Stochastic optimal control with a probability constraint

PGMO Conference

L. Pfeiffer

Inria-Saclay and Ecole Polytechnique

Members of the PGMO project: M. Bernhart (EDF), F. Bonnans (Inria), N. Frikha (Paris 7), O. Klopfenstein (EDF), X. Tan (Paris 9)

October 4th, 2013

Introduction

We are interested in:

- a stochastic optimal control problem
- a probability constraint on the final state.

Motivations:

- it may be impossible to satisfy the constraint a.s.
- even if possible, it may be very expensive.

Our approach combines

- dynamic programming
- Lagrange relaxation.

Settings

- Time steps: $0, 1, \ldots, j, \ldots, T$.
- Random events: $\omega_1, \ldots, \omega_j, \ldots, \omega_T$. They are independent and to simplify, identically distributed, such that:

$$\forall i \in \{1, \ldots, I\}, \ \mathbb{P}[\omega_j = i] = p_i, \ \text{where } \sum_{i=1}^{I} p_i = 1.$$

For an initial time j and an initial state $x \in \mathbb{R}^n$, we define:

- Control space: $\mathcal{U}_j$, the set of adapted processes $u = (u_j, \ldots, u_{T-1})$ in a compact $U \subset \mathbb{R}^m$.

NB: u is adapted $\iff u_k =$ function of $\omega_{j+1}, \ldots, \omega_k$.

- State variable: $X^{j,x,u} = (X_j^{j,x,u}, \ldots, X_T^{j,x,u})$, solution to

$$X_{k+1}^{j,x,u} = F(X_k^{j,x,u}, u_k, \omega_{k+1}), \quad X_j^{j,x,u} = x.$$

Settings

To sum up, the decision process is:

$$x \to \text{decision of } u_j \to \text{discovery of } \omega_{j+1} \to X_{j+1}^{j,x,u} \to \ldots X_T^{j,x,u}.$$

Let ϕ and $g : \mathbb{R}^n \to \mathbb{R}$. For all j, $x \in \mathbb{R}^n$, $z \in \mathbb{R}$, we define the value function $V_j(x,z)$:

$$V_j(x,z) = \inf_{u \in \mathcal{U}_j} \mathbb{E}\left[\phi(X_T^{j,x,u})\right], \text{ such that } \underbrace{\mathbb{E}\left[g(X_T^{j,x,u})\right] \geq z}_{(\star)}.$$

If $g(x) = \mathbf{1}_K \circ h(x) = \begin{cases} 1 & \text{if } h(x) \in K \\ 0 & \text{otherwise} \end{cases}$, then

$$(\star) \iff \mathbb{P}\left[h(X_T^{j,x,u}) \in K\right] \geq z.$$

Dynamic programming

Proposition

Let j, x, z, let $u \in \mathcal{U}_j$. Then, $(\star)$ holds if and only if there exists a martingale $Z = (Z_j, Z_{j+1}, ..., Z_T)$ such that

$$Z_j = z \quad \text{and} \quad Z_T \leq g(X_T^{j,x,u}) \text{ almost surely.}$$

We say that Z is an associated martingale (to j, x, u and $(\star)$).

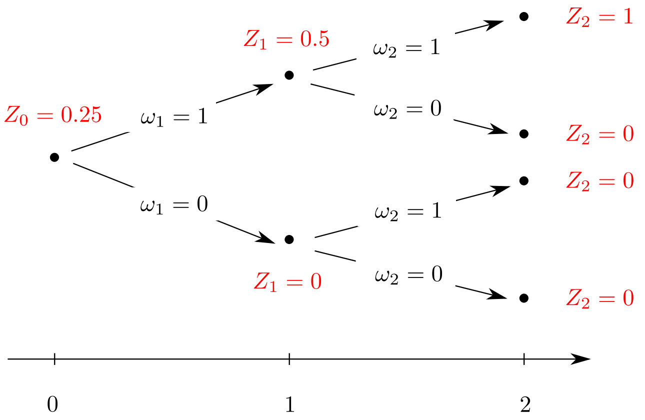
Proposition

The following dynamic programming equation holds: $\forall j, x, z$,

$$V_j(x,z) = \inf_{\substack{u \in U, (z_i)_i \in \mathbb{R}^I \\ s.t. \ \sum_{i=1}^I p_i z_i = z}} \left\{ \sum_{i=1}^I p_i V_{j+1}(F(x,u,i), z_i) \right\},$$

$$V_T(x,z) = \phi(x) \text{ if } z \leq g(x), \ +\infty \text{ otherwise.}$$

Dynamic programming



Example of a probability constraint of 0.25, where $p_1 = p_2 = 0.5$.

Dynamic programming

From a numerical point of view, several difficulties arise:

- The state variable (x, z) is $n+1$-dimensional $\rightarrow$ curse of dimensionality.
- How to discretize x ?
- How to discretize z ? Seems difficult when transition probabilities $(p_i)_i$ are not fractional.
- The value function is discontinuous w.r.t. z, and may be discontinuous w.r.t. x.
- The value function may be infinite for the highest values of z.

Boundary of the domain

The boundary of the domain of V is described as follows:

$$\mathcal{Z}_j(x) = \sup_{z \in \mathbb{R}} \{z,\ V_j(x,z) < +\infty\} = \sup_{u \in \mathcal{U}_j} \{\mathbb{E}[g(X_T^{j,x,u})]\}.$$

It can be computed by dynamic programming:

$$\mathcal{Z}_j(x) = \sup_{u \in U} \Big\{ \textstyle\sum_{i=1}^{I} p_i \mathcal{Z}_{j+1}(f(x,u,i)) \Big\}, \quad \mathcal{Z}_T(x) = g(x).$$

Proposition

Let j, x, let $u \in \mathcal{U}_j$ be such that $\mathbb{E}[g(X_T^{j,x,u})] = \mathcal{Z}_j(x)$. Let Z be an associated martingale.
Then, for all $k \in \{j, j+1, ..., T\}$, $Z_k = \mathcal{Z}_k(X_k^{j,x,u})$.

Thus, we can compute $V_j(x, \mathcal{Z}_j(x))$ (by dynamic programming).

Relaxation

Relaxation consists in playing with mixed strategies. An example:

$$\inf_{u \geq 0} \mathbb{E}[u] \text{ such that } \mathbb{P}[u=1] \geq z.$$

Then, $V(z) = \begin{cases} 1 & \text{if } z \in (0,1] \\ 0 & \text{if } z = 0. \end{cases}$

Now, let ω be a random value on $[0,1]$, of uniform law. A mixed strategy for the problem is:

$$u = \begin{cases} 1 & \text{if } \omega \in (0,z) \\ 0 & \text{if } \omega \in (z,1). \end{cases}$$

This strategy is feasible, and we deduce that $V^{\mathrm{r}}(z) = z$.

Relaxation

We use this idea for the problem under study. For all j, before playing u_j, we are allowed to observe a new random value ω'_j on $[0,1]$ of uniform law, independent of all the other random values.

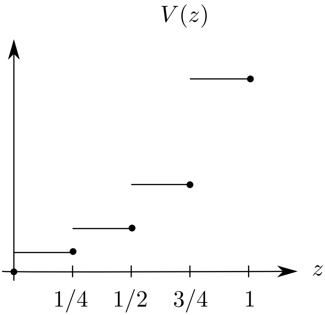
The space of mixed strategies (starting at j) is denoted $\mathcal{U}_j^r$.
We define the relaxed value function:

$$V_j^r(x,z) = \inf_{u \in \mathcal{U}_j^r} \mathbb{E}\left[\phi(X_T^{j,x,u})\right] \text{ s.t. } \left[g(X_T^{j,x,u})\right] \geq z.$$

Proposition

For all j, x, the relaxed value function $z \mapsto V_j^r(x,z)$ is the convex envelop of $z \mapsto V_j(x,z)$

Relaxation



Example of graphs for the value function and the relaxed value function.

An increasing subdifferential

We define the subdifferential (w.r.t. z):

$$\partial_z V_j^r(x,\bar{z}) = \{\lambda \in \mathbb{R} \,:\, \forall z,\ V_j(x,z) \geq V_j(x,\bar{z}) + \lambda(z-\bar{z})\}.$$

NB: if $z \mapsto V_j^r(x,z)$ is diff. at $\bar{z}$, then $\partial_z V_j^r(x,\bar{z}) = \{D_z V_j(x,\bar{z})\}$.

Theorem

Let j, x, z, let $u \in \mathcal{U}_j$ be an optimal control, let Z be an associated martingale. Let $\lambda \in \partial_z V_j(x,z)$, then,

$$\text{for all } k \geq j, \quad \lambda \in \partial V_k(X_k^{j,x,u}, Z_k) \text{ almost surely.}$$

In other words, the sensitivity of the value function w.r.t. to z is constant along optimal trajectories. This constant value may be considered as the Lagrange multiplier associated with $(\star)$.

Lagrange relaxation

Let us fix j, x, and $\bar{z}$. For all z, we have:

$$V_j(x,z) = \inf_{u\in\mathcal{U}_j} \sup_{\lambda\geq 0} \Big\{ \mathbb{E}\big[\phi(X_T^{j,x,u}) - \lambda(g(X_T^{j,x,u}) - z)\big] \Big\}. \qquad (P(z))$$

By Fenchel-Moreau-Rockafellar theorem,

$$V_j^r(x,z) = \sup_{\lambda\geq 0} \Big\{ \lambda z + \underbrace{\inf_{u\in\mathcal{U}_j} \mathbb{E}\big[\phi(X_T^{j,x,u}) - \lambda g(X_T^{j,x,u})\big]}_{\text{Problem } D(\lambda)} \Big\}.$$

Problem $D(\lambda)$ is an unconstrained problem, that can be solved by dynamic programming. Note that:

$$-\mathrm{Val}(D(\lambda)) = V_j^*(x,\lambda),$$

therefore, $\lambda \mapsto \mathrm{Val}(D(\lambda))$ is concave.

Lagrange relaxation

Proposition

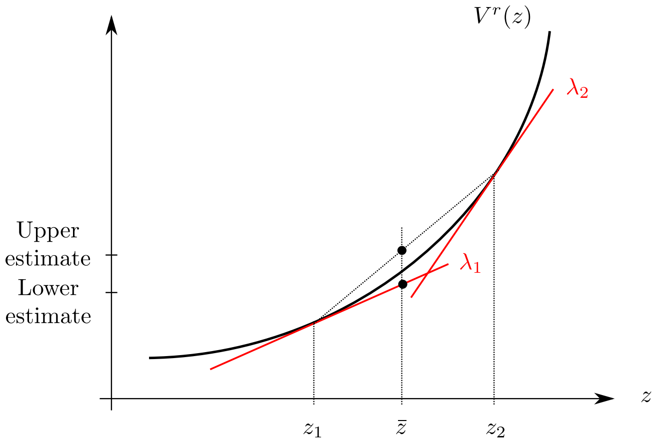
Let $0 \leq \lambda_1$. Let u_1 be an optimal solution to $D(\lambda_1)$. Then,

1. *The process u_1 is a solution to $P(z_1)$, where* $z_1 = \mathbb{E}\left[g(X_T^{j,x,u_1})\right]$.
2. *Let $\lambda_2 > \lambda_1$, u_2 a solution to $D(\lambda_2)$ and $z_2 = \mathbb{E}\left[g(X_T^{j,x,u_2})\right]$. Then, $z_2 \geq z_1$.*

A numerical method to solve $P(\bar{z})$:

1. Compute two values of λ, say $\lambda_1 < \lambda_2$, the associated u_1 and u_2, z_1 and z_2 and assume that $z_1 \leq \bar{z} \leq z_2$.
2. Let $\lambda = (\lambda_1 + \lambda_2)/2$. Compute the associated u and z.
3. If $\bar{z} \in [z_1, z]$, then go back to step 1 with λ_1 and λ, otherwise, use λ and λ_2.

Lagrange relaxation



Equation of the lines: $\mathrm{Val}(D(\lambda_i)) + \lambda_i z$.

Lagrange relaxation

The computation of $\mathrm{Val}(D(\lambda))$ is easier:

- the dimension of the state x is smaller: n,
- the value is finite.

Note that:

- the number of iterations on λ will be small,
- the method computes the relaxed value function.

Settings

Let $(W_t)_{t\in[0,T]}$ be a Brownian motion, let $t \in [0, T]$, let $\mathcal{U}_t$ be the set of adapted control processes in a given compact U.

Let $(X_s^{t,x,u})_{s\in[t,T]}$ be the solution to the following SDE:

$$\begin{cases} dX_s^{t,x,u} = & f(X_s^{t,x,u}, u_s)\,ds + \sigma(X_s^{t,x,u}, u_s)\,dW_s, \\ X_t^{t,x,u} = & x. \end{cases}$$

We consider:

$$V(t,x,z) = \inf_{u\in\mathcal{U}_t} \mathbb{E}\left[\phi(X_T^{t,x,u})\right] \text{ s.t. } \mathbb{E}\left[g(X_T^{t,x,u})\right] \geq z.$$

Settings

We assume that $\exists L > 0$ such that $\forall x, y \in \mathbb{R}^n$, $\forall u \in U$,

$$|f(x,u)| + |\sigma(x,u)| + |\phi(x)| \leq L(1 + |x|)$$
$$|f(x,u) - f(y,u)| + |\sigma(x,u) - \sigma(y,u)| + |\phi(x) - \phi(y)| \leq L|y - x|.$$

Theorem

Assume that g is Lipschitz. Then, for all t, for all $x \in \mathbb{R}^n$, $z \mapsto V(t,x,z)$ is convex.

NB: this does not take into account probability constraints.

Numerical method

For $\lambda \geq 0$, the dual problem $D(\lambda)$ reads:

$$\min_{u \in \mathcal{U}_t} \left\{ \mathbb{E}\left[\phi(X_T^{t,x,u}) - \lambda g(X_T^{t,x,u}) \right] \right\}.$$

We solve it with classical numerical methods, e.g.

- a semi-Lagrangian scheme
- a finite-difference scheme.

We compute an approximation of $V^{\mathrm{r}}(t,x,z)$ by computing:

$$\sup_{\lambda \in \Lambda} \left\{ \lambda z + \mathrm{Val}(D(\lambda)) \right\},$$

where Λ is a sampling of $\mathbb{R}_+$.

Numerical results

We work with a simple example:

$$dX_t = u_t\, dt + dW_t,$$

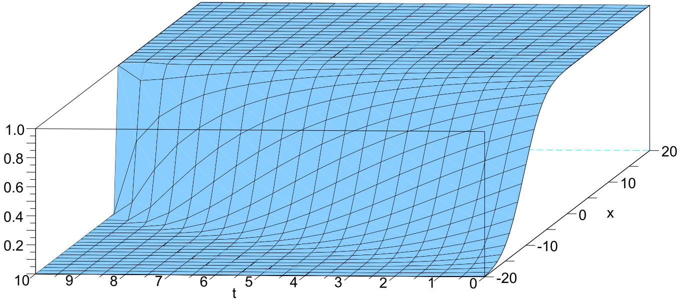
where $u_t \in [0,1]$. The problem is:

$$V(t,x,z) = \inf_{u \in \mathcal{U}_t} \left\{ \mathbb{E}\left[\int_t^T u_s^2\, ds \right] \right\} \text{ s.t. } \mathbb{P}\left[X_T^{t,x,u} \geq 0\right] \geq z.$$

Parameters for the discretization:

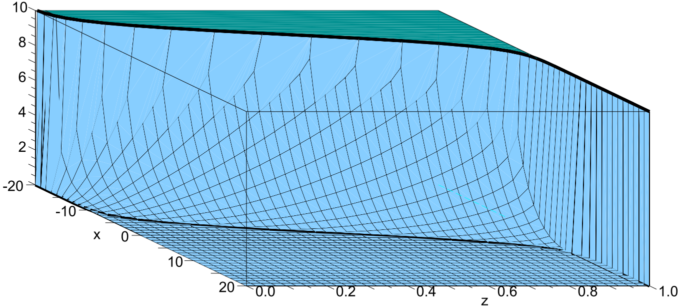
- Time steps: 20; $T = 10$
- Discretized control space: $\{0, 1/5, ...1\}$
- Number of space steps: 40, state space [-20,20]
- Probability steps: 40, $\Lambda = \{0, 1, ..., 100\}$.

Numerical results



Graph of the boundary (the maximum of probability)

Numerical results



Graph of the value function at time $t = 0$.

Bibliography I

L. Andrieu, G. Cohen, F.J. Vásquez-Abad,
Gradient-based simulation optimization under probability constraints.
EJOR, 2011.

B. Bouchard, R. Elie, C. Imbert,
Optimal control under stochastic target constraints,
SIAM J. on Control and Optimization, 2010.

J.B. Hiriart-Urruty, C. Lemaréchal,
Convex analysis and Minimization Algorithms,
Springer, 1993.

P. Carpentier, J.P. Chancelier, G. Cohen, M. De Lara, P. Girardeau,
Dynamic consistency for stochastic optimal control problems,
Annals of Operations Research, 2012.

Thank you for your attention!