Central limit theorems for transportation cost in general dimension

Eustasio del Barrio

Universidad de Valladolid. IMUVA.

Oaxaca, May, 2017

joint work with Jean-Michel Loubes

Outline

1. Empirical optimal transportation & matching
2. Uniqueness and stability of optimal transportation potentials
3. Variance bounds
4. CLTs for empirical transportation cost

Empirical transportation cost

P, Q probabilities on $\mathbb{R}^d$ and $c(x,y) = \|x-y\|^p$, $p \geq 1$.

$$\mathcal{W}_p^p(P,Q) = \min_{\pi \in \Pi(P,Q)} \int \|x-y\|^p d\pi(x,y)$$

$\Pi(P,Q)$ probabilities on $X \times Y$ with marginals P and Q

$\mathcal{W}_p$ is a metric on $\mathcal{F}_p$, probabilities on $\mathbb{R}^d$ with finite p-th moment

$X_1, \ldots, X_n \in \mathbb{R}^d$, $P_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$

Empirical transportation cost: $\mathcal{W}_p^p(P_n, Q)$

Empirical transportation cost

P, Q probabilities on $\mathbb{R}^d$ and $c(x,y) = \|x-y\|^p$, $p \geq 1$.

$$\mathcal{W}_p^p(P,Q) = \min_{\pi \in \Pi(P,Q)} \int \|x-y\|^p d\pi(x,y)$$

$\Pi(P,Q)$ probabilities on $X \times Y$ with marginals P and Q

$\mathcal{W}_p$ is a metric on $\mathcal{F}_p$, probabilities on $\mathbb{R}^d$ with finite p-th moment

$X_1, \ldots, X_n \in \mathbb{R}^d$, $P_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$

Empirical transportation cost: $\mathcal{W}_p^p(P_n, Q)$

What is the transportation cost from a (large) set of points to a fixed target?

Assume $X_1, \ldots, X_n$ i.i.d. P

Optimal matching

$X_1, \ldots, X_n \in \mathbb{R}^d$, $Y_1, \ldots, Y_n \in \mathbb{R}^d$

Cost of matching X_i to Y_j: $\|X_i - Y_j\|^p$

Optimal matching minimizes $\frac{1}{n}\sum_{i=1}^n \|X_i - Y_{\sigma(i)}\|^p$
σ permutation of $\{1, \ldots, n\}$.

Optimal matching cost $= \mathcal{W}_p^p(P_n, Q_n)$,

$$P_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}, \quad Q_n = \frac{1}{n}\sum_{i=1}^n \delta_{Y_i}$$

Optimal matching

$X_1, \ldots, X_n \in \mathbb{R}^d$, $Y_1, \ldots, Y_n \in \mathbb{R}^d$

Cost of matching X_i to Y_j: $\|X_i - Y_j\|^p$

Optimal matching minimizes $\frac{1}{n}\sum_{i=1}^n \|X_i - Y_{\sigma(i)}\|^p$
σ permutation of $\{1, \ldots, n\}$.

Optimal matching cost $= \mathcal{W}_p^p(P_n, Q_n)$,

$$P_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}, \quad Q_n = \frac{1}{n}\sum_{i=1}^n \delta_{Y_i}$$

What is the cost of matching two (large) sets of points?

Assume $X_1, \ldots, X_n$ i.i.d. P, $Y_1, \ldots, Y_n$ i.i.d. Q, independent of X_i's

$\mathcal{W}_p(P_n, P) \to 0$ iff $P_n \underset{w}{\to} P$ and $\int \|x\|^p dP_n \to \int \|x\|^p dP$.

P with finite p-th moment, P_n empirical measure $\Rightarrow \mathcal{W}_p(P_n, P) \to 0$ a.s.

Hence, $\mathcal{W}_p(P_n, Q) \to \mathcal{W}_p(P, Q)$ a.s., $\mathcal{W}_p(P_n, Q_n) \to \mathcal{W}_p(P, Q)$ a.s.

$\mathcal{W}_p(P_n, P) \to 0$ iff $P_n \underset{w}{\to} P$ and $\int \|x\|^p dP_n \to \int \|x\|^p dP$.

P with finite p-th moment, P_n empirical measure $\Rightarrow \mathcal{W}_p(P_n, P) \to 0$ a.s.

Hence, $\mathcal{W}_p(P_n, Q) \to \mathcal{W}_p(P, Q)$ a.s., $\mathcal{W}_p(P_n, Q_n) \to \mathcal{W}_p(P, Q)$ a.s.

How fast? Rates of convegence

$\mathcal{W}_p(P_n, P) \to 0$ iff $P_n \underset{w}{\to} P$ and $\int \|x\|^p dP_n \to \int \|x\|^p dP$.

P with finite p-th moment, P_n empirical measure $\Rightarrow \mathcal{W}_p(P_n, P) \to 0$ a.s.

Hence, $\mathcal{W}_p(P_n, Q) \to \mathcal{W}_p(P, Q)$ a.s., $\mathcal{W}_p(P_n, Q_n) \to \mathcal{W}_p(P, Q)$ a.s.

How fast? Rates of convegence

Description of fluctuation?

$\mathcal{W}_p(P_n, P) \to 0$ iff $P_n \underset{w}{\to} P$ and $\int \|x\|^p dP_n \to \int \|x\|^p dP$.

P with finite p-th moment, P_n empirical measure $\Rightarrow \mathcal{W}_p(P_n, P) \to 0$ a.s.

Hence, $\mathcal{W}_p(P_n, Q) \to \mathcal{W}_p(P, Q)$ a.s., $\mathcal{W}_p(P_n, Q_n) \to \mathcal{W}_p(P, Q)$ a.s.

How fast? Rates of convegence

Description of fluctuation?

The case $P = Q$

For $d = 2$, (Ajtai-Komlos-Tusnady, 1984; Talagrand & Yukich, 1993)

$$c(p)\left(\frac{\log n}{n}\right)^{1/2} \leq E(\mathcal{W}_p(P_n, U([0,1]^2))) \leq C(p)\left(\frac{\log n}{n}\right)^{1/2}.$$

For $d \geq 3$, Talagrand, Yukich, 1992-1994

$$E(\mathcal{W}_p(P_n, U([0,1]^d))) \leq C(d,p)\frac{1}{n^{1/d}}.$$

Extensions to compactly supported P with 'regular' density

If $d = 1$ and $P \sim f$ s.t. $\int_0^1 \left(\frac{(t(1-t))^{1/2}}{f(F^{-1}(t))}\right)^p dt < \infty$

$$\sqrt{n}\mathcal{W}_p(P_n, P) \to_w \left[\int_0^1 \left(\frac{B(t)}{f(F^{-1}(t))}\right)^p dt\right]^{1/p},$$

$B(t)$ Brownian bridge on $[0,1]$

No results for $P \neq Q$

An exception: Sommerfeld and Munk (2016) for the case P, Q with finite support; possibly nonnormal limits

Here CLTs for $\mathcal{W}_2^2(P_n, Q)$ and $\mathcal{W}_2^2(P_n, Q_m)$ for general P, Q and d

Valid CLTs, with normal limits under moment assumptions $(4 + \delta)$ and a bit of smoothness (on Q) asymptotic variances easily described in terms of dual formulation of OT

No results for $P \neq Q$

An exception: Sommerfeld and Munk (2016) for the case P, Q with finite support; possibly nonnormal limits

Here CLTs for $\mathcal{W}_2^2(P_n, Q)$ and $\mathcal{W}_2^2(P_n, Q_m)$ for general P, Q and d

Valid CLTs, with normal limits under moment assumptions $(4+\delta)$ and a bit of smoothness (on Q) asymptotic variances easily described in terms of dual formulation of OT

Beyond theoretical interest,

> *[the transportation cost distance] is an attractive tool for data analysis but statistical inference is hindered by the lack of distributional limits*

Sommerfeld and Munk (2016)

The Kantorovich duality

Denote

$$I[\pi] = \int_{\mathbb{R}^d \times \mathbb{R}^d} x \cdot y d\pi(x, y),$$

$\Phi = \{(\varphi, \psi) \in L_1(P) \times L_1(Q) : \varphi(x) + \psi(y) \geq x \cdot y$ for all $x, y\}$, and

$$J(\varphi, \psi) = \int_{\mathbb{R}^d} \varphi dP + \int_{\mathbb{R}^d} \psi dQ.$$

Then,

$$\min_{(\varphi,\psi)\in\Phi} J(\varphi, \psi) = \max_{\pi\in\Pi(P,Q)} \tilde{I}[\pi]$$

Maximizing pair for J can be chosen as pair of lsc, proper convex conjugate functions $\varphi(x) = \psi^*(x) \sup_{y\in\mathbb{R}^d}(x \cdot y - \psi(y))$

By Kantorovich duality, (ψ^*, ψ) is a minimizer of J and π is a maximizer of $\tilde{I}$ iff

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} (\psi^*(x) + \psi(y) - x \cdot y) d\pi(x, y) = 0,$$

iff $\psi^*(x) + \psi(y) - x \cdot y$ vanishes π-almost surely

Now $\psi^*(x) + \psi(y) - x \cdot y = 0 \iff x \in \partial\psi(y) \iff y \in \partial\psi^*(x)$,

$$\partial\psi(y) = \{z \in \mathbb{R}^d : \psi(y') - \psi(y) \geq z \cdot (y' - y) \text{ for all } y' \in \mathbb{R}^d\}$$

$\partial\psi(y)$ nonempty if $y \in \text{int}(\text{dom}(\psi))$; if ψ differentiable at y, $\partial\psi(y) = \{\nabla\psi(y)\}$

From this (Knott, Smith, Brenier,...) (ψ^*, ψ) a minimizing pair for J iff $Q \circ (\nabla\psi)^{-1} = P$; then $\pi = Q \circ (\nabla\psi, Id)^{-1}$ maximizes $\tilde{I}$.

$T = \nabla\psi$ optimal transportation map from Q to P; it is Q-a.s. unique:

Optimal transportation potential: lsc convex ψ s.t. (ψ^*, ψ) minimizes J (equivalently, lsc convex ψ s.t. such that $Q \circ (\nabla\psi)^{-1} = P$

Optimal transportation potentials not unique $(J(\psi^* - C, \psi + C) = J(\psi^*, \psi))$

Lemma

Assume ψ_1 and ψ_2 finite convex functions on nonempty convex, open $A \subset \mathbb{R}^d$ s.t.

$$\nabla\psi_1(x) = \nabla\psi_2(x) \quad \text{for a.e. } x \in A.$$

Then $\psi_1(x) = \psi_2(x) + C$ for all $x \in A$

As a consequence

Corollary

Assume P, $Q \in \mathcal{F}_2$ and

$$Q \text{ has a positive density in the interior of its convex support.} \tag{1}$$

Then, if ψ_1, ψ_2 are lsc convex and $J(\psi_1^*, \psi_1) = J(\psi_2^*, \psi_2) = \min_{(\varphi,\psi)\in\Phi} J(\varphi, \psi)$ $\psi_2 = \psi_1 + C$ in int(supp(Q))*. In particular, $\psi_2 = \psi_1 + C$ Q-a.s..*

Uniqueness of optimal transportation potential fails without (1)
(Take $P = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, Q_ε is the uniform on $(-\varepsilon - 1, -\varepsilon) \cup (\varepsilon, 1 + \varepsilon)$, $\varepsilon > 0$; $\psi_{\varepsilon,L}(x) = -x$, $x \leq -\frac{L}{2}$, $\psi_{\varepsilon,L}(x) = x + L$, $x \geq -\frac{L}{2}$, $0 < L < \varepsilon$, are optimal transportation potentials, but $\psi_{\varepsilon,L_2} \neq \psi_{\varepsilon,L_1} + C$)

Stability of optimal transportation potentials

Assume Q with a density, $\mathcal{W}_2(P_n, P) \to 0$,

If $\nabla\psi_n$ is o.t.p. from Q to P_n, $\nabla\psi$ is o.t.p. from Q to P, then

$$\nabla\psi_n \to \nabla\psi \quad Q-\text{a.s.}$$

How about ψ_n?

Approach based on Painlevé-Kuratowski convergence: if C_n subsets of $\mathbb{R}^d$

$$\limsup_{n\to\infty} C_n = \Big\{x \in \mathbb{R}^d : \, x = \lim_{j\to\infty} x_{n_j} \text{ for some } x_{n_j} \in C_{n_j}\Big\},$$

$$\liminf_{n\to\infty} C_n = \Big\{x \in \mathbb{R}^d : \, x = \lim_{n\to\infty} x_n \text{ with } x_n \in C_n \text{ if } n \geq n_0\Big\}$$

$C_n \to C$ in P-K sense if $C = \liminf_{n\to\infty} C_n = \limsup_{n\to\infty} C_n$

If T multivalued map from $\mathbb{R}^d$ to $\mathbb{R}^d$ (for each $x \in \mathbb{R}^d$, $T(x)$ is a subset of $\mathbb{R}^d$),

$$\text{gph}(T) = \Big\{(x,t) \in \mathbb{R}^d \times \mathbb{R}^d : \, t \in T(x)\Big\}.$$

Multivalued maps identified with subsets of $\mathbb{R}^d \times \mathbb{R}^d$
If T_n, T multivalued maps, $T_n \to T$ graphically if $\text{gph}(T_n) \to \text{gph}(T)$ in P-K sense
Some useful results

Theorem

(a) *Assume that for some $\varepsilon > 0$ and some subsequence $\{n_j\}$ $C_{n_j} \cap B(0,\varepsilon) \neq \emptyset$ for every $j \geq 1$. Then there exists a subsequence $\{n_{j_k}\}$ and a nonempty subset $C \subset \mathbb{R}^d$ such that $C_{n_{j_k}} \to C$ in P-K sense.*

(b) *Assume $\{T_n\}_{n\geq 1}$ multivalued maps such that for some bounded sets $C, D \subset \mathbb{R}^d$ and some subsequence $\{n_j\}$ there exist $x_{n_j} \in C$ with $T_{n_j}(x_{n_j}) \cap D \neq \emptyset$ for all $j \geq 1$. Then there exists a subsequence $\{n_{j_k}\}$ and a multivalued map, T, from $\mathbb{R}^d$ to $\mathbb{R}^d$, with nonempty domain s.t. $T_{n_{j_k}}$ converges graphically to T.*

Recall that π optimal (a maximizer of I) iff supp$(\pi) \subset$ gph$(\partial\psi)$ for some lsc convex ψ

Subgradients of convex maps characterized in terms of cyclical monotonicity:

T monotone if $(t_1 - t_0) \cdot (x_1 - x_0) \geq 0$ whenever $t_i \in T(x_i)$, $i = 0, 1$.

T cyclically monotone if for every choice of $m \geq 1$, points $x_0, \ldots, x_m$ and $t_i \in T(x_i)$, $i = 0, \ldots, m$

$$t_0 \cdot (x_1 - x_0) + t_1 \cdot (x_2 - x_1) + \cdots + t_m \cdot (x_0 - x_m) \leq 0.$$

Rockafellar's Theorem: $T = \partial\psi$ for some lsc convex ψ iff T maximal cyclically monotone

Theorem

If T_n cyclically monotone maps $\{T_n\}$ and $T_n \to T$ graphically then T is cyclically monotone. If T_n are maximal cyclically monotone then T is also maximal cyclically monotone.
If $\{\psi_n\}$ lsc, convex maps s.t. for some bounded $C, D \subset \mathbb{R}^d$ and some $\{n_j\}$ there exist $x_{n_j} \in C$ with $\partial\psi_{n_j}(x_{n_j}) \cap D \neq \emptyset$ for all $j \geq 1$, then there exist $\{n_{j_k}\}$ and a lsc convex ψ with dom$(\partial\psi) \neq \emptyset$ s.t. $\partial\psi_{n_{j_k}} \to \partial\psi$ graphically

If $\partial\psi_n \to \partial\psi$ graphically and for some (x_n, t_n) with $t_n \in \partial\psi_n(x_n)$ and (x_0, t_0) with $t_0 \in \partial\psi(x_0)$

$$(x_n, t_n) \to (x_0, t_0) \text{ and } \psi_n(x_n) \to \psi(x_0),$$

then

$$\lim_{n\to\infty} \psi_n(\tilde{x}_n) = \psi(x)$$

if $x \in \text{int}(\text{dom}(\psi))$

Theorem (Stability of optimal transportation potentials)

Assume Q satisfies (1) and $\mathcal{W}_2(P_n, P) \to 0$ and $\mathcal{W}_2(Q_n, Q) \to 0$. If ψ_n (resp. ψ) optimal transportation potentials from Q_n to P_n (resp. from Q to P) then there exist constants a_n such that if $\tilde{\psi}_n = \psi_n - a_n$ then $\tilde{\psi}_n(x) \to \psi(x)$ for every x in the interior of the support of Q (hence, for Q-almost every x)

Proof: If π_n, π o.t.plans $\pi_n \to_w \pi$; supp$(\pi_n) \subset$ gph$(\partial\psi_n)$ supp$(\pi) \subset$ gph$(\partial\psi) \Rightarrow \partial\psi_n \to \partial\rho$ graphically (along subsequences); $\rho = \psi(+C)$ in int(dom(ψ)); re-center to conclude.

If $Q_n = Q$ and (1) holds ψ_n differentiable at a.e. $x \in A$; from graphical convergence of $\partial\psi_n$ to $\partial\rho$ with $\rho = \psi$ in A conclude

$$\nabla\psi_n(x) \to \nabla\psi(x) \text{ at a.e. } x \in A$$

$\nabla\psi_n \to \nabla\psi$ Q-a.s
Recover known stability of o.t.maps

Theorem

Assume $Q, P, \{P_n\}_{n\geq 1} \in \mathcal{F}_4$ and Q satisfies (1); ψ_n, ψ optimal transportation potentials s.t. $\psi_n \to \psi$ Q-a.s. Then

$$\psi_n \to \psi \text{ in } L_2(Q)$$

Efron-Stein inequality

Assume $X_1, \ldots, X_n$ independent r.v.'s; $(X'_1, \ldots, X'_n)$ independent copy of $(X_1 \ldots, X_n)$
If $Z = f(X_1, \ldots, X_n)$ then

$$\mathrm{Var}(Z) \leq \frac{1}{2} \sum_{i=1}^{n} E(Z - Z_i)^2 = \sum_{i=1}^{n} E(Z - Z_i)_+^2,$$

with $Z_i = f(X_1, \ldots, X'_i, \ldots, X_n)$

If f symmetric in $x_1, \ldots, x_n$ and $X_1, \ldots, X_n$ i.i.d. then

$$\mathrm{Var}(Z) \leq nE(Z - Z_1)_+^2$$

Control of (one-sided) decrease of Z when X_1 replaced by X'_1 enough for control of $\mathrm{Var}(Z)$

Perfect for minimization functionals of empirical measure

Variance bounds for $\mathcal{W}_2^2(P_n, Q)$

If Q smooth $\mathcal{W}_2^2(P_n, Q) = \sum_{i=1}^n \int_{C_i} \|y - X_i\|^2 dQ(y)$ with

$$C_i = \Big\{ y : \nabla \psi_n(y) = X_i \Big\},$$

ψ_n optimal transportation potential from Q to P_n

P_n' empirical measure on $X_1', X_2, \ldots, X_n$; ψ_n' optimal transportation potential from Q to P_n'

Set $T(y) = X_i$ if $\nabla \psi_n'(y) = X_i$, $i = 2, \ldots, n$, $T(y) = X_1$ if $\nabla \psi_n'(y) = X_1'$

T suboptimal, but maps Q to P_n; hence,

$$\begin{aligned} \mathcal{W}_2^2(P_n, Q) - \mathcal{W}_2^2(P_n', Q) &\leq \int \|y - T(y)\|^2 dQ(y) - \int \|y - \nabla \psi_n'(y)\|^2 dQ(y) \\ &= \int_{C_1'} \Big(\|y - X_1\|^2 - \|y - X_1'\|^2 \Big) dQ(y) \end{aligned}$$

Consequence:

Theorem

If $P, Q \in \mathcal{F}_4$ and Q has a density

$$\mathrm{Var}(\mathcal{W}_2^2(P_n, Q)) \leq \frac{C(P,Q)}{n},$$

where

$$C(P,Q) = 8\Big(E(\|X_1 - X_2\|^2\|X_1\|^2) + (E\|X_1 - X_2\|^4)^{1/2}\Big(\int_{\mathbb{R}^d}\|y\|^4 dQ(y)\Big)^{1/2}\Big).$$

Alternative bound: if (φ_n, ψ_n) minimizers of J

$$\mathcal{W}_2^2(P_n, Q) = \int_{\mathbb{R}^d} (\|x\|^2 - 2\varphi_n(x))dP_n(x) + \int_{\mathbb{R}^d} (\|y\|^2 - 2\psi_n(y))dQ(y)$$

Similar for $\mathcal{W}_2^2(P_n', Q)$; by optimality,

$$\mathcal{W}_2^2(P_n', Q) \geq \int_{\mathbb{R}^d} (\|x\|^2 - 2\varphi_n(x))dP_n'(x) + \int_{\mathbb{R}^d} (\|y\|^2 - 2\psi_n(y))dQ(y).$$

Hence,

$$\begin{aligned} \mathcal{W}_2^2(P_n, Q) - \mathcal{W}_2^2(P_n', Q) &\leq \int_{\mathbb{R}^d} (\|x\|^2 - 2\varphi_n(x))dP_n(x) \\ &\quad - \int_{\mathbb{R}^d} (\|x\|^2 - 2\varphi_n(x))dP_n'(x) \\ &= \frac{1}{n}\left[(\|X_1\|^2 - \varphi_n(X_1)) - (\|X_1'\|^2 - \varphi_n(X_1'))\right] \end{aligned}$$

Consequence,

$$\mathrm{Var}(\mathcal{W}_2^2(P_n, Q)) \leq \frac{E(\left[(\|X_1\|^2 - \varphi_n(X_1)) - (\|X_1'\|^2 - \varphi_n(X_1'))\right)^2}{n} := \frac{C_n}{n}$$

C_n harder to control; however, if $P, Q \in \mathcal{F}_{4+\delta}$ and satisfy (1) $C_n \to C < \infty$ (sharp constants)

More important, linearization bounds:

Theorem

If $P, Q \in \mathcal{F}_{4+\delta}$ and satisfy (1), φ_0 o.t. potential from P to Q and

$$R_n = \mathcal{W}_2^2(P_n, Q) - \int_{\mathbb{R}^d} (\|x\|^2 - 2\varphi_0(x)) dP_n(x),$$

then

$$n\mathrm{Var}(R_n) \to 0$$

CLTs for empirical transportation cost

Theorem

If $P, Q \in \mathcal{F}_{4+\delta}$ and satisfy (1), φ_0 o.t. potential from P to Q and P_n empirical measure on $X_1, \ldots, X_n$, i.i.d. P r.v.'s then

$$n\mathsf{Var}(\mathcal{W}_2^2(P_n, Q)) \to \sigma^2(P, Q)$$

with

$$\sigma^2(P,Q) = \int_{\mathbb{R}^d} (\|x\|^2 - 2\varphi_0(x))^2 dP(x) - \Big(\int_{\mathbb{R}^d} (\|x\|^2 - 2\varphi_0(x)) dP(x) \Big)^2$$

and

$$\sqrt{n}\big(\mathcal{W}_2^2(P_n, Q) - E\mathcal{W}_2^2(P_n, Q)\big) \underset{w}{\to} N(0, \sigma^2(P, Q))$$

Furthermore, if Q_m empirical measure on $Y_1, \ldots, Y_m$ i.i.d. Q r.v.'s, independent of the X_i's, $n \to \infty$, $m \to \infty$ with $\frac{n}{n+m} \to \lambda \in (0,1)$, then

$$\tfrac{nm}{n+m}\mathsf{Var}(\mathcal{W}_2^2(P_n, Q_m)) \to (1-\lambda)\sigma^2(P,Q) + \lambda\sigma^2(Q,P)$$

- Limiting variances well-defined (independent of choice of o.t. potentials)
- Covers optimal matching setup
- Dimension free (but dimension plays a role on centering constants)
- No assumption of compact support
- If $P = Q$, $\sigma^2(P,P) = 0$;

$$\sqrt{n}\big(\mathcal{W}_2^2(P_n, P) - E\mathcal{W}_2^2(P_n, P)\big) \to 0$$

  in probability
- Smoothness of P not really important; with a different approach

Theorem

If P has finite support, $Q \in \mathcal{F}_4$ and satisfies (1) then

$$\sqrt{n}\big(\mathcal{W}_2^2(P_n, Q) - \mathcal{W}_2^2(P, Q)\big) \underset{w}{\to} N(0, \sigma^2(P,Q))$$

Open problems

- Most of approach works for other costs $c(x, y) = \|x - y\|^p$, $p > 1$; need for stability results for optimal c-concave potentials
- What if c not stricly convex? If $c(x, y) = \|x - y\|$ nonnormal limits may happen $(d = 1)$
- Related functionals: optimal partial transportation and matching, variation around empirical Wasserstein barycenters

References

Ajtai , M., Komlós , J. and Tusnády , G. (1984). On optimal matchings. *Combinatorica*, **4** 259–264.

del Barrio, E. and Loubes, J.-M. (2017). Central limit theorems for the empirical transportation cost in general dimension. *Submitted.*

Dobrić, V. and Yukich , J. E. (1995). Asymptotics for transportation cost in high dimensions. *J. Theoret. Probab.*, **8**, 97–118.

Sommerfeld, M. and Munk, A. (2016). Inference for Empirical Wasserstein Distances on Finite Spaces. *Preprint.* https://arxiv.org/abs/1610.03287v1

Talagrand, M. (1992). Matching random samples in many dimensions. *Ann. Appl. Probab.*, **2**, 846–856.

Talagrand , M. (1994). The transportation cost from the uniform measure to the empirical measure in dimension ≥ 3. *Ann. Probab.*, **22**, 919–959.