Optimal control of the Fokker-Planck equation

Laurent Pfeiffer
Institute for Mathematics and Scientific Computing,
University of Graz

Introduction

- Goal: solving some stochastic optimal control problems.
- Specificity: cost function involves the probability distribution of the state variable
  $\rightarrow$ Mean-field-type optimal control.
- Tool: Fokker-Planck equation
- Applications: finance, resource management.

Fokker-Planck equation

Consider the stochastic differential equation (SDE):

$$\mathrm{d}X_s = f(X_s)\,\mathrm{d}s + \sigma(X_s)\,\mathrm{d}W_s, \quad X_0 = x_0.$$

with $f : \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^n$, $(W_s)_{s\geq 0}$ a Brownian motion, and x_0 a random variable in $\mathbb{R}^n$ with probability distribution m_0.

Let $m(s,\cdot) \in \mathcal{P}(\mathbb{R}^n)$ be the probability distribution of X_s:

$$\mathbb{P}[X_s \in \Omega] = \int_\Omega m(s,x)\,\mathrm{d}x, \quad \forall \Omega \subset \mathbb{R}^n.$$

It is a weak solution to the Fokker-Planck equation (FP):

$$\partial_s m = -\nabla \cdot (mf) + \frac{1}{2} \underbrace{\Delta(m\sigma\sigma^T)}_{\sum_{i,j=1}^n \partial^2_{x_i x_j}\left(m(s,\cdot)\sigma_i(\cdot)\sigma_j(\cdot)\right)} =: \mathcal{A}(m(s,\cdot)), \quad m(0,\cdot) = m_0.$$

Note: the operators $\nabla\cdot$, ∇, ∇^2, and Δ are spatial operators.

Feynman-Kac formula

Let $T > 0$, $\phi : \mathbb{R}^n \to \mathbb{R}$, $(X_s^{t,x})_{s\in[t,T]}$ the solution to the SDE with initial condition $X_t^{t,x} = x$. We define $\forall t \in [0, T]$, $x \in \mathbb{R}^n$:

$$V(t,x) = \mathbb{E}\big[\phi(X_T^{t,x})\big] = \int_{\mathbb{R}^n} \phi(x')m(T,x')\,\mathrm{d}x' = \langle \phi(\cdot), m(T,\cdot)\rangle,$$

where $m(s,\cdot)$ is the distribution associated with $X_s^{t,x}$.

The Feynman-Kac formula states that:

$$-\partial_t V(t,\cdot) = \mathcal{A}^*(V(t,\cdot)), \quad V(T,\cdot) = \phi(\cdot),$$

where $\mathcal{A}^*\big(V(t,\cdot)\big)(x) = \nabla V(t,x)^\top f(x) + \frac{1}{2}\sigma(x)^\top \nabla^2 V(t,x)\sigma(x)$.

It is usually derived by dynamic programming, using Itô's formula.

Formulation

Let U be a compact subset of $\mathbb{R}^m$, let $\mathcal{U}$ be the set of adapted control processes to $(W_s)_{s\geq 0}$. For all $t \in [0,T]$, $x \in \mathbb{R}^n$, $u \in \mathcal{U}$, let $(X_s^{t,x,u})_{s\in[t,T]}$ be the solution to:

$$\mathrm{d}X_s = f(X_s,u_s)\,\mathrm{d}s + \sigma(X_s,u_s)\,\mathrm{d}W_s, \quad X_t = x,$$

where $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \times U \to \mathbb{R}^n$ are given.

Remark: We could try to restrict ourselves to feedback controls, described by functions $\mathbf{u} : [0,T] \times \mathbb{R}^n \to U$, so that the SDE reads:

$$\mathrm{d}X_s = f(X_s,\mathbf{u}(s,X_s))\,\mathrm{d}s + \sigma(X_s,\mathbf{u}(s,X_s))\,\mathrm{d}W_s, \quad X_t = x,$$

But without regularity assumptions on $\mathbf{u}$, not a well-posed SDE!

Formulation

For all $u \in \mathcal{U}$, we denote by $m^{t,x,u}(s,\cdot)$ the probability distribution of $X_s^{t,x,u}$. We aim at solving:

$$\boxed{\min_{u\in\mathcal{U}} \chi\left(m^{0,x_0,u}(T,\cdot)\right) \qquad (P)}$$

where the cost $\chi : \mathcal{P}(\mathbb{R}^n) \to \mathbb{R}$ and the initial state x_0 are given.

Remark: attempt of a PDE-constrained problem formulation:

$$\min_{\mathbf{u}:[0,T]\times\mathbb{R}^n\to U} \chi(m(T,\cdot)), \quad \text{subject to:}$$

$$\begin{cases} \partial_t m(t,\cdot) = -\nabla \cdot \left[m(t,\cdot)f(\cdot,\mathbf{u}(t,\cdot))\right] + \frac{1}{2}\Delta\left[m(t,\cdot)\sigma\sigma^\top(\cdot,\mathbf{u}(t,\cdot))\right] \\ m(0,\cdot) = \delta_{x_0}. \end{cases}$$

But well-posedness of the Fokker-Planck equation is not ensured.

Formulation

Possible application: risk-averse optimization ($n = 1$).

- Penalization of the variance:

$$\chi(m) = \int_{\mathbb{R}} x m(x)\,\mathrm{d}x + \varepsilon \int_{\mathbb{R}} \Big(x - \int_{\mathbb{R}} y m(y)\,\mathrm{d}y \Big)^2 m(x)\,\mathrm{d}x.$$

- Conditional Value at Risk:

$$\mathrm{CVaR}_\beta = \frac{1}{1-\beta} \int_{\mathbb{R}} x \mathbf{1}_{x \geq \mathrm{VaR}_\beta} m(x)\,\mathrm{d}x$$

where: $\mathrm{VaR}_\beta = \sup \Big\{ z \in \mathbb{R} \,|\, \int_{\mathbb{R}} \mathbf{1}_{x \leq z} m(x)\,\mathrm{d}x \leq \beta \Big\}$.

Optimality conditions

For standard problems, χ is linear, i.e. $\exists \phi : \mathbb{R}^n \to \mathbb{R}$ such that:

$$\chi(m^{0,x_0,u}(T,\cdot)) = \langle m^{0,x_0,u}(T,\cdot), \phi(\cdot)\rangle = \mathbb{E}\left[\phi(X_T^{0,x_0,u})\right].$$

The corresponding problem is solved by dynamic programming.

$$\inf_{u\in\mathcal{U}} \langle m^{0,x_0,u}(T,\cdot), \phi(\cdot)\rangle \qquad (P(\phi))$$

Theorem

The value function: $V(t,x) = \inf_{u\in\mathcal{U}} \langle m^{t,x,u}(T,\cdot), \phi(\cdot)\rangle$ *is the solution to the Hamilton-Jacobi-Bellman (HJB) equation:*

$$-\partial_t V(t,x) = \inf_{u\in U} \left\{ \nabla V(t,x)^\top f(x,u) + \frac{1}{2} tr\left[\nabla^2 V(t,x)\sigma\sigma^\top(x,u)\right]\right\}$$

$$V(T,x) = \phi(x).$$

$\to$ *Provides a characterization of the optimal control.*

Optimality conditions

Assumption

1. *χ is continuous for the Wasserstein d_1-distance*
2. *χ is differentiable: $\forall m_1 \in \mathcal{P}(\mathbb{R}^n)$, $\exists D\chi(m_1,\cdot) \in C(\mathbb{R}^n,\mathbb{R})$ such that for all $m_2 \in \mathcal{P}(\mathbb{R}^n)$, for all $\theta \in [0,1]$,*

$$\chi\big((1-\theta)m_1 + \theta m_2\big) = \chi(m_1) + \theta\langle D\chi(m_1,\cdot), m_2 - m_1\rangle + o(\theta).$$

Theorem

If $\bar{u} \in \mathcal{U}$ is a solution to (P), then $\bar{u}$ is a solution to $P(D\chi(\bar{m},\cdot))$, where $\bar{m} = m^{0,x_0,\bar{u}}(T,\cdot)$.

Remarks:

- The associated value function $V(t,x)$ may be seen as a Lagrange multiplier for the Fokker-Planck equation.
- Motivates a fixed-point method.

Optimality conditions

Lemma

The closure (for the d_1-distance) of the set of reachable measures is convex:

$$cl(\{ m^{0,x_0,u}(T,\cdot) \,|\, u \in \mathcal{U} \}).$$

Remark: does not hold with feedback controls!

Proof of the theorem. Let $u \in \mathcal{U}$, $m = m^{0,x_0,u}(T,\cdot)$, $(\theta_k)_k \to 0$. For all k, let u^k be such that:

$$d_1(m^{0,x_0,u_k}(T,\cdot),(1-\theta_k)\bar{m} + \theta_k m)) = o(\theta_k).$$

Then,

$$\begin{aligned}\chi(\bar{m}) \leq \chi(m^k) &= \chi\big((1-\theta_k)\bar{m} + \theta_k m\big) + o(\theta_k) \\ &= \chi(\bar{m}) + \theta_k \underbrace{\langle D\chi(\bar{m}), m - \bar{m}\rangle}_{\leq 0} + o(\theta_k).\end{aligned}$$

Algorithm

- Set $k = 0$, choose $m^0 \in \mathcal{P}(\mathbb{R}^n)$, fix $\varepsilon > 0$.
- While $\chi(m^k) < \chi(m^{k-1}) - \varepsilon$, do:
  1. Backward phase (HJB): solve $P(D\chi(m^k))$, optimal sol.: u^k.
  2. Forward phase (FP): compute $m(\cdot) = m^{0,x_0,u^k}(T,\cdot)$.
  3. Solve: $\min_{\theta \in [0,1]} \chi(\theta m^k + (1-\theta)m)$, solution: θ^k.
     Set: $m^{k+1} = \theta^k m^k + (1-\theta^k)m$.
  4. Set $k = k + 1$.

Technical lemma: the closure (for the d_1-distance) of the set of reachable probability distributions is convex.

Remark: does not provide a feedback optimal solution, unless $\theta^k = 0$ at the last iteration.

Algorithm

Backward phase:

- Discretization of the SDE (Semi-Lagrangian scheme) with a controlled Markov chain
- Resolution of the HJB equation (discrete dynamic programming principle)

Forward phase:

- Resolution of the FP equation (adjoint equation to the Markov chain $\rightarrow$ Chapman-Kolmogorov equation.)

Remarks:

- Curse of dimensionality
- Computational effort in the backward phase.

Algorithm

Modified version to obtain a feedback solution:

- At the iteration k, we solve, for some penalization coefficient $\alpha > 0$:

$$\inf_{u \in \mathcal{U}} \mathbb{E}\Big[D\chi(m^k, X_T^{0,x_0,u}) + \alpha \int_0^T |u_t - u_t^{k-1}|^2 \,\mathrm{d}t \Big].$$

- If α is large enough, the (feedback) solution u^k and the measure $m^{k+1} = m^{0,x_0,u^k}(T,\cdot)$ associated satisfy: $\chi(m^{k+1}) \leq \chi(m^k)$.

- At iteration k, need a new loop to find a suitable value of α.

Results

Example considered:

- SDE: $dX_s = u_s ds + dW_s$, $X_0 = 0$, with final time 1.
- Controls: $u_s \in U = [-2, 2]$
- Cost: $\chi(m) = d_2(m, m_{\text{ref}})$, with: $m_{\text{ref}} = \frac{1}{3}(\delta_{-2} + \delta_0 + \delta_2)$.

Discretization:

- Semi-Lagrangian scheme
- 200×200 points in $[0, 1] \times [-10, 10]$, 100 points for the control

Convergence:

ε	Nb. iterations	Time	Cost
10^{-1}	18	11 s	0.5685
10^{-3}	62	36 s	0, 5199
10^{-5}	70	47 s	0, 5198

Results

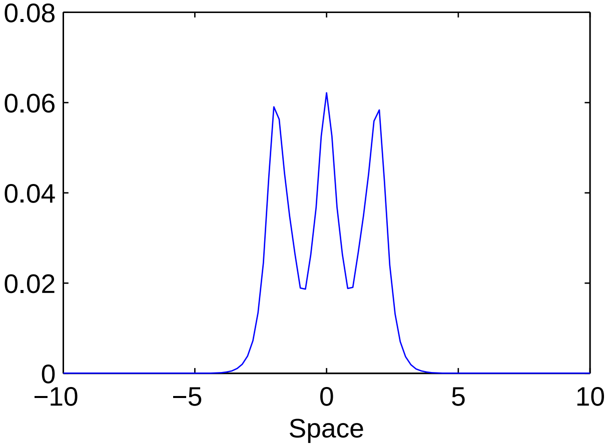

Figure: Final distribution

Results

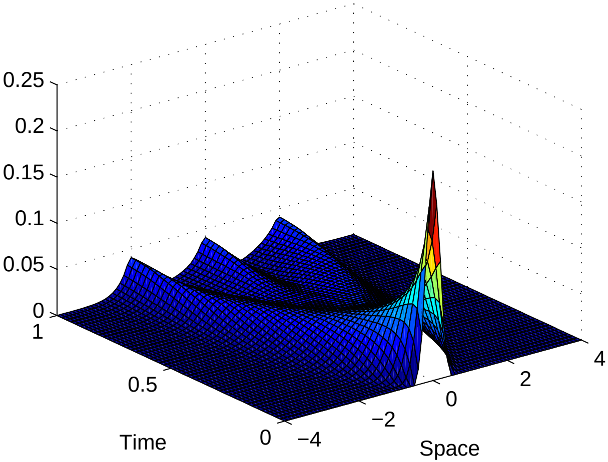

Figure: Distribution along time

Results

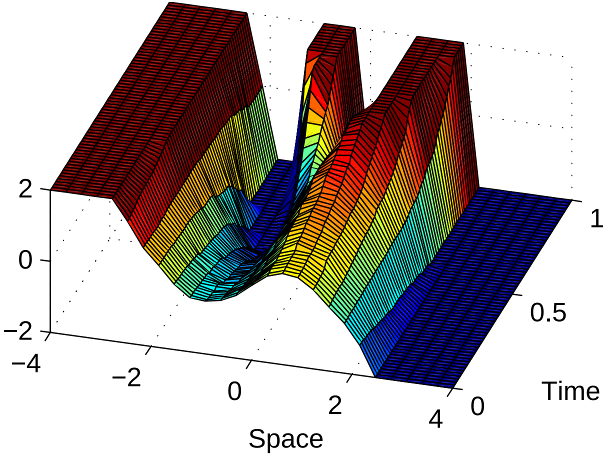

Figure: Control

Results

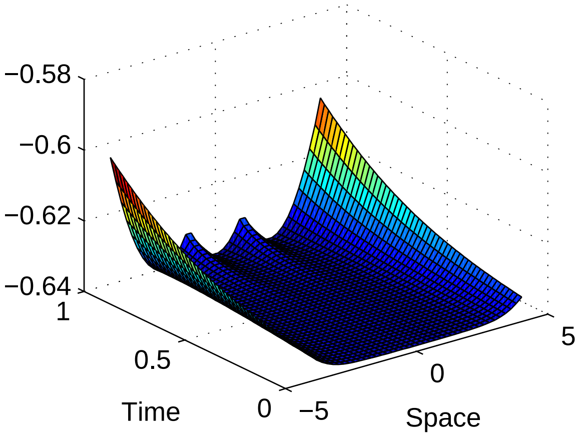

Figure: Value function

Bibliography

References:

A. Bensoussan, J. Frehse, and P. Yam. Mean-field games and mean-field type control theory. Springer, 2013.

M. Laurière and O. Pironneau. Dynamic programming for mean-field type control. CRAS, 2014.

L. Pfeiffer, Two approaches to constrained stochastic optimal control problems. *Preprint*, 2015.

Thank you for you attention.

Algorithm

Let $\phi : \mathbb{R}^n \to \mathbb{R}$, the usual approach to solve $P(\phi)$ is based on a consistent discretization of the SDE.

- Time discretization: $[0, \delta t, 2\delta t, ..., T = N\delta t]$
- Space discretisation $\{x_1, ..., x_k, ..., x_K\}$.
- Feedback control $u : (j, k) \in \{0, ..., N-1\} \times \{1, ..., K\} \to U$.
- Value function $V : (j, k) \in \{0, ..., N-1\} \times \{1, ..., K\} \to \mathbb{R}$.
- Prob. distribution $m : (j, k) \in \{0, ..., N-1\} \times \{1, ..., K\} \to \mathbb{R}_+$.

Discretization of the SDE with a Markov chain: for all time j, states x_k and x_l, control $v \in U$,

$$\mathbb{P}\left[X_{(j+1)\delta t} = x_l | X_{j\delta t} = x_k, u(j,k) = v\right] = P(k, l, v) \geq 0.$$

It holds: $\sum_{l=1}^{K} P(k, l, v) = 1$.

Algorithm

Dynamic programming for the value function:

$$V(j,k) = \inf_{v\in U}\left\{\sum_{l=1}^{K} P(k,l,v)V(j+1,l)\right\} \to \text{solution: } u(j,k),$$
$$V(N,k) = \phi(N\delta t, x_k).$$

This can be summarized as:

$$V(j,\cdot) = P_j V(j+1,\cdot), \quad \text{where } P_j = \big(P(k,l,u(j,k))\big)_{k,l=1,\dots,K}.$$

The matrix P_j is stochastic: $P_j e = e$, where $e = (1,\dots,1)^\top$.
The probability distribution is obtained from forwards:

$$m(j+1,\cdot) = P_j^\top m(j,\cdot), \quad m(0,\cdot) = \delta_{x_0}.$$

Observe that: $e^\top m(j+1,\cdot) = (P_j e)^\top m(j,\cdot) = e^\top m(j,\cdot) = \dots = 1.$