Part One: Optimal Transport

Katy Craig and Ashwin Trisal

May 30, 2020

Contents

7 Lecture 7 1

7.1 Subdifferentials ................................. 1

7.2 Convex optimization ............................ 2

7 Lecture 7

Last time: we did a crash course on convex analysis.

7.1 Subdifferentials

First: “convex \implies bowl-shaped”: recall that \( f \) is convex if, for any \( x_0, x_1 \in X, \alpha \in \mathbb{R} \), we have the inequality \( f((1-\alpha)x_0 + \alpha x_1) \leq (1-\alpha)f(x_0) + \alpha f(x_1) \). Given sufficient regularity on \( f \), which is not too restrictive, this is the same as asking for \( f'' \geq 0 \).

Definition 7.1. Given \( f : X \to \mathbb{R} \cup \{+\infty\} \), we define
\[
    f^*(y) = \sup_{x \in X} \langle y, x \rangle - f(x).
\]

Definition 7.2. Given \( f : X \to \mathbb{R} \cup \{+\infty\} \), we define the subdifferential
\[
    \partial f(x) := \{ y \in X^* \mid f(x') \geq f(x) + \langle y, x' - x \rangle + o(\|x' - x\|) \text{ as } x' \to x \}.
\]

We need \( f \) to be proper for the above definition: if \( f \) is constantly +\( \infty \), the definition is nonsensical.

If \( f \) is convex, \( \partial f(x) = \{ y \in X^* \mid f(x') \geq f(x) + \langle y, x' - x \rangle \forall x' \in X \} \).

The function \( x' \mapsto f(x) + \langle y, x' - x \rangle \) is affine in \( x' \); in real space, this is asking for the line through \( (x, f(x)) \) with slope \( y \). Saying \( f(x') \geq f(x) + \langle y, x' - x \rangle \) means that the line must lie below \( f \), so it is a supporting hyperplane.

If \( f \) is differentiable at \( x_0 \), \( \partial f(x_0) = \{ \nabla f(x) \} \).

*Based on lectures given by Katy Craig in Math 260L: Optimal Transport, spring quarter 2020, the University of California Santa Barbara
Consider \( f(x) = |x| \). This is convex, but not differentiable at 0. In fact, in this case, \( \partial f(0) = [-1, 1] \).

Side note: “subdifferential flows” are the right generalization of gradient flows.

Subdifferentials provide insight into convex conjugate functions via the following:

**Theorem 7.3.** Suppose \( f \) is proper, convex, and lower semicontinuous. Then \( y \in \partial f(x) \iff x \in \partial f^*(y) \).

**Proof.** It suffices to show \( \implies \), because \( f^{**} = f \). If \( y \in \partial f(x) \), then

\[
\begin{align*}
  f(x') &\geq f(x) + \langle x' - x, y \rangle \\
  \langle x, y \rangle - f(x) &\geq \langle x', y \rangle - f(x')
\end{align*}
\]

and by taking the supremum over \( x' \), we get \( \langle x, y \rangle - f(x) \geq f^*(y) \). By Young’s inequality, \( f^*(y') + f(x) \geq \langle y', x \rangle \). Summing these inequalities, we get \( f^*(y') \geq f^*(y) + \langle y' - y, x \rangle \forall y' \in X^* \). So \( x \in \partial f^*(y) \). \( \square \)

A similar theorem holds, even when we only have pointwise regularity information about the function.

**Theorem 7.4.** Suppose \( f \) is proper, convex, and lower semicontinuous at \( x_0 \). Then \( x_0 \in \partial f^*(y) \implies y \in \partial f^*(x_0) \).

**Example 7.5** (Mental image of a convex conjugate). Take \((X, \| \cdot \|) = (\mathbb{R}, | \cdot |) \cong (X^*, \| \cdot \|_{X^*})\).

<table>
<thead>
<tr>
<th>( f(x) )</th>
<th>( \partial f(x) )</th>
<th>( \partial f^*(x) )</th>
<th>( f^*(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{1}{2}</td>
<td>x</td>
<td>^2 )</td>
<td>( { x } )</td>
</tr>
</tbody>
</table>
| \( e^x \) | \( \{ e^x \} \) | \( \begin{cases} 
  \{ \ln x \} & \text{if } x > 0 \\
  \emptyset & \text{if } x \leq 0
\end{cases} \) | \( \begin{cases} 
  x \ln x - x & \text{if } x \geq 0 \\
  +\infty & \text{if } x < 0.
\end{cases} \) |
| \( |x| \) | \( 1_{x>0} + [-1, 1] \cdot 1_{x=0} - 1_{x<0} \) empty outside of \([-1, 1]\) | \( \chi_{[-1,1]}(x) = \begin{cases} 
  0 & \text{if } |x| \leq 1 \\
  +\infty & \text{if } |x| > 1
\end{cases} \) |

A sanity check to make sure we got the constants correct:

\[
f^*(0) = \sup_{x \in X} \langle x, 0 \rangle - f(x) = -\inf_{x \in X} f(x).
\]

**7.2 Convex optimization**

In convex optimization, we are given \( f \) a convex function on \( C \) a convex region. We want to solve

\[
\inf_{x \in C} f(x)
\]

with our questions being:

(i) What is the infimum?
(ii) Is the infimum attained?

(iii) If the minimizer exists, is it unique?

(iv) Can we characterize the minimizer (as a solution of a differential equation, perhaps?)

By defining \( \tilde{f}(x) = \begin{cases} f(x) & x \in C \\ \infty & x \notin C \end{cases} \), we can rewrite the problem as an infimum over all of \( X \), rather than the convex region \( C \).

The key tool, now, is observing how the problem changes under perturbations.

**Definition 7.6** (Primal and dual problems). Given \( X, U \) normed spaces, and \( F : X \times U \to \mathbb{R} \cup \{+\infty\} \), we define the

**primal problem**: \( \mathcal{P}_0 := \inf_{x \in X} f(x), \quad f(x) = F(x, 0) \)

and the

**dual problem**: \( \mathcal{D}_0 := \sup_{v \in U^*} g(v), \quad g(v) = -F^*(0, v) \).

\( F(x, u) \) encodes the perturbations of \( f(x) \) that we consider. We want to find simple, convex functions \( F \) so that either \( \mathcal{P}_0 \) or \( \mathcal{D}_0 \) match the original problem.

By Young’s inequality, \( \forall x \in X, y \in X^*, u \in U, v \in U^*, F(x, u) + F^*(y, v) \geq \langle y, x \rangle + \langle v, u \rangle \). In particular,

\[ F(x, 0) + F^*(0, v) \geq 0 \Leftrightarrow f(x) \geq g(v). \]

So the primal problem is always bigger than or equal to the dual problem.

We now seek conditions to prove \( \mathcal{P}_0 = \mathcal{D}_0 \): that is, we seek to close the “duality gap”.