

Soutenance mi-parcours PFE

Timothée Schmoderer

INSA de Rouen Normandie

14 décembre 2017

- 1 Introduction
 - Définitions
- 2 Reformulation du problème
- 3 Introduction aux opérateurs proximaux
- 4 Attaque numérique du problème
 - la plus simple
 - la plus brutale
 - la plus dure
- 5 Exemple
 - Exemple 1
 - Exemple 2
 - Exemple 3

Introduction

- ▶ Gaspard Monge 1871 - *Mémoire sur la théorie des déblais et des remblais*
- ▶ Leonid Kantorovitch 1942
- ▶ JD. Benamou et Y. Brenier 2000

Definition

Transport Soient 2 mesures de probabilités μ et ν sur \mathcal{R}^N .

Un transport est une application $T : \mathbb{R}^N \rightarrow \mathbb{R}^N$ qui envoie μ sur ν ,
càd. :

$$\forall B \in \mathcal{B}(\mathbb{R}^N) \quad \mu(T^{-1}(B)) = \nu(B) \quad (1)$$

C'est une relation de conservation de masse. On note $T_{\#}\mu = \nu$.

Supposons que $\mu = f_0 dx$, $\nu = f_1 dx$ et $T \in \mathcal{C}^1$ alors (1) devient :

$$\int_B f_1(y) dy = \int_{T^{-1}(B)} f_0(x) dx \quad (2)$$

$$= \int_B \sum_{x \in T^{-1}(y)} \left(\frac{f_0(x)}{|\nabla T(x)|} \right) dy \quad (3)$$

Définition : Problème de transport optimal

Notons $\mathcal{T}(f_0, f_1)$ l'ensemble des applications transport qui vérifient (3).

On se donne un coût : $C : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}^+$. Le problème de transport optimal est alors de trouver $T \in \mathcal{T}(f_0, f_1)$ qui réalise le

$$\min_{T \in \mathcal{T}(f_0, f_1)} \int C(x, T(x)) dx$$

Dans notre cas, $C(x, y) = \|x - y\|^2$. La valeur minimale est alors appelée distance L^2 de Wasserstein entre f_0 et f_1 .

Reformulation du problème

Théorème 2.1 : Benamou et Brenier

Soient deux densités de probabilités $f_0 dx$ et $f_1 dx$ alors

$$\min_{T \in \mathcal{T}(f_0, f_1)} \int \|x - T(x)\|^2 dx =$$

$$\min_{(\rho, v) \in C_v} \frac{1}{2} \int_{\mathbb{R}^N} \int_0^1 \rho(t, x) \|v(t, x)\|^2 dt dx$$

- ▶ $\rho : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}$ est la densité > 0 .
- ▶ $v : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ est un champ de vitesses.
- ▶ $C_v = \left\{ (\rho, v) \mid \partial_t \rho + \operatorname{div}_x(\rho v) = 0, \rho(0, \cdot) = f_0, \rho(1, \cdot) = f_1 \right\}$

On pose, $m = \rho v$ et :

$$J(t, x) = \begin{cases} \frac{\|m(t, x)\|^2}{\rho(t, x)} & \text{si } \rho(t, x) > 0 \\ 0 & \text{si } (\rho(t, x), m(t, x)) = (0, 0) \\ +\infty & \text{sinon} \end{cases}$$

De sorte que :

$$C_v = \{(\rho, mv) \mid \partial_t \rho + \operatorname{div}_x(m) = 0, \rho(0, \bullet) = f_0, \rho(1, \bullet) = f_1\}$$

On remarque que, si $\rho \rightarrow +\infty$ alors $J(t, x) \rightarrow 0$ donc la fonctionnelle n'est pas coercive ce qui rend l'existence minimiseurs non trivial. Et si $\rho \rightarrow 0$ alors $J(t, x) \rightarrow +\infty$ donc le gradient n'est pas lipschitz, ce qui nous empêche d'utiliser les algorithmes de minimisation classique comme (algorithme du gradient conjugué).

Introduction aux opérateurs proximaux

Dans ce qui suit, on se place en dimension 1 en espace.

Algorithme DR appliqué au problème de transport

Le problème se reformule sous la forme :

$$\min_{w=(m,\rho)\in\mathcal{H}} J(w) + \mathbf{i}_C(w)$$

Ainsi, étant donné $(z_0, w_0) \in \mathcal{H}^2$

$$\begin{cases} w_{n+1} &= w_n + \alpha (Prox_{\gamma J}(2z_n - w_n) - z_n) \\ z_{n+1} &= Prox_{\gamma \mathbf{i}_C}(w_{n+1}) \end{cases}$$

Grille

Discrétisation

On se donne deux entiers N et P . On divise le carré espace temps $[0, 1]^2$ en $(P + 1) \times (N + 1)$ points.

On note \mathcal{G}_c cette grille est w_{ij} la variable w discrétisée.

Attention : i parcourt l'axe spatial (les colonnes donc) et j l'axe temporel (les lignes).

Opérateur proximal de γJ

Proposition

$$\forall w \quad \text{Prox}_{\gamma J} = (\text{Prox}_{\gamma J}(w_{ij}))_{ij \in \mathcal{G}_c}$$

Où, pour tout $w_{ij} = (m_{ij}, f_{ij})$ on a :

$$\text{Prox}_{\gamma J} = \begin{cases} (\frac{f_{ij}^* m_{ij}}{f_{ij}^* + \gamma}, f_{ij}^*) & \text{si } f_{ij}^* > 0 \\ (0, 0) & \text{sinon} \end{cases}$$

Et, f_{ij}^* est la plus grande solution réelle du polynôme de degré 3 :

$$P(x) = (X - f_{ij})(X + \gamma)^2 - \frac{\gamma}{2} \|m_{ij}\|^2$$

Opérateur proximal de i_C

L'opérateur proximal de i_C est l'opérateur de projection sur l'ensemble convexe C . On reformule cette ensemble sous forme affine :

$$C = \{w \mid Aw = y\}$$

Où,

$$Aw = (\operatorname{div}(m) + \partial_t \rho, \rho(\bullet, 0), \rho(\bullet, 1)) \quad y = (0, f_0, f_1)$$

Finalement, on obtient :

$$\operatorname{Proj}_C(\bullet) = (Id - A^* \Delta^{-1} A) \bullet + A^* \Delta^{-1} y \quad \Delta^{-1} = (AA^*)^{-1}$$

Principe

Représentation des données sur la grille centrée. Méthode de gradient conjugué pour évaluer $Proj_C$.

Avantages

- ▶ rapidité de mise en oeuvre
- ▶ extensible facielement en dimension N

Inconvénients

- ▶ asymétrie de la dérivée

Exemple 1

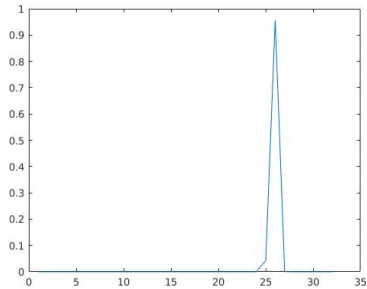


Figure – f_0

Exemple 1

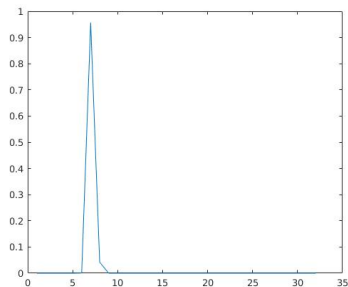


Figure – f_1

Exemple 1

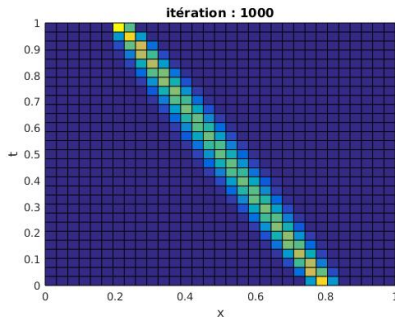


Figure – Transport f_0 sur f_1

Exemple 2

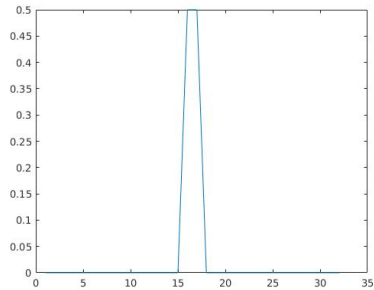


Figure – f_0

Exemple 2

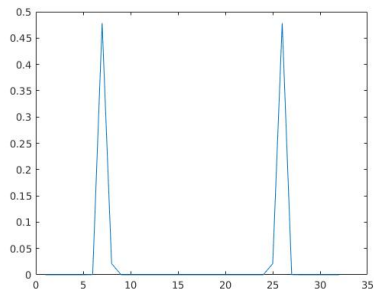


Figure – f_1

Exemple 2

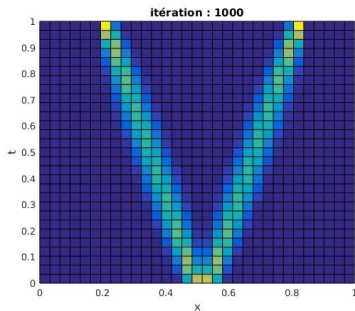


Figure – Transport f_0 sur f_1

Exemple 3

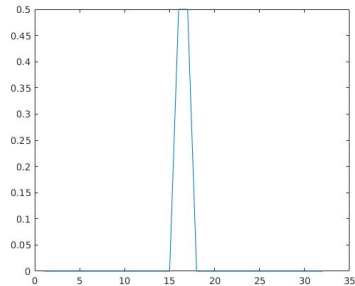


Figure – f_0

Exemple 3

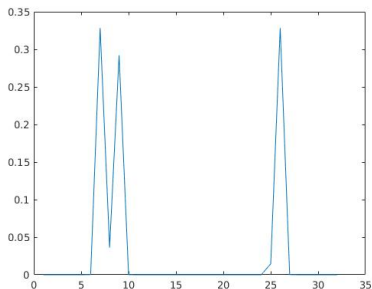


Figure – f_1

Exemple 3

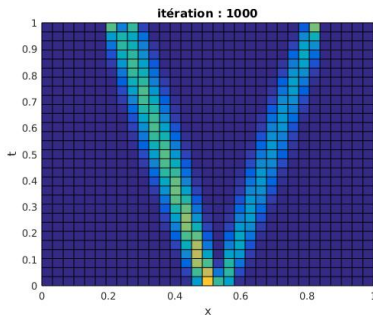


Figure – Transport f_0 sur f_1