Exercise Sheet 1 Solutions

Intro to OT Lent 2018–2019

1.1 Let $B \subseteq \mathbb{R}$. Then

$$\cancel{\text{pf}} \quad [T_1]_\# \mu(B) = \mu(T_1^{-1}(B))$$

$$= \int_{T_1^{-1}(B)} \chi_{(0,1]}(x)\, dx$$

$$= \int_{T_1^{-1}(B)\cap(0,1]} 1 \, dx$$

$$= \int_{B\cap(1,2]} 1 \, dy \qquad [\text{change of variables, } y = T_1(x) = x+1]$$

$$= \int_B \chi_{(1,2]}(y)\, dy$$

$$= \nu(B).$$

To show $S_\# \mu \neq \nu$ choose $B = (0,1]$. Then

$$S_\# \mu(B) = \mu(S^{-1}(B)) = \mu((0,\tfrac{1}{2}]) = \tfrac{1}{2}$$

and $\nu(B) = \nu((0,1]) = 0$.

1.2 Assume $\pi = (\mathrm{Id} \times f)_\# \mu$. Let $P(x,y) = x$. Then for $B \subseteq$

$$P_\# \pi(B) = \pi(P^{-1}(B)) = \pi(\{(x,y) \in X\times Y : P(x,y) \in B\})$$

$$= \pi(\{(x,y)\in X\times Y : x \in B\}) = \pi(B \times Y)$$

$$= \mu\left((\mathrm{Id}\times f)^{-1}\left(\underbrace{\cancel{\{(x,y)\in X\times Y : x\in B\}}}_{B\times Y}\right)\right)$$

$$= \mu\left(\{ x : (\mathrm{Id} \times f)(x) \in B \times Y \} \right)$$

$$= \mu\left(\{ x : x \in B \text{ and } f(x) \in Y \} \right)$$

$$= \mu(B)$$

By disintegration of measures (Thm 1.4) $\exists$ a family $\pi(\cdot|x)$, $x \in X$ s.t.

$$\int_{X \times Y} F(x,y)\, d\pi(x,y) = \int_X \int_Y F(x,y)\, d\pi(y|x)\, d\mu(x)$$

$\forall$ measurable $F : X \times Y \to [0, +\infty]$. Choose $F(x,y) = G(x)\chi_B(y)$. Then we have

$$\int_{X \times Y} G(x)\chi_B(y)\, d\pi(x,y) = \int_X G(x)\, \pi(B|x)\, d\mu(x) \qquad (1)$$

By the change of variable formula (Prop 2.3) the left hand side can be written

$$\int_{X \times Y} G(x)\chi_B(y)\, d\pi(x,y) = \int_{X \times Y} G(x)\chi_B(y)\, d(\mathrm{Id} \times f)_\# \mu(x,y)$$

$$= \int_X G(x)\chi_B(f(x))\, d\mu(x) \qquad (2)$$

comparing (1) and (2) we have $\pi(B|x) = \chi_B(f(x))$ for μ-a.e. $x \in X$.

Hence $\pi(\cdot|x) = \delta_{f(x)}$ as required.

[Alternative proof: check $\pi(\cdot|x) = \delta_{f(x)}$ satisfies the Disintegration of measures theorem.]

1.~~4~~3 Assume $\upsilon = T_\# \mu$. Let $\varphi : X \to \mathbb{R}$ be measurable + bounded

$$\int_Y \varphi(y)\, d\upsilon(y) = \int_Y \varphi(y)\, d(T_\# \mu)(y)$$

by prop 1.5

$$= \int_X \varphi(T(x))\, d\mu(x)$$ as claimed.

On the other hand, if $\int_Y \varphi(y)\, d\upsilon(y) = \int_X \varphi(T(x))\, d\mu(x)$

for all bounded and measurable φ then we choose

$\varphi = \chi_B$. So

$$\upsilon(B) = \int_Y \varphi(y)\, d\upsilon(y) = \int_X \varphi(T(x))\, d\mu(x)$$

$$= \int_X \chi_B(T(x))\, d\mu(x)$$

$$= \mu(\{x : T(x) \in B\})$$

$$= T_\# \mu(B).$$

Hence $\upsilon = T_\# \mu$ as required.

1.8.4 Assume $\nu = T_{\#}\mu$. Then

$$\int_Y F(y)\, d\nu(y) = \int_Y F(y)\, dT_{\#}\mu(y) \qquad \forall F: Y \to [0,+\infty] \text{ measurable.}$$

By change of variables (Prop 2.3)

$$\int_Y F(y)\, dT_{\#}\mu(y) = \int_X F(T(x))\, d\mu(x)$$

$$= \int_X F(T(x))\, f(x)\, dx$$

On the other hand

$$\int_Y F(y)\, d\nu(y) = \int_Y F(y)\, g(y)\, dy$$

$$= \int_{T^{-1}(Y)} F(T(x))\, g(T(x))\, |\det(\nabla T(x))|\, dx \quad \text{by using a change of variable } y = T(x)$$

$$= \int_X F(T(x))\, g(T(x))\, |\det(\nabla T(x))|\, dx \quad \text{since } X = T^{-1}(Y)$$

Since this holds for all F measurable, we have

$$f(x) = g(T(x))\, |\det(\nabla T(x))| \qquad \text{for } \mu \text{ a.e. } x \in X.$$

Since the argument holds in reverse we have that

$$f(x) = g(T(x))\, |\det(\nabla T(x))| \quad \text{implies} \quad \nu = T_{\#}\mu.$$

1.5 Since, (when $X = \mathbb{R} = Y$)

$$\int_X \varphi(T_2(x))\, d\mu(x) = \int_0^1 \varphi(2-x)\, dx$$

$$= \int_1^2 \varphi(y)\, dy \qquad (\text{c.o.v.s } y = T_2(x) = 2-x)$$

$$= \int_Y \varphi(y)\, d\nu(y)$$

$$\int_X \varphi(T_3(x))\, d\mu(x) = \int_0^{1/2} \varphi(x + 3/2)\, dx + \int_{1/2}^1 \varphi(2-x)\, dx$$

$$= \int_{3/2}^2 \varphi(y)\, dy + \int_1^{3/2} \varphi(y)\, dy$$

$$= \int_1^2 \varphi(y)\, dy$$

$$= \int_Y \varphi(y)\, d\nu(y)$$

So $(T_2)_{\#}\mu = (T_3)_{\#}\mu = \nu$.

1.6

$$S_\# \mu(A) = \mu(\{x : S(x) \in A\})$$

$$= \mu(\{x : -x \in A\})$$

$$= \mu(-A)$$

$$= \mu(A)$$

1.7

Assume $\pi = (Id \times T)_\# \mu$ and $\nu = T_\# \mu$.

Then let $P^X(x,y) = x$ and $P^Y(x,y) = y$.

$$P^X_\# \pi(A) = \pi(\{(x,y) : P^X(x,y) \in A\})$$

$$= (Id \times T)_\# \mu(\{(x,y) : x \in A\})$$

$$= (Id \times T)_\# \mu(A \times Y)$$

$$= \mu(\{x : (Id \times T)\cancel{\ldots}(x) \in A \times Y\})$$

$$= \mu(\{x : x \in A \text{ and } T(x) \in Y\})$$

$$= \mu(A)$$

Similarly,
$$P^Y_\# \pi(B) = (Id \times T)_\# \mu(X \times B)$$

$$= \mu(\{x : (Id \times T)(x) \in X \times B\})$$

$$= \mu(\{x : x \in X \text{ and } T(x) \in B\})$$

$$= \mu(\{x : T(x) \in B\})$$

$$= T_\# \mu(B)$$

$$= \nu(B)$$

Then $\pi \in \Pi(\mu, \nu)$.

On the other hand, if $\pi \in \Pi(\mu,\nu)$ then we have

$$\nu(B) = P^Y_\# \pi(B) = T_\# \mu(B)$$

Then $\nu = T_\# \mu$.

<u>1.8</u> (a) $c(x,y) = |x-y|^2$

$$M(T_1) = \int_0^1 |x - T_1(x)|^2 \, dx$$

$$= \int_0^1 |x - x - 1|^2 \, dx$$

$$= 1$$

$$M(T_2) = \int_0^1 |x - T_2(x)|^2 \, dx$$

$$= \int_0^1 |x - x + x|^2 \, dx$$

$$= \int_0^1 (x-1)^2 \, dx$$

$$= \int_{-1}^0 y^2 \, dy \qquad \text{letting } y = x-1$$

$$= \frac{24}{3}\left[y^3\right]_{-1}^{0}$$

$$= 4/3 .$$

$$M(f_3) = \int_0^1 |x - f_3(x)|^2\, dx$$

$$= \int_0^{1/2} |x - x - 3/2|^2\, dx + \int_{1/2}^1 |x - 2 + x|^2\, dx$$

$$= \int_0^{1/2} (3/2)^2\, dx + 4\int_{1/2}^1 (1-x)^2\, dx$$

$$= \frac{9}{8} + 4\int_{-1/2}^{0} y^2\, dy$$

$$= 9/8 + 4/3\left[y^3\right]_{-1/2}^{0}$$

$$= 9/8 + 4/3 \times 1/2^3$$

$$= \frac{27}{24} + \frac{4}{24}$$

$$= 31/24 .$$

(6) $M(f_1) = \int_0^1 |x - x - 1|\, dx = 1$

$M(f_2) = 2\int_0^1 |x-1|\, dx = 2\int_0^1 (1-x)\, dx \overset{y=1-x}{=} 2\int_0^1 y\, dy = [y^2]_0^1 = 1$

$M(f_3) = \int_0^{1/2} |3/2|\, dx + 2\int_{1/2}^1 1-x\, dx$

$= \frac{3}{4} + 2/8 = 1 .$

(c) $\mathcal{M}(T_1) = \int_0^1 \sqrt{1} \, dx = 1$

$$\mathcal{M}(T_2) = \sqrt{2} \int_0^1 \sqrt{1-x} \, dx = \sqrt{2} \int_0^1 \sqrt{y} \, dy$$

$$= \frac{2\sqrt{2}}{3} \left[y^{3/2} \right]_0^1 = \frac{2\sqrt{2}}{3}$$

$$\mathcal{M}(T_3) = \int_0^{1/2} \sqrt{3/2} \, dx + \sqrt{2} \int_{1/2}^1 \sqrt{1-x} \, dx$$

$$= \frac{1}{2} \sqrt{3/2} + \sqrt{2} \int_0^{1/2} \sqrt{y} \, dy$$

$$= \frac{1}{2} \sqrt{3/2} + \frac{2\sqrt{2}}{3} \left[y^{3/2} \right]_0^{1/2}$$

$$= \frac{1}{2} \sqrt{3/2} + \frac{2\sqrt{2}}{3} \frac{1}{2\sqrt{2}}$$

$$= \frac{\sqrt{3}}{2\sqrt{2}} + \frac{1}{3}$$

<u>1.9</u> Let $T : X \to Y$ then with $T_\# \mu = \nu$

$$\mathcal{M}(T) = \int_0^1 h(|x - T(x)|) \, dx$$

$$\geq h\left(\left| \int_0^1 x - T(x) \, dx \right| \right) \quad \text{by Jensens inequality}$$

Now
$$\int_0^1 x - T(x) \, dx = \int_0^1 x \, dx - \int_0^1 T(x) \, dx$$

$$= 1/2 - \int_0^1 T(x) \, d\mu(x) = 1/2 - \int_1^2 y \, d\nu(y)$$

$$= 1/2 - \frac{1}{2} \left[y^2 \right]_1^2 = 1/2 - 1/2 (4 - 1) = 1/2 - 3/2 = -1.$$

Hence $\mathcal{M}(T) \geq h(|-1|)$

Now,

$$\mathbb{M}(T_1) = \int_0^1 h(|x - x - 1|)\, dx$$

$$= \int_0^1 h(1)\, dx$$

$$= h(1).$$

Hence $\mathbb{M}(T_1) \le \mathbb{M}(T)$ $\forall T : X \to Y$ with $T_\# \mu = \nu$.

Since $(T_1)_\# \mu = \nu$ we have that T_1 is a minimiser of the MOT problem.

1.10

Claim h is convex, if so then by 1.9, $T_1(x) = x+1$ is optimal.

$h'(s) = 1 + \log(s+1)$

$h''(s) = \frac{1}{s+1} > 0$ for $s > 0$. $\therefore$ h is convex.

1.11

Let $T_\# \mu = \nu$. Then

$$\mathbb{M}(T) = \int_X h(T(x) - x)\, d\mu(x) = h\left(\int_X (T(x) - x)\, d\mu(x)\right)$$

$$= h\left(\int_X T(x)\, d\mu(x)\right) - h\left(\int_X x\, d\mu(x)\right)$$ (Since X, Y are bounded then we avoid the case where we have $\infty - \infty$).

Now $\int_X T(x)\, d\mu(x) = \int_Y y\, d(T_\# \mu)(y) = \int_Y y\, d\nu(y)$

So $\mathbb{M}(T) = h\left(\int_Y y\, d\nu(y)\right) - h\left(\int_X x\, d\mu(x)\right)$ which is independent of T.

1.12. Ex. Let $h(s) = s$, then by Exam 1.9 T_1 is optimal.

~~By exam 1.6 (b)~~ $\mathbb{M}(T_1) = 1$

~~Now~~ $\mathbb{M}(T_\varphi) = \int_0^1 |x - x - 2| \, dx$

Now,

$$\mathbb{M}(T_1) = \tfrac{1}{2} \int_0^2 |x - T_1(x)| \, dx$$

$$= \tfrac{1}{2} \int_0^2 |x - x - 1| \, dx$$

$$= 1$$

And

$$\mathbb{M}(T_\varphi) = \tfrac{1}{2} \int_0^2 |x - T_\varphi(x)| \, dx$$

$$= \tfrac{1}{2} \int_0^1 |x - x - 2| \, dx + \tfrac{1}{2} \int_1^2 0 \, dx$$

$$= \tfrac{1}{2} \int_0^1 2 \, dx = 1$$

To check $(T_1)_\# \mu = \nu$ we have

$$\underbrace{\cancel{\int \varphi(x) \, d(T_1)}}_{\int_X \varphi(T_1(x)) \, d\mu(x)} \; \tfrac{1}{2} \int_0^2 \varphi(T_1(x)) \, dx = \tfrac{1}{2} \int_0^2 \varphi(x+1) \, dx$$

$$= \tfrac{1}{2} \int_1^3 \varphi(y) \, dy = \int_Y \varphi(y) \, d\nu(y) \quad \text{and apply Ex 1.3.}$$

To check $(T_a)_\# \mu = \nu$ we have

$$\int_X \varphi(T_a(x))\, d\mu(x) = \tfrac12 \int_0^1 \varphi(x+2)\, dx + \tfrac12 \int_1^2 \varphi(x)\, dx$$

$$= \tfrac12 \int_2^3 \varphi(y)\, dy + \tfrac12 \int_1^2 \varphi(y)\, dy$$

$$= \tfrac12 \int_1^3 \varphi(y)\, dy$$

$$= \int_Y \varphi(y)\, d\nu(y)$$ and apply Ex 1.3.

1.13

(a) ~~Fix x~~ By Prop 1.5 $S_\# ((\tau_a)_\# \mu) = (S \circ \tau_a)_\# \mu$

So $$\int_X \varphi((S\circ \tau_a)(x))\, d\mu(x) = \int_X \varphi(T(x) - b)\, d\mu(x)$$

$$= \int_X \varphi(y-b)\, d(T_\# \mu)(y) \quad \text{by Prop 1.5}$$

$$= \int_X \varphi(y-b)\, d\nu(y)$$

$$= \int_X \varphi(\tau_b(y))\, d\nu(y)$$

If $S_\#((\tau_a)_\# \mu) = \omega$ then we have

$$\int_Y \varphi(y)\, d\omega(y) = \int_X \varphi((S\circ\tau_a)(x))\, d\mu(x) = \int_X \varphi(\tau_b(y))\, d\nu(y)$$

So $(\tau_b)_\# \nu = \omega$ and vice versa.

(b) Let $T^{\mu\nu}$ be an optimal map between μ and ν, i.e.

$\tau(\mu,\nu) = \mathbb{M}(T^{\mu\nu})$ Define $S = \tau_b(T^{\mu\nu} \circ \tau_a)$. ~~By part~~

~~(c) $S_\#(\tau_a$~~ Let $\mu_a = (\tau_a]_\# \mu$ and ~~$\cancel{k}$~~ $\nu_b = (\tau_b]_\# \nu$.

By part (c) $S_\# \mu_a = \nu_b$ so

$$\tau(\mu_a, \nu_b) \leq \mathbb{M}(S)$$

$$= \int_{-\infty}^{\infty} |S(x) - x|^2 \, d\mu_a(x)$$

$$= \int_{-\infty}^{\infty} |T^{\mu\nu}(x+a) - b - x|^2 \, d[\tau_a]_\# \mu(x)$$

$$= \int_{-\infty}^{\infty} |T^{\mu\nu}(\tau_a(x)+a) - b - \tau_a(x)|^2 \, d\mu(x) \quad \text{by Prop 1.5}$$

$$= \int_{-\infty}^{\infty} |T^{\mu\nu}(x) - b - x + a|^2 \, d\mu(x)$$

$$= \int_{-\infty}^{\infty} |T^{\mu\nu}(x) - x|^2 \, d\mu(x) + (a-b)^2 \int_{-\infty}^{\infty} d\mu(x)$$

$$+ 2\int_{-\infty}^{\infty} (T^{\mu\nu}(x) - x)(a-b) \, d\mu(x)$$

$$= \tau(\mu,\nu) + (a-b)^2 + 2(a-b)\left(\int_{-\infty}^{\infty} T^{\mu\nu}(x) \, d\mu(x) - \int_{-\infty}^{\infty} x \, d\mu(x) \right)$$

$$= \tau(\mu,\nu) + (a-b)^2 + 2(a-b)\left(\int_{-\infty}^{\infty} y \, d(\underbrace{T^{\mu\nu}_\# \mu}_{=\nu})(y) - \int_{-\infty}^{\infty} x \, d\mu(x) \right)$$

$$= \tau(\mu,\nu) + (a-b)^2 + 2(a-b)(M_\nu - M_\mu). \quad (*)$$

Conversely, let T^{μ_a,ν_b} be an optimal map between μ_a and ν_b.

Define $S = \tau_{-b} \circ (T^{\mu_a,\nu_b} \circ \tau_a)$. By part (a) (exchange $a \mapsto -a$, $b \mapsto -b$, $\mu \mapsto \mu_a$, $\nu \mapsto \nu_b$)

we have $S_\#([\tau_a]_\# \mu_a) = [\tau_{-b}]_\# \nu_b$.

Since $[\tau_a]_\# \mu_a = [\tau_a]_\# ([\tau_a]_\# \mu)$

$= (\tau_a \circ \tau_a)_\# \mu$

$= \mu$

and similarly $[\tau_{-b}]_\# \nu_b = \nu$ we have

that $S_\# \mu = \nu$, so S is a candidate map between

μ and ν. In particular

$$\begin{aligned}
\mathcal{T}(\mu,\nu) &\le M(S) \\
&= \int_{\mathbb{R}} |S(x) - x|^2 \, d\mu(x) \\
&= \int_{\mathbb{R}} | T^{\mu_a,\nu_b}(x-a) + b - x |^2 \, d(\tau_a)_\# \mu_a(x) \\
&= \int_{\mathbb{R}} | T^{\mu_a,\nu_b}(\tau_a(x) - a) + b - \tau_a(x) |^2 \, d\mu_a(x) \\
&= \int_{\mathbb{R}} | T^{\mu_a,\nu_b}(x) + b - x - a |^2 \, d\mu_a(x) \\
&= \int_{\mathbb{R}} | T^{\mu_a,\nu_b}(x) - x |^2 \, d\mu_a(x) + (b-a)^2 + 2(b-a) \int_{\mathbb{R}} (T^{\mu_a,\nu_b}(x) - x) \, d\mu_a(x) \\
&\cancel{= \mathcal{T}(\mu_a,\nu_b) + (b-a)^2 - 2(b-a)}
\end{aligned}$$

Now,

$$\int_{\mathbb{R}} x \; d\mu_a(x) = \int_{\mathbb{R}} x \; d[\tau_a]_{\#}\mu(x) = \int_{\mathbb{R}} \tau_a(x) \; d\mu(x)$$

$$= M_\mu - a$$

And

$$\int_{\mathbb{R}} T^{\mu_a, \nu_b}(x) \; d\mu_a(x) = \cancel{\int_{\mathbb{R}} T^{\mu_a,\nu_b}(x) \; d[\tau_a]_{\#}\mu(x)} = \int_{\mathbb{R}}$$

$$= \int_{\mathbb{R}} y \; d\, T^{\mu_a,\nu_b}_{\#}\mu_a(y)$$

$$= \int_{\mathbb{R}} y \; d\nu_b(y) = \int_{\mathbb{R}} y \; d[\tau_b]_{\#}\nu(y)$$

$$= \int_{\mathbb{R}} \tau_b(y) \, d\nu(y)$$

$$= M_\nu - b$$

So,

$$\tau(\mu,\nu) \le \tau(\mu_a,\nu_b) + (b-a)^2 + 2(b-a)(M_\nu - b - M_\mu + a)$$

$$= \tau(\mu_a,\nu_b) + (b-a)^2 - 2(b-a)^2 + 2(b-a)(M_\nu - M_\mu)$$

$$= \tau(\mu_a,\nu_b) - (b-a)^2 - 2(a-b)(M_\nu - M_\mu) \qquad (**)$$

Combining (*) and (**) implies

$$\tau(\mu,\nu) = \tau(\mu_a,\nu_b) - (b-a)^2 - 2(a-b)(M_\nu - M_\mu).$$

[Alternative proof: use explicit construction of transport maps in $\mathbb{R}$ in terms of the cdf's.]