Exercise Sheet 3 Solutions

Intro to OT

Lent 2018-2019

3.1

Given $\varphi(x) = -2x\cdot x_0$ and $\nu = T_\# \mu$

then $\psi(y) = \inf_x \left(|x-y|^2 - \varphi(x) \right)$

$$= \inf_x \big(\underbrace{|x-y|^2 + 2x\cdot x_0}_{F(x)} \big)$$

$$\cancel{= \sup_x |x|^2}$$

$\nabla F(x) = 2(x-y) + 2x_0$

so $\nabla F(x) = 0 \Rightarrow x = y - x_0$. thus

$$\psi(y) = |x_0|^2 + 2(y-x_0)\cdot x_0$$

One can check $\varphi(x) + \psi(y) \le |x-y|^2$

Now $\min_{T: T_\#\mu = \nu} M(T) \ge J(\varphi,\psi) = -2\int_{\mathbb{R}^d} x\cdot x_0 \, d\mu(x) + \int_{\mathbb{R}^d} |x_0|^2 + 2(y-x_0)\cdot x_0 \, d\nu(y)$

by duality

$$= -2\int_{\mathbb{R}^d} x\cdot x_0 \, d\mu(x) + \int_{\mathbb{R}^d} |x_0|^2 \, d(T_\#\mu)(x) + 2\int_{\mathbb{R}^d} (y-x_0)\cdot x_0 \, d(T_\#\mu)(y)$$

$$= -2\int_{\mathbb{R}^d} x\cdot x_0 \, d\mu(x) + \int_{\mathbb{R}^d} |x_0|^2 \, d\mu(x) + 2\int_{\mathbb{R}^d} (T(x)-x_0)\cdot x_0 \, d\mu(x)$$

$$= -2\int_{\mathbb{R}^d} x\cdot x_0 \, d\mu(x) + \int_{\mathbb{R}^d} |x_0|^2 \, d\mu(x) + 2\int_{\mathbb{R}^d} x\cdot x_0 \, d\mu(x)$$

$$= \int_{\mathbb{R}^d} |x_0|^2 \, d\mu(x) = \int_{\mathbb{R}^d} |T(x) - x|^2 \, d\mu(x) = M(T)$$

thus T is optimal.

3.2

(a) $$\varphi_1^*(y) = \sup_{x \in \mathbb{R}^d} \left(x \cdot y - \varphi_1(x)\right)$$

$$= \sup_{x \in \mathbb{R}^d} \Big(\underbrace{x \cdot y - x \cdot x}_{F_1(x) :=}\Big)$$

$$\nabla F_1(x) = y - 2x \quad \text{so} \quad \nabla F_1(x) = 0 \implies x = \tfrac{1}{2} y$$

so $$\varphi_1^*(y) = \tfrac{1}{2} y \cdot y - \tfrac{1}{4} y \cdot y = \tfrac{1}{4} y \cdot y$$

(b) $$\varphi_2^*(y) = \sup_{x \in \mathbb{R}^d} \left(x \cdot y - x \cdot x_0\right)$$

$$= \sup_{x \in \mathbb{R}^d} x \cdot (y - x_0)$$

$$= \begin{cases} 0 & \text{if } y = x_0 \\ +\infty & \text{else} \end{cases}$$

(c) $$\varphi_3^*(y) = \sup_{x \in \mathbb{R}^d} \left(x \cdot y - \varphi_3(x)\right)$$

$$= x_0 \cdot y$$

(d) $$\varphi_4^*(y) = \sup_{x \in \mathbb{R}^d} \Big(\underbrace{x \cdot y - \tfrac{1}{p}|x|^p}_{F_4(x)}\Big)$$

$$\nabla F_4(x) = y - x|x|^{p-2} \quad \text{so} \quad \nabla F_4(x) = 0 \implies y = x|x|^{p-2}$$

So $x \cdot y = |x|^p$, hence $F_4(x) = |x|^p - \frac{1}{p}|x|^p = \left(1 - \frac{1}{p}\right)|x|^p$

Moreover $|y| = |x|^{p-1}$ so $|x|^p = |y|^{\frac{p}{p-1}}$, then $F_4(x) = \left(1 - \frac{1}{p}\right)|y|^{\frac{p}{p-1}}$

So $$\varphi_\varphi^*(y) = \left(1-\frac{1}{p}\right)|y|^{\frac{p}{p-1}}$$

Note that if we let $q = \frac{1}{1-\frac{1}{p}} = \frac{p}{p-1}$ then
$$\varphi_\varphi^*(y) = \frac{1}{q}|y|^q \quad \text{and} \quad \frac{1}{p}+\frac{1}{q} = \frac{1}{p}+1-\frac{1}{p} = 1.$$

3.3

Let $x \in \mathbb{R}^d$, then
$$y \in \partial\varphi(x) \underset{\text{prop 6.5}}{\iff} x\cdot y = \varphi(x) + \varphi^*(y) \underset{\text{prop 6.7}}{\iff} x\cdot y = \varphi^{**}(x) + \varphi^*(y)$$

$$\underset{\text{prop 6.5}}{\iff} x \in \partial\varphi^*(y)$$

3.4

(1) Injectivity of the gradient.

Assume $x, y \in \mathbb{R}^d$ are such that $\nabla\varphi(x) = \nabla\varphi(y)$ and $x \neq y$.

By strict convexity $\quad \nabla\varphi(x)\cdot(y-x) < \varphi(y) - \varphi(x)$ $(*)$

$$\nabla\varphi(y)\cdot(x-y) < \varphi(x) - \varphi(y) \quad (+)$$

Using $\nabla\varphi(x) = \nabla\varphi(y)$ we have $(*) \Rightarrow \nabla\varphi(y)\cdot(y-x) < \varphi(y) - \varphi(x)$

$\Rightarrow \nabla\varphi(y)\cdot(x-y) > \varphi(x) - \varphi(y)$ a contradiction with $(+)$

$\Big[\nabla\varphi(x)\cdot(y-x) < \varphi(y) - \varphi(x)$ follows from $\varphi(y) > \varphi(x) + \nabla\varphi(x)\cdot(y-x)\Big]$.

(2) Surjectivity of the gradient. Fix y_0

Let $g(x) = \varphi(x) - x y_0$ ~~for fixed y~~.

Since g is cts then if $g(x) \to +\infty$ as $|x| \to +\infty$ then $\exists$ a minimum of g.

Now $g(x) = \varphi(x) - x \cdot y_0 \geq \varphi(x) - |x||y_0| = |x| \left(\underbrace{\frac{\varphi(x)}{|x|}}_{\to +\infty} - |y| \right)$

For $|x|$ large enough $\frac{\varphi(x)}{|x|} > |y|$. Then $g(x) \to +\infty$ as $|x| \to +\infty$.

~~Hence $\varphi^*(y) = \sup_{x \in \mathbb{R}^d} (-g(x)) = -\inf_{x \in \mathbb{R}^d} g(x) = -\min_{x \in \mathbb{R}^d} g(x)$.~~

~~As the minimum of C^1 function φ^* is C^1, let $x = \nabla \varphi^*(y)$ which by Ex 3.3 implies $y = \nabla \varphi(x)$~~

Since g is cts, and diff and coercive $\exists$ a min of g, at say x_0 and $\nabla g(x_0) = 0$. But $\nabla g(x_0) = \nabla \varphi(x_0) - y_0$

So $\exists x_0$ s.th. $\nabla \varphi(x_0) = y_0$. In particular $\nabla \varphi$ is surjective

3.5

Let $n \in \mathbb{N}$ and $(x_i, y_i) \in \Gamma$, $i = 1, \ldots, n$.

Then $y_i \in \partial \varphi(x_i) \Rightarrow \varphi(z) \geq \varphi(x_i) + y_i \cdot (z - x_i) \quad \forall z \in \mathbb{R}^d$

choose $z = x_{i+1}$ then

$$\varphi(x_{i+1}) \geq \varphi(x_i) + y_i \cdot (x_{i+1} - x_i) \quad \forall i$$

Summing over i

$$\sum_{i=1}^n \varphi(x_{i+1}) \geq \sum_{i=1}^n \varphi(x_i) + \sum_{i=1}^n y_i \cdot (x_{i+1} - x_i)$$

$$\Rightarrow \quad 0 \geq \sum_{i=1}^n y_i \cdot x_{i+1} - \sum_{i=1}^n y_i \cdot x_i \quad \text{as required.}$$

3.6 We have $(\hat\varphi, \hat\psi) \in \hat\Phi$ (ignoring integrability condition)

$$\iff \hat\varphi(x) + \hat\psi(y) \geq x\cdot y \quad \forall x,y$$

$$\iff |x|^2 - 2\hat\varphi(x) + |y|^2 - 2\hat\psi(y) \leq |x|^2 + |y|^2 - 2x\cdot y = |x-y|^2$$

$$\iff \varphi(x) + \psi(y) \leq |x-y|^2$$

$$\iff (\varphi, \psi) \in \Phi_c$$

And
$$J(\hat\varphi, \hat\psi) = \int \hat\varphi \, d\mu(x) + \int \hat\psi(y) \, d\nu(y)$$

$$= \int \left(|x|^2 - 2\hat\varphi(x)\right) d\mu(x) + \int |y|^2 - 2\hat\psi(y) \, d\nu(y)$$

$$= \int |x|^2 \, d\mu(x) + \int |y|^2 \, d\nu(y) - 2 J(\hat\varphi, \hat\psi).$$

So
$$\max_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi) = \int |x|^2 \, d\mu(x) + \int |y|^2 \, d\nu(y) - 2 \inf_{(\hat\varphi,\hat\psi)\in\hat\Phi} J(\hat\varphi, \hat\psi)$$

And moreover if $(\varphi, \psi) \in \Phi_c$ maximises J ~~the~~ in Φ_c

$\iff (\hat\varphi, \hat\psi) \in \hat\Phi$ minimises J in $\hat\Phi$.

By Corollary 6.3 we have $T = \nabla\hat\varphi$ is an OT plan where $(\hat\varphi, \hat\varphi^*)$ minimises J in $\hat\Phi$. Then $T = \nabla\left(\frac{1}{2}|\cdot|^2 - \frac{1}{2}\varphi\right) = Id - \frac{1}{2}\nabla\varphi$ where (φ, φ^c) maximises J in Φ_c (and where $\varphi^c = |\cdot|^2 - 2\hat\varphi^*$).

3.7

We see that $\varphi(x) = -2x\cdot x_0$ is the maximum of $\mathcal{J}$ in Φ_c.

By Ex 3.6 the OT map is given by

$$T(x) = x - \tfrac{1}{2}\nabla\varphi(x) = x + x_0 .$$

3.8

We use the ansatz $T(x) = \frac{x}{|x|^q}$. In order for

$T_\#\mu = \nu$ we need, by Ex 1.4,

$$(*) \quad f(x) = g(T(x))\,|\det(\nabla T(x))| .$$

Now since $\frac{\partial}{\partial x_i} T_j(x) = \frac{1}{|x|^q}\left(\delta_{ij} - \frac{q x_i x_j}{|x|^2}\right)$ we have

$$\det \nabla T(x) = \frac{1}{|x|^{2q}} \det \begin{pmatrix} 1 - \frac{q x_1^2}{|x|^2} & -\frac{q x_1 x_2}{|x|^2} \\ -q\frac{x_1 x_2}{|x|^2} & 1 - \frac{q x_2^2}{|x|^2} \end{pmatrix}$$

$$= \frac{1}{|x|^{2q}}\left(\left(1 - \frac{q x_1^2}{|x|^2}\right)\left(1 - \frac{q x_2^2}{|x|^2}\right) - \frac{q^2 x_1^2 x_2^2}{|x|^4}\right)$$

$$= \frac{1}{|x|^{2q}}\left(1 - \frac{q x_1^2}{|x|^2} - \frac{q x_2^2}{|x|^2} + \frac{q^2 x_1^2 x_2^2}{|x|^4} - \frac{q^2 x_1^2 x_2^2}{|x|^4}\right)$$

$$= \frac{1-q}{|x|^{2q}}$$

Then $(*) \iff \frac{1}{\pi} = \cancel{3|T(x)|} \; \frac{3|x|(1-q)}{2\pi\, |x|^q\, |x|^{2q}}$.

This is satisfied for $q = 1/3$.

Now we recall that if $(\tilde{\varphi}, \tilde{\varphi}^*)$ maximise $\mathcal{J}$ over $\tilde{\Phi}$ then we have $\nabla\hat{\varphi} = T$. Assume $\hat{\varphi}(x) = \alpha(|x|)$. Then

$$\nabla\hat{\varphi}(x) = \frac{\alpha'(|x|)\,x}{|x|}.$$

Hence $\nabla\hat{\varphi} = T$ implies

$$\frac{\alpha'(|x|)}{|x|} = \frac{1}{|x|^{\cancel{4}1/3}}$$

Hence,

$$\alpha'(r) = r^{2/3} \implies \alpha(r) = \frac{3r^{5/3}}{5}$$

$$\therefore \quad \hat{\varphi}(x) = \frac{3|x|^{5/3}}{5}$$ [One can check that $\nabla\hat{\varphi} = T$]

Now to compute $\hat{\varphi}^*$ we note by exercise 3.4 that

$$\hat{\varphi}^*(y) = x\cdot y - \hat{\varphi}(x) \quad \text{when } \nabla\hat{\varphi}(x) = y \iff \frac{x}{|x|^{1/3}} = y$$

$$= x\cdot y - \frac{3|x|^{5/3}}{5}$$

$$= |y|^2|x|^{1/3} - \frac{3}{5}|x|^{5/3}$$ since $\frac{x\cdot y}{|x|^{1/3}} = |y|^2$

$$= |y|^2|y|^{1/2} - \tfrac{3}{5}|y|^{5/2}$$ and $|x|^{2/3} = |y|$, $|x| = |y|^{3/2}$

$$= \tfrac{2}{5}|y|^{5/2}$$

By exercise 3.6 we have

$$\varphi(x) = |x|^2 - \cancel{4}2\hat{\varphi}(x) = |x|^2 - \frac{\cancel{3}6}{5}|x|^{5/3}$$

$$\psi(y) = |y|^2 - 2\hat{\varphi}^c(y) = |y|^2 - \cancel{4}\!8/5\,|y|^{5/2}$$

as candidate maximum to the dual problem.

First we check $(\varphi, \psi) \in \Phi_c$:

$$\varphi(x) + \psi(y) = |x|^2 + |y|^2 - 6\cancel{9}/5\,|x|^{5/3} - \cancel{3}4/5\,|y|^{5/2}$$

By Young's inequality $|x\cdot y| \leq \frac{|x|^p}{p} + \frac{|y|^q}{q}$ when $\frac{1}{p} + \frac{1}{q} = 1$, $p,q > 1$.

choosing $p = 5/3$ and $q = 5/2$ we have $\frac{1}{p} + \frac{1}{q} = 3/5 + 2/5 = 1$ and

$$-2x\cdot y \geq -\frac{6|x|^{5/3}}{5} - \frac{4|y|^{5/2}}{5}$$

Then $\varphi(x) + \psi(y) \leq |x|^2 + |y|^2 - 2x\cdot y = |x-y|^2$.

Now

$$\mathbb{J}(\varphi, \psi) = \int_{\mathbb{R}^2} \varphi(x)\,dx + \int_{\mathbb{R}^2} \psi(y)\,d\nu(y)$$

$$= \frac{2\pi}{\pi}\int_0^1 \left(r^2 - \tfrac{6}{5}r^{5/3}\right) r\,dr + \frac{3\times 2\pi}{2\pi}\int_0^1 \left(r^2 - \tfrac{4}{5}r^{5/2}\right) r^2\,dr$$

$$= \frac{-2\times 17}{220} + \frac{3\times 3}{55} = \frac{1}{110}$$

And

$$\mathcal{M}(T) = \int_{\mathbb{R}^2} |x - T(x)|^2 \, d\mu(x)$$

$$= \cancel{\frac{2\pi}{\pi}\int_0^1} \int_{\mathbb{R}^2} \left|x - \frac{x}{|x|^{1/3}}\right|^2 d\mu(x)$$

$$= \int_{\mathbb{R}^2} |x|^2 \left(1 - \frac{1}{|x|^{1/3}}\right)^2 d\mu(x)$$

$$= \frac{2\pi}{\pi} \int_0^1 r^3 \left(1 - \frac{1}{r^{1/3}}\right)^2 dr$$

$$= \frac{1}{110}.$$

Since $\mathcal{M}(T) = \mathcal{J}(\varphi, \psi)$ it follows that (from duality) that T and (φ, ψ) are optimal.

3.9 Let $\mu = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}$ and $\nu = \frac{1}{n}\sum_{j=1}^{n} \delta_{y_j}$. By Thm 3.7 $\exists$ a solution to the MOT problem between μ and ν with cost $c(x,y) = |x-y|$. i.e. $\exists\, T^\dagger : \{x_i\}_{i=1}^n \to \{y_j\}_{j=1}^n$ s.t.

$$\inf_{T : T_\# \mu = \nu} \mathcal{M}(T) = \mathcal{M}(T^\dagger), \quad T^\dagger_\# \mu = \nu$$

where $\mathcal{M}(T) = \frac{1}{n}\sum_{i=1}^{n} |x_i - T(x_i)|$.

Consider the lines l_i from x_i to $T^\dagger(x_i)$. For a contradiction we assume $\exists\, k \neq j$ s.t. l_k intersects l_j. Define a transport map $\hat{T}$ by

$$\hat{T}(x_i) = \begin{cases} T^\dagger(x_i) & \text{if } i \neq k,j \\ T^\dagger(x_k) & \text{if } i = j \\ T^\dagger(x_j) & \text{if } i = k. \end{cases}$$

We will show $M(T^\dagger) > M(\hat{T})$.

Indeed, let $\{x\} = \ell_k \cap \ell_j$. Then

$$|x_j - T^\dagger(x_j)| + |x_k - T^\dagger(x_k)| = |x_j - x| + |x - T^\dagger(x_j)| + |x_k - x| + |x - T^\dagger(x_k)|$$

$$\geq |x_j - T^\dagger(x_k)| + |x_k - T^\dagger(x_j)| \quad \text{by } \Delta \text{ inequality}$$

To show the inequality is in fact strict let us assume that

$$|x_j - T^\dagger(x_j)| + |x_k - T^\dagger(x_k)| = |x_j - T^\dagger(x_k)| + |x_k - T^\dagger(x_j)|$$

Then,
$$|x_j - T^\dagger(x_k)| = |x_j - x| + |x - T^\dagger(x_k)|$$
and
$$|x_k - T^\dagger(x_j)| = |x_k - x| + |x - T^\dagger(x_j)|$$

So ~~ante~~ x is also on the line between x_j and $T^\dagger(x_k)$ and x_k and $T^\dagger(x_j)$. This is only possible if pts are colinear.

Thus $|x_j - T^\dagger(x_j)| + |x_k - T^\dagger(x_k)| > |x_j - T^\dagger(x_k)| + |x_k - T^\dagger(x_j)|$

so $M(T^\dagger) > M(\hat{T})$

Since $\hat{T}$ is still a bijection and ~~$\hat{T}$~~ $\mu(\hat{T}^{-1}(x_i)) = \frac{1}{n}$ then $\hat{T}_\# \mu = \nu$. So $\hat{T}$ is admissible, and is a minimiser of the (MO) problem. This is a contradiction. So $T^\dagger$ satisfies the claimed property. Since $T^\dagger$ is bijective we can define $\sigma(i) = j$ ~~exp~~ when $T(x_i) = y_j$