SML Recitation Notes:
The subdifferential

Martin Azizyan

February 1, 2012

Throughout this document, we will assume $f : \mathbb{R}^p \to \mathbb{R}$ is convex.

Definition 1. The *subdifferential* of f at a point $x \in \mathbb{R}^p$ is

$$\partial f(x) = \{y : \forall z \in \mathbb{R}^p \ f(z) \geq f(x) + y^T(z - x)\}.$$

We see from the definition that the subdifferential is a closed set. The subdifferential is also always non-empty (**for convex functions only!**). This is a consequence of the Supporting Hyperplane Theorem, which states that at any point on the boundary of a convex set, there exists at least one supporting hyperplane (i.e. a plane such that one of the half-spaces it defines contains the entire set). Since the epigraph of a convex function is a convex set, we can apply the Supporting Hyperplane Theorem to the set of points $(x, f(x))$, which are exactly the boundary points of the epigraph.

Similar statements can be made if the domain of f is a subset of $\mathbb{R}^d$. In this case the definition must be changed accordingly, and the subdifferential won't be defined (or will be empty, depending on convention) for points outside the domain, or in its boundary (why do we have to exclude the boundary?).

Subdifferentials are useful in convex optimization problems. The following simple result shows why.

Theorem 2. $\mathbf{0} \in \partial f(x)$ *if and only if* x *minimizes* f.

Proof. **"If":** Suppose x is a minimum of f. By definition $\forall z \in \mathbb{R}^p$ we have $f(z) \geq f(x) = f(x) + \mathbf{0}^T(z - x)$, so $\mathbf{0} \in \partial f(x)$.
"Only if": Suppose $\mathbf{0} \in \partial f(x)$. Then by definition of the subdifferential, $\forall z \in \mathbb{R}^p$ we have $f(z) \geq f(x) + \mathbf{0}^T(z - x) = f(x)$, so x is a minimum of f. $\square$

The geometrical interpretation of $\partial f(x)$ as the set of supporting hyperplanes suggests similarity to the differential $\nabla f(x)$, which describes the tangent hyperplane to f (when it exists). The similarity can also be seen algebraically, which leads to the following result.

Theorem 3. *If* f *is differentiable at* x*, then* $\partial f(x) = \{\nabla f(x)\}$. *Conversely, if* $\partial f(x) = \{g\}$ *for some* $g \in \mathbb{R}^p$*, then* $\nabla f(x)$ *exists and is equal to* g.

We won't prove this statement here, though if you recall the definition of the differential it should be clear why it is true.

Calculating the subdifferential of a function can sometimes be tricky. Here are some results that are often useful.

Theorem 4. *1. For $\alpha \geq 0$, $\partial(\alpha f)(x) = \{\alpha y : y \in \partial f(x)\}$.*

2. Let $h(x) = f(Ax + b)$. Then $\partial h(x) = \{A^T y : y \in \partial f(Ax + b)\}$.

3. For any $x \in \mathbb{R}^p$, $\partial f(x)$ is a convex set.

4. Let $f_1, ..., f_m$ convex, and suppose $f(x) = \sum_{i=1}^{m} f_i(x)$. Then

$$\partial f(x) = \left\{ \sum_{i=1}^{m} y_i : y_i \in \partial f_i(x) \right\}$$

5. Let $f_1, ..., f_m$ convex, and suppose $f(x) = \max_{i=1,...,m} f_i(x)$. For any fixed x, if $k \in \{1, ..., m\}$ such that $f(x) = f_k(x)$ (not necessarily unique), then $\partial f_k(x) \subseteq \partial f(x)$. In fact

$$\partial f(x) = \operatorname{conv}\left(\bigcup_{k: f(x) = f_k(x)} \partial f_k(x) \right)$$

Proof. 1. Obvious since $f(z) \geq f(x) + y^T(z - x)$ if and only if $\alpha f(z) \geq \alpha f(x) + (\alpha y)^T(z - x)$.

2. Similar to 1.

3. Let $y_1, y_2 \in \partial f(x)$. Then for any $z \in \mathbb{R}^p$ and $\theta \in (0, 1)$,

$$\begin{aligned} f(z) &\geq \max\left\{ f(x) + y_1^T(z - x), f(x) + y_2^T(z - x) \right\} \\ &\geq \theta \left[f(x) + y_1^T(z - x) \right] + (1 - \theta) \left[f(x) + y_2^T(z - x) \right] \\ &= f(x) + (\theta y_1 + (1 - \theta) y_2)^T (z - x) \end{aligned}$$

so $\theta y_1 + (1 - \theta) y_2 \in \partial f(x)$.

4. Let $y_i \in \partial f_i(x)$ for $i = 1, ..., m$. Write

$$f(z) = \sum_{i=1}^{m} f_i(z) \geq \sum_{i=1}^{m} \left[f_i(x) + y_i^T(z - x) \right] = f(x) + \left(\sum_{i=1}^{m} y_i \right)^T (z - x)$$

so $\sum_{i=1}^{m} y_i \in \partial f(x)$, and $\left\{ \sum_{i=1}^{m} y_i : y_i \in \partial f_i(x) \right\} \subseteq \partial f(x)$. We will not prove equality here.

5. For some x, let k s.t. $f(x) = f_k(x)$, and $y \in \partial f_k(x)$. Then $f(z) \geq f_k(z) \geq f_k(x) + y^T(z - x) = f(x) + y^T(z - x)$, so $\partial f_k(x) \subseteq \partial f(x)$. Applying this fact for all k s.t. $f(x) = f_k(x)$, and using statement 3, we have

$$\operatorname{conv}\left(\bigcup_{k: f(x) = f_k(x)} \partial f_k(x) \right) \subseteq \partial f(x).$$

Again, we won't prove the other direction.

□

The missing parts of the proofs of statements 4 and 5 are somewhat complicated, and can be found in Convex Analysis by Rockafellar.