Solution of the Monge problem II: Local approach

In the previous chapter, we tried to establish the almost sure single-valuedness of the \(c\)-subdifferential by an argument involving “global” topological properties, such as connectedness. Since this strategy worked out only in certain particular cases, we shall now explore a different method, based on local properties of \(c\)-convex functions. The idea is that the global question “Is the \(c\)-subdifferential of \(\psi\) at \(x\) single-valued or not?” might be much more subtle to attack than the local question “Is the function \(\psi\) differentiable at \(x\) or not?” For a large class of cost functions, these questions are in fact equivalent; but these different formulations suggest different strategies. So in this chapter, the emphasis will be on tangent vectors and gradients, rather than points in the \(c\)-subdifferential.

This approach takes its source from the works by Brenier, Rachev and Rüschendorf on the quadratic cost in \(\mathbb{R}^n\), around the end of the eighties. It has since then been improved by many authors, a key step being the extension to Riemannian manifolds, first addressed by McCann in 2000.

The main results in this chapter are Theorems 10.28, 10.38 and (to a lesser extent) 10.42, which solve the Monge problem with increasing generality. For Parts II and III of this course, only the particular case considered in Theorem 10.41 is needed.

A heuristic argument

Let \(\psi\) be a \(c\)-convex function on a Riemannian manifold \(M\), and \(\phi = \psi^c\). Assume that \(y \in \partial_c \psi(x)\); then, from the definition of \(c\)-subdifferential,
one has, for all $\tilde{x} \in M$,

$$
\begin{cases}
\phi(y) - \psi(x) = c(x, y) \\
\phi(y) - \psi(\tilde{x}) \leq c(\tilde{x}, y).
\end{cases}
$$

(10.1)

It follows that

$$
\psi(x) - \psi(\tilde{x}) \leq c(\tilde{x}, y) - c(x, y).
$$

(10.2)

Now the idea is to see what happens when $\tilde{x} \to x$, along a given direction. Let $w$ be a tangent vector at $x$, and consider a path $\varepsilon \to \tilde{x}(\varepsilon)$, defined for $\varepsilon \in [0, \varepsilon_0)$, with initial position $x$ and initial velocity $w$. (For instance, $\tilde{x}(\varepsilon) = \exp_x(\varepsilon w)$; or in $\mathbb{R}^n$, just consider $\tilde{x}(\varepsilon) = x + \varepsilon w$.) Assume that $\psi$ and $c(\cdot, y)$ are differentiable at $x$, divide both sides of (10.2) by $\varepsilon > 0$ and pass to the limit:

$$
-\nabla \psi(x) \cdot w \leq \nabla_x c(x, y) \cdot w.
$$

(10.3)

If then one changes $w$ to $-w$, the inequality will be reversed. So necessarily

$$
\nabla \psi(x) + \nabla_x c(x, y) = 0.
$$

(10.4)

If $x$ is given, this is an equation for $y$. Since our goal is to show that $y$ is determined by $x$, then it will surely help if (10.4) admits at most one solution, and this will obviously be the case if $\nabla_x c(x, \cdot)$ is injective. This property (injectivity of $\nabla_x c(x, \cdot)$) is in fact a classical condition in the theory of dynamical system, where it is sometimes referred to as a twist condition.

Three objections might immediately be raised. First, $\psi$ is an unknown of the problem, defined by an infimum, so why would it be differentiable? Second, the injectivity of $\nabla_x c$ as a function of $y$ seems quite hard to check on concrete examples. Third, even if $c$ is given in the problem and a priori quite nice, why should it be differentiable at $(x, y)$? As a very simple example, consider the square distance function $d(x, y)^2$ on the 1-dimensional circle $S^1 = \mathbb{R}/(2\pi\mathbb{Z})$, identified with $[0, 2\pi)$:

$$
d(x, y) = \min(|x - y|, 2\pi - |x - y|).
$$

Then $d(x, y)$ is not differentiable as a function of $x$ when $|y - x| = \pi$, and of course $d(x, y)^2$ is not differentiable either (see Figure 10.1).

Similar problems would occur on, say, a compact Riemannian manifold, as soon as there is no uniqueness of the geodesic joining $x$ to $y$. 