An Invitation to Statistics in Wasserstein Space

Wang Chao 2020212429

Shanghai University of Finance and Economics

May 15, 2022

Outline

Content

Analogy with Procrustes Analysis

γ_1

$\{\mu_1, \ldots, \mu_N, \gamma_1\}$ compatible

γ_n

Recall the Algorithm 1

$\mu_1, \ldots, \mu_N$ $\quad [\frac{1}{N}\sum_{i=1}^{N} t_{\gamma_1}^{\mu^i}]\# \gamma_1$ $\quad J_2(\underline{\gamma})$

Algorithm 1 Steepest descent via Procrustes analysis

γ_1 $\quad 0 \le \tau \le 1$

(A) Set a tolerance threshold $\varepsilon > 0$.

(B) For $j = 0$, let γ_j be an arbitrary absolutely continuous measure.

(C) For $i = 1, \ldots, N$ solve the (pairwise) Monge problem and find the optimal transport map $\mathbf{t}_{\gamma_j}^{\mu^i}$ from γ_j to μ^i.

$\tau = 1$ $\quad F'(\gamma_j)$

(D) Define the map $T_j = N^{-1} \sum_{i=1}^{N} \mathbf{t}_{\gamma_j}^{\mu^i}$.

(E) Set $\gamma_{j+1} = T_j \# \gamma_j$, i.e. push-forward γ_j via T_j to obtain γ_{j+1}.

(F) If $\|F'(\gamma_{j+1})\| < \varepsilon$, stop, and output γ_{j+1} as the approximation of $\bar{\mu}$ and $\mathbf{t}_{\gamma_{j+1}}^{\mu^i}$ as the approximation of $\mathbf{t}_{\bar{\mu}}^{\mu^i}$, $i = 1, \ldots, N$. Otherwise, return to step (C).

karcher mean

$\gamma_1, \gamma_2 \cdots \gamma_n$ $\quad F'(\gamma) = 0$

$\{\gamma_j\}$ $\quad W_2(\gamma_j, \bar{\mu}) \to 0$, $\bar{\mu}$, Frechet mean

Analogy with Procrustes Analysis

Algorithm 1 iterates the two steps of registration and linear averaging given the current template γ_j. $\underline{\gamma_j}$

1. Registration: by finding the optimal transportation maps $t_{\gamma_j}^{\mu^i}$,we identify each μ^i with the element $t_{\gamma_j}^{\mu^i} - i \in \log_{\gamma_j}(\mu^i)$
2. Averaging:the registered measures are averaged linearly, using the common coordinate system of the registration step (1), as elements in the linear space γ_j. $\longrightarrow$ $\underline{\mathrm{Tan}_{\gamma_j}}$

Content

Convergence of Algorithm 1

$(\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow W_2(\mathbb{R}^d)$.

$P($ $\Lambda(\omega)$ is absolutely continuous with bounded density $) = q$

Theorem 5.3.1 (Limit Points are Karcher Means) $q_n \to q$

$\Lambda_1 \quad \cdots \quad \Lambda_n$

Let $\mu^1, \ldots, \mu^N \in \mathcal{W}_2(\mathbb{R}^d)$ be probability measures and suppose that one of them is absolutely continuous with a bounded density. Then, the sequence generated by Algorithm 1 stays in a compact set of the Wasserstein space $\mathcal{W}_2(\mathbb{R}^d)$, and any limit point of the sequence is a Karcher mean of $(\mu^1, \ldots, \mu^N)$.

$\{\gamma_j\}$

Convergence of Algorithm 1

Proposition 3.1.8 states that the Frechet mean $\bar{\mu}$ is a Karcher mean ,we obtain immediately

Corollary 5.3.2 (Wasserstein Convergence of Steepest Descent)

Under the conditions of Theorem 5.3.1, if F has a unique stationary point, then the sequence $\{\gamma_j\}$ generated by Algorithm 1 converges to the Frechet mean $\bar{\mu}$ of $\mu_1, \ldots, \mu_N$ in the Wasserstein metric,

$$W_2(\gamma_j, \bar{\mu}) \to 0$$

Convergence of Algorithm 1

$\mathcal{A}(\mathcal{W}_2(\mathbb{R}^d))$ $= \{\mathcal{A}(\mu) : \mu$ is absolutely continuous & has compact closure$\}$

$\mathcal{A}$ $\quad j \to j+1 \quad \mathcal{A}(\gamma_j) = \gamma_{j+1}$ $\quad$ tight

The proof of Theorem 5.3.1 is achieved by establishing the following facts:

1. The sequence $\{\gamma_j\}$ stays in a compact subset of $\mathcal{W}_2(\mathbb{R}^d)$ (Lemma 5.3.5)
2. Any limit of $\{\gamma_j\}$ is absolutely continuous (Proposition 5.3.6)
3. Algorithm 1 acts continuously on its argument (Corollary 5.3.8)

$\gamma_j \to \gamma \qquad \mathcal{A}(\gamma_j) \to \mathcal{A}(\gamma)$

$\gamma_{j+1} \to \gamma$

Lemma 5.3.5

The sequence generated by Algorithm 1 stays in a compact subset of the Wasserstein space $\mathcal{W}_2(\mathbb{R}^d)$

$\frac{1}{N}\sum_{i=1}^{N} t_\gamma^{\mu^i} = i. \quad \gamma$-a.s. $\Longleftarrow$ $\boxed{\mathcal{A}(\gamma) = [\frac{1}{N}\sum_{i=1}^{N} t_\gamma^{\mu^i}]_\# \gamma = \gamma = i_\# \gamma}$

$\Rightarrow \frac{1}{N}\sum_{i=1}^{N} t_\gamma^{\mu^i} - i = 0$

$F'(\gamma) = 0 \Rightarrow \gamma$ is Karcher mean

Convergence of Algorithm 1

Let $\mathcal{A}$ denote the steepest descent iteration,that is, $\mathcal{A}(\gamma_j) = \gamma_{j+1}$.
If a weekly convergent sequence $\{\gamma_j\}$ of absolutely continuous measures has densities with uniform bound C,then the limit γ is absolutely continuous measure.Indeed,for any open set $O \subset \mathbb{R}^d$,$\liminf \gamma_k(O) \leq CLeb(O)$, so $\gamma(O) \leq CLeb(O)$ by the portmanteau Lemma.It follows that γ is absolutely continuous with density bounded by C. $\forall O \subset \mathbb{R}^d$, $\gamma_k(O) \leq C\, Leb(O)$, $\forall k$

$\liminf \gamma_k(O) \leq C\, Leb(O)$

$\gamma_j \rightharpoonup \gamma$ implies $\Downarrow$ $\quad \gamma(O) \leq \liminf_{k\to\infty} \gamma_k(O) \leq C\, Leb(O)$

$\Rightarrow \gamma$ has density function bounded C.

Proposition 5.3.6 (Uniform Density Bound)

For each $i = 1, ..., N$ denote by g^i the density of μ^i (if it exists) and $||g^i||_\infty$ its supremum,taken to be infinite if g^i does not exist (or if g^i is unbounded).Let γ_0 be any absolutely continuous probability measure.Then the density of $\gamma_1 = \mathcal{A}(\gamma_0)$ is bounded by the $1/d$-th harmonic mean of $||g^i||_\infty$,

$$C_\mu = [\frac{1}{N}\sum_{i=1}^{N}\frac{1}{||g_i||_\infty^{1/d}}]^{-d}$$

$g^1 \cdots g^N$

$< \infty$

The constant C_μ depends only on the measures $(\mu^1, ..., \mu^N)$,and is finite as long as one μ^i has a bounded density,since $C_\mu \le N^d ||g^i||_\infty$ for any i.

Proof of Proposition 5.3.6

let h_i denote the density function of φ_i

$$\varphi_1 = \Big[\frac{1}{N}\sum_{i=1}^{N} t_{\gamma_0}^{\mu_i}\Big]_{\#}\gamma_0 = (t_{\gamma_0}^{\varphi_1})_{\#}\gamma_0$$

$$\int f(t_{\gamma_0}^{\varphi_1}(x))\,d\gamma_0 = \int f(x)\,d\,t_{\gamma_0 \#}^{\varphi_1}\gamma_0 = \int f(x)\,d\varphi_1$$

$x \to t_{\gamma_0}^{\varphi_1}(x)$

$\gamma_0 \qquad \varphi_1$

$$h_1(t_{\gamma_0}^{\varphi_1}(x)) = \frac{h_0(x)}{\det \nabla t_{\gamma_0}^{\varphi_1}(x)}$$

$x \to t_{\gamma_0}^{\mu_i}(x)$

$\gamma_0 \qquad \mu_i$

$$g^i(t_{\gamma_0}^{\mu_i}(x)) = \frac{h_0(x)}{\det \nabla t_{\gamma_0}^{\mu_i}(x)}$$

$$\Rightarrow \det \nabla t_{\gamma_0}^{\varphi_1}(x) = \det \nabla\Big(\frac{1}{N}\sum_{i=1}^{N} t_{\gamma_0}^{\mu_i}(x)\Big) = \frac{1}{N^d}\det \sum_{i=1}^{N} \nabla t_{\gamma_0}^{\mu_i}(x)$$

invoke a Lemma

Brunn-Minkowski inequality

$$[\det(A+B)]^{1/d} \geq [\det A]^{1/d} + (\det B)^{1/d}$$

$$[\det \nabla t_{\gamma_0}^{\varphi_1}(x)]^{1/d} = \frac{1}{N}\Big[\det \sum_{i=1}^{N} \nabla t_{\gamma_0}^{\mu_i}(x)\Big]^{1/d} \geq \frac{1}{N}\Big[\sum_{i=1}^{N} \det \nabla t_{\gamma_0}^{\mu_i}(x)\Big]^{1/d}$$

$$\frac{1}{h_i^{1/d}(t_{\varphi_0}^{\varphi_1}(x))} = \frac{[\det \sum_{i=1}^{N} \nabla t_{\varphi_0}^{\mu^i}(x)]^{1/d}}{N\, h_0^{1/d}(x)} \geqslant \frac{\frac{1}{N}[\sum_{i=1}^{N} \det \nabla t_{\varphi_0}^{g^i}(x)]^{1/d}}{h_0^{1/d}(x)}$$

$$= \frac{1}{N}\sum_{i=1}^{N} \frac{1}{[g^i(t_{\gamma_0}^{\mu^i}(x))]^{1/d}} \geqslant \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\|g^i\|_\infty^{1/d}} = C_\mu^{-1/d}$$

Proposition 5.3.7

Let $\{\gamma_n\}$ be a sequence of absolutely continuous measures with uniformly bounded densities,suppose that $W_2(\gamma_n,\gamma) \to 0$,and let

$\varphi_n = \bar{X}_\# \eta_j$ $\quad J_j = (t_{\gamma_j}^{\mu^1}, \cdots, t_{\gamma_j}^{\mu^n})_\# \varphi_j \Rightarrow J_j \to J$

$$\eta_j = (t_{\gamma_j}^{\mu^1}, \ldots, t_{\gamma_j}^{\mu^n}, i)_{\#\gamma_j}, \quad \eta = (t_\gamma^{\mu^1}, \ldots, t_\gamma^{\mu^n}, i)_{\#\gamma}$$

$\gamma = \bar{x}_\# J$ $\quad \rightsquigarrow \mathbb{R}^d \to (\mathbb{R}^d)^{N+1}$ $\quad \Rightarrow A(\varphi_j) \to A(\varphi)$

Then $\eta_j \to \eta$ in $\mathcal{W}((\mathcal{R}^d)^{N+1})$.

φ_0, a.s. $\quad \Sigma \subset \mathbb{R}^d$, $\quad \varphi_0(\Sigma) = 1$.

$$\varphi_1(t_{\varphi_0}^{\varphi_1}(\Sigma)) = \varphi_0[(t_{\varphi_0}^{\varphi_1})^{-1}(t_{\varphi_0}^{\varphi_1}(\Sigma))] \geqslant \varphi_0(\Sigma) = 1$$

#

The sketch of proof

Theorem 2.21: $\forall\ h: (\mathbb{R}^d)^{N+1} \to \mathbb{R}$

$$|h(t_1, \dots, t_N, y)| \le \frac{2}{N}\sum_{i=1}^{N} \|t_i\|^2 + 2\|y\|^2$$

a.d. $\int h\, d\gamma_n \to \int h\, d\gamma \Rightarrow \gamma_n \to \gamma$ in $W_2(\mathbb{R}^{d\times(N+1)})$

$g_n(x): \mathbb{R}^d \to \mathbb{R}$ $\quad g_n(x) = h(t_{\varphi_n}^{\mu^1}(x), \dots, t_{\varphi_n}^{\mu^N}(x), x)$

$g(x) = h(t_{\varphi}^{\mu^1}(x), \dots, t_{\varphi}^{\mu^N}(x), x)$

$$\int_{\mathbb{R}^d} g_n(x)\, d\varphi_n = \int_{\mathbb{R}^d} h(t_{\varphi_n}^{\mu^1}(x), \dots, x)\, d\varphi_n$$

$$= \int_{\mathbb{R}^d} h\, d\underbrace{(t_{\varphi_n}^{\mu^1}(x), \dots, t_{\varphi_n}^{\mu^N}(x), x)_{\#}\varphi_n}_{\gamma_n}$$

$$= \int_{\mathbb{R}^d} h\, d\gamma_n$$

$\int_{\mathbb{R}^d} g\, d\varphi = \int_{\mathbb{R}^d} h\, d\gamma \Rightarrow \boxed{\int_{\mathbb{R}^d} g_n\, d\varphi_n \to \int_{\mathbb{R}^d} g\, d\varphi.}$

Convergence of Algorithm 1

① $g_{n,R} = \min\{g_n, 4R\}$ approximate g_n

$\gamma_n \to \gamma$ $\quad \sup_n \int [g_n(x) - g_{n,R}(x)]\, d\gamma_n(x) \to 0$

② $\int g_{n,R}\, d\gamma_n - \int g_R\, d\gamma = \int g_R\, d(\gamma_n - \gamma) + \int_\Omega (g_{n,R} - g_R)\, d\gamma_n$

$+ \int_{R^d \setminus \Omega} (g_{n,R} - g_R)\, d\gamma_n$

Corollary 5.3.8 (Continuity of $\mathcal{A}$)

If $W_2(\gamma_n, \gamma) \to 0$ and γ_n have uniformly bounded densities, then $\mathcal{A}(\gamma_n) \to \gamma$.

$\gamma_n = [\frac{1}{N}\sum_{i=1}^{N} \mu^i_{\gamma_{n-1}}]\gamma_{n-1}$

$\Omega_1 \subseteq \operatorname{supp} \gamma$ $\quad g_n \to g$ uniformly on Ω

③ portmanteau lemma $\Rightarrow \int g_R\, d(\gamma_n - \gamma) \to 0$

$\int_\Omega (g_{n,R} - g_R)\, d\gamma_n \to 0$ since uniformly converge

$\gamma[R^d / \Omega] < \varepsilon$

Content

The optimal maps between centred nondegenerate Gaussian measures with covariances A and B

$A \to B$

$$t_A^B(x) = \left[A^{-1/2}[A^{1/2}BA^{1/2}]^{1/2}A^{-1/2}\right]x, \quad x \in \mathbb{R}^d$$

$T x$

The Frechet mean of a collection of Gaussian measures does not admit a closed-form formula and is only known to be a Gaussian measure whose covariance matrix Γ is the unique invertible root of the matrix equation

$$\boxed{\Gamma = \frac{1}{N}\sum_{i=1}^{N}[\Gamma^{1/2}S_i\Gamma^{1/2}]^{1/2}}$$

where S_i is the covariance of μ^i.

Theorem 5.4.1 (Convergence in Gaussian Case)

Let $\mu^1, ..., \mu^N$ be Gaussian measures with zero means and covanriance matrices S_i with S_1 nonsingular,and let the initial point γ_0 be $N(0, \Gamma_0)$ with Γ_0) nonsingular.Then the sequences of iterates generated by Algorithm 1 converges to the unique Frechet mean of $(\mu^1, ..., \mu^N)$

Proof of Theorem 5.4.1

$\gamma_0 \sim N(0, \Gamma_0)$, $\quad \gamma_1 = [\frac{1}{N}\sum_{i=1}^{N} t_{\gamma_0}^{\mu^i}(x)]\# \gamma_0$

$= [\frac{1}{N}\sum_{i=1}^{N} T_1(x)]\#\gamma_0$

$\gamma_K \sim N(0, \Gamma_K)$, where Γ_K is nonsingular,

$\gamma_K \Rightarrow \gamma$ It suffices to show γ is Gaussian measure

$\Gamma_K \to \Gamma$, $\Rightarrow \phi_K(t) \to \phi(t)$, $= e^{-\frac{1}{2}t^T\Gamma t}$

$e^{-\frac{1}{2}t^T\Gamma_K t}$

→ Levy's continuity Theorem $\Rightarrow N(0, \Gamma_K) \to N(0, \Gamma)$ weakly

$\mu \sim N(0,S)$, $W_2^2(\mu, \delta_0) = tr(S)$

$0 \le tr(\Gamma_K) = W_2^2(\gamma_K, \delta_0)$ $\quad C = \sup_K tr(\Gamma_K) < \infty$

$2|A_{ij}| \le A_{ii} + A_{jj} \Rightarrow \{\Gamma_K\}$ is bounded sequence. #

Figure 5.1 shows density plots of N = 4 centred Gaussian measures on $\mathbb{R}^2$ with covariances $\boxed{S_i}$ ~ Wishart(I_2, 2)

Wishart distribution:

$X_1 \cdots X_n \sim N_p(0, V)$

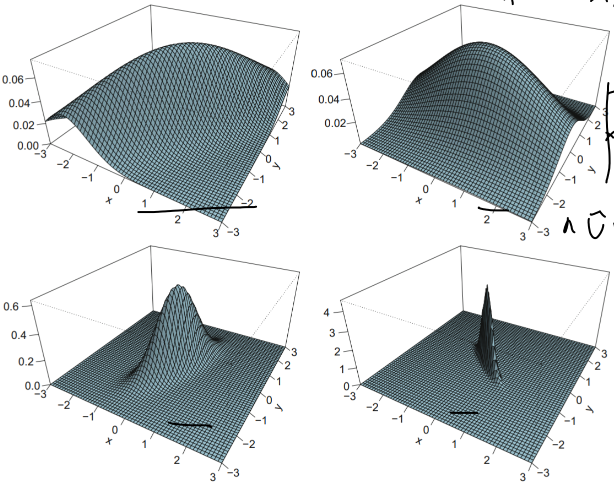
$\hat{V} = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T$

$p=1$,

$x_1 \cdots x_n \sim N(0, \sigma^2)$,

$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$

$n\hat{V} \sim \text{Wishart}(V, n)$



Gaussian Measures

Figure 5.2 shows the density of the resulting Frechet mean. The corresponding optimal maps are displayed in Figure 5.3.

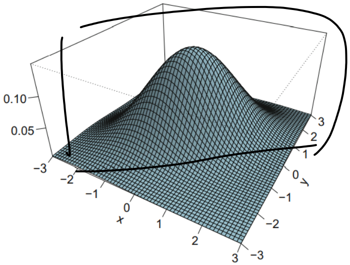
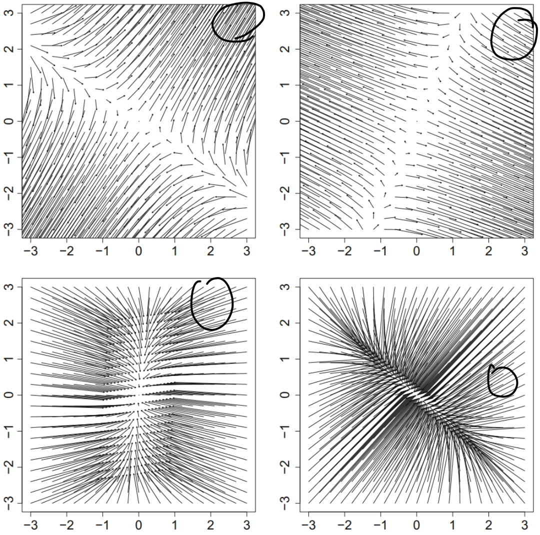


Fig. 5.2: Density plot of the Fréchet mean of the measures in Fig. 5.1

Gaussian Measures

The corresponding optimal maps are displayed in Figure 5.3.



Compatible Measures

Recall that a collection $\mathcal{C} \subset \mathcal{W}_2(\mathcal{X})$ is compatible if for all $\gamma, \rho, \mu \in \mathcal{C}, t_\mu^\rho \circ t_\gamma^\mu = t_\gamma^\rho$ in $L_2(\gamma)$.

Lemma 5.4.2 (Compatibility and Convergence)

If $\gamma_0 \cup \{\mu^i\}$ is compatible, then Algorithm 1 converges to the Frechet mean of $\{\mu^i\}$ after a single step

Proof of Lemma 5.4.2

By definition,the next iterate is

$\mu_1 \cdots \mu_N$

$$\gamma_1 = [\frac{1}{N}\sum_{i=1}^{N} t_{\gamma_0}^{\mu^i}]_{\#\gamma_0}$$

$\gamma_1 = [\frac{1}{N}\sum_{i=1}^{N} t_{\mu_1}^{\mu^i}]_{\#}\mu_1$

$\boxed{N-1}$

$t_{\mu_1}^{\mu_1} = i$

which is the Frechet mean by Theorem 3.1.9.

$\Rightarrow t_\mu^\nu \circ t_\gamma^\mu = F_\nu^{-1} \circ F_\mu \circ F_\mu^{-1} \circ F_\gamma$

μ, ν, γ

$t_\mu^\nu = F_\nu^{-1} \circ F_\mu$ $\quad = F_\nu^{-1} \circ F_\gamma$

$t_\gamma^\mu = F_\mu^{-1} \circ F_\gamma$ $\quad = t_\gamma^\nu$

The One-Dimensional Case

1.
$$f^i(x) = \frac{1}{2}\phi\left(\frac{x - m_1^i}{\sigma_1^i}\right) + \frac{1}{2}\phi\left(\frac{x - m_2^i}{\sigma_2^i}\right)$$

where ϕ is the standard normal density,and the parameters are generated independently as

γ_0 $\quad f^1, \cdots, f^4$

$$m_1^i \sim U[-13, -3], \quad m_1^i \sim U[3, 13], \quad \sigma_1^i, \sigma_2^i \sim Gamma(4, 4)$$

$x > m_3^i$

2.
$$f^i(x) = \frac{3}{5}\frac{\beta_i^3}{\Gamma(3)}(x - m_3^i)^2 e^{-\beta_i(x - m_3^i)} + \frac{2}{5}\phi(x - m_4^i)$$

where $\beta^i \sim Gamma(4, 1), \quad m_3^i \sim U[1, 4], \quad m_4^i \sim U[-4, -1].$

Figure 5.4 plots univariate densities and Frechet mean respectively.

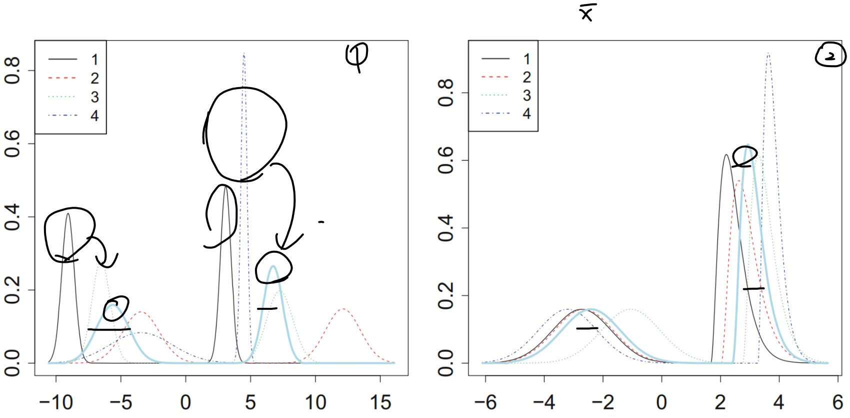


Fig. 5.4: Densities of a bimodal Gaussian mixture (left) and a mixture of a Gaussian with a gamma (right), with the Fréchet mean density in light blue

Compatible Measures

Figure 5.5 shows the optimal maps pushing the Frechet mean $\bar{\mu}$ to the measures $(\mu^1, \dots, \mu^N)$ in each case.

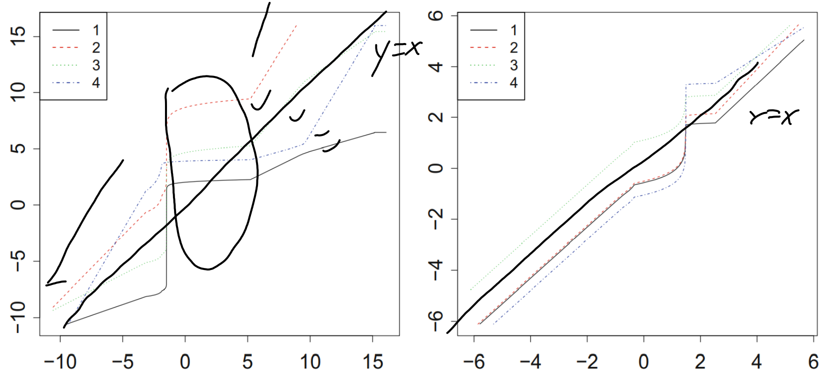


Fig. 5.5: Optimal maps $\mathbf{t}_{\bar{\mu}}^{\mu^i}$ from the Fréchet mean $\bar{\mu}$ to the four measures $\{\mu^i\}$ in Fig. 5.4. The left plot corresponds to the bimodal Gaussian mixture, and the right plot to the Gaussian/gamma mixture

Content

Population Version of Algorithm 1

Let $\Lambda \in \mathcal{W}_p(\mathbb{R}^d)$ be a random measure with finite Frechet functional.We assume that

$$q = \mathbb{P}(\Lambda \quad \text{absolutely continuous with density by} \quad M) > 0$$

$\{\Lambda_1 \cdots \Lambda_n\} \leftarrow q_n$

$q_n \to q$ almost sure

we can define a population version of Algorithm 1 with the iteration function

$$\mathcal{A}(\gamma) = [\mathbb{E}t_\gamma^\Lambda]_{\#\gamma}$$

where $\gamma \in \mathcal{W}_2(\mathbb{R}^d)$.

$$\mathcal{A}(\gamma) = \left[\frac{1}{N}\sum_{i=1}^{N} t_\gamma^{\mu^i}\right]_{\#\gamma}$$

γ is absolutely continuous

Population Version of Algorithm 1

Lemma 5.5.1

Any descent iterate γ has density bounded by $q^{-d}M$, where q and M are as in (5.7).

$\mathcal{L}(\varphi)$

fix $\omega \in \Omega$ $(\Omega, \mathcal{F}, \mathbb{P})$

Proof

$\Lambda_1(\omega), \dots \Lambda_n(\omega)$

Let $\Lambda_1, \dots$ be a sample from Λ and let q_n be the proportion of measures in $(\Lambda_1, \dots, \Lambda_n)$ that have density bounded by M. By the law of large numbers, we have

$$\frac{1}{n}\sum_{i=1}^{n} t_\gamma^{\Lambda_i} \to \mathbb{E}t_\gamma^{\Lambda}, \qquad q_n \to q$$

$\{t_n\}$ 且 $t_n \to t$

almost sure.

$\mu \in W_p(\mathcal{X})$

Recall the Lemma 2.4.5 states that $t_{\#\mu}$ is a continuous mapping from $\mathcal{L}_p(\mu)$ to $\mathcal{W}_p(\mathcal{X})$.

$\lim_{n\to\infty}(t_{n\#\mu}) = (\lim_{n\to\infty} t_n)_{\#\mu}$

Population Version of Algorithm 1

Continued

Therefore

Empirical case

$\mathcal{A}(\gamma)$

$$\mathcal{A}(\gamma) = [\mathbb{E}t_\gamma^\Lambda]_{\#\gamma} = [\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n t_\gamma^{\Lambda_i}]_{\#\gamma} = \lim_{n\to\infty}[\frac{1}{n}\sum_{i=1}^n t_\gamma^{\Lambda_i}]_{\#\gamma}$$

Let $\Lambda_n = [\frac{1}{n}\sum_{i=1}^n t_\gamma^{\Lambda_i}]_{\#\gamma}$, Proposition 5.3.6 (empirical case) gives that λ_n has density bounded by $q_n^{-d}M$ which goes to $q^{-d}M$ almost sure.
So for any $C > q^{-d}M$, we have λ_n has density bounded by C as long as n large enough. By the portmanteau Lemma, $\mathcal{A}(\gamma) = \lim_{n\to\infty}\lambda_n$ has density bounded by C. Let $C \searrow q^{-d}M$ gives the desired conclusion.

$\mathcal{A}(\gamma)$ has density function bounded by $q^{-d}M$

Population Version of Algorithm 1

$\mathcal{A}(\gamma) = [\mathbb{E} t_\gamma^\Lambda]_{\#\gamma} = i_{\#\gamma} = \gamma$

By the Proposition 3.2.14, γ is a Karcher mean of Λ iff $\mathbb{E}t_\gamma^\Lambda = i$. Therefore Lemma 5.5.1 ensures the Karcher mean has a bounded density since $\gamma = i_{\#_\gamma} = [\mathbb{E}t_\gamma^\Lambda]_{\#_\gamma}$ The similar conclusion can be done for Frechet mean as following Theorem.

Theorem 5.5.2 (Bounded Density for Population Frechet Mean)

Let $\Lambda \in \mathcal{W}_p(\mathbb{R}_d)$ be a random measure with finite Frechet functional. If Λ has a bounded density with positive probability, then the Frechet mean of Λ is absolutely continuous with a bounded density.

Population Version of Algorithm 1

Proof

Denote the empirical Frechet mean of the sample $(\Lambda_1, ..., \Lambda_n)$ by λ_n,which has density bounded by $q_n^{-d}M$. Fix $\omega \in \Omega$, almost sure
Proposition 3.2.7 ensures that there is a unique Frechet mean λ of Λ,and $\boxed{\lambda_n \to \lambda}$ in $\mathcal{W}_2(\mathbb{R}^d)$ almost sure by Corollary 3.2.10.
Let $C > q^{-d}M$.Since $q_n^{-d}M \to q^{-d}M$ almost sure,we have λ_n has density bounded by C when n is large enough.So by portmanteau Lemma,λ has density bounded by $q^{-d}M$.Let $C \searrow q^{-d}M$ gives the desired conclusion.

We have discussed the empirical version of this result in Proposition 3.1.8.

Population Version of Algorithm 1

Theorem 3.1.9 states that If $\{\gamma, \mu^1, ..., \mu^N\}$ are compatible measures,then $[\frac{1}{N}\sum_{i=1}^{N} t_\gamma^{\mu^i}]_{\#_\gamma}$ is the Frechet mean of $\{\mu^1, ..., \mu^N\}$.

The population version of Theorem 3.1.9 is

Theorem 5.5.3 (Frechet Mean of Compatible Measures)

Let $\Lambda : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathcal{W}_2(\mathbb{R})$ be a random measure with finite Frechet functional,and suppose that with positive probability Λ is absolutely continuous and has bounded density.If the collect $\{\gamma\} \cup \Lambda(\Omega)$ is compatible and γ is absolutely continuous,then $[\mathbb{E}t_\gamma^\Lambda]_{\#_\gamma}$ is the Frechet mean of Λ.