Approximation of optimal control

Jesper Carlsson
Mattias Sandberg
Anders Szepessy

GOAL

$$\frac{dX}{dt} = f(X_t, \alpha_t),$$

$$\inf_{\alpha \in \mathcal{A}} \Big(g(X_T) + \int_0^T h(X_s, \alpha_s) ds \Big),$$

given $X_0, f, g, h, \qquad \mathcal{A} = \{\alpha : [0, T] \to B\}.$

Ex: Inverse problem: Implied volatility estimation from option F

$$\partial_t F = \sigma \partial_{xx} F$$
$$\min_{\sigma} \int_0^T \int_{\mathbb{R}_+} (F - F_m)^2 dx dt.$$

Ex: Topology optimization: Find $\sigma : \Omega \to \{\sigma_-, \sigma_+\}$

$$\text{div}(\sigma \nabla \varphi(x)) = 0 \quad x \in \Omega, \qquad \sigma \frac{\partial \varphi}{\partial n}\Big|_{\partial\Omega} = I$$
$$\min_{\sigma} \Big(\int_{\partial\Omega} I \varphi ds + \eta \int_{\Omega} \sigma dx \Big).$$

TWO ALTERNATIVES

- Minimization with ODE constraint.

- Dynamic programming and Hamilton-Jacobi PDE.

ODE Motivation ($h = 0$)

Lagrange principle

$$\mathcal{L} = \int_0^T \lambda_t \cdot \Big(f(X_t, \alpha_t) - X_t' \Big)\, dt + g(X_T)$$

$$\begin{aligned}
\delta\lambda: &\quad X_t' = f(X_t, \alpha_t), \quad X_0 \text{ given}, \\
\delta X: &\quad -\lambda'(t)_i = \partial_{x_i} f(X_t, \alpha_t) \cdot \lambda_t, \quad \lambda_T = g'(X_T), \\
\delta\alpha: &\quad \alpha_t^* \in \operatorname{argmin}_{\alpha \in B}\ \lambda_t^* \cdot f(X_t^*, \alpha).
\end{aligned}$$

Euler approx. $\bar{X}_n, \bar{\lambda}_n$, Gradient method for α:

$$\min_{\bar{\alpha}} g(\bar{X}_T;\ \bar{X}(t_n) = x),$$

$$0 = \frac{\partial g}{\partial \bar{\alpha}_n} = \frac{\partial g}{\partial \bar{X}_{n+1}} \cdot \frac{\partial \bar{X}_{n+1}}{\partial \bar{\alpha}_n} = \bar{\lambda}_{n+1} \cdot \frac{\partial f}{\partial \bar{\alpha}_n} \Delta t,$$

$$\bar{\alpha}^{i+1} = \bar{\alpha}^i + \beta \bar{\lambda} \cdot \frac{\partial f}{\partial \bar{\alpha}^i} \Delta t.$$

PDE Motivation ($h = 0$)

Dynamic Programming

$$u(x,t) \equiv \inf_{\alpha} g(X_T;\ X_t = x)$$

$$u(\cdot, T) = g,$$

$$\frac{d}{dt}u(X_t^*, t) = 0 \implies \partial_t u + f^* \cdot \nabla u = 0,$$

$$g_1, g_3 \geq g_2 \implies \frac{d}{dt}u(X_t, t) \geq 0.$$

Therefore

$$\partial_t u(x,t) + \underbrace{\min_{\alpha \in B}\Big(f(x,\alpha)\cdot\nabla u(x,t)\Big)}_{H(\nabla u(x,t),x)} = 0, \quad t < T,$$

$$u(\cdot, T) = g.$$

APPROXIMATION

Hamilton-Jacobi by FEM or FD:

\+ Global minimum.

\− Not $d \gg 1$.

Minimization with ODE constraint:

\− Local minimum.

\+ $d \gg 1$!

PLAN

The talk presents

- Convergence for ODE method with
- non smooth control; two reasons
- regularization by consistency with Hamilton-Jacobi PDE and Pontryagin principle.

Pontryagin Principle and the Symplectic Euler

$$\bar{X}_{n+1} = \bar{X}_n + \Delta t \; H_\lambda(\bar{\lambda}_{n+1}, \bar{X}_n), \quad \bar{X}_0 = X(0),$$

$$\bar{\lambda}_n = \bar{\lambda}_{n+1} + \Delta t \; H_x(\bar{\lambda}_{n+1}, \bar{X}_n), \quad \bar{\lambda}_N = g'(\bar{X}_N).$$

Motivation

These Hamilton equations are the bi-characteristics for the Hamilton-Jacobi PDE and

$$f\Big(X_s, \alpha^*(X_s, \lambda_s)\Big) = H_\lambda(\lambda_s, X_s)$$
$$h = H - \lambda H_\lambda$$

Non Smooth Control

Two reasons:

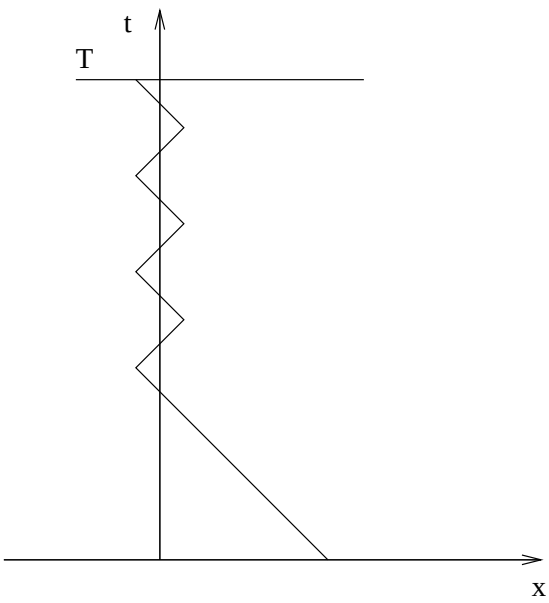
- H only Lipschitz

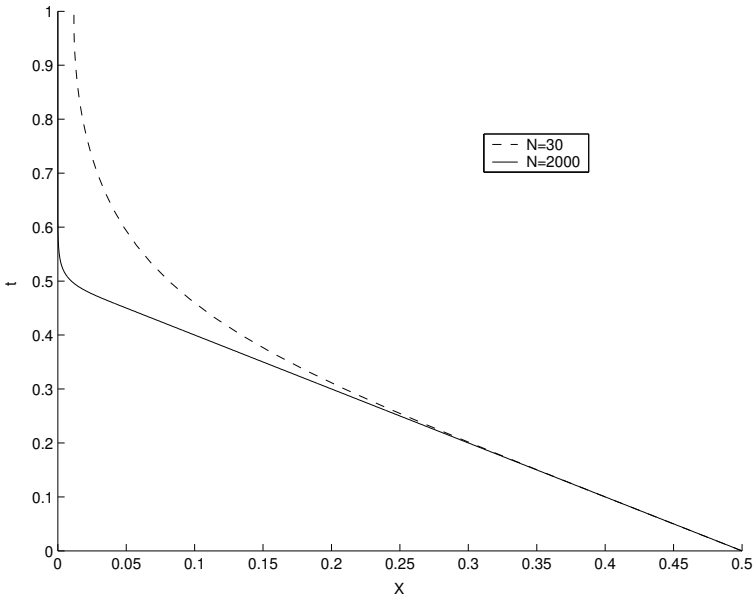
- Colliding backward paths X

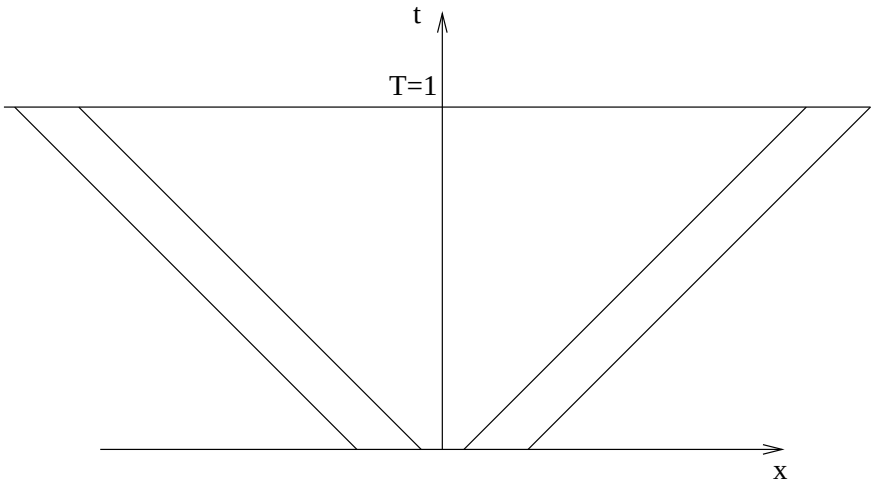
Regularized H^δ:

Ex: $f = \alpha \implies H(\lambda) = \min_{\alpha \in [-1,1]} (\lambda\, \alpha) = -|\lambda|.$

Examples


t
T
x


N=30
N=2000
t
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.05
0.15
0.25
0.35
0.45


t
T=1
x

Convergence for $d \gg 1$

Theorem. Assume f, g, h Lipschitz,

$$\|\partial_{\bar{X}(t_n)} \bar{\lambda}_{n+1}\|_{L^\infty(\Omega_- \cup \Omega_+)} \leq K,$$

then

$$|u - \bar{u}| = \mathcal{O}(\delta + \Delta t + \Delta t^2/\delta).$$

Proof.

0. $\bar{u}(x, t_n) = \min_{\bar{X}_n = x} \Big(g(\bar{X}_N) + \sum_{m \geq n} h(\bar{X}_m, \bar{\lambda}_{m+1}) \Delta t \Big)$

1. $\partial_{\bar{X}_n} \bar{u} = \bar{\lambda}_n$

2. $\partial_t \bar{u}(\bar{X}_t, t) + H(\partial_x \bar{u}, \bar{X}_t) = \mathcal{O}(\delta + \Delta t + \frac{(\Delta t)^2}{\delta})$,

3. 1 & 2 and maximum principle for viscosity solutions

4. H concave, subdifferential empty at shocks.

Topology Optimization

Find $\sigma : \Omega \to \{\sigma_-, \sigma_+\}$

$$\operatorname{div}\Big(\sigma \nabla \varphi(x)\Big) = 0 \quad x \in \Omega, \qquad \sigma \frac{\partial \varphi}{\partial n}\Big|_{\partial\Omega} = I$$

$$\min_{\sigma}\Big(\int_{\partial\Omega} I\varphi ds + \eta \int_{\Omega} \sigma dx\Big)$$

$$\mathcal{L} = \int_{\partial\Omega} I(\varphi + \lambda) ds + \int_{\Omega} \sigma \underbrace{(\eta - \nabla\varphi \cdot \nabla\lambda)}_{v} dx$$

$$H = \min_{\sigma} \int_{\Omega} \sigma v dx + \ldots = \int_{\Omega} \underbrace{\min_{\sigma} \sigma v}_{s(v)} dx + \ldots$$

Regularized:

$$H^{\delta}(\lambda, \varphi) = \int_{\Omega} s_{\delta}(\eta - \nabla\varphi \cdot \nabla\lambda) dx + \int_{\partial\Omega} I(\varphi + \lambda) ds$$

Symmetry implies $\varphi = \lambda$ and the Pontryagin principle yields

$$\mathrm{div}\Big(s'_\delta(\eta - |\nabla\varphi|^2)\nabla\varphi(x)\Big) = 0 \;\; x \in \Omega, \qquad s'_\delta \frac{\partial\varphi}{\partial n}|_{\partial\Omega} = I.$$

Solved by FEM and the Newton-method with decreasing δ or by nonlinear multigrid with two tricks:

- adaptive damping in smoothing steps; help for $s'_\delta \simeq 0$,

- increasing number of smoothing steps on coarser meshes; to improve the corrections.

Consistency with Infinite Dimensional H-J eq.

$$u(\varphi,\tau)=\int_{\tau}^{T}(\int_{\partial\Omega}I\varphi ds+\eta\int_{\Omega}\sigma dx)dt$$

$$\int_{\Omega}\Big(\partial_t\varphi w+s'(\eta-\nabla\varphi\cdot\nabla\lambda)\nabla\varphi\cdot\nabla w\Big)dx=\int_{\partial\Omega}Iwds$$
$$\int_{\Omega}\Big(-\partial_t\lambda w+s'(\eta-\nabla\varphi\cdot\nabla\lambda)\nabla\lambda\cdot\nabla w\Big)dx=\int_{\partial\Omega}Iwds$$

$$\lim_{T\to\infty}\frac{u(\varphi,t)-\bar{u}(\varphi,t)}{T}=\mathcal{O}(\delta+\Delta x)$$

Using the Gateaux derivative $\partial_\varphi u=\lambda$ and

$$\partial_t u+H(\partial_\varphi u,\varphi)=0$$

Implied volatility estimation from option F

$$\partial_t F = \sigma \partial_{xx} F$$
$$\min_{\sigma} \int_0^T \int_{\mathbb{R}} (F - F_m)^2 dx dt$$

Discrete:

$$\begin{aligned} H &= \min_{\sigma} \left(\sigma D^2 F \cdot \lambda + |F - F_m|^2 \right) \\ &= \sum_i \min_{\sigma_i} \sigma_i (D^2 F)_i \lambda_i + |F - F_m|^2 \\ &= \sum_i s\Big((D^2 F)_i \lambda_i \Big) + |F - F_m|^2 \end{aligned}$$

implies

$$\begin{aligned} \partial_t F &= s'(\lambda D^2 F) D^2 F \\ \partial_t \lambda &= -D^2 \Big(s'(\lambda D^2 F) \lambda \Big) - 2(F - F_m) \end{aligned}$$