On subdifferential calculus – highlights

September 2012

Fundamentals:

Working with $+\infty$ and $-\infty$:

- $\forall \alpha \in (-\infty, +\infty] \; \alpha + (+\infty) = (+\infty) + \alpha = +\infty$.
- $\forall \alpha \in [-\infty, +\infty) \; \alpha - (+\infty) = \alpha + (-\infty) = -\infty$.
- neither $(+\infty) - (+\infty)$ nor $(+\infty) + (-\infty)$ etc. defined!
- careful! $2 + (+\infty) = 3 + (+\infty) \not\Rightarrow 2 = 3$
- $\forall \alpha \in (0, +\infty] \; \alpha \cdot (+\infty) = +\infty$
- $\forall \alpha \in [-\infty, 0) \; \alpha \cdot (+\infty) = -\infty$
- By definition: $0 \cdot (+\infty) = 0 \cdot (-\infty) = 0$.
- $\forall \alpha \in \mathbb{R} \; \alpha/(+\infty) = \alpha/(-\infty) = 0$.
- neither $(+\infty)/(+\infty)$ nor $(+\infty)/(-\infty)$ etc. defined!
- $(+\infty)/(+\infty)$, etc. undefined.
- careful! $2/(+\infty) = 3/(+\infty) \not\Rightarrow 2 = 3$
Convex sets in $\mathbb{R}^n$:

**Definition A.1**: $S \subset \mathbb{R}^n$ is convex if

$$\forall x_1, x_2 \in S \forall \lambda \in [0,1] \lambda x_1 + (1 - \lambda) x_2 \in S.$$ 

Convex functions:

**Definition 2.1**: Let $S \subset \mathbb{R}^n$ be convex. Then $f : S \to (-\infty, +\infty]$ is convex on $S$ if

$$\forall x_1, x_2 \in S \forall \lambda \in [0,1] f(\lambda x_1 + (1 - \lambda) x_2) \leq \lambda f(x_1) + (1 - \lambda) f(x_2).$$

Also, $f$ is strictly convex on $S$ if

$$\forall x_1, x_2 \in S, x_1 \neq x_2 \forall \lambda \in (0,1) f(\lambda x_1 + (1 - \lambda) x_2) < \lambda f(x_1) + (1 - \lambda) f(x_2).$$

**Remark**: $f \neq -\infty$, so $\lambda(+\infty) + (1 - \lambda)(-\infty)$ cannot confuse us.

**Associated definition**: Let $f : S \to [-\infty, +\infty)$. Then: $f$ is (strictly) concave on $S \iff -f$ is (strictly) convex on $S$.

**Example**: $f_1(x) := p^t x + \alpha$ is affine, i.e., both convex and concave, on $\mathbb{R}^n$ for any $p \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. It is neither strictly convex nor strictly concave.

**Example**: $f_2(x) := \beta |x|^2$ is strictly convex on $\mathbb{R}^n$ if $\beta > 0$. It is strictly concave on $\mathbb{R}^n$ if $\beta < 0$. 


Example (Exercise 2.1c): Let $S := \mathbb{R}_+$. Define $f_3 : S \to (-\infty, +\infty]$ by $f_3(x) := 1/x$ if $x > 0$ and by $f_3(0) := \gamma$. Then $f_3$ can only be made convex on $S$ by setting $\gamma = +\infty$.

Example (Exercise 2.7b): Define $f_4 : \mathbb{R} \to (-\infty, +\infty]$ by $f_4(x) := 1 - \sqrt{1 - x^2}$ if $|x| \leq 1$ and $f_4(x) = +\infty$ if $|x| > 1$. Then $f_4$ is convex on $\mathbb{R}$.

Definition (Exercise 2.2): Let $S \subset \mathbb{R}^n$ be convex. Then $f : S \to (-\infty, +\infty]$ is quasiconvex on $S$ if

$$\forall \alpha \in \mathbb{R} S_\alpha := \{x \in S : f(x) \leq \alpha\}$$

is convex

Every convex function on $\mathbb{R}^n$ is quasiconvex, but not conversely.

Domain extension by adding values $+\infty$:

Exercise 2.5: Let $S \subset \mathbb{R}^n$ be convex. Let $f : S \to (-\infty, +\infty]$. Define $\hat{f} : \mathbb{R}^n \to (-\infty, +\infty]$ by

$$\hat{f}(x) := \begin{cases} f(x) & \text{if } x \in S \\ +\infty & \text{if } x \notin S. \end{cases}$$

Exercise: $\hat{f}$ convex on $\mathbb{R}^n \iff f$ convex on $S$.

Consequence: From now on we mainly consider convex functions on $\mathbb{R}^n$. This is thanks to working with $+\infty$!
**New habit:** Speak of “convex functions” instead of “convex functions on $\mathbb{R}^n$.

**Definition 2.2:** Let $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$. The *essential domain* of $f$ is defined by

$$\text{dom } f := \{x \in \mathbb{R}^n : f(x) < +\infty\}.$$ 

Exercise: $f$ convex $\Rightarrow$ dom $f$ is convex, but not conversely.

Connections between convex sets and convex functions:

*From convex sets to convex functions:*

**Definition 2.3:** Let $S \subset \mathbb{R}^n$. The *indicator function* $\chi_S$ of $S$ is defined by

$$\chi_S(x) := \begin{cases} 
0 & \text{if } x \in S \\
+\infty & \text{if } x \notin S.
\end{cases}$$

Exercise: $S$ convex set $\iff$ $\chi_S$ convex function.

*From convex functions to convex sets:*

**Definition 2.4:** Let $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$. The *epigraph* $\text{epi } f \subset \mathbb{R}^{n+1}$ is defined by

$$\text{epi } f := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq y\}.$$
Exercise: $f$ convex function $\iff$ epi $f$ convex set.

**Remark:** Many proofs of results for convex functions “work” on their convex epigraphs by means of separation results (see Appendix A).

**Example:** For $S \subset \mathbb{R}^n$ let $f := \chi_S$. Then $epi f = S \times \mathbb{R}_+.$

*From convex functions to more convex functions:*

Easy: Let $f_i : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ be convex and let $\alpha_i \in [0, +\infty]$ for $i = 1, \ldots, m$. Then $f(x) := \sum_{i=1}^{m} \alpha_i f_i(x)$ defines a convex function, as does $f(x) := \max_{1 \leq i \leq m} \alpha_i f_i(x)$.

**Exercise 2.6:** Let $S \subset \mathbb{R}^n$ be convex. Let $f : S \rightarrow \mathbb{R}$ be convex and let $g : D \rightarrow \mathbb{R}$ be convex and nondecreasing on a convex interval $D \subset \mathbb{R}$, with $D \supset f(S)$. Then $h(x) := g(f(x))$ defines a convex function $h : S \rightarrow \mathbb{R}$.

**Example (Exercise 2.7):** a. If $f : \mathbb{R}^n \rightarrow [0, +\infty]$ is convex on $\mathbb{R}^n$, then so is $f^2$. However, $f^2$ need not be convex if $f$ can also take negative values.

b. $f(x) := 1 - \sqrt{1 - x^2}$ is convex on $[-1, +1]$.

c. $f(x) := \exp(x^2)$ is convex on $\mathbb{R}$. 
Subdifferentials and subgradients of convex functions

**Definition 2.5:** Let $f : \mathbb{R}^n \to (-\infty, +\infty]$, $f \neq +\infty$, and let $x_0 \in \text{dom} f$ (so $f(x_0) \in \mathbb{R}$).

a. A *subgradient of $f$ at $x_0$* is a $\xi \in \mathbb{R}^n$ with
   $$f(x) \geq f(x_0) + \xi^t(x - x_0) \text{ for all } x \in \mathbb{R}^n.$$  

b. The *subdifferential of $f$ at $x_0$* is the set
   $$\partial f(x_0) := \{\xi \in \mathbb{R}^n : \xi \text{ is subgradient of } f \text{ at } x_0\}.$$  

This set may be empty!

**Observation:** If $x_0 \notin \text{dom} f$ (so $f(x_0) = +\infty$) then $\partial f(x_0) = \emptyset$. But $\partial f(x_0) = \emptyset$ is also possible for $x_0 \in \text{dom } f$.

**Example:** a. Let $f(x) := 1 - \sqrt{1 - x^2}$ on $[-1, +1]$ and define $f(x) := +\infty$ if $x < -1$ or $x > 1$. Then $f$ is convex and $1 \in \text{dom } f$. However, $\partial f(1) = \emptyset$.

b. Let $f(x) := |x|$ on $\mathbb{R}$. Then $\partial f(2) = \{1\}$, $\partial f(-3) = \{-1\}$ and $\partial f(0) = [-1, +1]$.

*For differentiable convex functions: “subgradient = gradient”:

**Proposition 2.6:** Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be convex. If $f$ is differentiable at $x_0 \in \text{int dom } f$, then $\partial f(x_0) = \{\nabla f(x_0)\}$. 
Here: “int” means “interior”.

**Example (Exercise 2.9b):** In previous example with \( f(x) = 1 - \sqrt{1 - x^2} \) on \([-1, +1]\) and \( f(x) = +\infty \) if \( x < -1 \) or \( x > 1 \), one has \( \partial f(x) = \{x/\sqrt{1 - x^2}\} \) for every \( x \in (-1, 1) \).

**How to determine convexity of functions:**

**Proposition 2.7:** Let \( S \subset \mathbb{R}^n \) be open and convex. Let \( f : S \to \mathbb{R} \).

1. If \( f \) is differentiable, then \( f \) is convex on \( S \) \( \iff \) \( \forall x_1, x_2 \in S (\nabla f(x_1) - \nabla f(x_2))^t (x_1 - x_2) \geq 0 \).

2. If \( f \) is differentiable, then \( f \) is strictly convex on \( S \) \( \iff \) \( \forall x_1, x_2, x_1 \neq x_2 (\nabla f(x_1) - \nabla f(x_2))^t (x_1 - x_2) > 0 \).

3. If \( f \) is twice continuously differentiable, then \( f \) is convex on \( S \) \( \iff \) the Hessian matrix

\[
H_f(x) := \left( \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{i,j}
\]

is positive semidefinite at every point \( x \) of \( S \).

4. If \( f \) is twice continuously differentiable, then \( H_f(x) \) is positive definite at every point \( x \) of \( S \) \( \Rightarrow \) \( f \) is strictly convex on \( S \).
**Definition:** An $n \times n$ matrix $M$ is *positive semidefinite* if $d^t M d \geq 0$ for all $d \in \mathbb{R}^n$. And $M$ is *positive definite* if $d^t M d > 0$ for all $d \in \mathbb{R}^n$, $d \neq 0$.

**Corollary 2.8:** Let $S \subset \mathbb{R}$ be open and convex. Let $f : S \to \mathbb{R}$.

(i) If $f$ is differentiable, then $f$ is convex [strictly convex] on $S$ $\iff$ $f'$ is nondecreasing [increasing] on $S$.

(ii) If $f$ is twice continuously differentiable, then $f$ is convex [strictly convex] on $S$ $\iff$ $[\implies] f''(x) \geq 0$ $[\implies] f''(x) > 0$ for all $x \in S$. 

MR-theorem and “small” KKT-theorem

**Theorem 2.9 (Moreau-Rockafellar)** Let $f, g : \mathbb{R}^n \to (-\infty, +\infty]$ be convex. Then
\[
\forall x_0 \in \mathbb{R}^n \partial f(x_0) + \partial g(x_0) \subset \partial (f + g)(x_0).
\]
Moreover, if $\text{int dom } f \cap \text{dom } g \neq \emptyset$. Then also
\[
\forall x_0 \in \mathbb{R}^n \partial (f + g)(x_0) \subset \partial f(x_0) + \partial g(x_0).
\]

**Comment:** First part is trivial. Proof of second part goes by separating hyperplane Theorem A.4, applied to disjoint convex sets $\Lambda_f$ and $\Lambda_g$ that are “epigraph-like” – see syllabus.

**Theorem 2.10 (“small KKT”):** Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and let $S \subset \mathbb{R}^n$ be nonempty convex. Consider the optimization problem
\[
(P) \quad \inf_{x \in S} f(x).
\]
Then
\[
\bar{x} \in S \text{ optimal for } (P) \iff \exists \xi \in \partial f(\bar{x}) \forall x \in S \xi^t(x - \bar{x}) \geq 0.
\]

**Sketch of proof.** Observe
\[
\bar{x} \in S \text{ optimal for } (P) \iff 0 \in \partial (f + \chi_S)(\bar{x}).
\]
Then apply MR-theorem to right side.