Augmented Lagrangian methods for gradient flows

M. Laborde,
joint work with J-D. Benamou and G. Carlier.

Ceremade,
Paris-Dauphine University

SMAI-MODE2016, March 23, Toulouse

Plan

1. JKO scheme and PDEs
2. Augmented Lagrangian method
3. Porous medium equation
4. Interactions in the potential energy
5. Interactions in the internal energy

Table of contents

JKO scheme ([Jordan, Kinderlehrer and Otto,1998])

- We define $(\rho_h^k)_{k\in\mathbb{N}}$ by induction such that $\rho_h^0 = \rho_0$ and for all $k \in \mathbb{N}$,

$$\rho_h^{k+1} \in \underset{\rho\in\mathcal{P}(\Omega)}{\operatorname{argmin}} W_2^2(\rho, \rho_h^k) + 2h\mathcal{F}(\rho), \tag{1}$$

and

$$\rho_h(t,\cdot) := \rho_h^k \qquad \text{if } t \in ((k-1)h, kh].$$

JKO scheme ([Jordan, Kinderlehrer and Otto,1998])

- We define $(\rho_h^k)_{k\in\mathbb{N}}$ by induction such that $\rho_h^0 = \rho_0$ and for all $k \in \mathbb{N}$,

$$\rho_h^{k+1} \in \underset{\rho\in\mathcal{P}(\Omega)}{\text{argmin}}\, W_2^2(\rho, \rho_h^k) + 2h\mathcal{F}(\rho), \tag{1}$$

and

$$\rho_h(t,\cdot) := \rho_h^k \qquad \text{if } t \in ((k-1)h, kh].$$

- Under some asumptions on $\mathcal{F}$, $\rho_h \to \rho \in \mathcal{C}^{0,1/2}([0,T],\mathcal{P}(\Omega))$ and ρ is a weak solution, at least formally, of

$$\partial_t \rho + \text{div}(\rho\nabla\frac{\delta\mathcal{F}}{\delta\rho}) = 0, \qquad \rho_{|t=0} = \rho_0.$$

where $\frac{\delta\mathcal{F}}{\delta\rho}$ is the first variation of $\mathcal{F}$.

Examples

- If $\mathcal{F}(\rho) := \begin{cases} \int_\Omega \frac{1}{m-1}\rho^m & \text{if } \rho \ll \mathcal{L}, \\ +\infty & \text{otherwise,} \end{cases}$ for $m > 1$, then ρ solves

$$\partial_t \rho = \frac{m}{m-1} \operatorname{div}(\rho \nabla \rho^{m-1}) = \Delta \rho^m.$$

- If $\mathcal{V}(\rho) := \int_\Omega V(x)\, d\rho(x)$ then ρ solves

$$\partial_t \rho = \operatorname{div}(\rho \nabla V).$$

- If $\mathcal{W}(\rho) := \frac{1}{2} \int_{\Omega\times\Omega} W(x-y)\, d\rho(x) d\rho(y)$ then ρ solves

$$\partial_t \rho = \operatorname{div}(\rho \nabla W * \rho).$$

Table of contents

Benamou-Brenier formulation

- The Benamou-Brenier formula,

$$W_2^2(\rho,\nu) = \inf_{\mu_t, m_t} \int_0^1 \int_\Omega \frac{|m_t|^2}{\mu_t}\, dxdt,$$

subject to constraints that $\mu \geqslant 0$, $m = 0$ when $\mu = 0$ and

$$\partial_t \mu + \mathrm{div}(m) = 0,\ \mu_{|t=0} = \rho, \mu_{|t=1} = \nu.$$

Benamou-Brenier formulation

- The Benamou-Brenier formula,

$$W_2^2(\rho,\nu) = \inf_{\mu_t, m_t} \int_0^1 \int_\Omega \frac{|m_t|^2}{\mu_t}\, dxdt,$$

subject to constraints that $\mu \geqslant 0$, $m = 0$ when $\mu = 0$ and

$$\partial_t \mu + \mathrm{div}(m) = 0,\ \mu_{|t=0} = \rho, \mu_{|t=1} = \nu.$$

- As was observed in [Benamou,Brenier],

$$\Psi(\mu, m) = \begin{cases} \frac{|m|^2}{\mu} & \text{if } \mu \geqslant 0, \\ 0 & \text{if } \mu = 0, m = 0, \\ +\infty & \text{otherwise,} \end{cases}$$

is a convex, lsc, 1-homogenous function and can be rewrite as

$$\Psi(\mu, m) = \sup\left\{a\mu + b\cdot m \,:\, (a,b) \in K\right\},$$

where $K := \left\{(a,b) \in \mathbb{R}\times\mathbb{R}^n \,:\, a + \frac{1}{2}|b|^2 \leqslant 0\right\}$.

JKO scheme

- Let us consider one step of JKO where $\mathcal{F}$ is energy of the form

$$\mathcal{F}(\rho) = \int_\Omega F(\rho(x))\, dx + \int_\Omega V(x)\rho(x)\, dx.$$

JKO scheme

- Let us consider one step of JKO where $\mathcal{F}$ is energy of the form

$$\mathcal{F}(\rho) = \int_\Omega F(\rho(x))\,dx + \int_\Omega V(x)\rho(x)\,dx.$$

- Using the Benamou-Brenier formula, one step of JKO scheme can be rewrite as a convex minimization problem:

$$\inf_{\mu_t, m_t} \int_0^1 \int_\Omega \Psi(\mu_t, m_t)\,dxdt + 2h\mathcal{F}(\mu_1),$$

subject to constraints that

$$\partial_t \mu + \mathrm{div}(m) = 0,\ \mu_{|t=0} = \rho_h^k.$$

dual problem

- Then the dual formulation

$$\inf_{\Phi(t,x)} \left\{ \int_\Omega \Phi(0,x)\rho_h^k + \mathcal{F}^*(-\Phi(1,.)) \,:\, (\partial_t \Phi, \nabla \Phi) \in K \right\},$$

where

$$\mathcal{F}^*(c) := \sup_{\mu \geqslant 0} \left\{ \int_\Omega ((c(x) - hV(x))\mu(x) - hF(\mu(x)))\, dx \right\}.$$

dual problem

- Then the dual formulation

$$\inf_{\Phi(t,x)} \left\{ \int_\Omega \Phi(0,x)\rho_h^k + \mathcal{F}^*(-\Phi(1,.)) \, : \, (\partial_t \Phi, \nabla\Phi) \in K \right\},$$

where

$$\mathcal{F}^*(c) := \sup_{\mu \geqslant 0} \left\{ \int_\Omega ((c(x) - hV(x))\mu(x) - hF(\mu(x)))\, dx \right\}.$$

- Another formulation is

$$\inf_{\phi,q} \left\{ F(\Phi) + G(q) \, : \, \Lambda\Phi = q \right\},$$

where $\Lambda\Phi = (\partial_t\Phi, \nabla\Phi, -\Phi(1,\cdot))$, $q = (a,b,c) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$ and

$$F(\Phi) := \int_\Omega \Phi(0,\cdot)\rho_h^k, \; G(q) := \int_0^1 \chi_K(a,b)dxdt + h\mathcal{F}^*\left(\frac{c}{h}\right).$$

Relaxation problem and augmented Lagrangian method

- Then when we relax the problem, the JKO scheme is equivalent to find a saddle-point of the Lagrangian

$$L(\Phi, q, \sigma) := F(\Phi) + G(q) + \sigma \cdot (\Lambda\Phi - q),$$

where $\sigma := (\mu, m, \mu_1)$, $q := (a, b, c) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$,
$\Lambda\Phi := (\partial_t \Phi, \nabla\Phi, -\Phi(1, \cdot))$,

Relaxation problem and augmented Lagrangian method

- Then when we relax the problem, the JKO scheme is equivalent to find a saddle-point of the Lagrangian

$$L(\Phi, q, \sigma) := F(\Phi) + G(q) + \sigma \cdot (\Lambda\Phi - q),$$

where $\sigma := (\mu, m, \mu_1)$, $q := (a, b, c) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$,
$\Lambda\Phi := (\partial_t \Phi, \nabla\Phi, -\Phi(1, \cdot))$,

- Now for $r > 0$, we consider the augmented Lagrangian function

$$L_r(\Phi, q, \sigma) := F(\Phi) + G(q) + \sigma \cdot (\Lambda\Phi - q) + \frac{r}{2}|\Lambda\Phi - q|^2.$$

We note that being a saddle-point of L is equivalent to being a saddle-point of L_r.

Augmented Lagrangian algorithm ALG2 splitting scheme (1)

This algorithm consists, starting from (Φ^0, q^0, σ^0), to generate inductively a sequence as follows:

- **Step 1:** minimization w.r.t Φ:

$$\Phi^{n+1} := argmin\left\{F(\Phi) + \sigma^n \cdot \Lambda\Phi + \frac{r}{2}|\Lambda\Phi - q^n|^2\right\},$$

- **Step 2:** minimization w.r.t q:

$$q^{n+1} := argmin\left\{G(q) - \sigma^n \cdot q + \frac{r}{2}|\Lambda\Phi^{n+1} - q|^2\right\},$$

- **Step 3:** update the multiplier by the gradient ascent formula

$$\sigma^{n+1} = \sigma^n + r(\Lambda\Phi^{n+1} - q^{n+1}).$$

Augmented Lagrangian algorithm ALG2 splitting scheme (2)

- Step 1 corresponds to solve an elliptic problem

$$-\Delta_{t,x}\Phi^{n+1} = \mathrm{div}_{t,x}((\mu^n, m^n) - r(a^n, b^n)), \text{ in } (0,1)\times\Omega,$$

with the boundary conditions

$$\begin{aligned} r\partial_t\Phi^{n+1}(0,\cdot) &= \rho_h^k - \mu^n(0,\cdot) + ra^n(0,\cdot), \\ r(\partial_t\Phi^{n+1}(1,\cdot) + \Phi^{n+1}(1,\cdot)) &= \mu_1^n - \mu^n(1,\cdot) + r(a^n(1,\cdot) - c^n(\cdot)), \\ (r\nabla\Phi^{n+1} + m^n - rb^n)\cdot\nu &= 0, \text{ on } \partial\Omega. \end{aligned}$$

Augmented Lagrangian algorithm ALG2 splitting scheme (2)

- Step 1 corresponds to solve an elliptic problem

$$-\Delta_{t,x}\Phi^{n+1} = \mathrm{div}_{t,x}((\mu^n, m^n) - r(a^n, b^n)), \text{ in } (0,1)\times\Omega,$$

with the boundary conditions

$$\begin{aligned} r\partial_t\Phi^{n+1}(0,\cdot) &= \rho_h^k - \mu^n(0,\cdot) + ra^n(0,\cdot), \\ r(\partial_t\Phi^{n+1}(1,\cdot) + \Phi^{n+1}(1,\cdot)) &= \mu_1^n - \mu^n(1,\cdot) + r(a^n(1,\cdot) - c^n(\cdot)), \\ (r\nabla\Phi^{n+1} + m^n - rb^n)\cdot\nu &= 0, \text{ on } \partial\Omega. \end{aligned}$$

- Step 2 splits into two convex pointwise problems

$$(a^{n+1}(t,x), b^{n+1}(t,x)) = P_K\left(D_{t,x}\phi^{n+1}(t,x) + \frac{1}{r}(\mu^n(t,x), m^n(t,x))\right),$$

and

$$c^{n+1}(x) = argmin_{c\in\mathbb{R}}\left\{\frac{r}{2}|\Phi^{n+1}(1,x) - \frac{1}{r}\mu^n(1,x) + c|^2 + h\mathcal{F}^*\left(\frac{c}{h}\right)\right\}.$$

Table of contents

Porous media equation

- We want to solve, for $m > 1$,

$$\partial_t \rho - \Delta \rho^m - \mathrm{div}(\rho \nabla V) = 0,$$

Porous media equation

- We want to solve, for $m > 1$,

$$\partial_t \rho - \Delta \rho^m - \mathrm{div}(\rho \nabla V) = 0,$$

- In this case

$$\mathcal{F}^*(c) = \frac{1}{rh^{\frac{1}{m-1}}} \left(\frac{m}{m-1} \right)^{\frac{m-1}{m}} \int_\Omega \left((c(x) - hV(x))_+ \right)^{\frac{m-1}{m}} \, dx.$$

Then

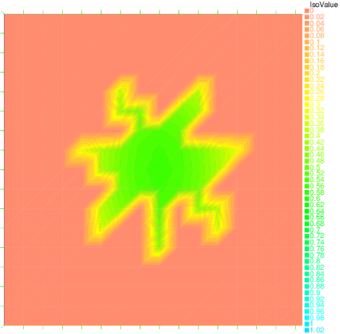
$$c^{n+1}(x) = \begin{cases} \frac{1}{r}\mu^n(1,x) - \Phi^{n+1}(1,x) & \text{if } \overline{c} \leqslant hV(x), \\ \text{the root in } (hV(x), +\infty) \text{ of } (2) & \text{otherwise,} \end{cases}$$

where (2) is the equation

$$c - \overline{c} + \frac{1}{rh^{\frac{1}{m-1}}} \left(\frac{m-1}{m} \right)^{\frac{1}{m-1}} (c - hV(x))^{\frac{1}{m-1}} = 0 \qquad (2)$$

Simulation

$$m = 3, \qquad V(x) = \frac{|x|^2}{2}.$$



The algorithm converges to the Barenblatt profile

$$\rho_\infty = \frac{m-1}{2m}(1-|x|^2)_+^{1/(m-1)}.$$

Table of contents

System with nonlocal interactions

- We consider I species solving, for all $i \in \{1, \dots, I\}$,

$$\begin{cases} \partial_t \rho_i = \alpha_i \Delta \rho_i + \mathrm{div}(\rho_i \nabla V_i[\boldsymbol{\rho}]) \\ \rho_{i|t=0} = \rho_{i,0}, \end{cases}$$

  with no-flux boundary condition, and where $\boldsymbol{\rho} := (\rho_1, \dots, \rho_I)$.

System with nonlocal interactions

- We consider I species solving, for all $i \in \{1, \ldots, I\}$,

$$\begin{cases} \partial_t \rho_i = \alpha_i \Delta \rho_i + \mathrm{div}(\rho_i \nabla V_i[\boldsymbol{\rho}]) \\ \rho_{i|t=0} = \rho_{i,0}, \end{cases}$$

  with no-flux boundary condition, and where $\boldsymbol{\rho} := (\rho_1, \ldots, \rho_I)$.
- The proof of existence is based on a semi-implicit JKO scheme (defined by [DiFrancesco and Fagioli]): we construct by induction $(\rho_{i,h}^k)$ such that

$$\rho_{i,h}^{k+1} \in \mathrm{argmin}\, W_2^2(\rho, \rho_{i,h}^k) + 2h \left(\alpha_i \int_\Omega \rho \log \rho + \int_\Omega V_i[\boldsymbol{\rho}_h^k] \rho \right).$$

Proximal

- We note $Prox_{\mathcal{F}}(\overline{c})$ the solution of

$$\inf_{c\in\mathbb{R}} \left\{ \frac{r}{2}|c-\overline{c}|^2 + \mathcal{F}(c) \right\}.$$

Moreover, $Prox_{\mathcal{F}^*}(\overline{c}) = \overline{c} - Prox_{\mathcal{F}}(\overline{c})$.

Proximal

- We note $Prox_{\mathcal{F}}(\overline{c})$ the solution of

$$\inf_{c \in \mathbb{R}} \left\{ \frac{r}{2} |c - \overline{c}|^2 + \mathcal{F}(c) \right\}.$$

  Moreover, $Prox_{\mathcal{F}^*}(\overline{c}) = \overline{c} - Prox_{\mathcal{F}}(\overline{c})$.
- Then $c_i^{n+1}(x) = \overline{c}_i - \tilde{c}_i$, where

$$\overline{c}_i = \frac{1}{r} \mu_i^n(1, x) - \Phi_i^{n+1}(1, x),$$

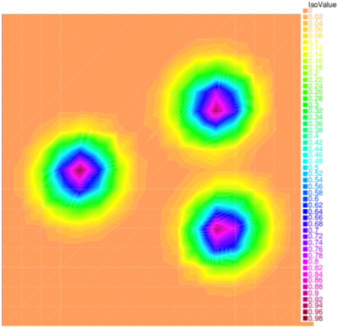
  and $\tilde{c}_i$ is the root of

$$\tilde{c}_i - \overline{c}_i + h V_i[\boldsymbol{\rho}_h^k] + h \log(\tilde{c}_i) = 0.$$

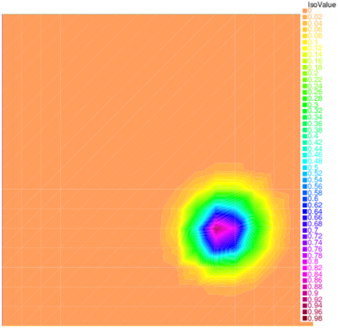
Simulations for 3 species

$$V_1[\rho_1, \rho_2, \rho_3] = |x|^2 * \rho_2 - |x|^2 * \rho_3, \; V_2[\rho_1, \rho_2, \rho_3] = |x|^2 * \rho_3 - |x|^2 * \rho_1,$$

and $V_3[\rho_1, \rho_2, \rho_3] = |x|^2 * \rho_1 - |x|^2 * \rho_2$.



$\rho_1 + \rho_2 + \rho_3$



ρ_1

Table of contents

Crowd motion

- We consider two species solving

$$\begin{cases} \partial_t \rho_1 - \Delta \rho_1 - \mathrm{div}(\rho_1(\nabla V_1 + \nabla p)) = 0 \\ \partial_t \rho_2 - \Delta \rho_2 - \mathrm{div}(\rho_2(\nabla V_2 + \nabla p)) = 0 \\ p \geqslant 0,\ \rho_1 + \rho_2 \leqslant 1,\ p(1 - \rho_1 - \rho_2) = 0, \\ \rho_{1|t=0} = \rho_{1,0}, \qquad \rho_{2|t=0} = \rho_{2,0} \end{cases}$$

with no-flux boundary conditions (see for one density [Mészá ros,Santambrogio] and [Maury, Roudneff-Chupin, Santambrogio] without diffusion).

Crowd motion

- We consider two species solving

$$\begin{cases} \partial_t \rho_1 - \Delta \rho_1 - \operatorname{div}(\rho_1(\nabla V_1 + \nabla p)) = 0 \\ \partial_t \rho_2 - \Delta \rho_2 - \operatorname{div}(\rho_2(\nabla V_2 + \nabla p)) = 0 \\ p \geqslant 0, \ \rho_1 + \rho_2 \leqslant 1, \ p(1 - \rho_1 - \rho_2) = 0, \\ \rho_{1|t=0} = \rho_{1,0}, \qquad \rho_{2|t=0} = \rho_{2,0} \end{cases}$$

  with no-flux boundary conditions (see for one density [Mészá ros,Santambrogio] and [Maury, Roudneff-Chupin, Santambrogio] without diffusion).

- Gradient flow structure with

$$\mathcal{F}(\rho_1, \rho_2) := \sum_{i=1}^{2} \int_\Omega (\rho_i \log \rho_i + V_i \rho_i) + \chi_{[0,1]}(\rho_1 + \rho_2).$$

Proximal

Using, the formualtion $Prox_{\mathcal{F}^*}(\overline{c}) = \overline{c} - Prox_{\mathcal{F}}(\overline{c})$ we compute

$$\tilde{c}_i := \operatorname{argmin} \frac{1}{2}|c - \overline{c}_i|^2 + c(hV_i + h\log(c)).$$

- If $\tilde{c}_1 + \tilde{c}_2 \leqslant 1$, then

$$c_i^{n+1}(x) = \overline{c}_i - \tilde{c}_i,$$

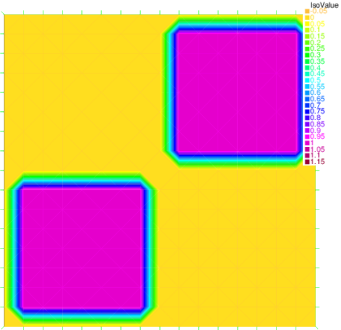
Proximal

Using, the formualtion $Prox_{\mathcal{F}^*}(\overline{c}) = \overline{c} - Prox_{\mathcal{F}}(\overline{c})$ we compute

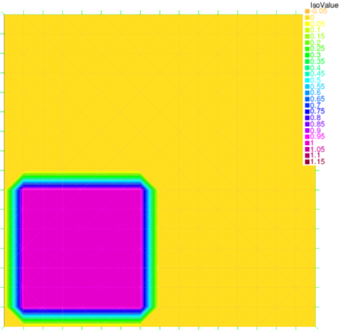
$$\tilde{c}_i := \operatorname{argmin} \frac{1}{2}|c - \overline{c}_i|^2 + c(hV_i + h\log(c)).$$

- If $\tilde{c}_1 + \tilde{c}_2 \leqslant 1$, then
$$c_i^{n+1}(x) = \overline{c}_i - \tilde{c}_i,$$
- Otherwise, the constraint is saturated then we write $\tilde{c}_1 = u$ and $\tilde{c}_2 = 1 - u$ where u minimizes
$$\frac{1}{2}\left(|u - \overline{c}_1|^2 + |1 - u - \overline{c}_2|^2\right) + h\left(u(\log(u) + V_1) + (1-u)(\log(1-u) + V_2)\right).$$
And
$$c_1^{n+1}(x) = \overline{c}_1 - u, \qquad c_2^{n+1}(x) = \overline{c}_2 - 1 + u.$$

Simulation



$\rho_1 + \rho_2$



ρ_1

Thank you for your attention

- L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Math., ETH Zurich, 2005.
- J.-D. Benamou, Y. Brenier,*A computational fluid mechanics solution to the Monge-Kantorovich mass transfert* problem. Numer. Math., 84 (3)(2000), 375-393.
- J.-D. Benamou, G. Carlier and M. L, *An augmented Lagrangian approach to Wasserstein gradient flows and applications*, accepted to publication in ESAIM.
- G. Carlier and M. L, On systems of continuity equations with nonlinear diffusion and nonlocal drifts, preprint 2015.
- R. Jordan, D. Kinderlehrer and F. Otto, *The Variational Formulation of the Fokker-Plank Equation*, SIAM J. of Math. Anal. 29 (1998), 1-17.
- M. Di Francesco and S. Fagioli, *Measure solutions for nonlocal interaction PDEs with two species*, 2013 Nonlinearity 26 2777.