Numerical Approximations for Minimax Markov Control Problems

Tomás Prieto-Rumeau. UNED, Madrid, Spain

Joint work with François Dufour. INRIA, Bordeaux, France

Stochastic control, computational methods, and applications
IMA, Minneapolis, May 2018

Motivation

- This talk is part of a "general approach" to obtain numerical approximations for control problems, developed by FD and TPR since 2010.
- The goal is to provide **explicit** numerical approximations for a large class of MDPs.
- ⚠ Most results in the literature address MDPs with discrete (sometimes, compact) state or action space.
- Our results cover
  - discrete-time and continuous-time discounted MDPs, constrained and unconstrained.
  - discrete-time and continuous-time average MDPs.
  - **games against nature**, zero sum-games.
- Our ideas are based on discrete approximations to "general probability measures": quantization, concentration inequalities, approximations in the Wasserstein metric...

Motivation

- This talk is part of a “general approach” to obtain numerical approximations for control problems, developed by FD and TPR since 2010.
- The goal is to provide **explicit** numerical approximations for a large class of MDPs.

⚠ Most results in the literature address MDPs with discrete (sometimes, compact) state or action space.

- Our results cover
  - discrete-time and continuous-time discounted MDPs, constrained and unconstrained.
  - discrete-time and continuous-time average MDPs.
  - **games against nature**, zero sum-games.
- Our ideas are based on discrete approximations to “general probability measures”: quantization, concentration inequalities, approximations in the Wasserstein metric...

Statement of the problem

- Approximate the value function and an optimal policy of a minimax Markov control process (a.k.a. robust control problem or game against nature).
- Control model: general state and action spaces X, A, B, discrete-time, discounted cost criterion.
- When using the DP approach to the infinite horizon discounted minimax problem, the Bellman equation reads:

$$v(x) = \min_{a \in A} \max_{b \in B} \left\{ c(x,a,b) + \alpha \int v(y) Q(dy|x,a,b) \right\}$$

for $x \in X$.
- The optimal value function is v. An action $a \in A$ attaining the min is optimal.
- It is not possible to solve, in general, this equation.

Statement of the problem

- Approximate the value function and an optimal policy of a minimax Markov control process (a.k.a. robust control problem or game against nature).
- Control model: general state and action spaces X, A, B, discrete-time, discounted cost criterion.
- When using the DP approach to the infinite horizon discounted minimax problem, the Bellman equation reads:

$$v(x) = \min_{a \in A} \max_{b \in B} \left\{ c(x,a,b) + \alpha \int v(y) Q(dy|x,a,b) \right\}$$

  for $x \in X$.
- The optimal value function is v. An action $a \in A$ attaining the min is optimal.
- It is not possible to solve, in general, this equation.

Motivation

- The idea is to discretize X, A, and B, so that

$$v(x) = \min_{a\in A} \max_{b\in B} \left\{ c(x,a,b) + \alpha \int v(y) Q(dy|x,a,b) \right\}, \ x \in X$$

becomes

$$v'(x) = \min_{a\in A'} \max_{b\in B'} \left\{ c(x,a,b) + \alpha \int v'(y) Q'(dy|x,a,b) \right\}, \ x \in X'.$$

- Basic questions:
  - How to discretize the elements of the control problem?
  - Is it possible to get bounds on $||v - v'||$?

Our approach

- Discretization of the action spaces A, B is made by finite approximations A', B' in the Hausdorff metric (a **geometric** criterion).
- Discretization of the state space is based on approximating probability measures Q by p.m.'s Q' with finite support (a **probabilistic** criterion).
- We show that

$$||v - v'|| \sim d_H(A, A') + d_H(B, B') + d(Q, Q').$$

- The choice of the Hausdorff metric is quite "natural". What about measuring the distance between two p.m.'s?

Our approach

- Discretization of the action spaces A, B is made by finite approximations A', B' in the Hausdorff metric (a **geometric** criterion).
- Discretization of the state space is based on approximating probability measures Q by p.m.'s Q' with finite support (a **probabilistic** criterion).
- We show that

$$||v - v'|| \sim d_H(A, A') + d_H(B, B') + d(Q, Q').$$

- The choice of the Hausdorff metric is quite "natural". What about measuring the distance between two p.m.'s?

The Wasserstein metric

- 1-*Wasserstein metric.* For probability measures in $\mathscr{P}_1(X)$ with finite first moment

$$W_1(\lambda,\mu) = \inf_{\{\nu:\nu_1=\lambda,\nu_2=\mu\}} \int_{X\times X} \rho(x_1,x_2)\nu(dx_1,dx_2).$$

- The dual Kantorovich-Rubinstein characterization gives

$$W_1(\lambda,\mu) = \sup_{f\in\mathbb{L}_1(X)} \left| \int f d\mu - \int f d\lambda \right|$$

  for all 1-Lipschitz continuous functions.

- $W_1(\mu_n,\mu) \to 0$ if and only if $\mu_n \Rightarrow \mu$ (weak convergence) and

$$\int_X \rho(x,x_0)\mu_n(dx) \to \int_X \rho(x,x_0)\mu(dx)$$

  for some (and then for all) $x_0 \in X$.

Why use the Wasserstein metric?

- Given a function f and a p.m. measure μ on a Polish space X, we want to approximate

$$\int_X f d\mu \quad \text{with} \quad \int_X f d\lambda$$

for some p.m. λ with finite support.

- If $d_{LP}(\mu,\lambda) < \epsilon$ and f is, e.g., bounded and continuous,

$$\left| \int_X f d\mu - \int_X f d\lambda \right| < ??$$

- If $W_1(\mu,\lambda) < \epsilon$ and f is L-Lipschitz continuous, then

$$\left| \int_X f d\mu - \int_X f d\lambda \right| < L\epsilon.$$

- This forces us to "live" in a Lipschitz continuous world

Why use the Wasserstein metric?

- Given a function f and a p.m. measure μ on a Polish space X, we want to approximate

$$\int_X f d\mu \quad \text{with} \quad \int_X f d\lambda$$

for some p.m. λ with finite support.

- If $d_{LP}(\mu,\lambda) < \epsilon$ and f is, e.g., bounded and continuous,

$$\left| \int_X f d\mu - \int_X f d\lambda \right| < ??$$

- If $W_1(\mu,\lambda) < \epsilon$ and f is L-Lipschitz continuous, then

$$\left| \int_X f d\mu - \int_X f d\lambda \right| < L\epsilon.$$

- This forces us to "live" in a Lipschitz continuous world

Why use the Wasserstein metric?

- Given a function f and a p.m. measure μ on a Polish space X, we want to approximate

$$\int_X f d\mu \quad \text{with} \quad \int_X f d\lambda$$

for some p.m. λ with finite support.

- If $d_{LP}(\mu,\lambda) < \epsilon$ and f is, e.g., bounded and continuous,

$$\left| \int_X f d\mu - \int_X f d\lambda \right| < ??$$

- If $W_1(\mu,\lambda) < \epsilon$ and f is L-Lipschitz continuous, then

$$\left| \int_X f d\mu - \int_X f d\lambda \right| < L\epsilon.$$

- This forces us to "live" in a Lipschitz continuous world

Approximation in the Wasserstein metric

Theorem

Let $\mu \in \mathscr{P}_1(X)$. For every $\epsilon > 0$ there exists a p.m. λ on X with finite support such that $W_1(\mu, \lambda) < \epsilon$.

The construction of λ is "theoretical" and it can pose a computational challenge.

For n i.i.d. draws with distribution μ, let μ_n be the empirical probability measure.

Theorem (Boissard, 2011)

Let $\mu \in \mathscr{P}_1(X)$ have a finite exponential moment. For each $\gamma > 0$ there exist $C_1, C_2 > 0$ such that

$$\mathbb{P}\{W_1(\mu_n, \mu) > \gamma\} \leq C_1 \exp\{-C_2 n\} \quad \text{for all } n \geq 1.$$

Approximation in the Wasserstein metric

Theorem

Let $\mu \in \mathscr{P}_1(X)$. For every $\epsilon > 0$ there exists a p.m. λ on X with finite support such that $W_1(\mu, \lambda) < \epsilon$.

The construction of λ is "theoretical" and it can pose a computational challenge.

For n i.i.d. draws with distribution μ, let μ_n be the empirical probability measure.

Theorem (Boissard, 2011)

Let $\mu \in \mathscr{P}_1(X)$ have a finite exponential moment. For each $\gamma > 0$ there exist $C_1, C_2 > 0$ such that

$$\mathbb{P}\{W_1(\mu_n, \mu) > \gamma\} \leq C_1 \exp\{-C_2 n\} \quad \text{for all } n \geq 1.$$

Dynamics of the control model

- Two players, the controller and the opponent (or nature), act on a stochastic dynamic system $\{x_k\}$.
  - The system is in state x_k,
  - The controller plays first and takes an action a_k,
  - The opponent observes the controller and then takes an action b_k,
  - The controller incurs a cost $c(x_k, a_k, b_k)$,
  - The system makes a transition $x_{k+1} \sim Q(\cdot|x_k, a_k, b_k)$.
- The game is played on an infinite time horizon $t \geq 0$ with costs discounted at a constant factor $0 < \alpha < 1$.
- The controller wants to optimize the "worst scenario"

$$\min_{\pi} \sup_{\gamma} E\Big[\sum_{t=0}^{\infty} \alpha^t c(x_t, a_t, b_t)\Big].$$

Dynamics of the control model

- Two players, the controller and the opponent (or nature), act on a stochastic dynamic system $\{x_k\}$.
  - The system is in state x_k,
  - The controller plays first and takes an action a_k,
  - The opponent observes the controller and then takes an action b_k,
  - The controller incurs a cost $c(x_k, a_k, b_k)$,
  - The system makes a transition $x_{k+1} \sim Q(\cdot | x_k, a_k, b_k)$.
- The game is played on an infinite time horizon $t \geq 0$ with costs discounted at a constant factor $0 < \alpha < 1$.
- The controller wants to optimize the "worst scenario"

$$\min_{\pi} \sup_{\gamma} E\Big[\sum_{t=0}^{\infty} \alpha^t c(x_t, a_t, b_t)\Big].$$

Definition of the control model

The control model $\mathcal{M}$

Consider a control model $\mathcal{M} = (X, A, B, \mathbb{K}_A, \mathbb{K}, Q, c)$ where

- A Borel state space X.
- Borel action spaces A and B.
- $A(x)$ is the set of available actions for the controller in state $x \in X$,

$$\mathbb{K}_A = \{(x,a) : x \in X, a \in A(x)\}.$$

- $B(x,a)$ is the set of available actions for the nature,

$$\mathbb{K} = \{(x,a,b), x \in X, a \in A(x), b \in B(x,a)\}.$$

- $Q \equiv Q(B|x,a,b)$ is a stochastic kernel on X given $\mathbb{K}$.
- $c : \mathbb{K} \to \mathbb{R}$ is the cost function.

Definition of the control model

- **Policies for the controller** Π_A. Based on the history

$$(x_0, a_0, b_0, \ldots, x_{t-1}, a_{t-1}, b_{t-1}, x_t)$$

the controller (randomly) chooses an action in $A(x_t)$.

- **Deterministic stationary policies** $\mathbb{F}_A$ are given by $f : X \to A$ such that $f(x) \in A(x)$ for $x \in X$.
- **Policies for the opponent** Π_B. Based on

$$(x_0, a_0, b_0, \ldots, x_{t-1}, a_{t-1}, b_{t-1}, x_t, a_t)$$

the opponent (randomly) chooses an action in $B(x_t, a_t)$.

- Given an initial state $x \in X$ and policies $(\pi, \gamma) \in \Pi_A \times \Pi_B$, there exists a unique p.m. $P^{\pi,\gamma,x}$ on $\mathbb{K}^\infty$ modeling the minimax problem.

Definition of the control model

Optimality criterion

- The total expected discounted cost is

$$V(x,\pi,\gamma) = E^{\pi,\gamma,x}\Big[\sum_{t=0}^{\infty} \alpha^t c(x_t,a_t,b_t)\Big].$$

- The minimax value function V^* is defined as

$$V^*(x) = \inf_{\pi\in\Pi_A} \sup_{\gamma\in\Pi_B} V(x,\pi,\gamma) \quad \text{for every } x\in X.$$

- A policy $\pi^*\in\Pi_A$ for the controller is a minimax policy if

$$\sup_{\gamma\in\Pi_B} V(x,\pi^*,\gamma) = V^*(x) \quad \text{for every } x\in X.$$

Hypotheses

- The action sets $A(x)$ and $B(x,a)$ are compact, and Lipschitz continuous; e.g.,

$$d_H(A(x),A(y)) \leq L_{\mathbb{A}} \cdot \rho(x,y) \quad \text{for all } x,y \in X,$$

- There exist a p.m. $\mu \in \mathscr{P}_1(X)$ and $q : X \times \mathbb{K} \to \mathbb{R}^+$ such that

$$Q(B|x,a,b) = \int_B q(y|x,a,b)\mu(dy)$$

  for all $B \in \mathscr{B}(X)$ and $(x,a,b) \in \mathbb{K}$.
- For some "weight function" $w : X \to [1,\infty)$

$$|c(x,a,b)| \leq \bar{c}w(x) \quad \text{for all } (x,a,b) \in \mathbb{K} \ .$$

- $Qw \leq \beta w$ with $\alpha\beta < 1$.

Hypotheses

Lipschitz continuity conditions

- The weight function w, the cost function c and the density q are Lipschitz continuous.
- The density function satifies
  - $q(y|x,a,b) \leq \overline{q}w(x)$.
  - $y \mapsto w(y)q(y|x,a,b)$ is $L_{wq}w(x)$-Lipschitz-continuous.

Some notation

- $u \in \mathbb{B}_w(X)$ if it is measurable and w-bounded:

$$|u(x)| \leq ||u||_w w(x) \quad \text{for } x \in X.$$

- $u \in \mathbb{L}_w(X)$ if it is Lipschitz-continuous and w-bounded.

Dynamic programming equation

- Given $u \in \mathbb{B}_w(X)$ define, for $x \in X$,

$$Tu(x) = \min_{a \in A(x)} \max_{b \in B(x,a)} \{c(x,a,b) + \alpha Qu(x,a,b)\}.$$

- The **dynamic programming** equation is

$$v(x) = Tv(x) \quad \text{for all } x \in X.$$

- Regularity of T:
  - If $u \in \mathbb{B}_w(X)$ then $Tu \in \mathbb{L}_w(X)$.
  - T is an $\alpha\beta$-contraction operator on $\mathbb{B}_w(X)$.

Dynamic programming equation

- Given $u \in \mathbb{B}_w(X)$ define, for $x \in X$,

$$Tu(x) = \min_{a \in A(x)} \max_{b \in B(x,a)} \{c(x,a,b) + \alpha Qu(x,a,b)\}.$$

- The **dynamic programming** equation is

$$v(x) = Tv(x) \quad \text{for all } x \in X.$$

- Regularity of T:
  - If $u \in \mathbb{B}_w(X)$ then $Tu \in \mathbb{L}_w(X)$.
  - T is an $\alpha\beta$-contraction operator on $\mathbb{B}_w(X)$.

Dynamic programming equation

Theorem

(i) The value function V^* is the unique solution in $\mathbb{B}_w(X)$ of the DP equation.

(ii) $f \in \mathbb{F}_A$ is a minimax policy if and only if it "attains the minimum" in the DP equation

$$V^*(x) = \max_{b \in B(x,f(x))} \{c(x,f(x),b) + \alpha QV^*(x,f(x),b)\}$$

for every $x \in X$, and such f indeed exists.

(iii) V^* is in $\mathbb{L}_w(X)$

$$L_{V^*} = \left(L_c + \frac{\alpha L_q \bar{c} \mu(w)}{1-\alpha\beta}\right) \cdot (1 + L_{\mathbb{A}}) \cdot (1 + L_{\mathbb{B}}).$$

Dynamic programming equation

Theorem

(i) The value function V^* is the unique solution in $\mathbb{B}_w(X)$ of the DP equation.

(ii) $f \in \mathbb{F}_A$ is a minimax policy if and only if it "attains the minimum" in the DP equation

$$V^*(x) = \max_{b \in B(x,f(x))} \{c(x,f(x),b) + \alpha QV^*(x,f(x),b)\}$$

for every $x \in X$, and such f indeed exists.

(iii) V^* is in $\mathbb{L}_w(X)$

$$L_{V^*} = \left(L_c + \frac{\alpha L_q \bar{c}\mu(w)}{1-\alpha\beta}\right) \cdot (1 + L_{\mathbb{A}}) \cdot (1 + L_{\mathbb{B}}).$$

Approximation of the control model

Discretization of the action space

For all small $\delta > 0$ construct finite set $A_\delta(x) \subseteq A(x)$ and $B_\delta(x,a) \subseteq B(x,a)$ with

- $x \mapsto A_\delta(x)$ and $(x,a) \mapsto B_\delta(x,a)$ are Lipschitz continuous.
- The finite action sets are close to the original action sets

$$d_H(A(x), A_\delta(x)) \le \delta w(x) \quad \text{and} \quad d_H(B(x,a), B_\delta(x,a)) \le \delta w(x).$$

- In this way, we obtain the feasible pairs $\mathbb{K}_A^\delta$ and triplets $\mathbb{K}^\delta$.

Approximation of the control model

Discretization of the state space

- Recall that the transition kernel is

$$Q(B|x,a,b)=\int_B q(y|x,a,b)\mu(dy).$$

- Replace $\mu\in\mathscr{P}_1(X)$ with $\nu\in\mathscr{P}_1(X)$ with finite support.
- Define the transition kernel

$$Q_\nu(B|x,a,b)=\frac{\int_B q(y|x,a,b)\nu(dy)}{\int_X q(y|x,a,b)\nu(dy)}.$$

Approximation of the control model

The approximating control model

- We approximate $\mathscr{M}$ with

$$\mathscr{M}_{\nu,\delta} = (X, A, B, \mathbb{K}_A^\delta, \mathbb{K}^\delta, Q_\nu, c).$$

- The discount factor is $\alpha > 0$.
- We can define policies for the controller and the opponent.
- The minimax value function $V^*_{\nu,\delta}$ is

$$V^*_{\nu,\delta}(x) = \inf_{\pi\in\Pi_A^\delta} \sup_{\gamma\in\Pi_B^\delta} V_{\nu,\delta}(x,\pi,\gamma) \quad \text{for every } x \in X.$$

- The definition of a minimax strategy is similar to that for $\mathscr{M}$.

Approximation of the control model

Solving $\mathscr{M}_{v,\delta}$

- The dynamic programming operator is, for $x \in X$,

$$T_{v,\delta} v(x) = \min_{a \in A_\delta(x)} \max_{b \in B_\delta(x,a)} \{c(x,a,b) + \alpha Q_v v(x,a,b).\}$$

- The operator $T_{v,\delta}$ inherits all properties from T except:
  - If $u \in \mathbb{B}_w(X)$ then $Tu \in \mathbb{L}_w(X)$
  - If $u \in \mathbb{B}_w(X)$ then $T_{v,\delta} u \in \mathbb{C}_w(X)$
- So, $\mathscr{M}_{v,\delta}$ is not a Lipschitz continuous control model!

Approximation of the control model

Solving $\mathscr{M}_{\nu,\delta}$

(i) the value function $V^*_{\nu,\delta}$ is the unique solution in $\mathbb{B}_w(X)$ of the DP equation $v = T_{\nu,\delta} v$.

(ii) The policy $f \in \mathbb{F}^\delta_A$ is a minimax policy if and only if

$$V^*_{\nu,\delta}(x) = \max_{b \in B_\delta(x,f(x))} \{c(x,f(x),b) + \alpha Q_\nu V^*_{\nu,\delta}(x,f(x),b)\}$$

for $x \in X$, and such f indeed exist.

Approximation of the value function

Theorem

There are constants $\mathbf{G}_1$ and $\mathbf{G}_2$ s.t. for every $\delta > 0$ and every $\nu \in \mathscr{P}_1(X)$ we have

$$\| V^* - V^*_{\nu,\delta} \|_w \le \mathbf{G}_1 \delta + \mathbf{G}_2 W_1(\mu,\nu).$$

The constants are $\mathbf{G}_1 = \frac{2(2+L_{\mathbb{B}})}{1-\alpha\beta}\Big(L_c + \frac{\alpha L_q \bar{c}}{1-\alpha\beta}\mu(w)\Big)$ and

$$\begin{aligned} \mathbf{G}_2 &= \frac{4\alpha\bar{c}}{(1-\alpha\beta)^2}(L_{wq} + \bar{q}L_w + L_q\beta) \\ &+ \frac{4\alpha\bar{q}}{(1-\alpha\beta)}\Big(L_c + \frac{\alpha L_q \bar{c}\mu(w)}{1-\alpha\beta}\Big)\cdot(1+L_{\mathbb{A}})\cdot(1+L_{\mathbb{B}}). \end{aligned}$$

Approximation of the value function

Sketch of the proof

- Compare the operators T and $T_{v,\delta}$ when applied to a Lipschitz continuous $u \in \mathbb{L}_w(X)$.
- Difficulty: double (nested) approximation of the actions sets.
- Then compare V^* and $V^*_{v,\delta}$ taking advantage that $V^* \in \mathbb{L}_w(X)$.

Approximation of a minimax strategy

Idea

- Find a minimax policy $f^*_{\nu,\delta} \in \mathbb{F}^\delta_A$ for the control problem $\mathcal{M}_{\nu,\delta}$.
- This policy is admissible for $\mathcal{M}$ because $\mathbb{F}^\delta_A \subseteq \mathbb{F}_A$.
- Plug this policy into the original model $\mathcal{M}$.
- Get upper bounds on

$$0 \le \sup_{\gamma \in \Pi_B} V(x, f^*_{\nu,\delta}, \gamma) - V^*(x) \le ??$$

Approximation of a minimax strategy

Theorem

There are constants $\mathbf{K}_1$ and $\mathbf{K}_2$ s.t. for every $\delta > 0$ and every $\nu \in \mathscr{P}_1(X)$ we have

$$||V^* - \sup_{\gamma \in \Pi_B} V(\cdot, f^*_{\nu,\delta}, \gamma)||_w \leq \mathbf{K}_1 \delta + \mathbf{K}_2 W_1(\mu, \nu).$$

Sketch of the proof:

- The value function $V^*_{\nu,\delta}$ is not Lipschitz continuous.
- We can prove that it is locally Lipschitz continuous.
- We can approximate $V^*_{\nu,\delta}$ in the w-norm with a Lipschitz continuous function.

An inventory management system

Consider the dynamics

$$x_{t+1} = \max\{x_t + a_t - \xi_t, 0\} \quad \text{for } t \in \mathbb{N}$$

where

- x_t is the stock level at the beginning of period t;
- a_t is the amount ordered at the beginning of period t;
- ξ_t is the random demand at the end of period t with distribution $F(\cdot, b)$.

The capacity of the warehouse is $M > 0$. Therefore,

$$X = A = [0, M] \quad A(x) = [0, M - x] \quad \text{and} \quad B = B(x, a) = [\underline{b}, \overline{b}].$$

An inventory management system

The controller incurs:

- a buying cost of $b>0$ for each unit;
- a holding cost $h>0$ for each period and unit;
- and receives $p>0$ for each unit that is sold.

The running cost function is

$$c(x,a,b) = ba + h(x+a) - pE[\min\{x+a,\xi\}].$$

Theorem

If the $\{\xi_t\}$ are i.i.d. with density function $f(\cdot,\cdot)$, which is Lipschitz continuous on $[0,M]\times B$ with $f(0,b)=0$, then the inventory management system satisfies our assumptions.

An inventory management system

Fix $0 < \mathfrak{p} < 1$. The probability measure μ is

$$\mu\{0\} = \mathfrak{p} \quad \text{and} \quad \mu(B) = \frac{1-\mathfrak{p}}{M}\lambda(B) \quad \text{for measurable } B \subseteq (0, M].$$

Given $n \geq 1$, approximate it with

$$\nu_n\{0\} = \mathfrak{p} \quad \text{and} \quad \nu_n\{M \cdot j/n\} = (1-\mathfrak{p})/n \quad \text{for } 0 < j \leq n$$

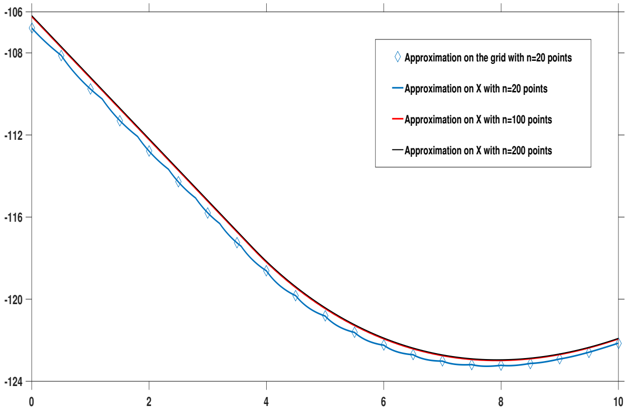
supported on the grid Γ_n.
The density function of the demand is a $\gamma(1/b, 2)$.
The approximating action sets are

$$A_\delta(x) = \left\{ \frac{(M-x)j}{n-1} \ : \ j = 0, 1, \ldots, n-1 \right\}.$$

$$B_\delta(x, a) = \left\{ \frac{(\overline{b} - \underline{b})j}{n-1} \ : \ j = 0, 1, \ldots, n-1 \right\}$$

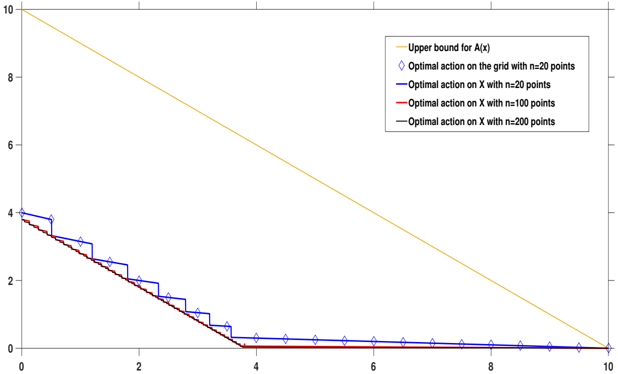
An inventory management system

For $n = 20, 100, 200$ we solve the minimax problem on Γ_n and then we extend it to the whole $X = [0, M]$.



An inventory management system

We determine a minimax $\widetilde{f}_{n,\delta}$ for $\mathscr{M}_{n,\delta}$ and we evaluate it for $\mathscr{M}$.



Intuitively, $f^*(x) = (M_0 - x)^+$ for some $M_0 \simeq 3.8$.

Conclusions

- We have proposed a general procedure to approximate a continuous state and action minimax control problem
- We can do this for a "Lipschitz-continuous" control model.
- The approximation error combines Hausdorff (for the actions) and Wasserstein (for the states) distances.
- In the application, our method provides very good approximations.

Dufour, F., Prieto-Rumeau, T. (2018).
Approximation of discounted minimax Markov control problems and zero-sum Markov games using Hausdorff and Wasserstein distances.

Dynamic Games and Applications. In press (online 19 march 2018) pp. 1–35.

Thank you for your attention.