Monge Formulation of Optimal Transport — Exercises

Exercise 1 (Discrete assignment)**.** Let $n = 3$ and $C = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 2 & 2 \\ 4 & 3 & 1 \end{pmatrix}$. Identify every optimal assignment.

Exercise 2 (Two points to two points: when does each permutation win?)**.** Fix $x_1, y_1, y_2 \in \mathbb{R}^d$ and let $x_2 = z \in \mathbb{R}^d$ vary. With quadratic cost $c(x, y) = \|x - y\|^2$, compare the two assignments

$$\text{Id: } x_1 \mapsto y_1,\ z \mapsto y_2 \qquad \text{vs} \qquad \text{Swap: } x_1 \mapsto y_2,\ z \mapsto y_1.$$

Show that the set of z where Id is better than Swap is a half-space bounded by a hyperplane orthogonal to $y_1 - y_2$. Give the equation of this hyperplane and explain that it divides $\mathbb{R}^d$ into two regions, one for each optimal permutation.

Exercise 3 (Discrete $\to$ continuous and continuous $\to$ discrete)**.** (1) In 1D let $\alpha = \sum_{i=1}^N a_i\, \delta_{x_i}$ (finite support) and let β be absolutely continuous. Show that there is no measurable map T with $T_\# \alpha = \beta$.

(2) Conversely, let α be absolutely continuous on $\mathbb{R}$ and let $\beta = \sum_{j=1}^M b_j\, \delta_{y_j}$ with $\sum_j b_j = 1$. Build many maps T with $T_\# \alpha = \beta$.

(3) Let $\alpha = \mathcal{U}[0, 1]$ and let $\beta = \sum_{j=1}^M b_j\, \delta_{y_j}$ on $\mathbb{R}$ with $b_j > 0$, $\sum_j b_j = 1$, and $y_1 < \cdots < y_M$. For $p \geq 1$, compute the W_p–optimal Monge map $T : [0, 1] \to \mathbb{R}$ from α to β.

(4) Suppose $b_j = 1/n$ for $j = 1, \ldots, n$ and $y_j = \frac{j-1/2}{n}$ (midpoints of the n equal subintervals of $[0, 1]$). Compute $W_2^2(\alpha, \beta)$.

Exercise 4 (Uniform laws on intervals, quadratic cost)**.** Let $\alpha = \mathcal{U}_{[a,b]}$ and $\beta = \mathcal{U}_{[c,d]}$ on $\mathbb{R}$ with $a < b$, $c < d$, and $c(x, y) = |x - y|^2$.

(i) Find the Monge optimal transport map T.

(ii) Compute $\mathsf{W}_2^2(\alpha, \beta)$ in closed form.

Exercise 5 (Many Monge solutions for $|x - y|$)**.** Let $\alpha = \mathcal{U}_{[0,2]}$, $\beta = \mathcal{U}_{[1,3]}$ and $c(x, y) = |x - y|$.

(i) Show $T_1(x) = x + 1$ is optimal and $\mathsf{W}_1(\alpha, \beta) = 1$.

(ii) Show

$$T_2(x) = \begin{cases} x + 2, & x \in [0, 1], \\ x, & x \in (1, 2], \end{cases}$$

is also optimal.

(iii) Build a family of additional Monge solutions.

Exercise 6 (Affine images)**.** (1) **1-D case.** Let $\nu = (x \mapsto ax + b)_\# \mu$ with $a > 0$. what is the Monge map from μ to ν for the quadratic cost ? Compute the W_2 distance in terms of the mean $m_\mu = \int x\, d\mu$ and variance $\mathrm{Var}(\mu) = \int (x - m_\mu)^2\, d\mu$.

(2) **Higer dimension.** Let $\nu = (x \mapsto Ax + b)_\# \mu$ with $A \in \mathbb{R}^{d \times d}$. Give a simple condition on A ensuring that $T(x) = Ax + b$ is the Brenier map (assume μ is absolutely continuous). Under this condition, compute the W_2 distance in terms of $m_\mu = \mathbb{E}[X]$ and the covariance $\Sigma_\mu = \mathbb{E}\big[(X - m_\mu)(X - m_\mu)^\top\big]$.

Exercise 7 (Piecewise constant densities on [0,1)**.** , quadratic cost] Let α have density $f = 1$ on $[0, \frac{1}{2})$ and $f = 2$ on $[\frac{1}{2}, 1]$, while β has density $g = 2$ on $[0, \frac{1}{2})$ and $g = 1$ on $[\frac{1}{2}, 1]$. Construct (and sketch) the optimal Monge map on $\mathbb{R}$ for the quadratic cost and compute the exact cost (for the corresponding probability measures obtained by normalising by the common mass 3/2).

Exercise 8 (Translation decomposition for W_2)**.** Let $\alpha, \beta \in \mathcal{P}_2(\mathbb{R}^d)$ (finite second moments). For $u \in \mathbb{R}^d$ write $T_u(x) = x + u$.

(a) Show there exist unique vectors a, b such that the *centred* measures

$$\alpha_0 := (T_{-a})_\#\alpha, \qquad \beta_0 := (T_{-b})_\#\beta$$

satisfy $\int x\, d\alpha_0(x) = \int x\, d\beta_0(x) = 0$, and

$$\alpha = (T_a)_\#\alpha_0, \qquad \beta = (T_b)_\#\beta_0.$$

(b) Prove the decomposition

$$\boxed{\mathsf{W}_2^2(\alpha, \beta) = \|a - b\|^2 + \mathsf{W}_2^2(\alpha_0, \beta_0)}.$$

(c) If α is absolutely continuous and T_0 is the Brenier map from α_0 to β_0, show that

$$T(x) := T_0(x - a) + b$$

is the Brenier map from α to β.

Exercise 9 (Adding separable terms; quadratic cost vs. $-\langle x, y\rangle$)**.** Fix probability measures α, β on $\mathbb{R}^d$ and consider Monge maps T with $T_\#\alpha = \beta$.

(a) Let $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a measurable cost and define $\tilde{c}(x, y) = c(x, y) + u(x) + v(y)$, with $u : \mathbb{R}^d \to \mathbb{R}$, $v : \mathbb{R}^d \to \mathbb{R}$. Show that the sets of Monge minimizers for c and for $\tilde{c}$ (with the same marginals) coincide.

(b) Let $\lambda > 0$ and define $\hat{c}(x, y) = \lambda\, c(x, y)$. Show that the Monge minimizers for c and $\hat{c}$ coincide.

(c) Take $c_{\text{quad}}(x, y) = \|x - y\|^2, \quad c_{\text{lin}}(x, y) = -\langle x, y\rangle$. Prove that c_{quad} and c_{lin} have the same Monge minimizers (for fixed α, β).

Exercise 10 (Quantile coupling: explicit computation)**.** Let $\mu = \mathcal{U}[0, 1]$ and let ν be supported on $[0, 1]$ with density $2y$. (a) Compute the cumulative distribution functions F_μ and F_ν, and their quantile functions F_μ^{-1} and F_ν^{-1}. (b) Using monotone transport in 1D, write the optimal map $T = F_\nu^{-1} \circ F_\mu$. (c) Compute $W_2^2(\mu, \nu)$; using both $\int_0^1 |x - T(x)|^2\, dx$ and the 1D quantile formula $W_2^2(\mu, \nu) = \int_0^1 |F_\mu^{-1}(t) - F_\nu^{-1}(t)|^2\, dt$.

Exercise 11 (Cumulative functions)**.** Let α be a probability measure on $\mathbb{R}$ and let its cumulative distribution function be

$$C_\alpha(x) = \alpha\big((-\infty, x]\big), \qquad x \in \mathbb{R}.$$

Consider the pushforward by C_α, namely $(C_\alpha)_\#\alpha$.

(1) Assume α has a density with respect to the Lebesgue measure (equivalently, C_α is continuous). Prove that

$$(C_\alpha)_\#\alpha = \mathcal{U}[0, 1].$$

In your proof, clearly point out *where* the density (continuity) assumption is used.

(2) What happens if α is purely discrete, say $\alpha = \sum_i p_i\, \delta_{x_i}$ with $\sum_i p_i = 1$ and $x_1 < x_2 < \cdots$? Describe the law of $C_\alpha(X)$ for $X \sim \alpha$, and show that for all $t \in [0, 1]$,

$$\mathbb{P}\big(C_\alpha(X) \le t\big) \le t,$$

with equality at the jump values of C_α.

Exercise 12 (Pushforward of a uniform law by a nonlinear map)**.** Let $\alpha = \mathcal{U}[-1, 1]$ and consider the map $T : \mathbb{R} \to \mathbb{R}$, $T(x) = x^2$.

(1) Compute the law $\beta = T_\#\alpha$, i.e. find its density with respect to Lebesgue measure.

(2) Check that this density integrates to 1 over its support.

(3) Sketch the graph of the density and comment on its behaviour near 0 and 1.

Optional – More difficult exercises

Exercise 13 (Radial symmetry)**.** Let α, β be *radial* probability measures on $\mathbb{R}^d$ (i.e. invariant under all orthogonal maps Q: $Q_\#\alpha = \alpha$, $Q_\#\beta = \beta$). For any probability measure μ on $\mathbb{R}^d$, define its *radial cumulative distribution function* by

$$F_\mu(r) := \mu\big(B(0,r)\big) = \mathbb{P}_{X\sim\mu}\big(\|X\| \le r\big), \qquad r \ge 0.$$

One can show that if μ is invariant under rotation, then its density f has a radial density satisfies $f(x) = h(\|x\|)$ w.r.t. Lebesgue. Assume in addition that α is absolutely continuous w.r.t. Lebesgue. Consider the quadratic cost $c(x,y) = \|x-y\|^2$.

(i) Show that the Brenier (Monge) optimizer T pushing α to β is *radial*, i.e.

$$T(x) = t(\|x\|)\,\frac{x}{\|x\|} \quad \text{for } x \neq 0, \qquad T(0) = 0,$$

for some nondecreasing and non-negative function $t : [0,\infty) \to [0,\infty)$. You will first show that T commutes with rotation, $Q \circ T \circ Q = T$ for any rotation Q.

(ii) Prove that t is characterised by mass conservation in radius:

$$F_\alpha(r) = F_\beta\big(t(r)\big) \qquad \text{for all } r \ge 0,$$

and hence $t(r) = F_\beta^{-1}\big(F_\alpha(r)\big)$ (generalised inverse).

(iii) If $\alpha = \mathcal{U}(B(0,R_\alpha))$ and $\beta = \mathcal{U}(B(0,R_\beta))$, insantiate the previous formula. Could you have guessed this formula directly?

Solution. **Solution to Exercise 13.** *(i) T is radial.* Let Q be any orthogonal map. Because α, β are radial,

$$(Q\circ T\circ Q^{-1})_\#\alpha = Q_\#\Big(T_\#\big(Q^{-1}_\#\alpha\big)\Big) = Q_\#\big(T_\#\alpha\big) = Q_\#\beta = \beta,$$

and the cost is invariant under orthogonal changes of variables:

$$\int \|x - T(x)\|^2\, d\alpha(x) = \int \|Qx - QT(x)\|^2\, d\alpha(x) = \int \|x - (Q\circ T\circ Q^{-1})(x)\|^2\, d\alpha(x).$$

Hence $Q\circ T\circ Q^{-1}$ is also an optimal map from α to β. By uniqueness (a.e.) of the Brenier map we must have

$$Q\circ T\circ Q^{-1} = T \quad \alpha\text{-a.e.}$$

Fix $x \neq 0$ and choose Q to be any rotation that fixes the direction $u := x/\|x\|$ (i.e. rotates within the orthogonal complement $u^\perp$), for instance, in 2-D Q is the axial symmetry about the axis u. Then $Qx = x$, so the above identity gives $QT(x) = T(Qx) = T(x)$, meaning $T(x)$ is fixed by all rotations around the axis $\mathbb{R}u$. The only vectors with that property are multiples of u. Therefore

$$T(x) = t(\|x\|)\,u = t(\|x\|)\,\frac{x}{\|x\|} \quad (x \neq 0).$$

We want to apply Brenier's theorem. By radial symmetry it is natural to look for a radial potential, say $\phi(x) = \psi(\|x\|)$. Differentiating gives $\nabla\phi(x) = \psi'(\|x\|)\,\frac{x}{\|x\|}$, so the transport map has the form $T(x) = t(\|x\|)\,x/\|x\|$ with $t = \psi'$. To check convexity of ϕ, we compute its Hessian. For $x \neq 0$ with $r = \|x\|$ and $u = x/r$, one finds $\nabla^2\phi(x) = \psi''(r)\,u\otimes u + \frac{\psi'(r)}{r}\,(I - u\otimes u)$. Thus the eigenvalue in the radial direction is $\psi''(r)$, and in each of the $d-1$ tangential directions the eigenvalue is $\psi'(r)/r$. Convexity requires all eigenvalues to be nonnegative, hence $\psi'(r) \ge 0$ and $\psi''(r) \ge 0$. Equivalently, $t(r) = \psi'(r)$ must be nonnegative and nondecreasing.

(ii) Characterisation of t via radial CDFs. Let $X \sim \alpha$ and write $R = \|X\|$. For $\rho \ge 0$, using that T is radial we have

$$\{\,\|T(X)\| \le \rho\,\} = \{\,t(\|X\|) \le \rho\,\} = \{\,R \le t^{-1}(\rho)\,\} \quad (\text{with } t^{-1}(\rho) := \inf\{r : t(r) \ge \rho\}).$$

Taking probabilities,

$$F_\beta(\rho) = \mathbb{P}(\|T(X)\| \le \rho) = \mathbb{P}(R \le t^{-1}(\rho)) = F_\alpha\big(t^{-1}(\rho)\big).$$

Equivalently,

$$F_\alpha(r) = F_\beta\big(t(r)\big) \qquad (r \ge 0),$$

so $t = F_\beta^{-1} \circ F_\alpha$ (generalised inverse). This determines t uniquely on the support of α. The monotonicity of t follows from the monotonicity of F_α, F_β.

(iii) Uniform measures on balls. If $\alpha = \mathcal{U}(B(0, R_\alpha))$ and $\beta = \mathcal{U}(B(0, R_\beta))$, then for $0 \le r \le R_\alpha$ and $0 \le \rho \le R_\beta$,

$$F_\alpha(r) = \Big(\frac{r}{R_\alpha}\Big)^d, \qquad F_\beta(\rho) = \Big(\frac{\rho}{R_\beta}\Big)^d.$$

Solving $F_\alpha(r) = F_\beta\big(t(r)\big)$ gives

$$\Big(\frac{r}{R_\alpha}\Big)^d = \Big(\frac{t(r)}{R_\beta}\Big)^d \quad \Rightarrow \quad t(r) = \frac{R_\beta}{R_\alpha}\, r.$$

Therefore

$$\boxed{T(x) = \frac{R_\beta}{R_\alpha}\, x \quad (\alpha\text{-a.e.}).}$$

At $x = 0$ set $T(0) = 0$.