Thesis Presentation

Methods for Stochastic Optimal Control under State Constraints

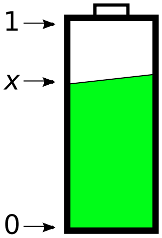
Per Rutquist

Chalmers University of Technology

Syscop Retreat, September 2017

Motivating Problem

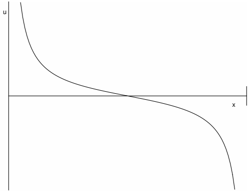
- Buffer control problem
- Minimize $E\{u^2\}$, subject to $\mathrm{d}x = u\mathrm{d}t + \mathrm{d}w$
- $x \in \Omega \subset \mathbb{R}$ (state), $u \in \mathbb{R}$ (control), $w \in \mathbb{R}$ (noise)
- $\Omega = \{ x \mid 0 \leq x \leq 1 \}$ (state constraints)



Motivating Problem
Solution

- Optimal control policy is known *exactly*. (Paper I)
- (re-scaled) trigonometric tangent function,

$$u^*(x) = -a\tan(bx + c)$$



- Optimal control actuation proportional to noise variance.
- Noise free case: $u^*(x) = 0$

Noise must be accounted for!

A noise free model cannot be used for optimal control in this case.

The Hamilton-Jacobi-Bellman equation

$$-\frac{\partial V}{\partial t} = \min_{\mathbf{u}} \left\{ l(\mathbf{x}) + (\nabla V)(\mathbf{f} + \mathbf{G}\mathbf{u}) + \mathbf{u}^T Q \mathbf{u} + \frac{1}{2}\mathrm{tr}\left[(\nabla^T \nabla V)\mathbf{G}\mathbf{W}\mathbf{G}^T \right] \right\}$$

In one dimension (Paper I)

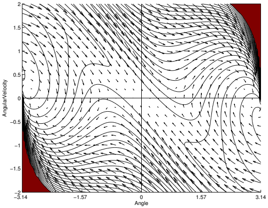
$$\nabla V = \kappa \frac{\nabla Z}{Z}$$

$\Rightarrow$ linear eigenvalue problem

In n dimensions (Paper III)

$$V = -2\kappa \log Z$$

$\Rightarrow$ linear eigenvalue problem
when control cost is $u^T Q u$
with $Q = \kappa W^{-1}$



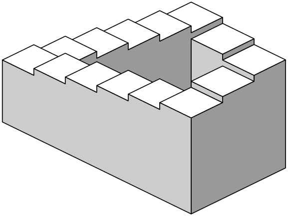
More generally (Paper V)

$$\nabla V = \frac{(\nabla Z)K}{Z}$$

does not require $Q = \kappa W^{-1}$

Caution!

Not every vector field is a gradient.



Finally (Paper VI)

Optimization under the Fokker Planck equation.

Fokker-Planck

The Fokker-Planck equation:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot J$$

where J is the "probability current"

$$J = -\rho f + \frac{1}{2}\nabla^T (W\rho)$$

Fokker-Planck vs. Hamilton-Jacobi-Bellman

- We can use optimization under the Fokker-Planck equation equation to deduce the Hamilton-Jacobi-Bellman equation.
- Example, minimize: $\int_\Omega \rho l(x,u,t)\mathrm{d}x$
- HJB is independent of the probability density ρ
  $\ldots = \min_u[l(x,u,t) + \nabla_x V(x,t)\cdot f(x,u,t)] + \ldots$
- FP has every term proportional to ρ
  $\ldots = \min_u \rho(x,t)[l(x,u,t) + \nabla_x V(x,t)\cdot f(x,u,t)] + \ldots$

HJB "forgets" about the $\rho = 0$ solution!

The HJB equation is ill-behaved close to state constraints, because it has been scaled by $1/\rho$.

Future Work

- Specialized solvers for this problem structure
- Adaptive meshes, radial-basis functions or neural networks.
- Exploring only the parts of the state-space with "nonzero" probability density