabla f(x_{k+1})\|^2}{\|\nabla f(x_k)\|^2}$$

5) After n points restart with $x_1 = x_{n+1}$, and terminate when $\nabla f(x_k) = 0$.

Theorem 4.6

Assume that $f \in C^1$ and there exists $x_1 \in R^n$ such that the set $\text{lev}_f(f(x_1))$ is compact. Then the sequence $\{x_k\}$ generated by the algorithm described above has the property that $f(x_k) > f(x_{k+1})$ of $\nabla f(x_k) \neq 0$. One has also global convergence with the generalized solution set $S = \{x \in R^n : \nabla f(x) = 0\}$.

Proof. Suppose that $\nabla f(x_k) \neq 0$. Then the vector

$$v_k = - \nabla f(x_k) + \beta_{k-1} v_{k-1} .$$

cannot be zero, because $\{v_1,\ldots,v_{k-1}\}$ are conjugate and $\nabla f(x_k)$ is orthogonal to them. This implies that

$$f(x_k) > f(x_{k+1}) .$$

Hence $\{x_k\}$ belongs to the compact set $\text{lev}_f(f(x_1))$. Take $\varphi = f$ as a descent function and observe that the algorithm is closed outside S. Now apply Theorem 1.4 to obtain the global convergence. ■

5.5 Newton and Quasi-Newton Methods

Newton Method

Let us consider the problem

$$\begin{aligned} &\min \quad f(x) \\ &\text{s.t. } x \in R^n . \end{aligned}$$

The idea behind the Newton method is to approximate f near a point x_k by a quadratic function whose minimum can be calculated analytically. Recall that $\nabla^2 f(x)$ denotes the Hessian of f at x. By the truncated Taylor series one has near x_k :

$$f(x) = f(x_k) + \nabla f(x_k)(x-x_k) + \frac{1}{2}(x-x_k)\nabla^2 f(x_k)(x-x_k) .$$

The function in the right hand side is quadratic.

If $\nabla^2 f(x)$ is positive definite, then the minimum x_{k+1} is obtained by

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k) .$$

The Newton algorithm is given as

$$A(x) = x - [\nabla^2 f(x)]^{-1} \nabla f(x) . \tag{5.1}$$

Proposition 5.1

If the Newton method converges, the order of convergence is two.

Proof. Suppose that x_* is a point with $\nabla f(x_*) = 0$ and $\nabla^2 f(x_*)$ is nonsingular. Then

$$x_{k+1} - x_* = A(x_k) - A(x_*) .$$

Hence one has by the mean-value theorem:

$$\begin{aligned} |x_{k+1} - x_*| &\leq |A(x_k) - A(x_*)| \\ &\leq |\nabla A(x_*)(x_k - x_*)| + \tfrac{1}{2} |\nabla^2 A(x_0)| \, |x_k - x_*|^2 \end{aligned}$$

where $x_0 \in (x_k, x_*)$. Since $\nabla A(x_*) = 0$, the above inequality implies

$$|x_{k+1} - x_*| \leq c|x_k - x_*|^2 ,$$

where c is a constant depending on $|f'''(x)|$ near x_*. ■

Proposition 5.2

If f is quadratic, its minimum can be reached in one step by the Newton method.

Proof. In this case, $\nabla^2 f(x) = Q$. The minimum of f is the solution of the equation

$$Qx - b = 0 .$$

Hence $x_* = Q^{-1}b$. Now, starting from any point x_k, by the algorithm (5.1),

$$\begin{aligned} x_{k+1} &= x_k - Q^{-1}(Qx_k - b) \\ &= Q^{-1}b \\ &= x_* , \end{aligned}$$

which completes the proof. ■

Note that if f is not quadratic, the method may diverge, and it may converge to saddlepoint or relative maxima.

Modified Newton Method

In order to guarantee convergence in the Newton method one modifies the algorithm by several ways. A general scheme can be given as follows:

$$x_{k+1} = x_k - \alpha_k H_k \nabla f(x_k) , \tag{5.2}$$

where H_k is a positive definite nxn-matrix, α_k is selected to minimize $f(x)$. The positive definiteness of H_k guarantees that

$$f(x_{k+1}) < f(x_k) ,$$

whenever $\nabla f(x_k) \neq 0$.

In the case of quadratic function

$$f(x) = \frac{1}{2} xQx - bx ,$$

α_k can be given explicitely by

$$\alpha_k = \frac{\nabla f(x_k) H_k \nabla f(x_k)}{\nabla f(x_k) H_k Q H_k \nabla f(x_k)} . \tag{5.3}$$

Theorem 5.3

Let x_* be the unique minimum of the quadratic function f and $f_0(x) = \frac{1}{2} (x-x_0)Q(x-x_0)$. Then for the algorithm (5.2) with α_k defined by (5.3) one has

$$f_0(x_{k+1}) \leq \left(\frac{\lambda_{nk}-\lambda_{1k}}{\lambda_{nk}+\lambda_{1k}}\right) f_0(x_k) ,$$

where λ_{nk} , λ_{1k} are the largest and smallest eigenvalues of the matrix H_kQ .

Proof. Let us calculate:

$$\frac{f_0(x_k) - f_0(x_{k+1})}{f_0(x_k)} = \frac{(\nabla f(x_k)\, H_k\, \nabla f(x_k))^2}{(\nabla f(x_k) H_k Q\, H_k \nabla f(x_k))(\nabla f(x_k) H_k Q^{-1} H_k \nabla f(x_k))} .$$

Letting

$$Q_k = H_k^{1/2}\, Q\, H_k^{1/2} ,$$

$$P_k = H_k^{1/2}\, \nabla f(x_k) ,$$

we obtain

$$\frac{f_0(x_k) - f_0(x_{k+1})}{f_0(x_k)} = \frac{< P_k, P_k >^2}{(P_k Q_k P_k)(P_k Q_k^{-1} P_k)} .$$

Using Lemma 3.3 one deduces the result at once because H_kQ is similar to Q_k ($H_k^{1/2}\, Q_k H_k^{-1/2} = H_kQ$) , hence they have the same eigenvalues. ∎

Variable Metric Method

(Davidon-Fletcher-Powell Method)

Most of quasi-Newton methods make use of differences in the gradient values to approximate $\nabla^2 f(x)$ in the Newton method. The Davidon-Fletcher-Powell method constructs H_k in (5.2) with the aid of H_{k-1}, $\nabla f(x_k)$, $\nabla f(x_{k+1})$. The procedure is as follows :

1) Start with an initial point x_1 and an approximation $H_1 = I$, the identity matrix. The iteration number is set to be k=1.

2) Compute $\nabla f(x_k)$ and set $v_k = - H_k \nabla f(x_k)$.

3) Find the minimum point x_{k+1} of f along $x_k + \lambda v_k$, $\lambda \geq 0$.

4) Test x_{k+1} for optimality. If it is the case, stop. Otherwise go to 5).

5) Calculate

$$y_k = x_{k+1} - x_k$$

$$\nabla_k = \nabla f(x_{k+1}) - \nabla f(x_k)$$

$$H_{k+1} = H_k + \frac{y_k^T y_k}{y_k \nabla_k} - \frac{(H_k \nabla_k)(H_k \nabla_k)^T}{\nabla_k H_k \nabla_k}$$

6) Set $k = k+1$ and go to 2).

Below qre some important properties of H_k.

Theorem 5.4

We have the following

i) All H_k constructed by the procedure above are symmetric positive definite

ii) If f is quadratic, then $v_1,\dots,v_n$ are conjugate and

$$H_k\nabla_j = y_j \quad , \text{ for all } j<k \ .$$

Moreover
$$H_kQy_j = y_j \quad , \text{ for } j<k \ .$$

In particular ,
$$H_{n+1} = Q^{-1} \ .$$

Proof. The symmetry of H_k is obvious. For the positive definiteness supposing by induction that H_k is positive definite, we prove it for H_{k+1} . Let z be an arbitrary nonzero vector in R^n . By definition of H_{k+1} ,

$$zH_{k+1}z = \frac{1}{\nabla_k H_k \nabla_k}\left[zH_kz\nabla_kH_k\nabla_k - (zH_k\nabla_k)^2\right] + \frac{(zy_k)^2}{y_k\nabla_k} \ . \qquad (5.4)$$

Since H_k is symmetric positive definite , there exists the matrix $H_k^{1/2}$. Denote

$$v = H_k^{1/2}z \ , \ u = H_k^{1/2}\nabla_k \ .$$

The expression under [...] in (5.4) can be written as .

$$zH_k z\nabla_k H_k \nabla_k - (zH_k \nabla_k)^2 = (v\ v)(u\ u) - (u\ v)^2 \ ,$$

which is nonnegative by the Cauchy-Schwatz inequality. The second term of the right hand side of (5.4) is also nonnegative because

$$\begin{aligned} y_k \nabla_k &= y_k(\nabla f(x_{k+1}) - \nabla f(x_k) \\ &= -y_k \nabla f(x_k) \quad \text{(by step 4), } y_k \nabla f(x_{k+1}) = 0) \\ &= -\lambda_k v_k \nabla f(x_k) \quad (\lambda_k \text{ is found by step 3)}) \\ &= \lambda_k \nabla f(x_k) H_k \nabla f(x_k) \quad \text{(by definition of } v_k) \\ &> 0 \quad \text{(by induction } H_k \text{ is positive definite).} \end{aligned}$$

In this way

$$zH_{k+1} z \geq 0 \ .$$

In fact, we have the strict inequality.

This is because the first term of the right hand side of (5.4) is zero if and only if two sectors v and u are proportional, which implies that z and ∇_k are proportional too. In other words, there is $t \neq 0$ such that $z = t\nabla_k$. Substituting this to the second term of (5.4) we obtain

$$\frac{(zy_k)^2}{y_k \nabla_k} = t^2\, y_k \nabla_k \ .$$

which is positive as we have proven.

Thus

$$zH_{k+1}z > 0 \quad \text{and the positive}$$

definiteness of H_{k+1} follows.

The second part of the theorem is proven by induction. We remember that

$$f(x) = \frac{1}{2}\, xQx - bx \,.$$

Observe first that

$$\begin{aligned} Qy_j &= Q(x_{j+1}-x_j) \\ &= Q(x_{j+1}) - b - (Q(x_j) - b) \\ &= \nabla f(x_{j+1}) - \nabla f(x_j) \\ &= \nabla_j \,, \text{ for all } j \,. \end{aligned} \tag{5.5}$$

For the beginning of induction, since $H_1 = I$, one has

$$\begin{aligned} H_2\nabla_1 &= H_1\nabla_1 + \frac{y_1^T y_1}{y_1\nabla_1}\,\nabla_1 - \frac{H_1\nabla_1(H_1\nabla_1)^T\nabla_1}{\nabla_1 H_1\,\nabla_1} \\ &= \nabla_1 + y_1 - \nabla_1 \\ &= y_1 \,. \end{aligned} \tag{5.6}$$

We show that v_1 , v_2 are conjugate. Let us calculate v_2Av_1 :

$$\begin{aligned} v_2Av_1 &= -\nabla f(x_2)H_2Qv_1 \\ &= -\nabla f(x_2)H_2Qy_1/\lambda_1 \quad \text{(because } y_1 = \lambda_1 v_1\text{)} \\ &= -\nabla f(x_2)H_2\nabla_1/\lambda_1 \quad \text{(by (5.5))} \\ &= -\nabla f(x_2)y_1 \quad \text{(by (5.6))} \\ &= 0 \qquad\qquad \text{(by step 3))} . \end{aligned}$$

Now we prove the assertions for $k+1$. Let us calculate $H_{k+1}\nabla_j$ by definition of H_{k+1} :

$$H_{k+1}\nabla_j = H_k\nabla_j + \frac{y_k^T y_k \nabla_j}{y_k\nabla_k} - \frac{H_k\nabla_k(H_k\nabla_k)^T\nabla_j}{\nabla_k H_k\nabla_k} . \tag{5.7}$$

For $j < k$, by induction

$$H_k\nabla_j = y_j ,$$

and by (5.5),

$$\begin{aligned} \nabla_k H_k \nabla_j &= \nabla_k y_j \\ &= y_k Q y_j \\ &= 0 . \end{aligned}$$

(We have used the fact that $v_1,\ldots,v_k$ are conjugate by induction, hence so are $y_1,\ldots,y_k$) .

Combine these relations and (5.7) to see that

$$H_{k+1}\nabla_j = y_j \text{ , for } j < k \text{ .}$$

In the case $j=k$, the equality

$$H_{k+1}\nabla_k = y_k$$

is obvious from (5.7) .

For the conjugacy of $v_1,\ldots,v_{k+1}$ it suffices to show that $v_{k+1}Qv_j = 0$ for $j \leq k$. Note first that since $x_{k+1} = x_j + \lambda_j v_j + \ldots + \lambda_k v_k$ and $y_j = \lambda_j v_j$,

$$\begin{aligned}
\nabla f(x_{k+1}(y_j) &= (Qx_{k+1} - b)y_j \\
&= (Qx_j - b + \lambda_j Qv_j + \ldots + \lambda_k Qv_k)y_j \\
&= \nabla f(x_j)y_j + \lambda_j v_j Qy_j \\
&= \nabla f(x_j)v_j + \lambda_j v_j Qv_j \text{ .}
\end{aligned}$$

Remembering that λ_j is a number minimizing f on $x_j + tv_j$, $t \geq 0$, one has

$$\lambda_j = -\nabla f(x_j)v_j \,/\, v_j Q v_j \,.$$

Hence

$$\nabla f(x_{k+1})y_j = 0 \quad \text{for} \quad j \le k \,. \tag{5.8}$$

With this in hand we are able to calculate $v_{k+1}Qv_j$:

$$\begin{aligned} v_{k+1}Qv_j &= -\nabla f(x_{k+1})H_{k+1}Qv_j \quad \text{(by definition of } v_{k+1}\text{)} \\ &= -\nabla f(x_{k+1})H_{k+1}Qy_j \,/\, \alpha_j \\ &= -\nabla f(x_{k+1})H_{k+1}\nabla_j \,/\, \alpha_j \quad \text{(by (5.5))} \\ &= -\nabla f(x_{k+1})y_j \,/\, \alpha_j \quad \text{(by property of } H_{k+1}\text{)} \\ &= 0 \qquad\qquad \text{(by (5.8))} \,. \end{aligned}$$

In the case $k=n$, since $v_1,\ldots,v_n$ are independent, $H_{n+1}Qv = v$ for all $v \in R^n$. Thus, $H_{n+1} = Q^{-1}$. ■

5.6 Problems

1. Is the result of Theorem 1.4 valid if instead of iii) we require A to be upper continuous?

   (Recall that A is upper continuous at x_0 if for any open set V containing $A(x_0)$, there is a neighborhood U of x_0 in X such that $A(x) \le V$, for all $x \in U$) .

2. Let $\{\lambda_k\}$ be a sequence of real numbers converging to 0. Suppose that $\{\lambda_k\}$ converges linearly. Find conditions of a sequence $\{t_k\}$ of positive numbers such that $\{t_k\lambda_k\}$ converges linearly to 0.

3. Use quadratic interpolation to estimate the location of the minimum of $f(x) = x^4 - 2x^2 + 1$ over $[0,2]$.

4. Suppose that f has continuous second partial derivatives and has a local minimum at x_*. Suppose further that the Hessian matrix $H(x_*)$ of f has the smallest eigenvalue $\lambda_1 > 0$ and largest eigenvalue λ_n. Prove that if $\{x_k\}$ is a sequence generated by the method of steepest descent and converges to x_*, then $\{f(x_k)\}$ converges to $f(x_*)$ linearly with convergence ratio no bigger than

$$\left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 .$$

5. What is the rate of convergence of the objective function

$$f(x,y) = x^2 + y^2 + xy - 3x$$

when applying the method of steepest descent.

6. What happens if we proceed the algorithm for $f \in C^1$ in Section 4 without restarting $x_1 = x_{n+1}$ at the step 5)?

7. Consider $f(x,y) = 2x^2 + 2xy + 5y^2$.

   Apply the conjugate gradient method to find the minimum.

8. Prove that in the variable metric method

$$Q^{-1} = \sum_{i=0}^{n-1} Q_i$$

where

$$Q_i = \frac{y_i^T y_i}{y_i \nabla i},$$

when f is the function $\frac{1}{2} xQx - b\,x$.

CHAPTER 6

CONSTRAINED OPTIMIZATION TECHNIQUES

6.1 Methods of Feasible Directions

We consider a minimization problem

$$(P) \qquad \min \quad f(x)$$
$$\text{s.t.} \quad x \in X ,$$

where X is a nonempty subset of R^n, f is a function on R^n which is assumed to have continuous partial derivatives.

Definition 1.1

Let $x \in X$. A direction v is said to be feasible at x if there is a positive ε such that

$$x + tv \in X \quad \text{for all} \quad t : 0 \leq t \leq \varepsilon .$$

A general scheme of a feasible direction algorithm is as follows

$$A(x) = \{y \in X : \ y \ \text{minimizes} \ f(x+tv) \ \text{over} \ t \in [0,\varepsilon]\} ,$$

where x is a feasible solution, v is a feasible direction. The feasible direction algorithm can be written as the composition of the maps D and M :

$$A = MD , \tag{1.1}$$

where D is a map of choosing feasible direction v , M is minimization of f along the chosen direction. Note that in general neither D , nor M is closed.

Definition 1.2

The set $\Omega = \{(x,v) \in R^n \times R^n : x \in X , v$ is a feasible direction at $x\}$ is called a set of uniformly feasible direction vectors if there exists $\delta > 0$ such that

$$x + tv \in X , \text{ for all } t \in [0,\delta] , (x,v) \in \Omega .$$

Let us define a map M_δ from Ω to R^n as follows

$$M_\delta(x,v) = \{y \in X : \ y \ \text{minimizes} \ f(x+tv) \ \text{over} \ t \in [0,\delta]\} .$$

Theorem 1.3

If $v \neq 0$, the map M_δ is closed at (x,v).

Proof. Suppose that $\{(x_k,v_k)\}$ converges to (x,v) with $v \neq 0$, and $y_k \in M_\delta(x_k,v_k)$ converges to y.

We want to show that $y \in M_\delta(x,v)$. Since

$$y_k = x_k + \alpha_k v_k .$$

one has

$$\alpha_k = \frac{\|y_k - x_k\|}{\|v_k\|} ,$$

where obviously, $\|v_k\| > 0$ if k is large enough. Hence α_k converges to

$$\alpha = \frac{\|y - x\|}{\|v\|} ,$$

which implies that $y = x + \alpha v$. Moreover,

$$f(y_k) \leq f(x_k + tv_k) , \text{ for all } t \in [0,\gamma] .$$

By the continuity of f, one obtains

$$f(y) \leq f(x + tv) \text{ , for all } t \in [0,\delta] \text{ .}$$

This shows that $y \in M_\delta(x,v)$. ■

An important consequence of Theorem 1.3 is that in order to develop a globally convergent algorithm, it is necessary to generate a map D which is closed and gives uniformly feasible directions.

Zoutendijk's Algorithm

Let us solve the following problem

$$\begin{aligned} &\min \ f(x) \\ &\text{s.t. } g_i(x) \leq 0 \text{ , } i,\ldots,m \text{ .} \end{aligned}$$

The algorithm is described as follows:

1) Starting from a point x_k which is feasible we solve the linear problem

$$\text{(LP)} \qquad \begin{aligned} &\min \ \alpha \\ &\text{s.t. } \nabla f(x_k)v - \alpha \leq 0 \\ &\qquad g_i(x_k) + \nabla g_i(x_k)v - \alpha \leq 0 \text{ , } i=1,\ldots,m \\ &\qquad\qquad -1 \leq v^i \leq 1 \text{ , } i=1,\ldots,n \text{ ,} \end{aligned}$$

where v^i denotes the ith component of v.

Let v_* is an optimal solution of (LP) and let t_k be the largest number in the interval $[0,1]$ such that $x_k + t_k v_*$ is feasible. Set $v_k = t_k v_*$.

2) Solve

$$\min_{t\in[0,1]} f(x_k + tv_k) .$$

Let the optimal solution be t_*. Take

$$x_{k+1} = x_k + t_* v_k .$$

3) Terminate when $v_k = 0$. If $v_k \neq 0$ repeat 1) for $x_k = x_{k+1}$.

Theorem 1.4

Assume that the set of feasible solutions $\{x \in R^n: g(x) \leq 0\}$ is convex compact and every feasible solution is regular in the sense that $\nabla g_1(x),\ldots,\nabla g_m(x)$ are linearly independent. Then the Zoutendijk algorithm has the global convergence property with the generalized solution set

$$S = \{x \in R^n : \nabla f(x) + \lambda \nabla g(x) = 0 , \lambda g(x) = 0 , \text{ some } \lambda \geq 0\} .$$

Consequently, any cluster point of the sequence $\{x_k\}$ generated by the above algorithm belongs to S (it is called a Kuhn-Tucker point).

Proof. It is obvious that the map $x_k \mapsto v_k$ in step 1) is closed. It generates uniformly feasible directions with $\delta = 1$. Hence the set

$$\Omega = \{(x,v) : x \text{ is feasible, } v \text{ is obtained by step 1)}\}$$

is compact. In view of Theorem 1.4 of Chapter 5, it suffices to show that it is a descent algorithm. Let $x \notin S$. We state that the optimal value of (LP) with $x_k = x$ is nonzero. In fact if that is not true, then dual program of (LP)

$$\begin{aligned} \max \quad & \xi_1 g_1(x) + \ldots + \xi_m g_m(x) \\ \text{s.t.} \quad & \xi_0 + \ldots + \xi_m = 1 \\ & \xi_0 \nabla f(x) + \xi_1 \nabla g(x) + \ldots + \xi_m \nabla g_m(x) = 0 \qquad (1.2) \\ & \xi_i \geq 0, \quad i=0,1,\ldots,m, \end{aligned}$$

has also zero optimal value, i.e.

$$\xi_1 g_1(x) + \ldots + \xi_m g_m(x) = 0,$$

hence only ξ_i's corresponding to active constraints are possibly nonzero. Remembering that x is regular, we conclude from $\xi_0 + \xi_1 + \ldots + \xi_m = 1$ that $\xi_0 > 0$. Dividing (1.2) by ξ_0 one sees that $x \in S$. This means that $x \notin S$ implies the optimal value of (LP), $\alpha \neq 0$. Actually $\alpha < 0$ because one can take v_i or $-v_i$ as well in the problem (LP). Hence, it follows from the inequality constraints of (LP) that v_* is a feasible direction with

$$\nabla f(x) v_* < 0 .$$

This implies that $f(x_{k+1}) < f(x_k)$ and the proof is complete. ■

Gradient Projection Method (Rosen's Method)

For the sake of simplicity of presentation, assume that $g_1,\ldots,g_m$ are linear. In this method feasible directions are obtained by projecting the steepest descent direction $-\nabla f(x)$ on the intersection of constraint hyperplanes, hence there is no need to solve problem (LP) at every iteration. How to calculate the projection of $-\nabla f(x)$?

Assume that at a feasible solution x_k, $g_1,\ldots,g_p$, $P \leq m$ are active constraints and $\nabla g(x_k),\ldots,\nabla g_p(x_k)$ are linearly independent. Denote

$$H = \{x \in R^n : \nabla g_i(x_k)\, x = 0 , \; i = 1,\ldots,p \},$$

then $x_k \in H$. The projection of $-\nabla f(x_k)$ on H can be expressed as

$$v_k = -\nabla f(x_k) - \sum_{i=1}^{P} \alpha_i \nabla g_i(x_k)$$

$$= -\nabla f(x_k) - \alpha A , \qquad (1.3)$$

where A is the matrix with columns $\nabla g_1(x_k),\ldots,\nabla g_p(x_k)$ and $\alpha = (\alpha_1,\ldots,\alpha_p)$. Moreover, since $v_k \in H$, it is orthogonal to all $\nabla g_i(x_k)$, $i = 1,\ldots, p$ and we have

$$A^T v_k = 0 .$$

Hence

$$\alpha = - \nabla f(x_k) A^T (AA^T)^{-1} ,$$

$$v_k = \nabla f(x_k) V , \qquad (1.4)$$

where

$$V = I - A^T(AA^T)^{-1}A$$

called the projection matrix .

We have two cases to consider :

i) $v_k \neq 0$. It follows from (1.3) that

$$\nabla f(x_k)v_k = - |\nabla f(x_k)|^2 ,$$

which means that v_k is a direction of descent.

ii) $v_k = 0$. Then $\nabla f(x_k) + \alpha A = 0$. One has two subcases

ii_a) $\alpha \geq 0$. Then x_k satisfies the Kuhn-Tucker condition, in other words, $x_k \in S$ where S is defined in Theorem 1.4.

ii_b) there is an index j with $\alpha_j < 0$, say $j=p$. Then denote A_p the matrix obtained from A by dropping the column $\nabla g_p(x_k)$, and use the projection v_k' of $-\nabla f(x_k)$ on the subspace

$$H_p = \{x \in R^n : x\nabla g_i(x_k) = 0 \ , \ i = 1,\ldots,p-1\}$$

to obtain the relations

$$-\nabla f(x_k) = \alpha A \tag{1.5}$$

$$-\nabla f(x_k) = v_k' + \alpha A_p \ . \tag{1.6}$$

It is clear that $v_k' \neq 0$ (because if $v_k' = 0$, (1.5) and (1.6) should imply $\alpha_p = 0$). Moreover,

$$0 > \nabla f(x_k)v_k' = - \alpha_p \nabla g_p(x_k)v_k'$$

(this is derived from (1.5) and the fact that $A_p v_k' = 0$), which implies that

$$\nabla g_p(x_k)v_k' < 0 \ .$$

In this way, v_k' is a feasible direction of descent.

Now the algorithm can be formulated as:

1) Start with an initial feasible solution x_1 . Set $k=1$.

2) Calculate

$$v_k = -\nabla f(x_k)V \ ,$$

where $V = I - A^T(AA^T)^{-1}A$

3) If $v_k \neq 0$, solve

$$\min\{f(x_k+tv_k) : t > 0 , x_k + tv_k \text{ is feasible}\} .$$

Let x_{k+1} be the optimal point.

If $v_k = 0$, find α by $\alpha = -\nabla f(x_k) A^T(AA^T)^{-1}$;

If $\alpha \geq 0$, stop : $x_k \in S$;

If not, go to 2) with the new A by deleting the column corresponding to the most negative component of α .

This algorithm is in general not closed. Therefore no convergence proof is available. Nevérthless the method has been used successfully for many nonlinear problems and no examples showing that it diverges in practice.

6.2 Penalty Function Methods

Suppose that we have a constrained problem

(P) $$\min \ f(x)$$
$$\text{s.t. } g_i(x) \leq 0 , \ i = 1,\ldots,m .$$

The crucial idea of penalty function methods is to solve an unconstrained problem with an augmented objective function $f(x) + P(x)$ instead of f, where the term $P(x)$ is to penalize infeasibility.

It is clear that if $P(x)$ is define by the rule

$$P(x) = \begin{cases} 0 & \text{if } x \text{ is feasible} \\ +\infty & \text{otherwise,} \end{cases}$$

then x_* is an optimal solution of (P) if and only if x_* solves the unconstrained problem

$$\begin{aligned} &\min \ f(x) + P(x) \\ &\text{s.t. } x \in R^n . \end{aligned}$$

However, since $P(x)$ is of bad quality, for instance it is discontinuous at the border of the feasible solution set, one tries to use other functions which approximate $P(x)$. The unconstrained problems obtained in this way will provide approximating solutions to (P). In the sequel it is assumed that f and g_i are continuous, and (P) has optimal solutions.

Interior Penalty Function Method

In this method one defines the augmented objective function by

$$\varphi(x,\alpha) = f(x) + \alpha \sum_{i=1}^{m} P_0(g_i(x)) ,$$

where $\alpha > 0$ is a penalty parameter and $P_0(y)$ is a continuous function with the property:

$$P_0(y) > 0 \qquad \text{for } y < 0$$

$$\lim_{y \to 0} P_0(y) = \infty \ .$$

It is common to take $P_0(y) = -1/y$, but sometimes one uses also $P_0(y) = \ln(-y)$.

The interior penalty function algorithm is described next:

1) Start with a feasible point x_0 which satisfies

$$g_i(x_0) < 0 \ , \ i = 1,\ldots,m \ .$$

Choose $\alpha_1 > 0$ and set the iteration number $k=1$.

2) Solve the problem

$$\min_{\text{s.t. } x \in R^n} \varphi(x,\alpha_k) = f(x) - \alpha \sum_{i=1}^{m} \frac{1}{g_i(x)}$$

by unconstrained minimization techniques with the initial point x_{k-1}, to obtain a minimum x_k.

3) Test the optimality of x_k. If it is an optimal solution of (P), stop. If not, go to the next step.

4) Set a new penalty parameter $\alpha_{k+1} = c\alpha_k$ with $c < 1$ and set $k = k+1$, $x_{k-1} = x_k$ and go to 2).

Theorem 2.1

The method described above has the following properties:

i) $\varphi(x,\alpha) \geq f(x)$ for all $\alpha > 0$, x feasible .

ii) $g_i(x_k) < 0$, $i = 1,\ldots,m$.

iii) $\{\varphi(x_k,\alpha_k)\}$ tends to the optimal value of (P) and any cluster point of $\{x_k\}$ is an optimal solution of (P) .

Proof. The first property is obvious. For the second property it suffices to note that by definition of φ , if $g_i(x_k) = 0$, then the value of $\varphi(x_k,\alpha_k)$ cannot be finite.

For the last property, let x_* be an optimal solution of (P) . We prove that

$$\lim_{k\to\infty} \varphi(x_k,\alpha_k) = f(x_*) .$$

Since $\{\alpha_k\}$ decreases to 0 , $\{\varphi(x_k,\alpha_k)\}$ is decreasing also. In view of i), it is bounded below, hence it converges to some number, say $\varphi_0 \geq f(x_*)$.

If $\varepsilon = \varphi_0 - f(x_*)$ is positive, then one can choose a point $\bar{x}$ such that

$$f(\bar{x}) < f(x_1) + \varepsilon/2$$

$$g_i(\bar{x}) < 0 \ , \ i = 1,\ldots,m \ .$$

Take k to be large enough such that

$$-\alpha_k \sum \frac{1}{g_i(\bar{x})} < \varepsilon/2$$

and

$$\varphi(x_k,\alpha_k) - \varphi_0 < \varepsilon/2 \ .$$

With these inequalities in hand we obtain:

$$\begin{aligned}
\varphi_0 &\leq \varphi(x_k,\alpha_k) \\
&\leq \varphi(\bar{x},\alpha_k) \\
&\leq f(\bar{x}) - \alpha_k \sum_{i=1}^{m} \frac{1}{g_i(\bar{x})} \\
&< f(\bar{x}) + \varepsilon/2 \\
&\leq f(x_*) + \varepsilon \\
&= \varphi_0 - f(x_*) + f(x_*) \\
&= \varphi_0 \ .
\end{aligned}$$

which is impossible. Hence $\varphi_0 = f(x_*)$.

Now, if $\bar{x}$ is any cluster point of $\{x_k\}$, say $\bar{x} = \lim x_{k_\ell}$, then by i)

$$\varphi(x_{k_\ell}, \alpha_{k_\ell}) \geq f(x_{k_\ell}) .$$

Hence

$$\varphi_0 = \lim \varphi(x_{k_\ell}, \alpha_{k_\ell}) \geq f(\bar{x}) .$$

This implies that $\bar{x}$ is also a minimum point of f on the feasible solution set. ■

Note that in this method minimal points of $\varphi(x, \alpha_k)$ are in the region of feasible solutions of (P) .

Exterior Penalty Function Method

In this method the augmented objective function is also defined by

$$\varphi(x, \alpha) = f(x) + \alpha \sum_{i=1}^{m} {}_0(g_i(x)).$$

The function $P_0(y)$ must be a continuous function with the property.

$$P_0(y) = 0 \quad \text{if and only if} \quad y \leq 0$$

$$P_0(y) \geq 0 \quad \text{for all } y \in R .$$

It is common to take

$$P_0(y) = \max[0,y] \text{ , or } P_0(y) = \max[0,y]^2 .$$

The algorithm can be described as follows:

1) Start with any $x_0 \in R^n$ and choose $\alpha_1 \in R^1$.
   Set the iterative number $k=1$.

2) Solve the unconstrained problem

$$\min f(x) + \alpha_k \sum_{i=1}^{m} \max [0,g_i(x)]^2$$

   to obtain an optimum point, say x_k .

3) Check the feasibility of x_k . If it is feasible, stop (because x_k is then also an optimal solution of (P)). Otherwise go to 4).

4) Choose $\alpha_{k+1} > \alpha_k$. Set $k=k+1$ and go to 2) .

Theorem 2.2

The following assertions are true:

i) $\varphi(x,\alpha) \geq f(x)$ for all $\alpha > 0$, $x \in R^n$.

ii) $\{\varphi(x_k,\alpha_k)\}$ is increasing if so is α_k ;

iii) $\{\varphi(x_k,\alpha_k)\}$ converges to the optimal value of (P) as α_k tends to ∞ , and any cluster point of $\{x_k\}$ is an optimal solution of (P) .

Proof. The proof is similar to that of the preceding theorem, so we omit it. ■

6.3 Cutting Plane Method

The basic idea of this method is to approximate the nonlinear objective and the constraints by linear functions and then solve the obtained linear problem by linear programming techniques.

Consider the following problem:

$$(P) \qquad \begin{aligned} &\min \quad f(x) \\ &\text{s.t.} \quad x \in X , \end{aligned}$$

where X is a closed convex set in R^n. This problem can be converted to a problem with a linear objective by introducing a new variable y:

$$\begin{aligned} &\min \quad y \\ &\text{s.t.} \quad x \in X , \; y \in R \\ &\qquad f(x) - y \leq 0 . \end{aligned}$$

It is evident that the two above problems are equivalent to each other. In the case f is convex, the constraint set of the latter problem is also convex. Thus, without loss of generality we may consider (P) with $f(x) = cx$, a linear function.

A general scheme of the cutting plane algorithm is as follows:

1) Choose a polytope P_k containing X.

Solve the problem

$$\min \ cx$$
$$\text{s.t. } x \in P_k .$$

Let x_k be an optimal solution.

If $x_k \in X$, it is also an optimal solution of (P). Stop. Otherwise go to 2)

2) Find a hyperplane

$$H_k = \{x \in R^n : a_k x \leq b_k\}$$

such that $X \subseteq H_k$ and $x_k \notin H_k$.

Update P_k to obtain P_{k+1} including as a constraint

$$a_k x \leq b_k .$$

Specific algorithms indicate how to choose H_k and how to update P_k.

Kelley's Algorithm

We solve the problem

$$(CP) \qquad \min \quad cx$$
$$\text{x.t.} \quad g_i(x) \leq 0 \ , \ i = 1,\ldots,m \ ,$$

where $g_1,\ldots,g_m$ are convex differentiable.

An important property of convex functions that we shall use is

$$g(x) \geq g(y) + \nabla g(y)(x-y) \ , \text{ for all } x,y \ . \tag{3.1}$$

The algorithm is as follows

1) Solve

$$\min \quad cx$$
$$\text{s.t.} \quad x \in P_k$$

where P_k is an initial polytope containing the feasible solution set of (CP).

Let x_k be an optimal solution. If x_k is a feasible solution of (CP), stop. Otherwise go to 2).

2) Let i be an index maximizing $g_i(x_k)$. Define

$$P_{k+1} = P_k \cap \{x \in R^n : g_i(x_k) + \nabla g_i(x_k)(x-x_k) \leq 0\} \ . \tag{3.2}$$

Return to step 1).

Proposition 3.1

Assume that the feasible solution set of (CP) is non-empty. Then P_{k+1} is a new polytope which contains the feasible solution set and does not contain x_k.

Proof. Observe first that $\nabla g_i(x_k) \neq 0$ because otherwise x_k would be a minimum point of g_i and this would imply that (CP) has no feasible solutions. Furthermore, if x is feasible, then $g(x) \leq 0$ and by (3.1),

$$g_i(x_k) + \nabla g_i(x_k)(x-x_k) \leq g_i(x) \leq 0 ,$$

which shows that $x \in P_{k+1}$. Finally, since $g_i(x_k) > 0$, one concludes that $x_k \notin P_{k+1}$. ∎

Theorem 3.2

Let $g_1,\ldots,g_m$ be continuously differentiable. Any limit point of the sequence $\{x_k\}$ obtained by Kelley's algorithm is an optimal solution of (CP).

Proof. Let $\{x_{k_j}\}$ be a subsequence of $\{x_k\}$ which converges to x_0. Without loss of generality, one can assume that the index i in the second step of the algorithm is the same throughout the subsequence. We show first that x_0 is feasible. It suffices to show that $g_i(x_0) \leq 0$ because i is the index maximizing $g_j(x_k)$, $j = 1,\ldots,m$. To this end, derive from (3.1) with $y = x_{k_j}$, $x = x_{k_{j'}}$ for $j' > j$ that

$$g_i(x_{k_j}) \leq \|\nabla g_i(x_{k_j})\| \; |x_{k_{j'}} - x_{k_j}| \; . \tag{3.3}$$

Since g_i is continuous differentiable and $x_{k_j} \to x_0$, $\|\nabla g_i(x_{k_j})\|$ is bounded with respect to j, and the right hand side of (3.3) tends to zero as j, j' tend to ∞. It follows from (3.3) that

$$g_i(x_0) = \lim_j g_i(x_{k_j}) \leq 0 \; .$$

Moreover, since the feasible solution set is contained in P_k, one has

$$cx_{k_j} \leq cx \text{ , for every feasible solution } x \text{ ,}$$

and for $j = 1,2,\ldots$. This implies that for any feasible solution x,

$$xc_0 \leq cx \text{ ,}$$

which completes the proof. ■

6.4 Problems

1. Give examples to show that the maps M and D in (1.1) are not closed.

2. Use the penalty function $\alpha/g(x)$ to solve the problem

$$\begin{aligned} &\min \quad x \\ &\text{s.t.} \quad 0 \leq x \leq 1 . \end{aligned}$$

3. Give a study of penalty function method by using a penalty function $\ln(-y)$.

Esta obra se terminó de imprimir en el mes de marzo de 1992, en la Sección de Offset del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional.
La edición consta de 300 ejemplares