Convergence of the iterates of Lloyd's algorithm

Léo Portales joint work with my supervisors Elsa Cazelles and Edouard Pauwels

IRIT, TSE, CNRS and Université de Toulouse

24/07/24

Discrete approximation of a continuous measure

μ ν

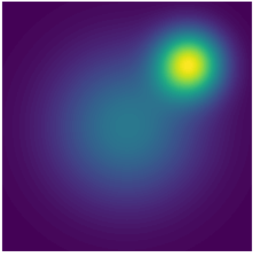
(Target Measure) $\frac{1}{N}\sum_{k=1}^{N}\delta_{y_k}$ $\sum_{k=1}^{N}\pi_k\delta_{y_k}$

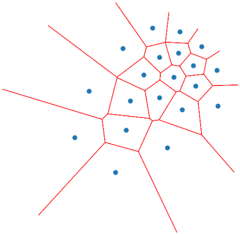
Simulations made using the library PyMongeAmpere by Quentin Mérigot.

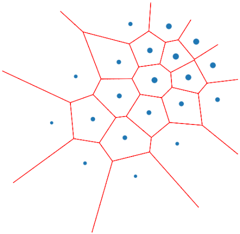
Discrete approximation of a continuous measure

μ ν

(Target Measure) $\frac{1}{N}\sum_{k=1}^{N}\delta_{y_k}$ $\sum_{k=1}^{N}\pi_k\delta_{y_k}$





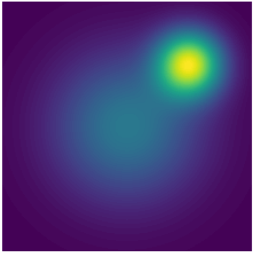


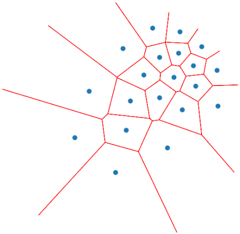
Simulations made using the library PyMongeAmpere by Quentin Mérigot.

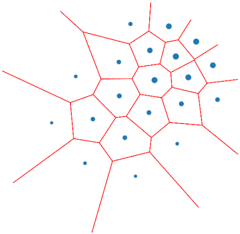
Discrete approximation of a continuous measure

μ ν

(Target Measure) $\frac{1}{N}\sum_{k=1}^{N}\delta_{y_k}$ $\sum_{k=1}^{N}\pi_k\delta_{y_k}$





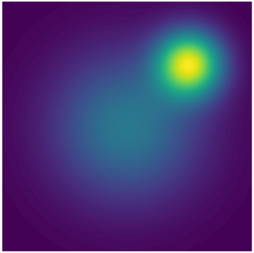


- μ is a compactly supported continuous probability measure on $\mathbb{R}^d$ with density f.

Simulations made using the library PyMongeAmpere by Quentin Mérigot.

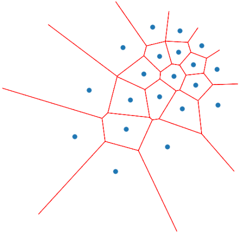
Discrete approximation of a continuous measure

μ (Target Measure)

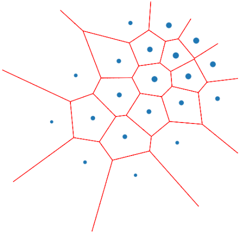


ν

$\frac{1}{N}\sum_{k=1}^{N}\delta_{y_k}$



$\sum_{k=1}^{N}\pi_k\delta_{y_k}$



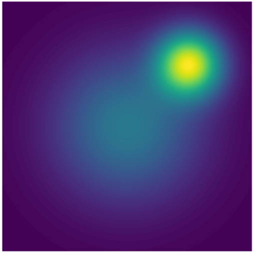
- μ is a compactly supported continuous probability measure on $\mathbb{R}^d$ with density f.
- $\nu(Y) := \sum_{k=1}^{N}\pi_k\delta_{y_k}$ where $Y = (y_1, \ldots, y_N) \in (\mathbb{R}^d)^N$ and $\pi \in \Delta_N$. N is fixed.

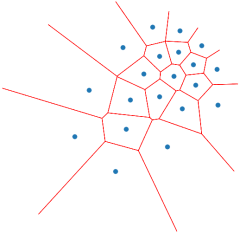
Simulations made using the library PyMongeAmpere by Quentin Mérigot.

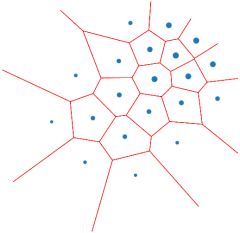
Discrete approximation of a continuous measure

μ ν

(Target Measure) $\frac{1}{N}\sum_{k=1}^{N}\delta_{y_k}$ $\sum_{k=1}^{N}\pi_k\delta_{y_k}$







- μ is a compactly supported continuous probability measure on $\mathbb{R}^d$ with density f.
- $\nu(Y) := \sum_{k=1}^{N}\pi_k\delta_{y_k}$ where $Y = (y_1, \dots, y_N) \in (\mathbb{R}^d)^N$ and $\pi \in \Delta_N$. N is fixed.

$$W_2(\mu, \nu(Y)) = \left(\inf_{\gamma\in\Pi(\mu,\nu(Y))}\int_{\mathbb{R}^d\times\mathbb{R}^d}\|x-y\|^2 d\gamma(x,y)\right)^{\frac{1}{2}}$$

Simulations made using the library PyMongeAmpere by Quentin Mérigot.

Uniform and optimal quantization

$$\boxed{\min_{Y:=(y_1,\dots,y_N)} \frac{1}{2} W_2^2 \left(\mu, \frac{1}{N} \sum_{i=1}^{N} \delta_{y_i} \right) = \min_{Y:=(y_1,\dots,y_N)} F_N(Y)}$$

Uniform and optimal quantization

$$\boxed{\min_{Y:=(y_1,\ldots,y_N)} \frac{1}{2} W_2^2\left(\mu, \frac{1}{N}\sum_{i=1}^{N}\delta_{y_i}\right) = \min_{Y:=(y_1,\ldots,y_N)} F_N(Y)}$$

where

$$F_N(Y) = \max_{w\in\mathbb{R}^N} \tfrac{1}{2}\left\{\int_{\mathbb{R}^d}\left(\min_{i=1,\ldots,N}\|x-y_i\|^2 - w_i\right)f(x)dx + \tfrac{1}{N}\sum_{i=1}^N w_i\right\}$$

Uniform and optimal quantization

$$\boxed{\min_{Y:=(y_1,\ldots,y_N)} \frac{1}{2} W_2^2\left(\mu, \frac{1}{N}\sum_{i=1}^{N} \delta_{y_i}\right) = \min_{Y:=(y_1,\ldots,y_N)} F_N(Y)}$$

where

$$F_N(Y) = \max_{w\in\mathbb{R}^N} \tfrac{1}{2}\left\{\int_{\mathbb{R}^d} \left(\min_{i=1,\ldots,N} \|x - y_i\|^2 - w_i\right) f(x)dx + \tfrac{1}{N}\sum_{i=1}^{N} w_i\right\}$$

(Uniform Quantization)

Uniform and optimal quantization

$$\boxed{\min_{Y:=(y_1,\dots,y_N)} \frac{1}{2} W_2^2\left(\mu, \frac{1}{N}\sum_{i=1}^N \delta_{y_i}\right) = \min_{Y:=(y_1,\dots,y_N)} F_N(Y)}$$

where

$$F_N(Y) = \max_{w\in\mathbb{R}^N} \tfrac{1}{2}\left\{\int_{\mathbb{R}^d} \left(\min_{i=1,\dots,N} \|x-y_i\|^2 - w_i\right) f(x)dx + \tfrac{1}{N}\sum_{i=1}^N w_i\right\}$$

(Uniform Quantization)

$$\boxed{\min_{Y,\pi>0,\, \sum_{i=1}^N \pi_i=1} \frac{1}{2} W_2^2(\mu, \sum_{i=1}^N \pi_i\delta_{y_i}) = \min_{Y:=(y_1,\dots,y_N)} G_N(Y)}$$

Uniform and optimal quantization

$$\boxed{\min_{Y:=(y_1,\ldots,y_N)} \frac{1}{2} W_2^2\left(\mu, \frac{1}{N}\sum_{i=1}^{N}\delta_{y_i}\right) = \min_{Y:=(y_1,\ldots,y_N)} F_N(Y)}$$

where
$F_N(Y) = \max_{w\in\mathbb{R}^N} \frac{1}{2}\left\{\int_{\mathbb{R}^d}\left(\min_{i=1,\ldots,N}\|x-y_i\|^2 - w_i\right) f(x)dx + \frac{1}{N}\sum_{i=1}^{N} w_i\right\}$

(Uniform Quantization)

$$\boxed{\min_{Y,\pi>0,\,\sum_{i=1}^N \pi_i=1} \frac{1}{2} W_2^2(\mu, \sum_{i=1}^{N}\pi_i\delta_{y_i}) = \min_{Y:=(y_1,\ldots,y_N)} G_N(Y)}$$

where $G_N(Y) = \frac{1}{2}\int_{\mathbb{R}^d} \min_{i=1,\ldots,N}\|x-y_i\|^2 f(x)dx$

Uniform and optimal quantization

$$\boxed{\min_{Y:=(y_1,\dots,y_N)} \frac{1}{2} W_2^2\left(\mu, \frac{1}{N}\sum_{i=1}^{N}\delta_{y_i}\right) = \min_{Y:=(y_1,\dots,y_N)} F_N(Y)}$$

where
$F_N(Y) = \max_{w\in\mathbb{R}^N} \frac{1}{2}\left\{\int_{\mathbb{R}^d}\left(\min_{i=1,\dots,N}\|x-y_i\|^2 - w_i\right) f(x)dx + \frac{1}{N}\sum_{i=1}^{N} w_i\right\}$

(Uniform Quantization)

$$\boxed{\min_{Y,\pi>0,\, \sum_{i=1}^{N}\pi_i=1} \frac{1}{2} W_2^2(\mu, \sum_{i=1}^{N}\pi_i\delta_{y_i}) = \min_{Y:=(y_1,\dots,y_N)} G_N(Y)}$$

where $G_N(Y) = \frac{1}{2}\int_{\mathbb{R}^d} \min_{i=1,\dots,N}\|x-y_i\|^2 f(x)dx$

(Optimal Quantization)

Voronoï and Laguerre tesselation

- i'th Voronoï cell:
  $V_i(Y) = \{x \in \mathbb{R}^d \mid \|x - y_i\|^2 < \|x - y_j\|^2 \, \forall j = 1, ..., N\}$

Optimal Quantization

Uniform Quantization

Voronoï and Laguerre tesselation

- i'th Voronoï cell:
  $V_i(Y) = \{x \in \mathbb{R}^d \mid \|x - y_i\|^2 < \|x - y_j\|^2 \, \forall j = 1, \dots, N\}$
- i'th Laguerre cell:
  $L_i(Y, w) = \{x \in \mathbb{R}^d \mid \|x - y_i\|^2 - w_i < \|x - y_j\|^2 - w_j \, \forall j = 1, \dots, N\}$

Optimal Quantization

Uniform Quantization

Voronoï and Laguerre tesselation

- i'th Voronoï cell:
  $V_i(Y) = \{x \in \mathbb{R}^d \mid \|x - y_i\|^2 < \|x - y_j\|^2 \; \forall j = 1, \dots, N\}$
- i'th Laguerre cell:
  $L_i(Y, w) = \{x \in \mathbb{R}^d \mid \|x - y_i\|^2 - w_i < \|x - y_j\|^2 - w_j \; \forall j = 1, \dots, N\}$

Optimal Quantization

$$\begin{aligned} G_N(Y) &= \frac{1}{2} \int_{\mathbb{R}^d} \min_{i=1,\dots,N} \|x - y_i\|^2 f(x) dx \\ &= \frac{1}{2} \sum_{i=1}^{N} \int_{V_i(Y)} \|x - y_i\|^2 f(x) dx \end{aligned}$$

Discrete

Uniform Quantization

Voronoï and Laguerre tesselation

- i'th Voronoï cell:
  $V_i(Y) = \{x \in \mathbb{R}^d \mid \|x - y_i\|^2 < \|x - y_j\|^2 \; \forall j = 1, \dots, N\}$
- i'th Laguerre cell:
  $L_i(Y, w) = \{x \in \mathbb{R}^d \mid \|x - y_i\|^2 - w_i < \|x - y_j\|^2 - w_j \; \forall j = 1, \dots, N\}$

Optimal Quantization

$$
\begin{aligned}
G_N(Y) &= \frac{1}{2} \int_{\mathbb{R}^d} \min_{i=1,\dots,N} \left\|x - y_i\right\|^2 f(x) dx \\
&= \frac{1}{2} \sum_{i=1}^{N} \int_{V_i(Y)} \left\|x - y_i\right\|^2 f(x) dx
\end{aligned}
$$

Discrete

Uniform Quantization

$$
\begin{aligned}
F_N(Y) &= \max_{w \in \mathbb{R}^N} \frac{1}{2} \left\{ \int_{\mathbb{R}^d} \left(\min_{i=1,\dots,N} \left\|x - y_i\right\|^2 - w_i \right) f(x) dx + \frac{1}{N} \sum_{i=1}^{N} w_i \right\} \\
&= \max_{w \in \mathbb{R}^N} \frac{1}{2} \left\{ \sum_{i=1}^{N} \int_{L_i(Y,w)} \left(\left\|x - y_i\right\|^2 - w_i \right) f(x) dx + \frac{1}{N} \sum_{i=1}^{N} w_i \right\}
\end{aligned}
$$

Lloyd's algorithm[1]: two equivalent formulations

Can be described in two ways

1. Choose an initial points cloud Y.

[1] Lloyd, Stuart P. Least squares quantization in PCM, 1982.

Lloyd's algorithm[1]: two equivalent formulations

Can be described in two ways

1. Choose an initial points cloud Y.
2. Construct either a Voronoï (OQ) or a Laguerre tessellation (UQ) of the support of μ.

[1] Lloyd, Stuart P. Least squares quantization in PCM, 1982.

Lloyd's algorithm[1]: two equivalent formulations

Can be described in two ways

1. Choose an initial points cloud Y.
2. Construct either a Voronoï (OQ) or a Laguerre tessellation (UQ) of the support of μ.
3. Set the new points cloud as the barycenters of every cell.

[1] Lloyd, Stuart P. Least squares quantization in PCM, 1982.

Lloyd's algorithm[1]: two equivalent formulations

Can be described in two ways

1. Choose an initial points cloud Y.
2. Construct either a Voronoï (OQ) or a Laguerre tessellation (UQ) of the support of μ.
3. Set the new points cloud as the barycenters of every cell.

Which translates to

- UQ:

$$\begin{cases} Y_0 \in (\mathbb{R}^d)^N \setminus D_N \\ Y_{n+1} = Y_n - N.\nabla F_N(Y_n) \end{cases}$$

[1] Lloyd, Stuart P. Least squares quantization in PCM, 1982.

Lloyd's algorithm[1]: two equivalent formulations

Can be described in two ways

1. Choose an initial points cloud Y.
2. Construct either a Voronoï (OQ) or a Laguerre tessellation (UQ) of the support of μ.
3. Set the new points cloud as the barycenters of every cell.

Which translates to

- UQ:

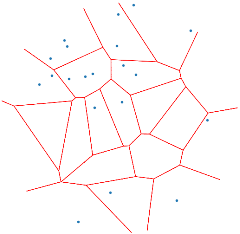
$$\begin{cases} Y_0 \in (\mathbb{R}^d)^N \setminus D_N \\ Y_{n+1} = Y_n - N.\nabla F_N(Y_n) \end{cases}$$

- OQ:

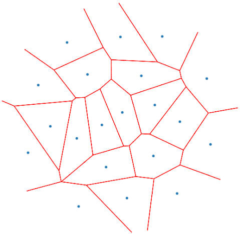
$$\begin{cases} Y_0 \in (\mathbb{R}^d)^N \setminus D_N \\ Y_{n+1} = Y_n - \frac{1}{\mu(V(Y_n))}.\nabla G_N(Y_n) \end{cases}$$

[1] Lloyd, Stuart P. Least squares quantization in PCM, 1982.

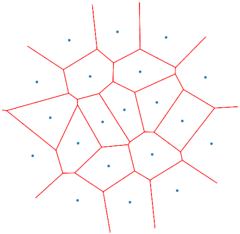
Lloyd's algorithm for (Uniform Quantization): an example with $\mu \sim \mathcal{N}_{\mathbb{S}^1}(0_2, \sigma^2 I_2)$



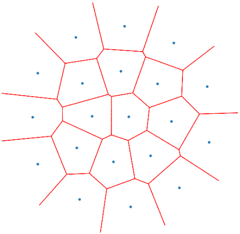
$[n = 0]$



$[n = 1]$



$[n = 2]$



$[n = 20]$

Simulations made using the library PyMongeAmpere by Quentin Mérigot.

Convergence of gradient schemes in non-convex setting

- F_N and G_N both non convex.

[2] H. Attouch, J. Bolte, and B.F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. 2013.

Convergence of gradient schemes in non-convex setting

- F_N and G_N both non convex.

We follow the geometric approach in optimization pioneered by J.Bolte.

[2] H. Attouch, J. Bolte, and B.F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. 2013.

Convergence of gradient schemes in non-convex setting

- F_N and G_N both non convex.

We follow the geometric approach in optimization pioneered by J.Bolte.

- The iterates of gradient sequences converge in this setting under **Lojasiewicz/Kurdyka-Lojasiewicz** inequality. [2]

[2] H. Attouch, J. Bolte, and B.F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. 2013.

Convergence of gradient schemes in non-convex setting

- F_N and G_N both non convex.

We follow the geometric approach in optimization pioneered by J.Bolte.

- The iterates of gradient sequences converge in this setting under **Lojasiewicz/Kurdyka-Lojasiewicz** inequality. [2]
- Such properties are true whenever the underlying objective function's graph belongs to an o-minimal structure.

[2] H. Attouch, J. Bolte, and B.F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. 2013.

Convergence of gradient schemes in non-convex setting

- F_N and G_N both non convex.

We follow the geometric approach in optimization pioneered by J.Bolte.

- The iterates of gradient sequences converge in this setting under **Lojasiewicz/Kurdyka-Lojasiewicz** inequality. [2]
- Such properties are true whenever the underlying objective function's graph belongs to an o-minimal structure.
- F_N and G_N do so under an analyticity assumption on the target density.

[2] H. Attouch, J. Bolte, and B.F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. 2013.

Convergence of Lloyd's algorithm for OQ and UQ

Uniform Quantization

$$F_N(Y) = \max_{w\in\mathbb{R}^N} \frac{1}{2}\left\{\int_{\mathbb{R}^d}\left(\min_{i=1,\dots,N}\|x-y_i\|^2 - w_i\right) f(x)dx + \frac{1}{N}\sum_{i=1}^N w_i\right\}$$

Optimal Quantization

$$G_N(Y) = \frac{1}{2}\int_{\mathbb{R}^d} \min_{i=1,\dots,N}\|x-y_i\|^2 f(x)dx$$

Convergence of Lloyd's algorithm for OQ and UQ

Uniform Quantization

$F_N(Y) = \max\limits_{w \in \mathbb{R}^N} \frac{1}{2} \left\{ \int_{\mathbb{R}^d} \left(\min_{i=1,\dots,N} \|x - y_i\|^2 - w_i \right) f(x) dx + \frac{1}{N} \sum_{i=1}^N w_i \right\}$

Let $(Y_n)_n$ be Lloyd iterates.

Optimal Quantization

$G_N(Y) = \frac{1}{2} \int_{\mathbb{R}^d} \min\limits_{i=1,\dots,N} \|x - y_i\|^2 f(x) dx$

Convergence of Lloyd's algorithm for OQ and UQ

Uniform Quantization

$F_N(Y) = \max_{w \in \mathbb{R}^N} \frac{1}{2} \left\{ \int_{\mathbb{R}^d} \left(\min_{i=1,\dots,N} \|x - y_i\|^2 - w_i \right) f(x) dx + \frac{1}{N} \sum_{i=1}^N w_i \right\}$

Let $(Y_n)_n$ be Lloyd iterates.

- Convergence to a set of critical points[3]: $\|\nabla F_N(Y_n)\| \underset{n \longrightarrow \infty}{\longrightarrow} 0$

Optimal Quantization

$G_N(Y) = \frac{1}{2} \int_{\mathbb{R}^d} \min_{i=1,\dots,N} \|x - y_i\|^2 f(x) dx$

[3] Quentin Merigot, Filippo Santambrogio, Clément Sarrazin. Non-asymptotic convergence bounds for Wasserstein approximation using point clouds. 2021

Convergence of Lloyd's algorithm for OQ and UQ

Uniform Quantization

$F_N(Y) = \max_{w \in \mathbb{R}^N} \frac{1}{2} \left\{ \int_{\mathbb{R}^d} \left(\min_{i=1,\ldots,N} \|x - y_i\|^2 - w_i \right) f(x)dx + \frac{1}{N} \sum_{i=1}^N w_i \right\}$

Let $(Y_n)_n$ be Lloyd iterates.

- Convergence to a set of critical points[3]: $\|\nabla F_N(Y_n)\| \underset{n \longrightarrow \infty}{\longrightarrow} 0$

Optimal Quantization

$G_N(Y) = \frac{1}{2} \int_{\mathbb{R}^d} \min_{i=1,\ldots,N} \|x - y_i\|^2 f(x)dx$

Let (Z_n) be Lloyd iterates.

[3] Quentin Merigot, Filippo Santambrogio, Clément Sarrazin. Non-asymptotic convergence bounds for Wasserstein approximation using point clouds. 2021

Convergence of Lloyd's algorithm for OQ and UQ

Uniform Quantization

$F_N(Y) = \max_{w \in \mathbb{R}^N} \frac{1}{2} \left\{ \int_{\mathbb{R}^d} \left(\min_{i=1,\ldots,N} \|x - y_i\|^2 - w_i \right) f(x)dx + \frac{1}{N} \sum_{i=1}^N w_i \right\}$

Let $(Y_n)_n$ be Lloyd iterates.

- Convergence to a set of critical points[3]: $\|\nabla F_N(Y_n)\| \underset{n \longrightarrow \infty}{\longrightarrow} 0$

Optimal Quantization

$G_N(Y) = \frac{1}{2} \int_{\mathbb{R}^d} \min_{i=1,\ldots,N} \|x - y_i\|^2 f(x)dx$

Let (Z_n) be Lloyd iterates.

- Convergence to a set of critical points[4]: $\|\nabla G_N(Z_n)\| \underset{n \longrightarrow \infty}{\longrightarrow} 0$

[3] Quentin Merigot, Filippo Santambrogio, Clément Sarrazin. Non-asymptotic convergence bounds for Wasserstein approximation using point clouds. 2021

[4] Maria Emelianenko, Lili Ju and Alexander Rand. Nondegeneracy and Weak Global Convergence of the Lloyd Algorithm in $\mathbb{R}^d$. 2008

Convergence of Lloyd's algorithm for OQ and UQ

Uniform Quantization

$F_N(Y) = \max_{w \in \mathbb{R}^N} \frac{1}{2} \left\{ \int_{\mathbb{R}^d} \left(\min_{i=1,\ldots,N} \|x - y_i\|^2 - w_i \right) f(x)dx + \frac{1}{N} \sum_{i=1}^N w_i \right\}$

Let $(Y_n)_n$ be Lloyd iterates.

- Convergence to a set of critical points[3]: $\|\nabla F_N(Y_n)\| \underset{n \longrightarrow \infty}{\longrightarrow} 0$

Optimal Quantization

$G_N(Y) = \frac{1}{2} \int_{\mathbb{R}^d} \min_{i=1,\ldots,N} \|x - y_i\|^2 f(x)dx$

Let (Z_n) be Lloyd iterates.

- Convergence to a set of critical points[4]: $\|\nabla G_N(Z_n)\| \underset{n \longrightarrow \infty}{\longrightarrow} 0$

Question:

What about $(Y_n)_{n \geq 1}$ and $(Z_n)_{n \geq 1}$?

[3] Quentin Merigot, Filippo Santambrogio, Clément Sarrazin. Non-asymptotic convergence bounds for Wasserstein approximation using point clouds. 2021

[4] Maria Emelianenko, Lili Ju and Alexander Rand. Nondegeneracy and Weak Global Convergence of the Lloyd Algorithm in $\mathbb{R}^d$. 2008

Main result

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is ???. Then the iterates of Lloyd for both uniform (UQ) and optimal quantization (OQ) converge.

Main result

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is ???. Then the iterates of Lloyd for both uniform (UQ) and optimal quantization (OQ) converge.

Remark

Convexity of Supp(μ) is not necessary for uniform quantization.

KL inequality

Kurdyka-Lojasiewicz Inequality

f verifies a KL inequality at x^* if there exists a neighborhood U of x^*, $c > 0$, $\eta > 0$ and a strictly increasing positive function $\Psi : [0, \eta[\longrightarrow \mathbb{R}$ such that:

$$\|\nabla(\Psi \circ f)(x)\| \geq c,\ \forall x \in U \text{ such that } 0 < f(x) < \eta. \tag{1}$$

KL inequality

Kurdyka-Lojasiewicz Inequality

f verifies a KL inequality at x^* if there exists a neighborhood U of x^*, $c > 0$, $\eta > 0$ and a strictly increasing positive function $\Psi : [0, \eta[\longrightarrow \mathbb{R}$ such that:

$$\|\nabla(\Psi \circ f)(x)\| \geq c, \ \forall x \in U \text{ such that } 0 < f(x) < \eta. \tag{1}$$

Main idea of KL: Composing f with Ψ makes it sharper around its critical points.

KL inequality

Kurdyka-Lojasiewicz Inequality

f verifies a KL inequality at x^* if there exists a neighborhood U of x^*, $c > 0$, $\eta > 0$ and a strictly increasing positive function $\Psi : [0, \eta[\longrightarrow \mathbb{R}$ such that:

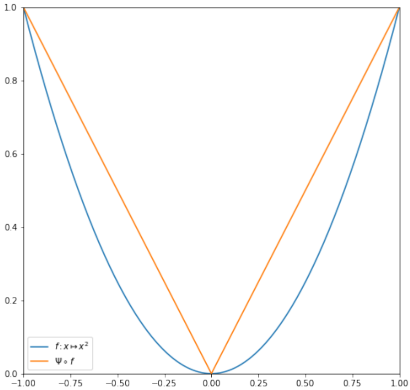
$$\|\nabla(\Psi \circ f)(x)\| \geq c, \; \forall x \in U \text{ such that } \; 0 < f(x) < \eta. \tag{1}$$

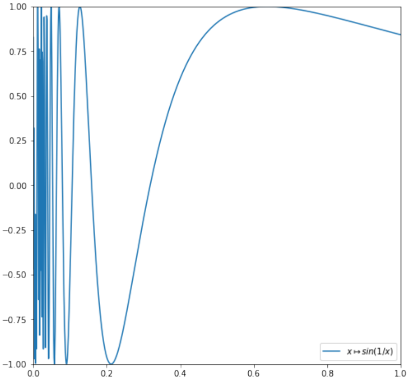
Main idea of KL: Composing f with Ψ makes it sharper around its critical points.

A function whose differential go to zero can have infinitely many oscillations as it goes to 0. The KL property prevents that from happening.

KL inequality

A case where the function is KL... and another one where it is not (at 0).





f:x$\mapsto x^2$, $\Psi : x \mapsto \sqrt{x}$

$g : x \mapsto \sin 1/x$.

Strong descent conditions

Strong descent conditions: Let $(Y_k)_{k\geq 1}$ be some sequence and F be some differentiable cost function.

Strong descent conditions

Strong descent conditions: Let $(Y_k)_{k\geq 1}$ be some sequence and F be some differentiable cost function.

- $F(Y_k) - F(Y_{k+1}) \geq \sigma \|\nabla F(Y_k)\|.\|Y_{k+1} - Y_k\|$

Strong descent conditions

Strong descent conditions: Let $(Y_k)_{k\geq 1}$ be some sequence and F be some differentiable cost function.

- $F(Y_k) - F(Y_{k+1}) \geq \sigma \|\nabla F(Y_k)\|.\|Y_{k+1} - Y_k\|$
- $F(Y_{k+1}) = F(Y_k) \implies Y_{k+1} = Y_k$.

Strong descent conditions

Strong descent conditions: Let $(Y_k)_{k\geq 1}$ be some sequence and F be some differentiable cost function.

- $F(Y_k) - F(Y_{k+1}) \geq \sigma \|\nabla F(Y_k)\|.\|Y_{k+1} - Y_k\|$
- $F(Y_{k+1}) = F(Y_k) \implies Y_{k+1} = Y_k$.

(We say that the sequence $(Y_k)_{k\geq 1}$ is SDC)

Strong descent conditions

Strong descent conditions: Let $(Y_k)_{k\geq 1}$ be some sequence and F be some differentiable cost function.

- $F(Y_k) - F(Y_{k+1}) \geq \sigma \|\nabla F(Y_k)\|.\|Y_{k+1} - Y_k\|$
- $F(Y_{k+1}) = F(Y_k) \implies Y_{k+1} = Y_k$.

(We say that the sequence $(Y_k)_{k\geq 1}$ is SDC)

- We show that G_N is SDC (largely thanks to Emelianenko, Ju, Rand [5]).

[5] M. Emelianenko, L. Ju, and A. Rand. Nondegeneracy and weak global convergence of the Lloyd algorithm in $\mathbb{R}^d$. 2008

Strong descent conditions

Strong descent conditions: Let $(Y_k)_{k\geq 1}$ be some sequence and F be some differentiable cost function.

- $F(Y_k) - F(Y_{k+1}) \geq \sigma \|\nabla F(Y_k)\| . \|Y_{k+1} - Y_k\|$
- $F(Y_{k+1}) = F(Y_k) \implies Y_{k+1} = Y_k$.

(We say that the sequence $(Y_k)_{k\geq 1}$ is SDC)

- We show that G_N is SDC (largely thanks to Emelianenko, Ju, Rand [5]).
- F_N SDC (can be deduced from Merigot, Santambrogio, Sarrazin [6]).

[5] M. Emelianenko, L. Ju, and A. Rand. Nondegeneracy and weak global convergence of the Lloyd algorithm in $\mathbb{R}^d$. 2008

[6] Quentin Merigot, Filippo Santambrogio, Clément Sarrazin. Non-asymptotic convergence bounds for Wasserstein approximation using point clouds. 2021

Idea of proof

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is ???. Then the iterates of Lloyd for both uniform (UQ) and optimal quantization (OQ) converge.

Sketch of proof

Idea of proof

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is ???. Then the iterates of Lloyd for both uniform (UQ) and optimal quantization (OQ) converge.

Sketch of proof

- F_N and G_N verify a KL inequality at all points of their domain.

Idea of proof

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is ???. Then the iterates of Lloyd for both uniform (UQ) and optimal quantization (OQ) converge.

Sketch of proof

- F_N and G_N verify a KL inequality at all points of their domain.
- SDC gradient methods on KL functions converge *(Bolte et.al, Absil et.al.)*

Idea of proof

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is ???. Then the iterates of Lloyd for both uniform (UQ) and optimal quantization (OQ) converge.

Sketch of proof

- F_N and G_N verify a KL inequality at all points of their domain.
- SDC gradient methods on KL functions converge *(Bolte et.al, Absil et.al.)*
- Lloyd's algorithm for OQ and UQ are SDC gradient methods on KL functions and thus the iterates converge.

Idea of proof

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is ???. Then the iterates of Lloyd for both uniform (UQ) and optimal quantization (OQ) converge.

Sketch of proof

- F_N and G_N verify a KL inequality at all points of their domain.
- SDC gradient methods on KL functions converge *(Bolte et.al, Absil et.al.)*
- Lloyd's algorithm for OQ and UQ are SDC gradient methods on KL functions and thus the iterates converge.

Question:

Why are F_N and G_N KL?

Idea of proof

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is ???. Then the iterates of Lloyd for both uniform (UQ) and optimal quantization (OQ) converge.

Sketch of proof

- F_N and G_N verify a KL inequality at all points of their domain.
- SDC gradient methods on KL functions converge *(Bolte et.al, Absil et.al.)*
- Lloyd's algorithm for OQ and UQ are SDC gradient methods on KL functions and thus the iterates converge.

Question:

Why are F_N and G_N KL? Because they are definable in an o minimal structure.

A (very quick) overview on o-minimality

$$\mathcal{SA} \subset \mathbb{R}_{\mathsf{an}} \subset \mathbb{R}_{\mathsf{an,exp}} \subset \mathcal{P}(\mathbb{R})$$

[7] H. Hironaka. Introduction to Real-analytic Sets and Real-analytic Maps. 1973

[8] L.van den Dries and C.Miller, On the real exponential field with restricted analytic functions. 1994

A (very quick) overview on o-minimality

$$\mathcal{SA} \subset \mathbb{R}_{\text{an}} \subset \mathbb{R}_{\text{an,exp}} \subset \mathcal{P}(\mathbb{R})$$

- $\mathcal{SA}$:

[7] H. Hironaka. Introduction to Real-analytic Sets and Real-analytic Maps. 1973

[8] L.van den Dries and C.Miller, On the real exponential field with restricted analytic functions. 1994

A (very quick) overview on o-minimality

$$\mathcal{SA} \subset \mathbb{R}_{\text{an}} \subset \mathbb{R}_{\text{an,exp}} \subset \mathcal{P}(\mathbb{R})$$

- $\mathcal{SA}$: $\bigcup_{i=1}^{N} \{x \in \mathbb{R}^d \mid P_{i_1}(x) \stackrel{>}{\geq} 0, \ldots, P_{i_l} \stackrel{>}{\geq} 0, (P_{i_k})_{k=1,\ldots,l} \subset \mathbb{R}_d[X]\}$

[7] H. Hironaka. Introduction to Real-analytic Sets and Real-analytic Maps. 1973

[8] L.van den Dries and C.Miller, On the real exponential field with restricted analytic functions. 1994

A (very quick) overview on o-minimality

$$\mathcal{SA} \subset \mathbb{R}_{\text{an}} \subset \mathbb{R}_{\text{an,exp}} \subset \mathcal{P}(\mathbb{R})$$

- $\mathcal{SA}$: $\bigcup_{i=1}^{N} \{x \in \mathbb{R}^d \mid P_{i_1}(x) \overset{>}{\geq} 0, ..., P_{i_l} \overset{>}{\geq} 0, (P_{i_k})_{k=1,...,l} \subset \mathbb{R}_d[X]\}$
  Definability stable by $+$, $\times$, $(.)^{-1}$, min, ...

[7] H. Hironaka. Introduction to Real-analytic Sets and Real-analytic Maps. 1973

[8] L.van den Dries and C.Miller, On the real exponential field with restricted analytic functions. 1994

A (very quick) overview on o-minimality

$$\mathcal{SA} \subset \mathbb{R}_{\text{an}} \subset \mathbb{R}_{\text{an,exp}} \subset \mathcal{P}(\mathbb{R})$$

- $\mathcal{SA}$: $\bigcup_{i=1}^{N} \{x \in \mathbb{R}^d \mid P_{i_1}(x) \overset{>}{\geq} 0, ..., P_{i_l} \overset{>}{\geq} 0, (P_{i_k})_{k=1,...,l} \subset \mathbb{R}_d[X]\}$
  Definability stable by $+$, $\times$, $(.)^{-1}$, min, ...
- $\mathbb{R}_{\text{an}}$: Same with the analytic maps restricted to a compact within their domain of definition [7] .

[7] H. Hironaka. Introduction to Real-analytic Sets and Real-analytic Maps. 1973
[8] L.van den Dries and C.Miller, On the real exponential field with restricted analytic functions. 1994

A (very quick) overview on o-minimality

$$\mathcal{SA} \subset \mathbb{R}_{\text{an}} \subset \mathbb{R}_{\text{an,exp}} \subset \mathcal{P}(\mathbb{R})$$

- $\mathcal{SA}$: $\bigcup_{i=1}^{N} \{x \in \mathbb{R}^d \mid P_{i_1}(x) \overset{>}{\geq} 0, ..., P_{i_l} \overset{>}{\geq} 0, (P_{i_k})_{k=1,...,l} \subset \mathbb{R}_d[X]\}$
  Definability stable by $+$, $\times$, $(.)^{-1}$, min, ...
- $\mathbb{R}_{\text{an}}$: Same with the analytic maps restricted to a compact within their domain of definition [7] example: $\exp|_{[-1,1]}$.

[7] H. Hironaka. Introduction to Real-analytic Sets and Real-analytic Maps. 1973

[8] L.van den Dries and C.Miller, On the real exponential field with restricted analytic functions. 1994

A (very quick) overview on o-minimality

$$\mathcal{SA} \subset \mathbb{R}_{\text{an}} \subset \mathbb{R}_{\text{an,exp}} \subset \mathcal{P}(\mathbb{R})$$

- $\mathcal{SA}$: $\bigcup_{i=1}^{N} \{x \in \mathbb{R}^d \mid P_{i_1}(x) \overset{>}{\geq} 0, ..., P_{i_l} \overset{>}{\geq} 0, (P_{i_k})_{k=1,...,l} \subset \mathbb{R}_d[X]\}$
  Definability stable by $+$, $\times$, $(.)^{-1}$, min, ...
- $\mathbb{R}_{\text{an}}$: Same with the analytic maps restricted to a compact within their domain of definition [7] example: $\exp|_{[-1,1]}$.
- $\mathbb{R}_{\text{an,exp}}$: Same with the exponential map.[8]

[7] H. Hironaka. Introduction to Real-analytic Sets and Real-analytic Maps. 1973

[8] L.van den Dries and C.Miller, On the real exponential field with restricted analytic functions. 1994

[9] Raf Cluckers, Daniel J. Miller. Stability under integration of sums of products of real globally subanalytic functions and their logarithms. 2009

A (very quick) overview on o-minimality

$$\mathcal{SA} \subset \mathbb{R}_{\text{an}} \subset \mathbb{R}_{\text{an,exp}} \subset \mathcal{P}(\mathbb{R})$$

- $\mathcal{SA}$: $\bigcup_{i=1}^{N} \{x \in \mathbb{R}^d \mid P_{i_1}(x) \overset{>}{\geq} 0, \ldots, P_{i_l} \overset{>}{\geq} 0, (P_{i_k})_{k=1,\ldots,l} \subset \mathbb{R}_d[X]\}$
  Definability stable by $+$, $\times$, $(.)^{-1}$, min, ...
- $\mathbb{R}_{\text{an}}$: Same with the analytic maps restricted to a compact within their domain of definition [7] example: $\exp|_{[-1,1]}$.
- $\mathbb{R}_{\text{an,exp}}$: Same with the exponential map.[8]
- If f is ??? then $Y \mapsto \int_{\mathcal{X}} f(x, Y)dx$ is definable in an o-minimal structure.[9]

[7] H. Hironaka. Introduction to Real-analytic Sets and Real-analytic Maps. 1973

[8] L.van den Dries and C.Miller, On the real exponential field with restricted analytic functions. 1994

[9] Raf Cluckers, Daniel J. Miller. Stability under integration of sums of products of real globally subanalytic functions and their logarithms. 2009

A (very quick) overview on o-minimality

$$\mathcal{SA} \subset \mathbb{R}_{\text{an}} \subset \mathbb{R}_{\text{an,exp}} \subset \mathcal{P}(\mathbb{R})$$

- $\mathcal{SA}$: $\bigcup_{i=1}^{N} \{x \in \mathbb{R}^d \mid P_{i_1}(x) \overset{>}{\geq} 0, \ldots, P_{i_l} \overset{>}{\geq} 0, (P_{i_k})_{k=1,\ldots,l} \subset \mathbb{R}_d[X]\}$
  Definability stable by $+$, $\times$, $(.)^{-1}$, min, ...
- $\mathbb{R}_{\text{an}}$: Same with the analytic maps restricted to a compact within their domain of definition [7] example: $\exp|_{[-1,1]}$.
- $\mathbb{R}_{\text{an,exp}}$: Same with the exponential map.[8]
- If f is GSA then $Y \mapsto \int_{\mathcal{X}} f(x, Y)dx$ is definable in an o-minimal structure.[9]
- Functions definable in an o minimal structure are KL.

[7] H. Hironaka. Introduction to Real-analytic Sets and Real-analytic Maps. 1973

[8] L.van den Dries and C.Miller, On the real exponential field with restricted analytic functions. 1994

[9] Raf Cluckers, Daniel J. Miller. Stability under integration of sums of products of real globally subanalytic functions and their logarithms. 2009

Examples of globally subanalytic (GSA) functions

- An image (regarded as a probability density.)

Examples of globally subanalytic (GSA) functions

- An image (regarded as a probability density.)
- Semi-algebraic densities (Uniform law on a semi algebraic set, the triangular distribution,...)

Examples of globally subanalytic (GSA) functions

- An image (regarded as a probability density.)
- Semi-algebraic densities (Uniform law on a semi algebraic set, the triangular distribution,...)
- Any probability density which is analytic and which is truncated on a compact semi algebraic subset (ex: Gaussians truncated on a sphere).

Examples of globally subanalytic (GSA) functions

- An image (regarded as a probability density.)
- Semi-algebraic densities (Uniform law on a semi algebraic set, the triangular distribution,...)
- Any probability density which is analytic and which is truncated on a compact semi algebraic subset (ex: Gaussians truncated on a sphere).
- Any mix of the previous ones.

Examples of globally subanalytic (GSA) functions

- An image (regarded as a probability density.)
- Semi-algebraic densities (Uniform law on a semi algebraic set, the triangular distribution,...)
- Any probability density which is analytic and which is truncated on a compact semi algebraic subset (ex: Gaussians truncated on a sphere).
- Any mix of the previous ones.

Think of it as analytic functions restricted to a semi-algebraic compact subset

Proving the definability of G_N and F_N

$$G_N(Y) = \frac{1}{2} \int_{\mathbb{R}^d} \min_{i=1,\dots,N} \|x - y_i\|^2 \underbrace{f(x)}_{\text{GSA}} dx$$

Proving the definability of G_N and F_N

$$G_N(Y) = \frac{1}{2} \int_{\mathbb{R}^d} \underbrace{\min_{i=1,\dots,N} \|x - y_i\|^2 \underbrace{f(x)}_{\text{GSA}} dx}_{\text{GSA}}$$

Proving the definability of G_N and F_N

$$G_N(Y) = \underbrace{\frac{1}{2}\int_{\mathbb{R}^d} \underbrace{\min_{i=1,\dots,N} \|x - y_i\|^2 \underbrace{f(x)}_{\text{GSA}}\, dx}_{\text{GSA}}}_{\text{definable}}$$

Proving the definability of G_N and F_N

$$G_N(Y) = \frac{1}{2} \underbrace{\int_{\mathbb{R}^d} \underbrace{\min_{i=1,\dots,N} \|x - y_i\|^2 \underbrace{f(x)}_{\text{GSA}} \, dx}_{\text{GSA}}}_{\text{definable}}$$

Conclusion: G_N is a KL function.

Proving the definability of G_N and F_N

Conclusion: G_N is a KL function.

$$F_N(Y) = \underbrace{\max_{w \in \mathbb{R}^N} \frac{1}{2} \int_{\mathbb{R}^d} \underbrace{\min_{i=1,\dots,N} (\|x - y_i\|^2 - w_i) \underbrace{f(x)}_{\text{GSA}}}_{\text{GSA}} dx + \frac{1}{N} \sum_{i=1}^{N} w_i}_{\text{definable}}$$

Proving the definability of G_N and F_N

Conclusion: G_N is a KL function.

$$F_N(Y) = \underbrace{\max_{w \in \mathbb{R}^N} \underbrace{\frac{1}{2} \int_{\mathbb{R}^d} \underbrace{\min_{i=1,\dots,N} (\|x - y_i\|^2 - w_i)\, \underbrace{f(x)}_{\text{GSA}}}_{\text{GSA}}\, dx + \frac{1}{N} \sum_{i=1}^{N} w_i}_{\text{definable}}}_{\text{definable}} .$$

Conclusion: F_N is a KL function.

Results

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is globally subanalytic. Then the iterates of Lloyd for both uniform and optimal quantization converge.

Results

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is globally subanalytic. Then the iterates of Lloyd for both uniform and optimal quantization converge.

Additional results:

We showed along the way the definability of $Y \mapsto D(\mu, \frac{1}{N}\sum_{k=1}^{N} \delta_{y_k})$ for

Results

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is globally subanalytic. Then the iterates of Lloyd for both uniform and optimal quantization converge.

Additional results:

We showed along the way the definability of $Y \mapsto D(\mu, \frac{1}{N}\sum_{k=1}^{N} \delta_{y_k})$ for

- $D = W_p$ (general Wasserstein)

Results

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is globally subanalytic. Then the iterates of Lloyd for both uniform and optimal quantization converge.

Additional results:

We showed along the way the definability of $Y \mapsto D(\mu, \frac{1}{N}\sum_{k=1}^{N} \delta_{y_k})$ for

- $D = W_p$ (general Wasserstein)
- $D = maxSW_2$ (max Sliced Wasserstein)

Results

Theorem

Let μ be a continuous probability measure supported on a compact and convex subset of $\mathbb{R}^d$. Suppose its density is globally subanalytic. Then the iterates of Lloyd for both uniform and optimal quantization converge.

Additional results:

We showed along the way the definability of $Y \mapsto D(\mu, \frac{1}{N}\sum_{k=1}^{N}\delta_{y_k})$ for

- $D = W_p$ (general Wasserstein)
- $D = maxSW_2$ (max Sliced Wasserstein)
- $D = W_\epsilon$ (Entropic regularization of the semi-discrete OT problem)

What to remember

- Continuous probabilities can be quantized with an arbitrary (OQ) or a uniform (UQ) measure on a points cloud.

What to remember

- Continuous probabilities can be quantized with an arbitrary (OQ) or a uniform (UQ) measure on a points cloud.
- This can be solved with Lloyd's algorithm.

What to remember

- Continuous probabilities can be quantized with an arbitrary (OQ) or a uniform (UQ) measure on a points cloud.
- This can be solved with Lloyd's algorithm.
- Sequential convergence: geometric approach to optimization and first order method (J. Bolte).

What to remember

- Continuous probabilities can be quantized with an arbitrary (OQ) or a uniform (UQ) measure on a points cloud.
- This can be solved with Lloyd's algorithm.
- Sequential convergence: geometric approach to optimization and first order method (J. Bolte).

Thank you for listening!

Let $Y \notin D_N$, then

$$\min_{\pi \in \Delta_N} W_2^2\left(\mu, \sum_{i=1}^N \pi_i \delta_{y_i}\right) = \int_{\mathbb{R}^d} \min_{i=1,\dots,N} \|x - y_i\|^2 d\mu(x).$$

Proof.

We denote $\mathcal{H}(Z, Y) = \sum_{i=1}^N \int_{V_i(Z)} \|x - y_i\|^2 d\mu(x)$.
We observe that $\mathcal{H}(Z, Y) = \int_{\mathbb{R}^d} \|x - T(x)\|^2 d\mu(x)$ with $T(x) = \sum_{i=1}^N y_i 1_{V_i(Z)}(x)$. This is the optimal transport map from μ to $\sum_{i=1}^N \mu(V_i(Z))\delta_{y_i}$, so:

$$\mathcal{H}(Z, Y) = W_2^2(\mu, \sum_{i=1}^N \mu(V_i(Z))\delta_{y_i}).$$

Since $\min_Z \mathcal{H}(Z, Y) = \mathcal{H}(Y, Y)$ we finally get

$$\min_Z \mathcal{H}(Z, Y) = \min_\pi W_2^2(\mu, \sum_{i=1}^N \pi_i \delta_{y_i}).$$