0.1 Subdifferentials

In this section X will denote a Banach space, X^* its dual and $\langle \cdot, \cdot \rangle$ the pairing between the elements of X^* and those of X. Monotone operators play the role of increasing or dicreasing functions from \mathbb{R} to \mathbb{R} .

Definition 1 An operator $T: X \to X^*$ is called monotone if

$$\langle Tx - Ty, x - y \rangle \ge 0,$$

strictly monotone if

$$\langle Tx - Ty, x - y \rangle > 0,$$

and strongly monotone if

$$\langle Tx - Ty, x - y \rangle \ge \alpha \left\| x - y \right\|^2$$

for some $\alpha > 0$ and all $x, y \in X$.

Definition 2 The graph of a multivalued operator $T: X \to 2^{X^*}$ is defined to be

$$G(T) = \{(x,\xi) : x \in X, \ \xi \in Tx\}.$$

Definition 3 A multivalued operator $T: X \to 2^{X^*}$ is called monotone if

$$\langle \xi - \eta, x - y \rangle \ge 0$$

strictly monotone if

$$\langle \xi - \eta, x - y \rangle > 0,$$

and strongly monotone if

$$\langle \xi - \eta, x - y \rangle \ge \alpha \|x - y\|^2$$

for some $\alpha > 0$, all $x, y \in X$ and all $\xi \in Tx, \eta \in Ty$.

Definition 4 $T: X \to 2^{X^*}$ is called maximal monotone if it is monotone and has no proper extension to a monotone operator. In other words, if $S: X \to 2^{X^*}$ is a monotone operator such that $G(S) \supseteq G(T)$ then G(S) = G(T).

Definition 5 An operator $F : X \to \mathbb{R}$ is called subdifferentiable at $x \in X$ if it has a linear minorant at F(x). In other words, if there exists a $\xi \in X^*$ such that

$$\langle \xi, y - x \rangle \le F(y) - F(x) \quad \forall y \in X$$

The set of all such ξ is calle the subdifferential of F at x and is denoted by $\partial F(x)$. If F does not have a subdifferential at x then $\partial F(x) = \phi$.

Roughly speaking, the subdifferential is the set of all "slopes" of lines that can be drawn supporting the epigraph of the operator F at the point (x, F(x)).

Proposition 6 $\partial F(x)$ is convex and w^* -closed.

Proof. Let $\xi, \eta \in \partial F(x), \alpha \in [0,1]$, then

$$\begin{aligned} \langle \alpha \xi + (1 - \alpha) \eta, y - x \rangle &= \alpha \langle \xi, y - x \rangle + (1 - \alpha) \langle \eta, y - x \rangle \\ &\leq \alpha \left(F(y) - F(x) \right) + (1 - \alpha) \left(F(y) - F(x) \right) \\ &= F(y) - F(x) \end{aligned}$$

hence, $\alpha \xi + (1 - \alpha) \eta \in \partial F(x)$. To show the w^* -closedness, let $\xi_n \in \partial F(x)$ such that $\xi_n \xrightarrow{w^*} \xi$, then, for all $y \in X$,

$$\langle \xi_n, y - x \rangle \le F(y) - F(x)$$

and hence,

$$\langle \xi, y - x \rangle \le F(y) - F(x),$$

i.e., $\xi \in \partial F(x)$.

Proposition 7 $\partial F(\cdot)$ is a monotone operator

Proof. Suppose $\xi \in \partial F(x)$ and $\eta \in \partial F(y)$, then

$$\begin{array}{rcl} \langle \xi, y - x \rangle + F\left(x \right) & \leq & F\left(y \right), \\ \langle \eta, x - y \rangle + F\left(y \right) & \leq & F\left(x \right) \end{array}$$

The second inequality can be rearranged as

$$\langle -\eta, y - x \rangle - F(x) \le -F(y)$$

which, when added to the first inequality yeilds

$$\langle \xi - \eta, y - x \rangle \leq 0$$

or

$$\langle \xi - \eta, x - y \rangle \ge 0.$$

In the proof of the following proposition we will need the following version of the Hahn-Banach Theorem.

Theorem 8 Let Y be a real locally convex space, A, B be two nonempty convex sets in Y such that A is open and $A \cap B = \phi$. Then there exists a closed hyperplane M that separates A and B. In other words, there exist $\xi \in Y^*$ and $\beta \in \mathbb{R}$ such that $\langle \xi, y \rangle + \beta > 0$ for all $y \in A$, $\langle \xi, y \rangle + \beta = 0$ for all $y \in M$ and $\langle \xi, y \rangle + \beta \leq 0$ for all $y \in B$.

Proposition 9 Suppose $F: X \to \overline{\mathbb{R}}$ is convex, $x \in \text{dom}(F)$ and F is continuous at x then $\partial F(x) \neq \phi$.

Proof. The idea is to obtain a hyperplane M that supports the epigraph of F. Since F is continuous at x, there exists an open set \mathcal{O} such that $x \in \mathcal{O}$ and F is bounded above, say by $c < \infty$ on \mathcal{O} . Notice that $\mathcal{O} \subset \text{dom}(F)$ and $\mathcal{O} \times (c, \infty) \subset \text{epi}(F)^{\circ}$ and thus $\text{epi}(F)^{\circ} \neq \phi$. $\text{epi}(F)^{\circ}$ is convex since F is convex. Also, $(x, F(x)) \notin \text{epi}(F)^{\circ}$. Hence, by the Hahn-Banach Theorem, there exists a closed hyperplane M in $X \times \mathbb{R}$ that supports $\text{epi}(F)^{\circ}$ and contains (x, F(x)). The equation of M can be written as $\langle \xi, y \rangle + mr + \beta = 0$ for some $\xi \in X^*$, $m, \beta \in \mathbb{R}$ (not all zeros) and all $(y, r) \in M$. Furthermore,

$$\langle \xi, y \rangle + mr + \beta > 0 \tag{1}$$

for all $(y, r) \in \operatorname{epi}(F)^{\circ}$ and

$$\langle \xi, x \rangle + mF(x) + \beta = 0.$$

Hence, $\beta = -\langle \xi, x \rangle - mF(x)$ and (1) becomes

$$\langle \xi, y - x \rangle + m \left(r - F \left(x \right) \right) > 0 \ \forall \left(y, r \right) \in \operatorname{epi}\left(F \right)^{\circ}.$$
⁽²⁾

If m = 0 we get $\langle \xi, y - x \rangle > 0$ for all $y \in \mathcal{O}$ (inparticular). Since \mathcal{O} is open, then $\xi = 0$ and $\beta = 0$, which is a contradiction. Hence, we may assume withour loss of generality that m > 0. In this case, (2) may be rearranged as

$$\left\langle -\frac{\xi}{m}, y - x \right\rangle + F(x) < r \ \forall (y, r) \in \operatorname{epi}(F)^{\circ}.$$

Therefore,

$$\left\langle -\frac{\xi}{m}, y - x \right\rangle + F(x) \le r \ \forall (y, r) \in \operatorname{epi}(F)$$

and, in particular

$$\left\langle -\frac{\xi}{m}, y - x \right\rangle + F(x) \le F(y).$$

Thus, $-\frac{\xi}{m} \in \partial F(x)$.

Exercise 10 If $F: X \to \overline{\mathbb{R}}$ is convex then $F(x+th)-F(x) \leq t (F(x+h)-F(x))$ for sufficiently small t > 0.

Exercise 11 If $F: X \to \overline{\mathbb{R}}$ is convex then

$$\frac{F\left(x+th\right)-F\left(x\right)}{t}$$

is increasing as a function of t. Hence, it has a limit as $t \to 0^+$ (which may be $\pm \infty$).

Exercise 12 Denote the limit in the previous exercise by d(x; h). Then for any $h \in X$

$$F(x) + \lambda d(x;h) \le F(x + \lambda h) \quad \forall \lambda \in \mathbb{R}.$$

Exercise 13 Let $x, y \in X$ and

$$\mathcal{L} = \{ x + \lambda y : \lambda \in \mathbb{R} \}.$$

If $\Pi : \langle \xi, u \rangle + \beta = 0$ is a hyperplane such that $\langle \xi, u \rangle + \beta \ge 0$ for all $u \in \mathcal{L}$, then $\langle \xi, u \rangle + \beta = 0$ for all $u \in \mathcal{L}$. In other words, if a hyperplane contains a straight line on one side then it must actually contain the line in it.

0.2 Relationship to the Gateaux Derivative

Notice that from the definitions of the Gateaux derivative and the subdifferentials, the F is assumed to be finite at the point of evaluation x. This will also be the case with convex functions.

Proposition 14 Suppose $F: X \to \overline{\mathbb{R}}$ is convex. If F has a Gateaux derivative at $x \in \text{dom } F$, then $\partial F(x) = \{F'(x)\}$. Conversely, if F is continuous at $x \in X$ and has a unique subdifferential $\{\xi\}$ at x, then F is Gateaux differentiable at x and $F'(x) = \xi$.

Proof. Suppose F has a Gateaux derivative at $x \in X$. we will show that $F'(x) \in \partial F(x)$.

$$\langle F'(x), th \rangle = F(x+th) - F(x) + o(t).$$

Thus, for sufficiently small t > 0, we have, using Exercise 10,

$$\langle F'(x),h\rangle \leq F(x+h) - F(x) + \frac{o(t)}{t}.$$

Taking the limit as $t \to 0^+$ we get

$$\langle F'(x),h\rangle \leq F(x+h) - F(x),$$

i.e., $F'(x) \in \partial F(x)$. On the other hand, suppose $\xi \in \partial F(x)$, then

$$\langle \xi, th \rangle + F(x) \leq F(x+th) \\ = \langle F'(x), th \rangle + o(t)$$

.

For t > 0 we get

$$\langle \xi, h \rangle \le \langle F'(x), h \rangle + \frac{o(t)}{t}$$

which gives

$$\langle \xi, h \rangle \le \langle F'(x), h \rangle.$$

For t < 0 we get

$$\langle \xi, h \rangle \ge \langle F'(x), h \rangle + \frac{o(t)}{t}$$

which gives

$$\langle \xi, h \rangle \ge \langle F'(x), h \rangle.$$

Combining the two inequalities gives

$$\langle \xi, h \rangle = \langle F'(x), h \rangle.$$

Hence $\xi = F'(x)$.

To show the converse, observe first that, for sufficiently small t > 0,

$$\langle \xi, th \rangle \le F(x+th) - F(x) \le t(F(x+h) - F(x))$$

Therefore,

$$\langle \xi, h \rangle \leq \frac{F(x+th) - F(x)}{t} \leq F(x+h) - F(x)$$

and, because F is continuous at x, it follows that d(x; h) is finite. Next observe that from Exercise 12, for sufficiently small $h \in X$

$$F(x) + \lambda d(x;h) \le F(x + \lambda h) \quad \forall \lambda \in \mathbb{R}.$$

Geometrically, this means that the line

$$L = \{(x, F(x)) + \lambda (h, d(x; h)) : \lambda \in \mathbb{R}\}$$

in $X \times \mathbb{R}$ does not intersect (epi F)[°]. Since the latter set is not empty we get by the Hahn-Banach theorem, the existence of a hyperplane $\Pi : \langle \eta, y \rangle + mr + \beta = 0$ that seperates (epi F)[°] and L. By Exercise 13, Π contains L. i.e.,

$$\langle \eta, x + \lambda h \rangle + m \left(F(x) + \lambda d(x; h) \right) + \beta = 0 \ \forall \lambda \in \mathbb{R}.$$

It follows that $m \neq 0$ and $d(x;h) = \left\langle \frac{\eta}{-m}, h \right\rangle$. i.e.,

$$d(x;h) = \lim_{t \to 0^+} \frac{F(x+th) - F(x)}{t}$$

is a continuous linear functional on X. Hence, d(x;h) = F'(x). From the uniqueness of the subdifferential of F, we get $F'(x) = \xi$.

Exercise 15 Give a direct proof that the operator d(x;h) is linear and continuous.

In terms of the Gateaux derivative, a convex function is charachterized as follows

Proposition 16 Suppose F is defined on the convex set $A \subset X$ inot \mathbb{R} . If F is Gateaux differentiable on A then the following are equivalent

- 1. F is convex
- 2. $F(y) \ge F(x) + \langle F'(x), y x \rangle$ for all $x, y \in \mathcal{A}$.

Proof. 1. \implies 2. is clear from the previous proposition since the Gateaux derivative is a subdifferentiable. To show the converse, rewrite the inequality with x replaced with $\alpha x + (1 - \alpha) y$, then

$$F(y) \ge F(\alpha x + (1 - \alpha)y) + \langle F'(\alpha x + (1 - \alpha)y), \alpha(y - x) \rangle$$

and similarly

$$F(x) \ge F(\alpha x + (1 - \alpha)y) + \langle F'(\alpha x + (1 - \alpha)y), (1 - \alpha)(x - y) \rangle.$$

Multiplying the first inequality by $(1 - \alpha)$ and the second by α and adding we get

$$(1 - \alpha) F(y) + \alpha F(x) \ge F(\alpha x + (1 - \alpha) y).$$

i.e., F is convex.

In terms of the monotonicity of the Gateaux derivative, the convexity of a function F is characterized as follows

Proposition 17 Suppose F is defined on the convex set $A \subset X$ inot \mathbb{R} . If F is Gateaux differentiable on A then the following are equivalent

- 1. F is convex
- 2. F'(x) is monotone on A.

Proof. If F is convex then its Gateaux derivative is a subdifferential and we saw before (Porposition 7) that the subdifferential is a monotone operator. On the other hand, suppose F'(x) is monotone on \mathcal{A} . The function

$$\varphi(\lambda) = F(x + \lambda(y - x)), \ \lambda \in [0, 1].$$

is differentiable and

$$\varphi'(\lambda) = \langle F'(x + \lambda(y - x)), y - x \rangle$$

If $\lambda_1 \leq \lambda_2$ then

$$\langle F'(x+\lambda_2(y-x)) - F'(x+\lambda_1(y-x)), (\lambda_2-\lambda_1)y - x \rangle \ge 0,$$

which implies that

$$\langle F'(x+\lambda_2(y-x)) - F'(x+\lambda_1(y-x)), y-x \rangle \ge 0,$$

i.e.,

$$\varphi'(\lambda_2) - \varphi'(\lambda_1) \ge 0.$$

Hence, $\varphi'(\cdot)$ is increasing. Therefore, $\varphi(\cdot)$ is convex (see exercise below) and

$$\varphi(\lambda) \le (1 - \lambda)\varphi(0) + \lambda\varphi(1).$$

i.e.,

$$F(x + \lambda (y - x)) \le (1 - \lambda) F(x) + \lambda F(y).$$

Exercise 18 If $\varphi' : I \to \mathbb{R}$ is increasing, then φ is convex.