From Partial Differential Equations and Applications (P Marcellini et al, eds) Dekker NY: 1996, pp 29-35

4. Allocation Maps with General Cost Functions

Luis A. Caffarelli Institute for Advanced Study, Princeton, New Jersey

Introduction: In the theory of allocations one is interested to transfer a domain Ω_1 into a domain Ω_2 (one can think on sending k points $X_1, \ldots, X_k$ in a one to one fashion to k points Y_1, Y_k), minimizing certain cost function $C(X, Y)$. (See for instance Rachev ([R]). For $C(X,Y) = \frac{1}{2}|X - Y|^2$, this problem has interest in several fields, and a complete existence theory was given by Brenier ([B]) and a regularity theory was developed by the author ([C1], [C2]).

Here we redo Brenier's theory for general strictly convex cost function $C(X - Y)$.

In a conversation R. McCann told us that he had obtained similar results in collaboration with W. Gangbo.

In all of this, note f and g will be two bounded functions with compact support Ω_1 and Ω_2 respectively, and the compatibility condition

$$\int_{\Omega_1} f(X)dX = \int_{\Omega_2} g(Y)dY.$$

We will consider a cost function $C(X - Y)$, smooth (say $C^{1,\alpha}$) and strictly convex.

In particular the map (for fixed Y_0)

$$X \longrightarrow \nabla_X C(X-Y)$$

is continuous with a continuous inverse.

We are interested to find, among all transformations $Y(X)$ that carry the density f onto the density g (i.e., for every continuous function h,

$$\int h(Y)g(Y)dY = \int h(Y(X))f(X)dX$$

the one that minimizes the allocation cost

$$\int C(X-Y(X))f(X)dX.$$

We will do that through a heuristically standard dual problem, following the ideas of Brenier ([B]) in the case of $C(X-Y) = |X-Y|^2$ providing the necessary details to complete this program. The dual problem is the following:

Problem 1: Among all pairs of continuous functions $\varphi(X), \psi(Y)$ defined in Ω_1, Ω_2, and satisfying

$$\varphi(X) + \psi(Y) \geq -C(X-Y)$$

minimize

$$\int \varphi(X)f(X)dX + \int \psi(Y)g(Y)dY.$$

We will show that this problem has a solution, that the map

$$Y = X + (\nabla C)^{-1}(\nabla\varphi(X))$$

is well defined and "nice" and that it solves the allocation problem. Of particular interest is the case

$$C(X-Y) = |X-Y|^p \quad (1 < p < \infty)$$

and $p = 1+\epsilon$, since, for ϵ going to zero, this provides a "smooth" approximation to the solution of the Monge mass transfer problem.

We start with the following simple

Theorem 1.

(a) *Problem 1 has a minimizing pair* φ_0, ψ_0.

(b) φ_0, ψ_0 *are* C*-convex, i.e.,*

$$\varphi_0(X) = \sup_X -C(X-Y) - \psi_0(Y)$$
$$\psi_0(Y) = \sup_X -C(X-Y) - \varphi_0(Y).$$

In particular φ_0, ψ_0 *are Lipschitz and "*C^1 *by below" (i.e., have a uniformly* C^1 *function supporting them by below).*

Proof. Given a pair φ, ψ, of admissible functions, its energy can be improved by replacing φ by

$$\varphi^* = \sup_Y[-C(X-Y) - \psi(Y)]$$

and vice versa.

Therefore we may restrict ourselves to C-convex pairs φ, ψ.

Such pairs must be normalized because if φ, ψ are admissible, then for any constant $\lambda, \varphi + \lambda, \psi - \lambda$ are admissible and with the same energy.

Thus, we may impose that $\varphi(X_0) = 0$.

This bounds ψ by below by

$$\psi(Y) \geq -C(X_0 - Y)$$

and φ by above and below by diam $(\Omega_1) \cdot [\sup_{Y\in\Omega_2} \|C(X,Y)\|_{Lip}]$. Further, if M is large φ and $\overline{\psi} = M$ form an admissible pair, as long as

$$M \geq 2 \sup_{X,Y\in\Omega_1,\Omega_2} |C(X-Y)|$$

Thus $\varphi, \min(\psi, M)$ are a new admissible pair.

We can restrict therefore our minimization to a family of uniformly bounded uniformly Lipschitz pairs φ, ψ, and thus we can extract a minimizing sequence φ_k, ψ_k that converges uniformly to a minimizer φ_0, ψ_0.

We now study the uniqueness and differentiability properties of φ_0, ψ_0.

For that purpose we will assign to every X_0 in Ω_1 the set of "images"

$$K(X_0) = \{Y \in \overline{\Omega}_2 : \varphi(Y_0) + \psi(Y) = -C(X_0, Y)\}.$$

If $Y \in K(X_0), -\psi(Y) - C(X,Y)$ supports (is tangent by below) $\varphi(X)$ at X_0. Thus heuristically

$$\nabla_X \varphi(X_0) = -\nabla C(X_0 - Y).$$

Since ∇C is a bicontinuous invertible map with inverse $\nabla C^*(P) \quad (P \in R^n)$ we should recuperate Y as $Y = (\nabla C^*)(\nabla\varphi(X_0) + X_0$. That is, always heuristically,

$$K(X_0) = \nabla C^*(\nabla\varphi(X_0)) + X_0.$$

Before the next theorem, a few elementary remarks about $\varphi(X)$ and $\psi(Y)$.

We say that a function w is $C^{1,\alpha}$ convex if

$$w(X) = \sup_{\lambda\in\Lambda} g_\lambda(X),$$

with $g_\lambda(X)$ in $C^{1,\alpha}$, uniformly in λ.

Lemma 1. *Let w be $C^{1,\alpha}$ convex, then*

(a) *at every point $X_0, w(X)$ is asymptotic to the convex cone*

$$\Gamma_{X_0} = \sup_{\substack{g_\lambda(X):\\ g_\lambda(X)=w(Y_0)}} \langle \nabla g_\lambda(X_0, X - Y_0\rangle + w(X_0).$$

More precisely

$$\Gamma_{Y_0}(X) - C|X - X_0|^{1+\alpha} \le w(X) \le \Gamma_{X_0}(X) + o(|X - X_0|).$$

(b) If, say, 0 is a supporting gradient to Γ_{X_0} at X_0, (i.e., $0 = \nabla g_{\lambda_1}(X_0)$, and $Z = \nabla g_{\lambda_2}(X_1)$ is a supporting gradient at X_1, then

$$\langle Z, X_1 - X_0\rangle \ge -|X_1 - X_0|^{1+\alpha}.$$

(c) A point of Lebesge differentiability for ∇w is a point of continuity.

(d) The sets of points that have more than one supporting g are a set of measure zero.

Proof. (a) is clear.

(b) follows from the fact that:

$$\langle Z, X_1 - X_0\rangle \ge w(X_1) - w(X_0) \ge$$
$$\Gamma_{X_0}(X_1) - w(X_0) \ge -C|X_1 - X_0|^{1+\alpha},$$

since 0 is a supporting gradient.

(c) If X_0 is not a point of continuity it has two supporting g'_s, with gradients (say) 0 and te_1.

Let us split $X - X_0 = \lambda e_1 + \mu e_2 \ (e_2 \perp e_1)$. If X stays in the cone

$$\lambda \le -M|\mu|,$$

from (b) we get

$$-\lambda(z_1 \pm \frac{\lambda}{M} z_2 \ge -C\lambda^{1+\alpha}$$

or

$$z_1 \le \frac{1}{M}|z_2| + C\lambda^\alpha.$$

On the other hdna, if X stays in the cone $\lambda \geq M|\mu|$ we get

$$z_1 - t \geq \frac{1}{M}|z_2| - C\lambda^{\alpha}.$$

For M large, λ small, depending on t and $\sup_{\lambda} \nabla g_{\lambda}$, these sets stay away from each other and X_0 cannot be a point of Lebesgue differentiability for ∇w. (d) follows from (c).

Corollary. *$K(X_0)$ is single valued and continuous almost everywhere in X.*

We are now ready to prove the change of variable formula.

Theorem. *Let h be a continuous function in the support of g. Then*

$$\int h(Y)g(Y)dY = \int h(K(X))f(X)dY.$$

Proof. We consider, in the optimization problem the perturbation

$$\psi_\epsilon(Y) = \psi(Y) + \epsilon h(Y)$$

$$\varphi_\epsilon(X) = \sup_{Y} -C(X-Y) - \psi(Y) - \epsilon h(Y).$$

The energy variation should thus be positive

$$0 \leq \delta E_\epsilon = \epsilon \int h(Y)g(Y)dY + \\ + \int (\varphi_\epsilon(X) - \varphi(X))f(X)dX.$$

Notice that by definition $|\varphi_\epsilon(X) - \varphi(X)| \leq \epsilon \sup_Y |h(Y)|$.

Therefore we divide by ϵ and we get

$$0 \leq \int h(Y)g(Y)dY + \int \frac{[\varphi_\epsilon(X) - \varphi(X)]}{\epsilon} f(X)dY.$$

The second term consists of uniformly bounded functions of which will be enough to compute the a.e. limit to obtain the limiting integrand.

For that we just consider the points X of continuity for $K(X)$.

We notice that

$$-\epsilon h(K(X)) \leq \varphi_\epsilon(X) - \varphi(X) \leq -\epsilon h(Y_\epsilon)$$

for Y_ϵ the point that realizes the value of φ_ϵ. But if a subsequence Y_ϵ converges to a Y_0 different to $(K(X))$ in the limit we will have a second supporting function $-C(X - Y_0) - \psi(Y_0)$ for φ at X, a contradiction. Thus the a.e. limit of $\frac{\varphi_\epsilon(X) - \varphi(X)}{\epsilon}$ is $h(K(X))$.

To complete this presentation, we show that the map $Y(X)$ minimizes the allocation integral.

Theorem. *Among all measurable maps $Y(X)$ that satisfy the change of variable formula, $K(X)$ is the unique minimizer of the cost function*

$$\int C(Y(X) - X)\ f(X)dX.$$

In particular, the pair φ, ψ is unique.

Proof. We write $I = \int \psi(Y)g(Y)dY$ using both changes of coordinates:

$$0 = \int [\psi(Y(X)) - \psi(K(X))]f(X)dX$$

$$\geq \int [-C(Y(X) - X) - \varphi(X)]$$

$$-[-C(K(X) - X) - \varphi(X)]f(X)dX.$$

(since for one we have inequality and the other equality) with equality in the integrals if and only if we have that $Y(X) \in K(X)$ a.e., that is $Y(X) = K(X)$, a.e.

References

[B] Brenier, Y., Polar Factorization and Monotone Rearrangement of Vector-Valued Functions, C.P.A.M., Vol. 44, 1991, pp. 375-417.

[C1] Caffarelli, L.A., The Regularity of Mappings with Convex Potentials, J. AMS 5, 1992, pp. 99-104.

[C2] Caffarelli, L.A., Boundary Regularity of Maps with Convex Potentials, C.P.A.M., Vol. 45, 1992, pp. 1141-1151.

[R] Rachev, S.T., Probability Metrics and the Stability of Stochastic Models, Wiley, 1991.