EPFL

Gradient Flows in Wasserstein Spaces

Variational Inference and Sampling

Student: L.Raffo

PostDoc. L.V.Santoro

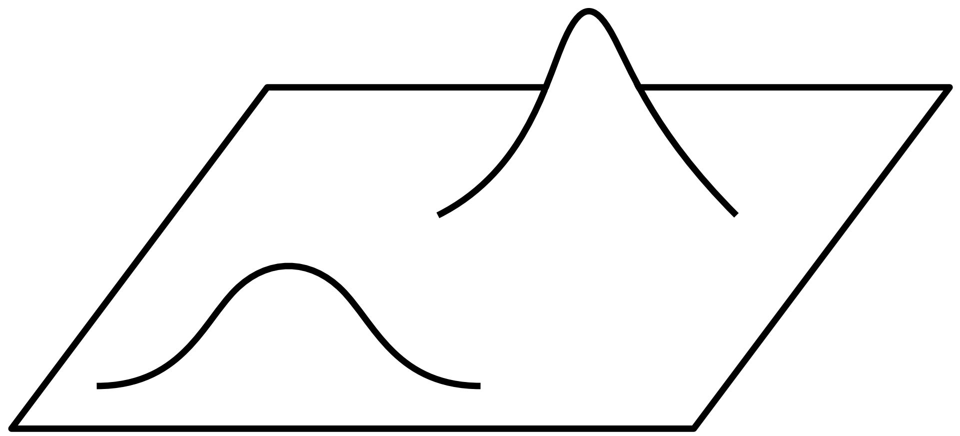
motivation

consider the Euclidean space $\mathbb{R}^d$.

motivation

consider the Euclidean space $\mathbb{R}^d$.

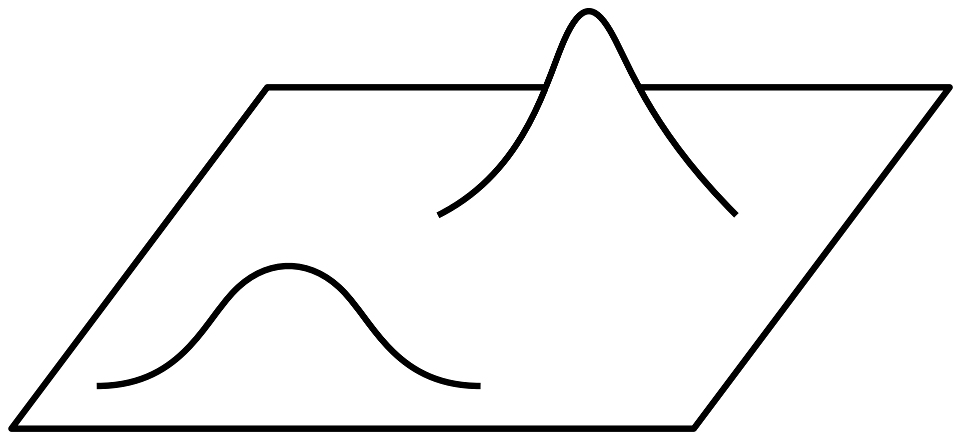
look at the family of probability measures $\mathcal{P}$ on $\mathbb{R}^d$.



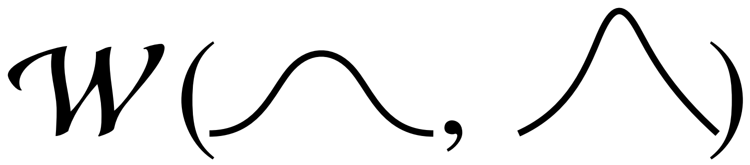
motivation

consider the Euclidean space $\mathbb{R}^d$.

look at the family of probability measures $\mathcal{P}$ on $\mathbb{R}^d$.



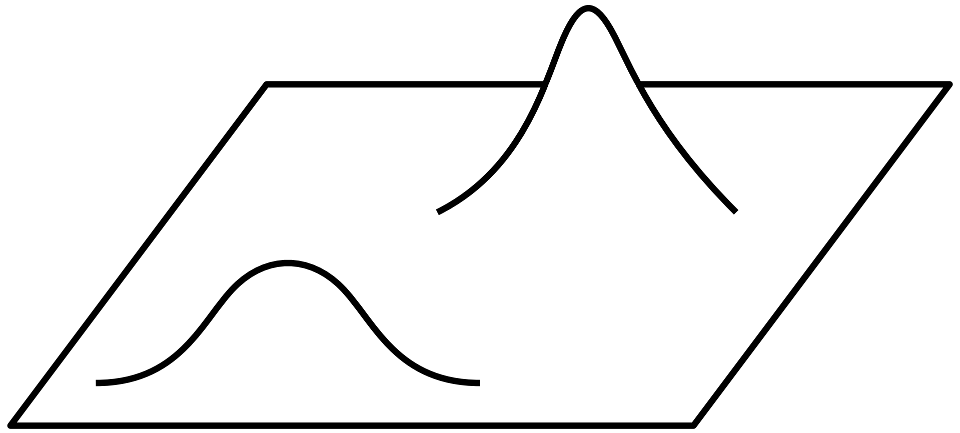
define a distance $\mathcal{W}$ between probability measures.



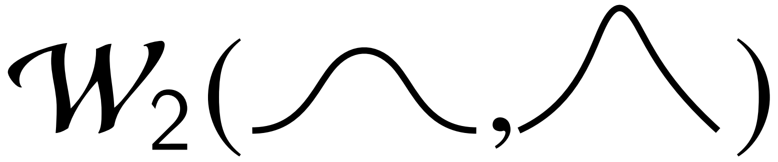
motivation

consider the Euclidean space $\mathbb{R}^d$.

look at the family of probability measures $\mathcal{P}_2$ on $\mathbb{R}^d$.



define a distance $\mathcal{W}_2$ between probability measures.



motivation

the metric space $(\mathcal{P}_2, \mathcal{W}_2)$ has nice geometric properties.

motivation

the metric space $(\mathcal{P}_2, \mathcal{W}_2)$ has nice geometric properties.

despite being infinite dimensional, it resembles a Riemannian manifold.

motivation

the metric space $(\mathcal{P}_2, \mathcal{W}_2)$ has nice geometric properties.

despite being infinite dimensional, it resembles a $\boxed{\text{Riemannian manifold.}}$

geodesics

tangent spaces

motivation

the metric space $(\mathcal{P}_2, \mathcal{W}_2)$ has nice geometric properties.

despite being infinite dimensional, it resembles a Riemannian manifold.

geodesics

midpoints

tangent spaces

Riemannian metric

motivation

we can study how measures evolve in $(\mathcal{P}_2, \mathcal{W}_2)$.

motivation

we can study how measures evolve in $(\mathcal{P}_2, \mathcal{W}_2)$.

gradient flows in $(\mathcal{P}_2, \mathcal{W}_2)$

motivation

we can study how measures $\boxed{\text{evolve}}$ in $(\mathcal{P}_2, \mathcal{W}_2)$.

gradient flows in $(\mathcal{P}_2, \mathcal{W}_2)$

variational inference

EPFL

motivation

we can study how measures $\boxed{\text{evolve}}$ in $(\mathcal{P}_2, \mathcal{W}_2)$.

gradient flows in $(\mathcal{P}_2, \mathcal{W}_2)$

variational inference

particles variational inference

motivation

we can give a geometrical interpretation to some SDEs.

motivation

we can give a geometrical interpretation to some SDEs.

Langevin diffusion

motivation

we can give a geometrical interpretation to some $\boxed{\text{SDEs.}}$

Langevin diffusion

sampling

plan

1. preliminaries. metric geometry, Monge and Kantorovich problems.

plan

1. preliminaries. metric geometry, Monge and Kantorovich problems.

2. Wasserstein spaces. pseudo-Riemannian geometry, evolution of measures, first variations, Wasserstein gradient flows.

plan

1. preliminaries. metric geometry, Monge and Kantorovich problems.

2. Wasserstein spaces. pseudo-Riemannian geometry, evolution of measures, first variations, Wasserstein gradient flows.

3. variational inference. KL divergence, geodesic convexity, hints on the JKO scheme.

plan

1. **preliminaries.** metric geometry, Monge and Kantorovich problems.
2. **Wasserstein spaces.** pseudo-Riemannian geometry, evolution of measures, first variations, Wasserstein gradient flows.
3. **variational inference.** KL divergence, geodesic convexity, hints on the JKO scheme.
4. **particles variational inference.** (RKHS), many particles systems, SVGD.

plan

1. preliminaries. metric geometry, Monge and Kantorovich problems.
2. Wasserstein spaces. pseudo-Riemannian geometry, evolution of measures, first variations, Wasserstein gradient flows.
3. variational inference. KL divergence, geodesic convexity, hints on the JKO scheme.
4. particles variational inference. (RKHS), many particles systems, SVGD.
5. sampling. Langevin diffusion as a gradient flow.

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

fix any metric space $(\mathcal{S}, d)$.

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

fix any $\boxed{\text{metric space}}$ $(\mathcal{S}, d)$.

positive definiteness
symmetry
triangle inequality

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

fix any $\boxed{\text{metric space}}$ $(\mathcal{S}, d)$.

positive definiteness
symmetry
triangle inequality

we can already define *paths*: $\omega : I \to \mathcal{S}$, ω continuous.

EPFL

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

fix any $\boxed{\text{metric space}}$ $(\mathcal{S}, d)$.

positive definiteness
symmetry
triangle inequality

$I \subseteq \mathbb{R}$

we can already define *paths*: $\omega : \boxed{I} \to \mathcal{S}$, ω continuous.

EPFL

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

fix any $\boxed{\text{metric space}}$ $(\mathcal{S}, d)$.

$I \subseteq \mathbb{R}$

positive definiteness
symmetry
triangle inequality

we can already define *paths*: $\omega : \boxed{I} \to \mathcal{S}$, ω continuous.

and lengths: $L(\omega) := \sup \sum_{i=1}^{n-1} d(\omega_{t_i}, \omega_{t_{i+1}}) \approx \int_I \| \frac{d\omega}{dt} \|_d dt$.

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

fix any $\boxed{\text{metric space}}$ $(\mathcal{S}, d)$.

positive definiteness
symmetry
triangle inequality

$I \subseteq \mathbb{R}$

we can already define *paths*: $\omega : \boxed{I} \to \mathcal{S}$, ω continuous.

and lengths: $L(\omega) := \sup \sum_{i=1}^{n-1} d(\omega_{t_i}, \omega_{t_{i+1}}) \approx \int_I \| \frac{d\omega}{dt} \|_d dt$.

we say that $\omega : [a, b] \to \mathcal{S}$ has constant speed if $L(\omega_{[s,t]}) = \frac{t-s}{b-a} L(\omega)$, for any $a \leq s \leq t \leq b$.

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

fix any $\boxed{\text{metric space}}$ $(\mathcal{S}, d)$.

positive definiteness
symmetry
triangle inequality

$I \subseteq \mathbb{R}$

we can already define *paths*: $\omega : \boxed{I} \to \mathcal{S}$, ω continuous.

and lengths: $L(\omega) :=$ sup $\sum_{i=1}^{n-1} d(\omega_{t_i}, \omega_{t_{i+1}}) \approx \int_I \| \frac{d\omega}{dt} \|_d dt$.

we say that $\omega : [a, b] \to \mathcal{S}$ has constant speed if $L(\omega_{[s,t]}) = \frac{t-s}{b-a} L(\omega)$, for any $a \leq s \leq t \leq b$.

$\omega_2 : I_2 \to \mathcal{S}$ is a *reparametrixation* of $\omega_2 : I_1 \to \mathcal{S}$ if $\exists\ \phi : I_1 \to I_2$ such that $\omega_1 = \omega_2 \circ \phi$.

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

fix any $\boxed{\text{metric space}}$ $(\mathcal{S}, d)$.

positive definiteness
symmetry
triangle inequality

$I \subseteq \mathbb{R}$

we can already define *paths*: $\omega : \boxed{I} \to \mathcal{S}$, ω continuous.

and lengths: $L(\omega) :=$ sup $\sum_{i=1}^{n-1} d(\omega_{t_i}, \omega_{t_{i+1}}) \approx \int_I \| \frac{d\omega}{dt} \|_d dt$.

we say that $\omega : [a, b] \to \mathcal{S}$ has constant speed if $L(\omega_{[s,t]}) = \frac{t-s}{b-a} L(\omega)$, for any $a \leq s \leq t \leq b$.

$\omega_2 : I_2 \to \mathcal{S}$ is a *reparametrixation* of $\omega_2 : I_1 \to \mathcal{S}$ if $\exists$ $\boxed{\phi : I_1 \to I_2}$ such that $\omega_1 = \omega_2 \circ \phi$.

continuous
non decreasing
bijective

EPFL

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

fix any $\boxed{\text{metric space}}$ $(\mathcal{S}, d)$.

positive definiteness
symmetry
triangle inequality

$I \subseteq \mathbb{R}$

we can already define *paths*: $\omega : \boxed{I} \to \mathcal{S}$, ω continuous.

and lengths: $L(\omega) := \sup \sum_{i=1}^{n-1} d(\omega_{t_i}, \omega_{t_{i+1}}) \approx \int_I \|\frac{d\omega}{dt}\|_d dt$.

we say that $\omega : [a, b] \to \mathcal{S}$ has constant speed if $L(\omega_{[s,t]}) = \frac{t-s}{b-a} L(\omega)$, for any $a \leq s \leq t \leq b$.

$\omega_2 : I_2 \to \mathcal{S}$ is a *reparametrixation* of $\omega_2 : I_1 \to \mathcal{S}$ if $\exists$ $\boxed{\phi : I_1 \to I_2}$ such that $\omega_1 = \omega_2 \circ \phi$.

continuous
non decreasing
bijective

fact: any ω_1 can be reparametrized to have constant speed.

preliminaries. metric geometry.

in short: abstraction of key ideas from differential geometry.

fix any $\boxed{\text{metric space}}$ $(\mathcal{S}, d)$.

positive definiteness
symmetry
triangle inequality

$I \subseteq \mathbb{R}$

we can already define *paths*: $\omega : \boxed{I} \to \mathcal{S}$, ω continuous.

and lengths: $L(\omega) :=$ sup $\sum_{i=1}^{n-1} d(\omega_{t_i}, \omega_{t_{i+1}}) \approx \int_I \| \frac{d\omega}{dt} \|_d dt$.

we say that $\omega : [a, b] \to \mathcal{S}$ has constant speed if $L(\omega_{[s,t]}) = \frac{t-s}{b-a} L(\omega)$, for any $a \leq s \leq t \leq b$.

$\omega_2 : I_2 \to \mathcal{S}$ is a *reparametrixation* of $\omega_2 : I_1 \to \mathcal{S}$ if $\exists$ $\boxed{\phi : I_1 \to I_2}$ such that $\omega_1 = \omega_2 \circ \phi$.

continuous
non decreasing
bijective

fact: any ω_1 can be reparametrized to have constant speed.
moreover, we can choose $I_2 = [0, 1]$.

preliminaries. metric geometry.

fix $x_0, x_1 \in \mathcal{S}$.

preliminaries. metric geometry.

fix $x_0, x_1 \in \mathcal{S}$.

a path $\omega : [0,1] \to \mathcal{S}$, with $\omega(0) = x_0$ and $\omega(1) = x_1$ is a *geodesic* if $d(x_0, x_1) = L(\omega)$.

preliminaries. metric geometry.

fix $x_0, x_1 \in \mathcal{S}$.

a path $\omega : [0,1] \to \mathcal{S}$, with $\omega(0) = x_0$ and $\omega(1) = x_1$ is a *geodesic* if $d(x_0, x_1) = L(\omega)$.

ω can be reparametrized to be a *constant speed geodesic*.

preliminaries. metric geometry.

fix $x_0, x_1 \in \mathcal{S}$.

a path $\omega : [0,1] \to \mathcal{S}$, with $\omega(0) = x_0$ and $\omega(1) = x_1$ is a *geodesic* if $d(x_0, x_1) = L(\omega)$.

ω can be reparametrized to be a *constant speed geodesic*.

$(\mathcal{S}, d)$ is said to be a *geodesic space* if for any given x_0, x_1 we can exhibit a geodesic.

preliminaries. metric geometry.

fix $x_0, x_1 \in \mathcal{S}$.

a path $\omega : [0,1] \to \mathcal{S}$, with $\omega(0) = x_0$ and $\omega(1) = x_1$ is a *geodesic* if $d(x_0, x_1) = L(\omega)$.

ω can be reparametrized to be a *constant speed geodesic*.

$(\mathcal{S}, d)$ is said to be a *geodesic space* if for any given x_0, x_1 we can exhibit a geodesic.

for us, constant speed geodesic.

EPFL

preliminaries. metric geometry.

fix $x_0, x_1 \in \mathcal{S}$.

a path $\omega : [0,1] \to \mathcal{S}$, with $\omega(0) = x_0$ and $\omega(1) = x_1$ is a *geodesic* if $d(x_0, x_1) = L(\omega)$.

ω can be reparametrized to be a *constant speed geodesic*.

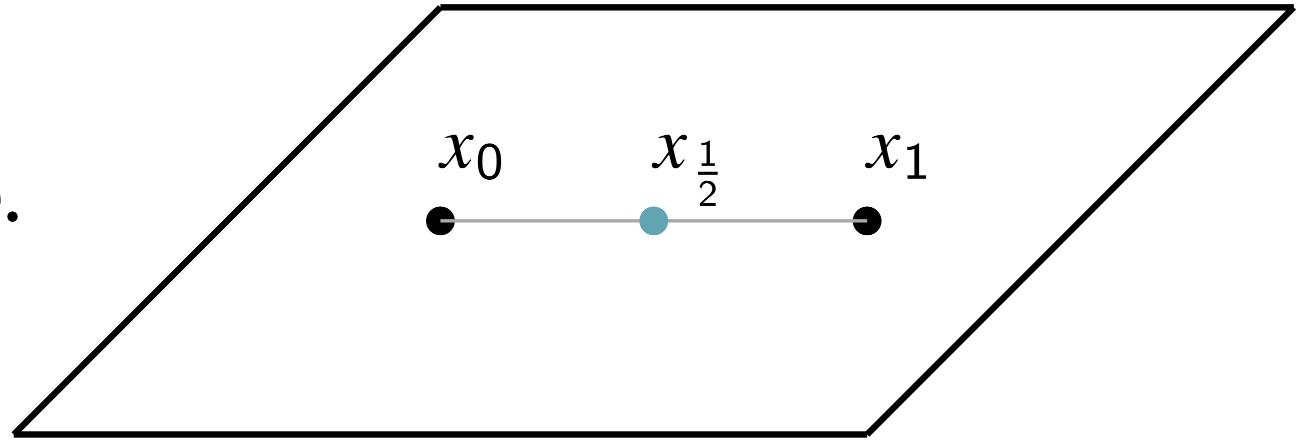
$(\mathcal{S}, d)$ is said to be a *geodesic space* if for any given x_0, x_1 we can exhibit a geodesic.

for us, constant speed geodesic.

if $(\mathcal{S}, d)$ is a geodesic space, for any $x_0, x_1 \in \mathcal{S}$ we can define the *midpoint* as $\omega_{\frac{1}{2}}$, where $\omega_{\frac{1}{2}} = \omega(\frac{1}{2})$ and ω is a constant speed geodesic between x_0 and x_1.

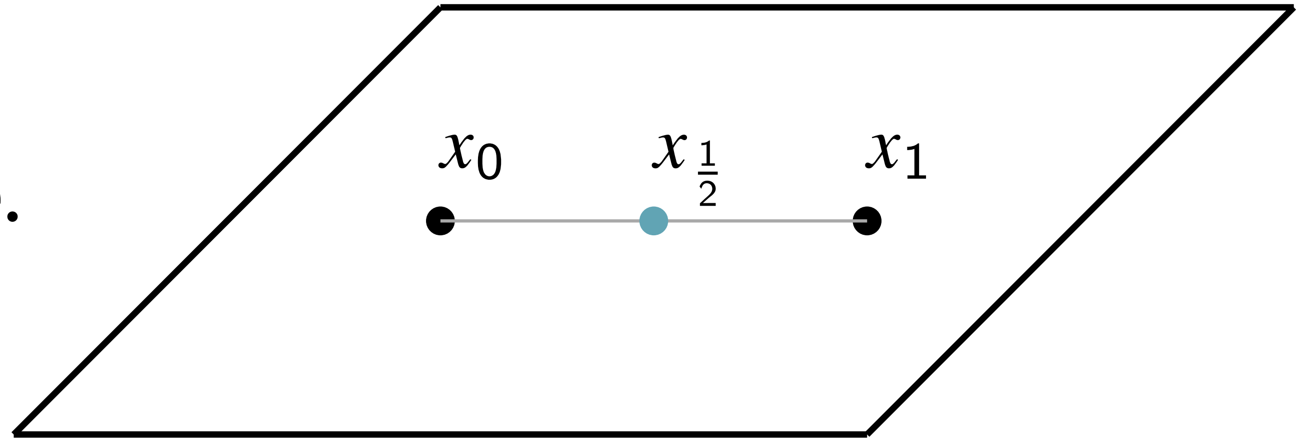
preliminaries. metric geometry.

$(\mathcal{S}, d) = (\mathbb{R}^2, \|\cdot\|_2)$ is a geodesic space.

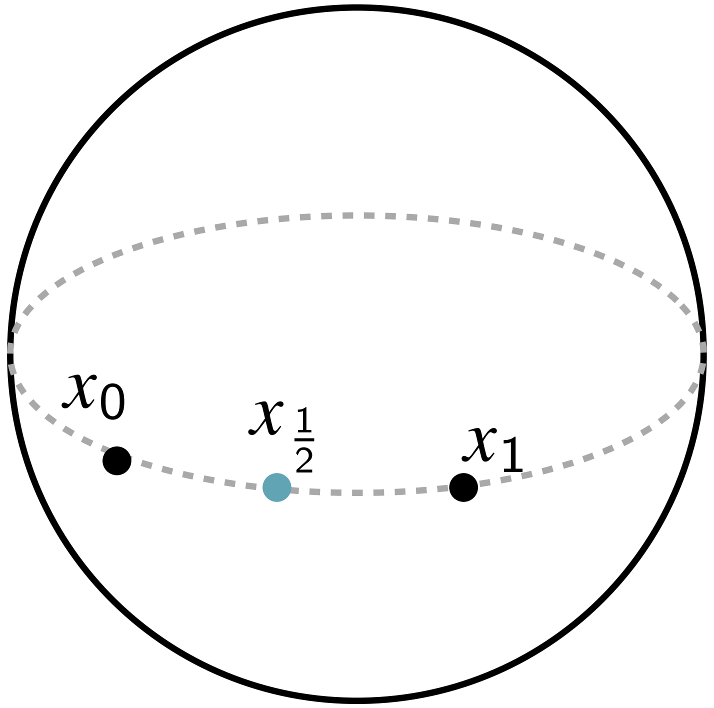


preliminaries. metric geometry.

$(\mathcal{S}, d) = (\mathbb{R}^2, \|\cdot\|_2)$ is a geodesic space.



$(\mathcal{S}, d) = \boxed{(\mathbb{S}^2, \|\cdot\|_2)}$ is a geodesic space.



$\|\cdot\|_2$ is induced by the inner product on tangent spaces of $\mathbb{S}^2$ induced by the Euclidean one.