Multivariate Tail Quantile Contours based on Optimal Transport

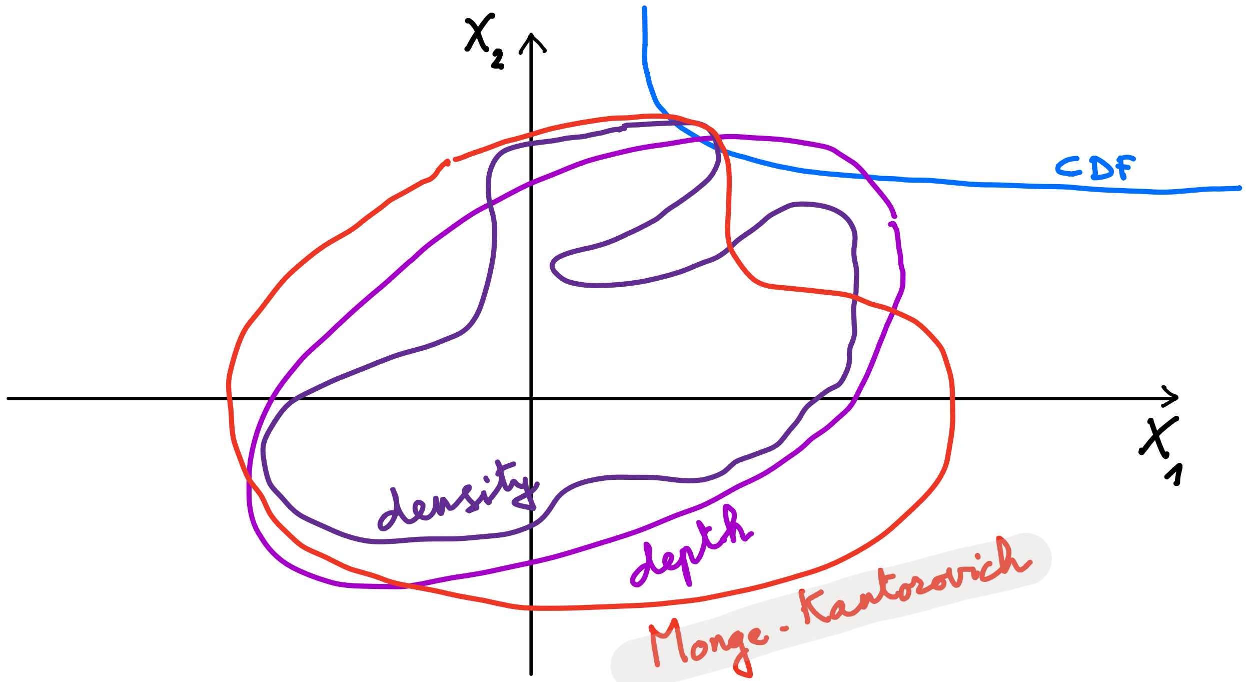
Cees de Valk
Royal Netherlands Meteorological Institute

Johan Segers
UCLouvain

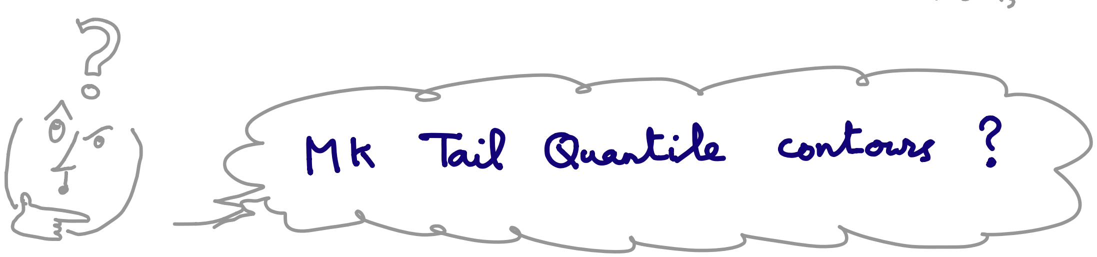
Annual Meeting of the RBSS

October 18–19, 2018

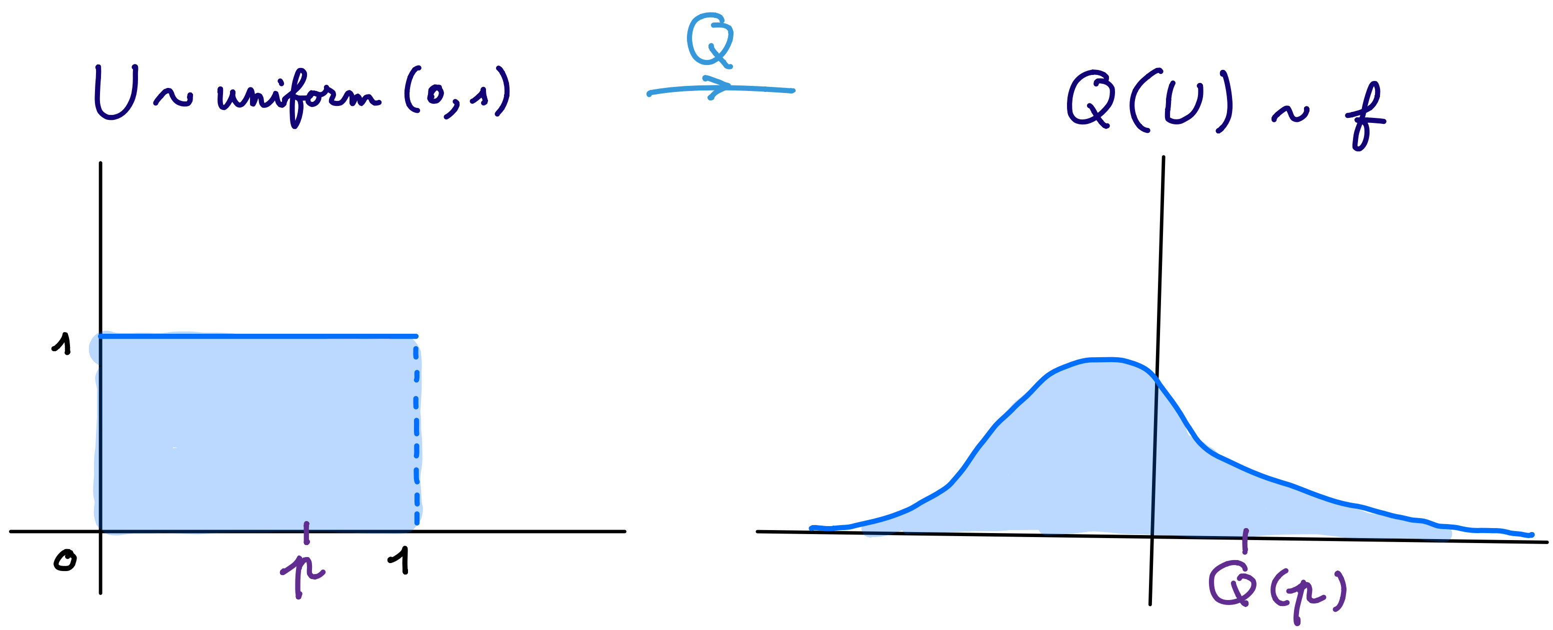
Quantile Contours



↳ Chernozhukov, Galichon, Hallin, Henry (AoS 2017)



QUANTILE TRANSFORM



Among all $T : \mathbb{R} \to \mathbb{R}$ such that $T(U) \sim f$:

- Q is the only monotone one
- Q minimizes $E[(U - T(U))^2]$

⟶ $d \geq 2$?

MONGE PROBLEM : OPTIMAL DETERMINISTIC COUPLING

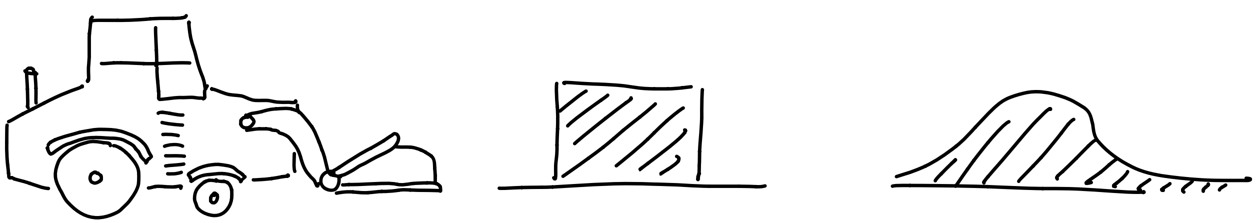
$$X \sim \mu \xrightarrow{T} T(X) = Y \sim \nu$$

Reference — target

Among all $T : \mathbb{R}^d \to \mathbb{R}^d$ such that $T(X) = Y$
find the one that minimizes $\mathbb{E}\left[|X - T(X)|^2\right]$

provided $\mathbb{E}\,|X|^2 < \infty$
$\mathbb{E}\,|Y|^2 < \infty$

G. Monge (1776) "... déblais ... remblais"



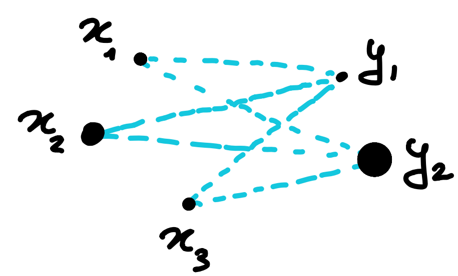
KANTOROVICH : OPTIMAL RANDOM COUPLING

$X \sim \mu$
$Y \sim \nu$ joint law $(X,Y) \sim \pi$

Among all distributions π on $\mathbb{R}^d \times \mathbb{R}^d$ with margins μ and ν, find the one that minimizes

$$\int |x-y|^2 \, d\pi(x,y) = E\left[|X-Y|^2\right]$$

- $Y = T(X)$: MONGE
- finite 2nd moments
- minimum : WASSERSTEIN distance
- discrete case: linear program



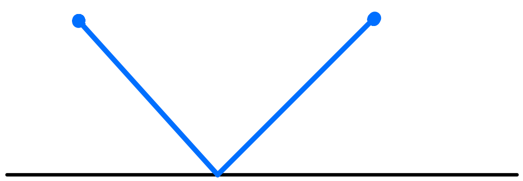
Subdifferential of a convex function

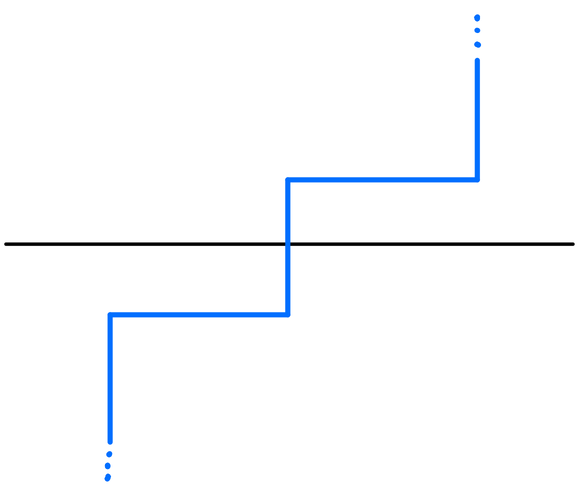
$\psi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$

$\partial\psi : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$

$\partial\psi(x) =$

$\{y : \forall z,\ \psi(z) \geq \psi(x) + \langle y, z-x \rangle\}$





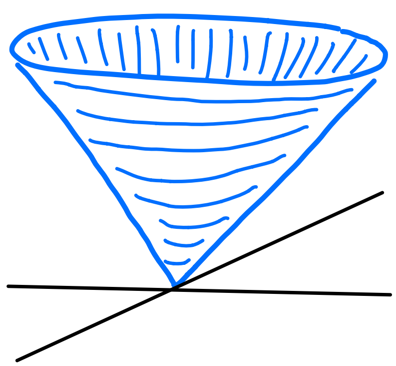
$d = 1$

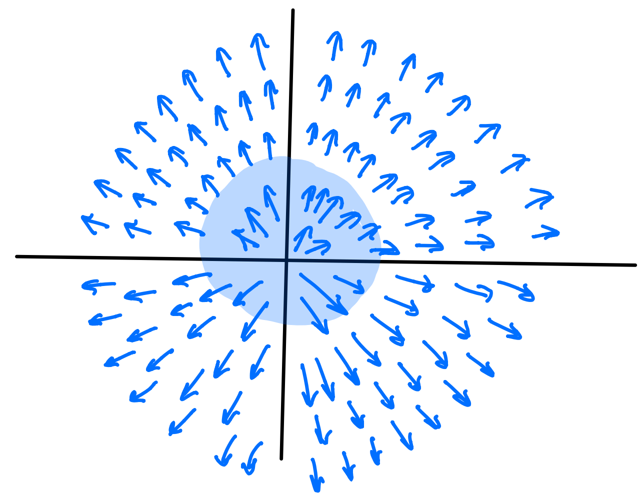
Subdifferential of a convex function

$$\psi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$$

$$\partial\psi : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$$

$$\partial\psi(x) = \{ y : \forall z,\ \psi(z) \geq \psi(x) + \langle y, z - x \rangle \}$$





$d = 2$

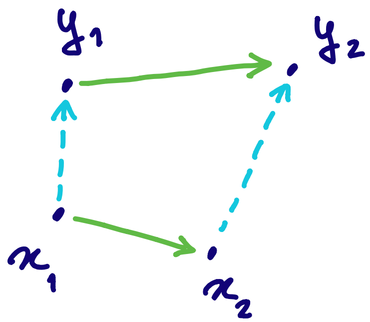
Monotone Mappings

$T : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is monotone if

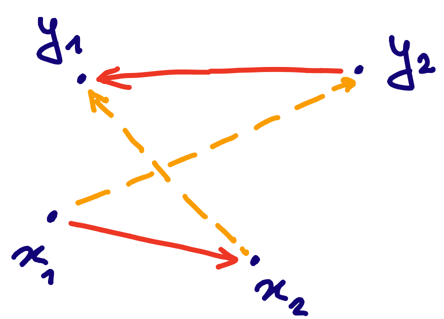
$\forall \; y_1 \in T(x_1), \quad y_2 \in T(x_2),$

$$\langle y_2 - y_1 , \; x_2 - x_1 \rangle \geq 0$$

monotone :



not monotone :



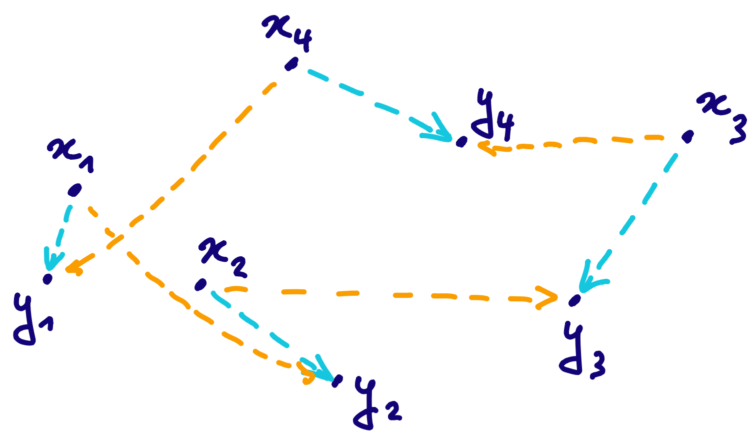
$$|x_1 - y_1|^2 + |x_2 - y_2|^2 \leq |x_1 - y_2|^2 + |x_2 - y_1|^2$$

CYCLICALLY MONOTONE MAPPINGS

$T : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is cyclically monotone if

$\forall\, y_1 \in T(x_1), \ldots, y_k \in T(x_k)$:

$$|x_1 - y_1|^2 + \ldots + |x_k - y_k|^2 \leq |x_1 - y_2|^2 + \ldots + |x_k - y_1|^2$$



- cyclically monotone $\Rightarrow$ monotone
- $d = 1$: ... ... $\Leftrightarrow$ ...

Rockafeller's Theorem

$T : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is maximal cyclically monotone

$$\Updownarrow$$

$$T = \partial \psi$$

for some l.s.c. convex $\psi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$

$\psi \not\equiv +\infty$

Rockafeller 1970
"Convex Analysis"

McCANN (1995)

μ, ν distributions on $\mathbb{R}^d$

- There exists a coupling π whose support is contained in the graph of $\partial\psi$ for some l.s.c. convex $\psi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$

- If μ vanishes on "small" sets, then π is the law of $(X, \nabla\psi(X))$, $X \sim \mu$

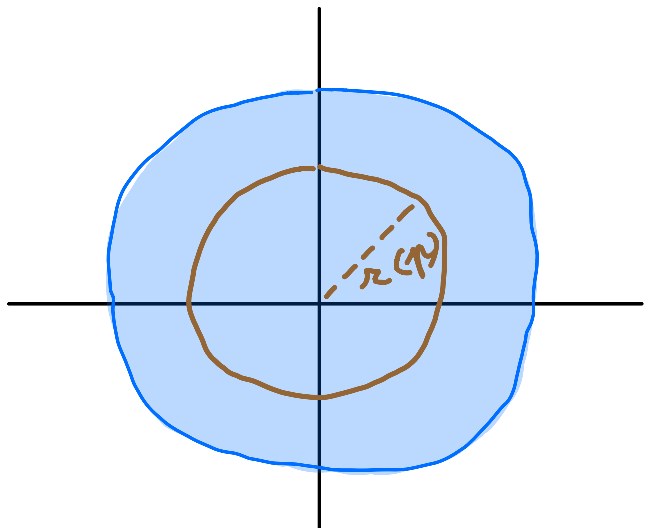
If finite 2nd moments, π and $\nabla\psi$ solve M-K problem

BRENIER (1987) : smooth case

M-K Quantile Contours

Reference

μ : (some) uniform distribution on $\{x \in \mathbb{R}^d : |x| \leq 1\}$

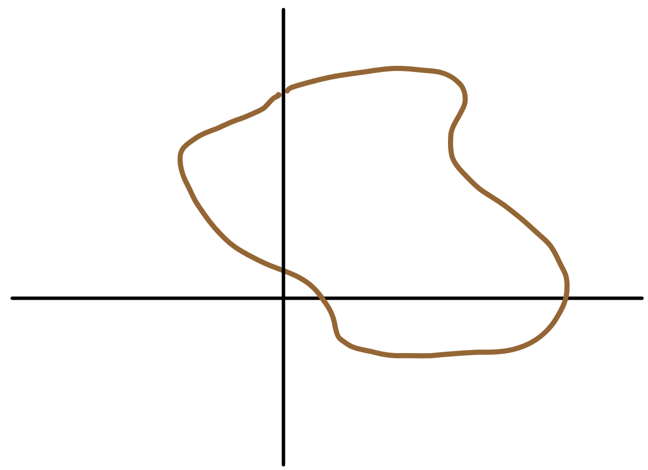


$$Q_\mu(p) = \{x : |x| = r(p)\}$$

Target

$$\nu = (\nabla\psi)_\# \mu$$

Brenier/McCann



$$Q_\nu(p) = \bigcup_{|x|=r(p)} \partial\psi(x)$$

Chernozhukov, Galichon, Hallin, Henry (2017)

G. MONGE

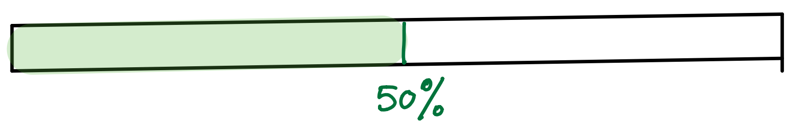
quantiles

KANTOROVICH

Cyclic monotonicity

So far so good...

Now let's go into the tails!



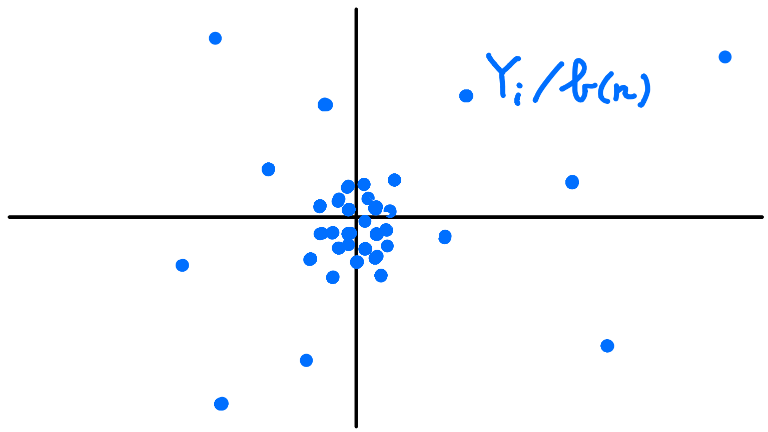
Zooming out on Large Observations

$$Y \sim \nu \quad \text{on} \quad \mathbb{R}^0$$

$$\nu_n(\cdot) = \mathbb{E}\left[\sum_{i=1}^{n} \mathbb{1}\left(\frac{1}{b(n)} Y_i \in \cdot \right) \right]$$

$$= n \, \mathbb{P}\left[\frac{1}{b(n)} Y \in \cdot \right]$$

scaling $b(n) \to \infty$



Intensity measure of point process, total mass $n \to \infty$

Regularly Varying Distributions

$Y \in \mathbb{R}^d$ with law ν is regularly varying if

$$\nu_n(\cdot) = n\,\mathbb{P}\left[\frac{1}{b(n)} Y \in \cdot\right] \xrightarrow{0} \bar{\nu}(\cdot)$$

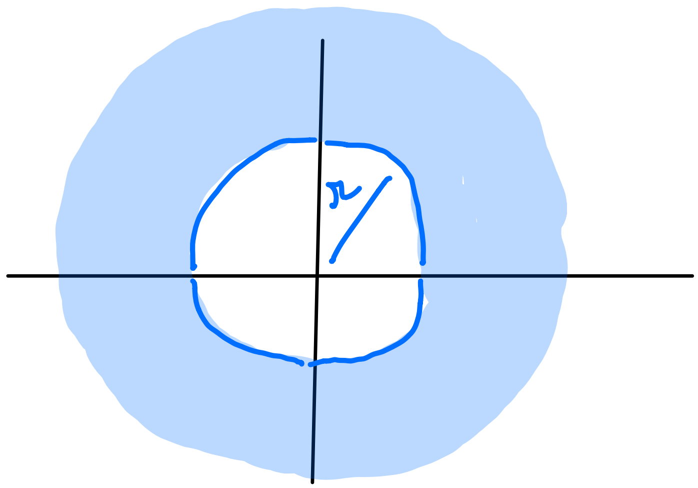
limit measure $\bar{\nu}$:

- infinite mass
- finite on sets bounded away from 0
- homogeneous : $\exists\, \alpha > 0, \quad \nu(\lambda\,\cdot) = \lambda^{-\alpha}\,\nu(\cdot)$

S.I. Resnick
(1987, 2007)

RADIAL DISTRIBUTION

$$\tilde{\nu}(\{y : |y| > r\}) = c \cdot r^{-\alpha}, \qquad r > 0$$



Power law behaviour

2nd moment ∞ if $0 < \alpha < 2$

Choice of Reference Distribution

μ = spherically symmetric version of ν

= law of $|Y|\,U$ with

$Y \sim \nu$

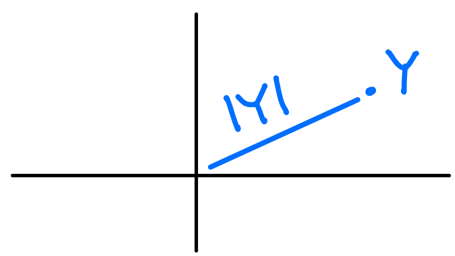
$U \sim$ uniform on $\{u : |u| = 1\}$

$Y \perp\!\!\!\perp U$

Radial distribution is preserved:

power law tail exponent $-\alpha$

If $X \sim \mu$ then $|X| \overset{d}{=} |Y|$



Coupling Reference & Target

Target $\quad Y \sim \nu$, regularly varying

Reference $\quad X \sim \mu$, $\quad |X| \overset{d}{=} |Y|$
spherically symmetric

McCann (1995) : $\exists$ coupling $(X, Y) \sim \pi$

$$\text{support}(\pi) \subset \text{graph}(\partial\psi) \subset \mathbb{R}^d \times \mathbb{R}^d$$

for some convex function ψ on $\mathbb{R}^d$

μ need not be smooth :
not guaranteed that $Y = \nabla\psi(X)$

Quantile Contours

$(X,Y) \sim \pi$

Reference

$X \sim \mu$

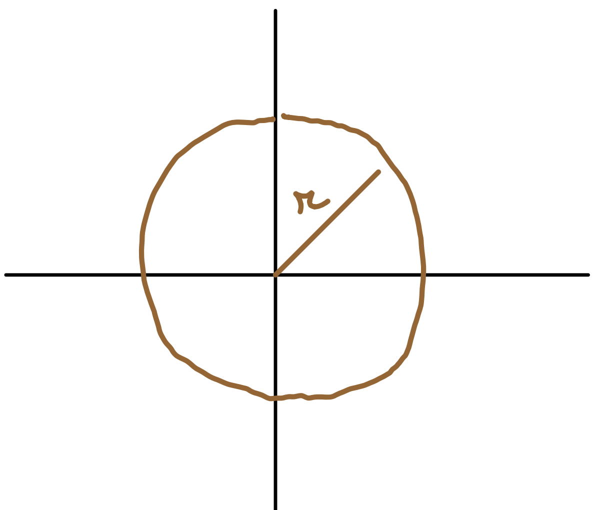
$|X| \stackrel{d}{=} |Y|$

X spher. sym.

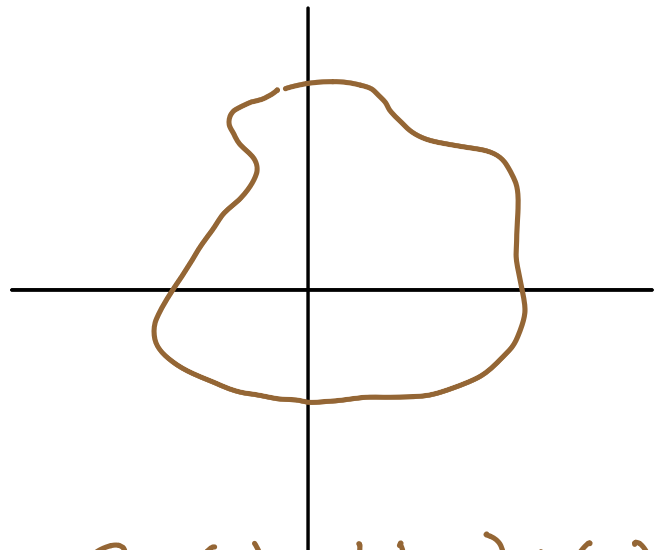
$\text{support}(\pi) \subset \{(x,y) : y \in \partial\psi(x)\}$

Target

$Y \sim \nu$



$Q_X(r) = \{x : |x| = r\}$

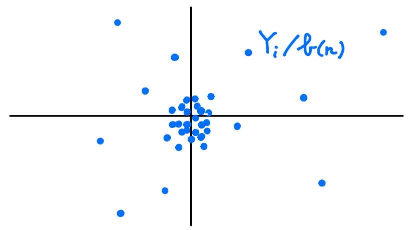


$Q_Y(r) = \bigcup_{|x|=r} \partial\psi(x)$

REFERENCE LIMIT INTENSITY MEASURE

Recall regular variation :

$$n \, \mathbb{P}\left[\frac{1}{b(n)} Y \in \cdot \right] \xrightarrow{\circ} \bar{\nu}(\cdot)$$



$\bar{\mu}$:= spherically symmetric such that

$$\bar{\mu}(\{x : |x| > r\}) = \bar{\nu}(\{y : |y| > r\})$$

$$= c \cdot r^{-\alpha}, \qquad r > 0$$

infinite measures $\Rightarrow$ McCANN theory does not apply !

TAIL STABILITY

Recall : support$(X,Y) \subset$ graph $(\partial\psi)$

$$Q_Y(r) = \bigcup_{|x|=r} \partial\psi(x) \qquad r \to \infty \ ?$$

$\exists!$ convex $\bar{\Psi}$ on $\mathbb{R}^d$ with $\bar{\Psi}(0) = 0$ and

- $\frac{1}{t}\,\partial\psi(t\,\cdot) \xrightarrow{g} \partial\bar{\Psi}(\cdot)$, $\quad t \to \infty$

  ↓ graphical convergence implies local uniform convergence at x where $\bar{\Psi}$ is cont. diff.

- $\bar{\nu}(\cdot) = \bar{\mu}(\{x : \nabla\bar{\Psi}(x) \in \cdot\})$

THE LIMIT IS HOMOGENEOUS

$\bar{\Psi}$ and $\partial \bar{\Psi}$ are homogeneous :

$$\bar{\Psi}(\lambda x) = \lambda^2 \, \bar{\Psi}(x)$$

$$\partial \bar{\Psi}(\lambda x) = \lambda \, \partial \bar{\Psi}(x)$$

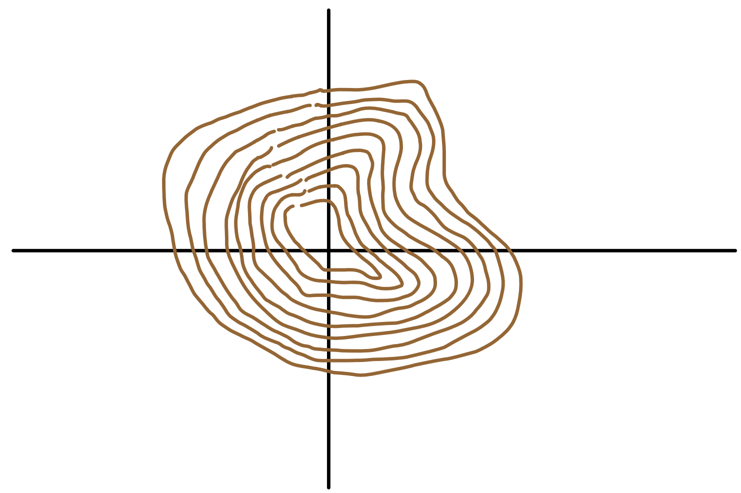
$\lambda > 0$

$x \in \mathbb{R}^d$

Limit contours are homothetic :

$$Q_{\bar{\Psi}}(r) = r \, Q_{\bar{\Psi}}(1)$$

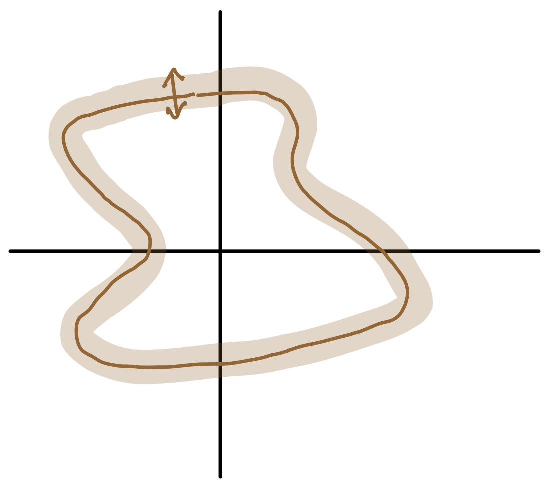
$$Q_{\bar{\Psi}}(1) = \bigcup_{|x|=1} \partial \bar{\Psi}(x)$$

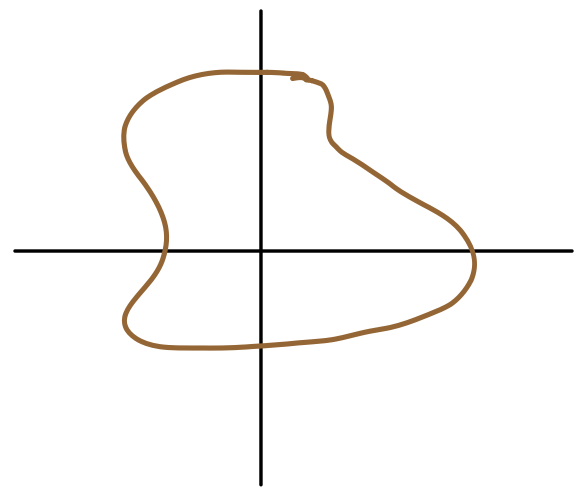


Convergence of Quantile Contours

For $0 < a < b < +\infty$:

$$\bigcup_{a \le \lambda \le b} \frac{1}{r\lambda} Q_{\gamma}(r\lambda) \xrightarrow[r \to \infty]{\text{Hausdorff distance}} Q_{\overline{\psi}}(1)$$





Nonparametric tail quantile contour estimator

Input: points $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ in $\mathbb{R}^d$

Wanted: estimator of $Q_{\bar{\psi}}(1) = \lim_{r\to\infty} \bigcup_{|\boldsymbol{x}|=r} r^{-1}\partial\psi(\boldsymbol{x})$

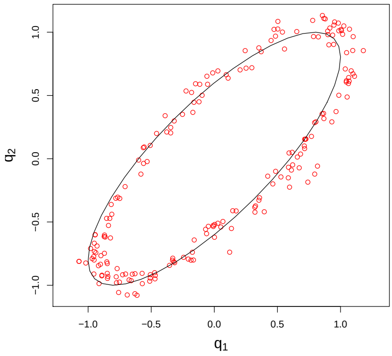
Algorithm:

Step 1 Find estimator $\partial\hat{\psi}_n$ of optimal plan $\partial\psi$ from μ to ν, with μ spherically symmetric version of ν

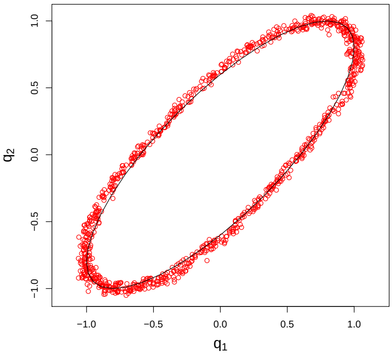
Step 2 Choose large r_n and put

$$\hat{Q}_n \quad := \quad \left\{ \frac{1}{|\boldsymbol{x}|} \partial\hat{\psi}_n(\boldsymbol{x}) : |\boldsymbol{x}| > r_n \right\}$$

Multivariate Student t distribution

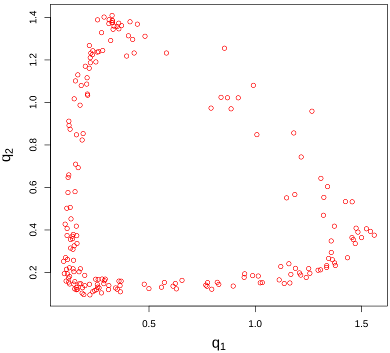


$n = 2000$
$k = 200$

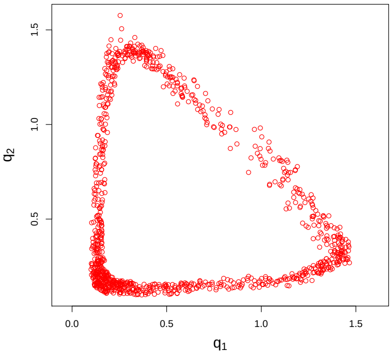


$n = 10000$
$k = 1000$

Fréchet margins, independence copula

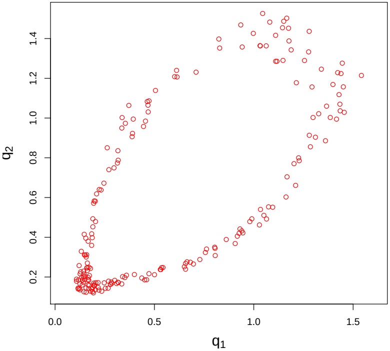


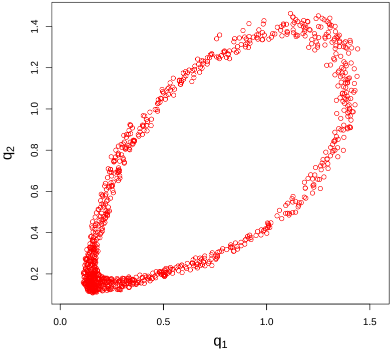
$n = 2000$
$k = 200$



$n = 10000$
$k = 1000$

Fréchet margins, Gumbel copula

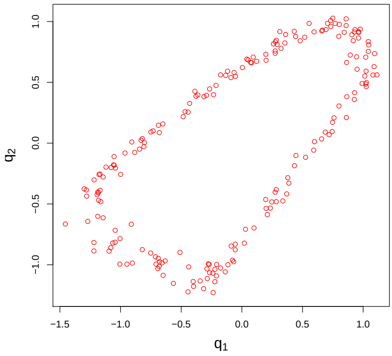




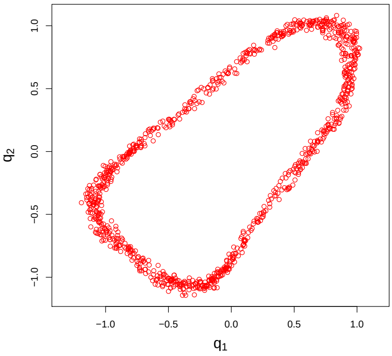
$n = 2000$
$k = 200$

$n = 10000$
$k = 1000$

Student t margins, Gumbel copula



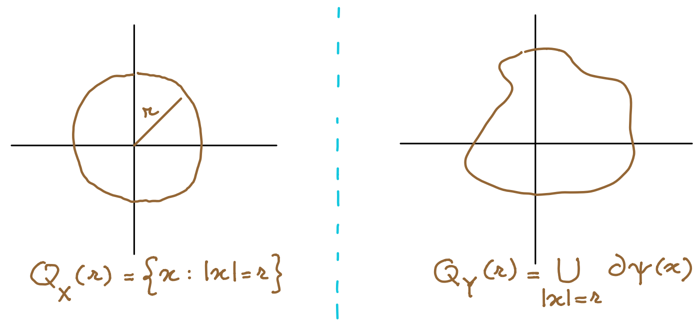
$n = 2000$
$k = 200$



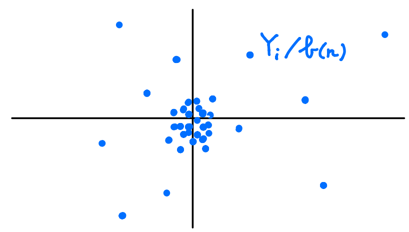
$n = 10000$
$k = 1000$

Tail Quantile Contours

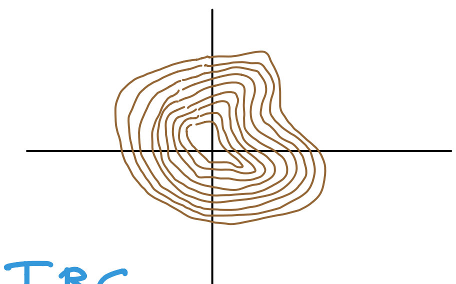
Quantile contours:
transporting spheres
via cyclically
monotone mappings



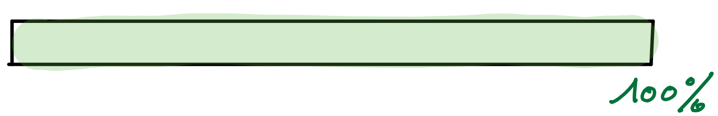
Regular variation:
zooming out, limiting
intensity measure



tail quantile contours:
homothetic to a single shape



Nonparametric estimation feasible: TBC...



Bibliography I

Bingham, N., C. Goldie, and J. Teugels (1987). *Regular variation*. Cambridge: Cambridge University Press.

Brenier, Y. (1987). Décomposition polaire et rérrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math. 305*, 805–808.

Chernozhukov, V., A. Galichon, M. Hallin, and M. Henry (2017). Monge–Kantorovich depth, quantiles, ranks and signs. *The Annals of Statistics 45*(1), 223–256.

Cuesta-Albertos, J. and M. C. (1989). Notes on the Wasserstein metric in Hilbert spaces. *The Annals of Probability 17*, 1264–1276.

Cuesta-Albertos, J., C. Matrán, and A. Tuero-Diaz (1997). Optimal transportation plans and convergence in distribution. *Journal of Multivariate Analysis 60*, 72–83.

Fisher, R. and L. Tippett (1928). On the estimation of the frequency distributions of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society 24*, 180–190.

Gnedenko, B. (1943). *Annals of Mathematics 44*, 423–453.

Hult, H. and F. Lindskog (2006). Regular variation for measures on metric spaces. *Publ. Inst. Math. (Beograd) (N.S.) 80(94)*, 121–140.

Bibliography II

Karamata, J. (1930). Sur un mode de croissance régulière des fonctions. *Mathematica (Cluj) 4*, 38–53.

McCann, R. J. (1995). Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J. 80*(2), 309–323.

Resnick, S. I. (1987). *Extreme values, regular variation, and point processes*. New York: Springer-Verlag.

Resnick, S. I. (2007). *Heavy-tail phenomena. Probabilistic and statistical modeling*. New York: Springer.

Rockafellar, R. and R. Wets (1998). *Variational Analysis*. New York: Springer.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton: Princeton University Press.

Rüschendorf, L. (1991). Fréchet bounds and their applications. In G. Dall'Aglio, S. Kotz, and G. Salinette (Eds.), *Advances in Probability Distributions with Given Marginals (Rome 1990)*, pp. 151–187. Dordrecht: Kluwer Academic Publishers.

Schuhmacher, D., B. Bähre, C. Gottschlich, F. Heinemann, and B. Schmitzer (2017). *transport: Optimal Transport in Various Forms*. R package version 0.9-4.

Villani, C. (2009). *Optimal Transport. Old and New*. Berlin: Springer-Verlag.