An Introduction to *Stochastic Dual Dynamic Programming* (SDDP).

V. Leclère (CERMICS, ENPC)

18/05/2018

Introduction

- Large scale stochastic optimization problems are hard to solve
- Different ways of attacking such problems:
  - decompose the problem and coordinate solutions
  - construct easily solvable approximations (Linear Programming)
  - find approximate value functions or policies
- Behind the name SDDP, *Stochastic Dual Dynamic Programming*, one finds three different things:
  - a class of algorithms, based on specific mathematical assumptions
  - a specific implementation of an algorithm
  - a software implementing this method, and developed by the PSR company

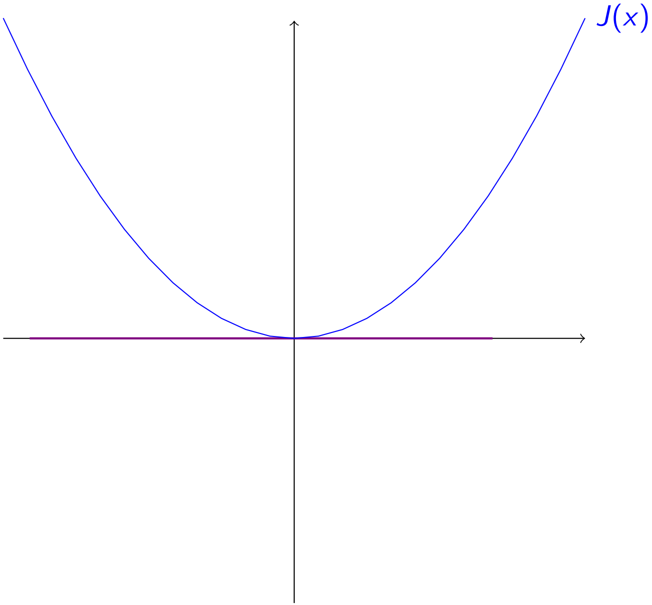
Introduction

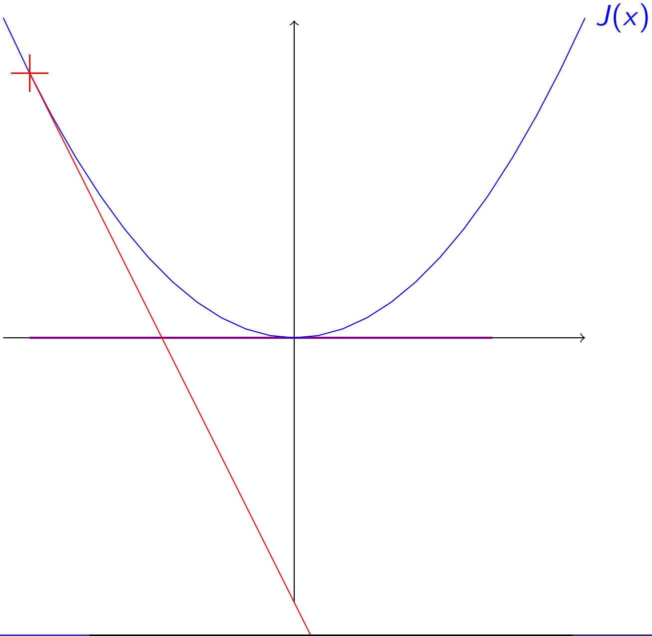
- Large scale stochastic optimization problems are hard to solve
- Different ways of attacking such problems:
  - decompose the problem and coordinate solutions
  - construct easily solvable approximations (Linear Programming)
  - find approximate value functions or policies
- Behind the name SDDP, *Stochastic Dual Dynamic Programming*, one finds three different things:
  - a class of algorithms,
    based on specific mathematical assumptions
  - a specific implementation of an algorithm
  - a software implementing this method,
    and developed by the PSR company

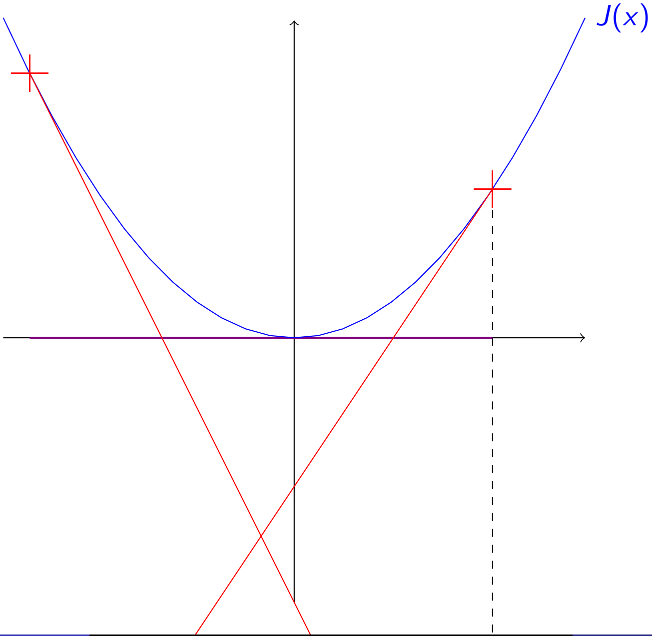
Setting

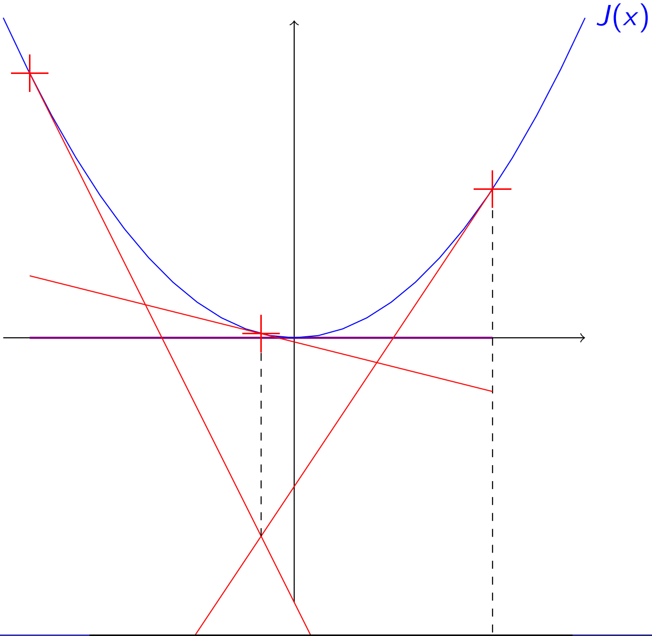
- Multi-stage stochastic optimization problems with finite horizon.
- Continuous, finite dimensional state and control.
- Convex cost, linear dynamic.
- Discrete, stagewise independent noises.

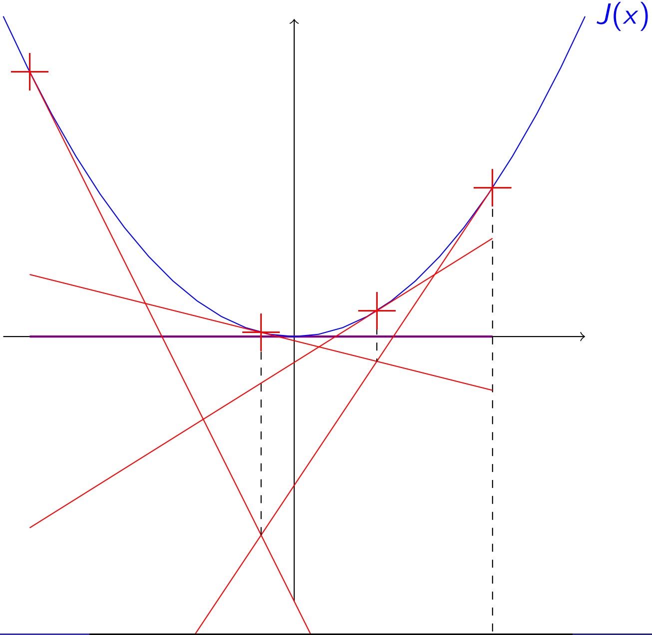
Contents


$J(x)$


$J(x)$


$J(x)$


$J(x)$


$J(x)$

Kelley algorithm

Data: Convex objective function J, Compact set X, Initial point $x_0 \in X$

Result: Admissible solution $x^{(k)}$, lower-bound $\underline{v}^{(k)}$

Set $J^{(0)} \equiv -\infty$;

for $k \in \mathbb{N}$ **do**

  Compute a subgradient $\lambda^{(k)} \in \partial J(x^{(k)})$;

  Define a cut $\mathcal{C}^{(k)} : x \mapsto J(x^{(k)}) + \langle \lambda^{(k)}, x - x^{(k)} \rangle$;

  Update the lower approximation $J^{(k+1)} = \max\{J^{(k)}, \mathcal{C}^{(k)}\}$;

  Solve $(P^{(k)}) : \quad \min_{x \in X} J^{(k+1)}(x)$;

  Set $\underline{v}^{(k)} = val(P^{(k)})$;

  Select $x^{(k+1)} \in sol(P^{(k)})$;

end

Algorithm 1: Kelley's cutting plane algorithm

Contents

Problem considered

We consider an optimal control problem in discrete time with finite horizon T

$$\min_{u \in \mathbb{U}^T} \quad \sum_{t=0}^{T-1} L_t(x_t, u_t) + K(x_T)$$

$$s.t. \quad x_{t+1} = f_t(x_t, u_t), \quad x_0 \text{ given}$$

$$u_t \in U_t(x_t)$$

- Where the variables are
  - $x_t \in \mathbb{X}$, the state at time t
  - $u_t \in \mathbb{U}$, the control applied at the beginning of $[t, t+1[$
- We assume that
  - the dynamics functions $(x_t, u_t) \mapsto f_t(x_t, u_t)$ are affine
  - $\mathbb{X}$ and $U_t(x)$ are compact convex
- the instantaneous costs $L_t(x_t, u_t)$ and the final cost $K(x_T)$ are convex

Problem considered

We consider an optimal control problem in discrete time with finite horizon T

$$\min_{u \in \mathbb{U}^T} \quad \sum_{t=0}^{T-1} L_t(x_t, u_t) + K(x_T)$$
$$s.t. \quad x_{t+1} = f_t(x_t, u_t), \quad x_0 \text{ given}$$
$$u_t \in U_t(x_t)$$

- Where the variables are
  - $x_t \in \mathbb{X}$, the state at time t
  - $u_t \in \mathbb{U}$, the control applied at the beginning of $[t, t+1[$
- We assume that
  - the dynamics functions $(x_t, u_t) \mapsto f_t(x_t, u_t)$ are affine
  - $\mathbb{X}$ and $U_t(x)$ are compact convex
- the instantaneous costs $L_t(x_t, u_t)$ and the final cost $K(x_T)$ are convex

Contents

Introducing Bellman's function

We look for solutions as policies, where a policy is a sequence of functions $\pi = (\pi_1, \dots, \pi_{T-1})$ giving for any state x a control u

This problem can be solved by dynamic programming, thanks to the Bellman function that satisfies

$$\begin{cases} V_T(x) &= K(x), \\ V_t(x) &= \min_{u_t \in U_t(x)} \left\{ L_t(x, u_t) + V_{t+1} \circ f_t(x, u_t) \right\} \end{cases}$$

Indeed, an optimal policy for the original problem is given by

$$\pi_t(x) \in \arg\min_{u_t \in \mathbb{U}} \left\{ L_t(x, u_t) + V_{t+1} \circ f_t(x, u_t) \right\}$$

Introducing Bellman's function

We look for solutions as policies, where a policy is a sequence of functions $\pi = (\pi_1, \dots, \pi_{T-1})$ giving for any state x a control u

This problem can be solved by dynamic programming, thanks to the Bellman function that satisfies

$$\begin{cases} V_T(x) & = & K(x), \\ V_t(x) & = & \min\limits_{u_t \in U_t(x)} \left\{ L_t(x, u_t) + V_{t+1} \circ f_t(x, u_t) \right\} \end{cases}$$

Indeed, an optimal policy for the original problem is given by

$$\pi_t(x) \in \arg\min_{u_t \in \mathbb{U}} \left\{ L_t(x, u_t) + V_{t+1} \circ f_t(x, u_t) \right\}$$

Introducing Bellman's operator

We define the Bellman operator

$$\mathcal{T}_t(A) : x \mapsto \min_{u_t \in U_t(x)} \big\{ L_t(x, u_t) + A \circ f_t(x, u_t) \big\}$$

With this notation, the Bellman Equation reads

$$\begin{cases} V_T &= K, \\ V_t &= \mathcal{T}_t(V_{t+1}) \end{cases}$$

Any approximate cost function $\tilde{V}_{t+1}$ induce an *admissible* policy

$$\pi_t^{\tilde{V}_{t+1}} : x \mapsto \arg\min \mathcal{T}_t\big(\tilde{V}_{t+1}\big)(x).$$

By Dynamic Programming, $\pi_t^{V_{t+1}}$ is optimal.

Introducing Bellman's operator

We define the Bellman operator

$$\mathcal{T}_t(A) : x \mapsto \min_{u_t \in U_t(x)} \left\{ L_t(x, u_t) + A \circ f_t(x, u_t) \right\}$$

With this notation, the Bellman Equation reads

$$\begin{cases} V_T & = K, \\ V_t & = \mathcal{T}_t(V_{t+1}) \end{cases}$$

Any approximate cost function $\tilde{V}_{t+1}$ induce an *admissible* policy

$$\pi_t^{\tilde{V}_{t+1}} : x \mapsto \arg\min \mathcal{T}_t\big(\tilde{V}_{t+1}\big)(x).$$

By Dynamic Programming, $\pi_t^{V_{t+1}}$ is optimal.

Introducing Bellman's operator

We define the Bellman operator

$$\mathcal{T}_t(A) : x \mapsto \min_{u_t \in U_t(x)} \left\{ L_t(x, u_t) + A \circ f_t(x, u_t) \right\}$$

With this notation, the Bellman Equation reads

$$\begin{cases} V_T & = K, \\ V_t & = \mathcal{T}_t(V_{t+1}) \end{cases}$$

Any approximate cost function $\tilde{V}_{t+1}$ induce an *admissible* policy

$$\pi_t^{\tilde{V}_{t+1}} : x \mapsto \arg\min \mathcal{T}_t\big(\tilde{V}_{t+1}\big)(x).$$

By Dynamic Programming, $\pi_t^{V_{t+1}}$ is optimal.

Properties of the Bellman operator

- Monotonicity:

$$V \leq \overline{V} \quad \Rightarrow \mathcal{T}_t\big(V\big) \leq \mathcal{T}_t\big(\overline{V}\big)$$

- Convexity: if L_t is jointly convex in (x, u), V is convex, and f_t is affine then

$$x \mapsto \mathcal{T}_t\big(V\big)(x) \quad \text{is convex}$$

- Polyhedrality: for any polyhedral function V, if L_t is also polyhedral, and f_t affine, then

$$x \mapsto \mathcal{T}_t\big(V\big)(x) \quad \text{is polyhedral}$$

Duality property

- Consider $J : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ jointly convex, and define

$$\varphi(x) = \min_{u \in \mathbb{U}} J(x, u)$$

- Then we can obtain a subgradient $\lambda \in \partial\varphi(x_0)$ as the dual multiplier of

$$\begin{aligned} \min_{x,u} \quad & J(x,u), \\ s.t. \quad & x_0 - x = 0 \qquad\qquad [\lambda] \end{aligned}$$

(This is the marginal interpretation of the multiplier)

- In particular, we have that

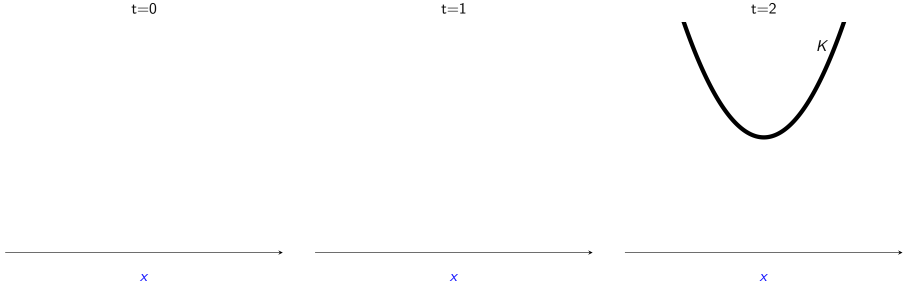
$$\varphi(\cdot) \geq \varphi(x_0) + \langle \lambda, \cdot - x_0 \rangle$$

Contents

General idea

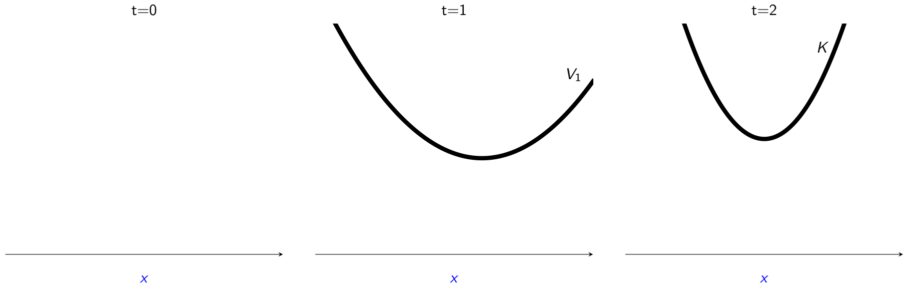
- The SDDP algorithm recursively constructs an approximation of each Bellman function V_t as the supremum of affine functions
- At stage k, we have a lower approximation $\underline{V}_t^{(k)}$ of V_t and we want to construct a better approximation
- We follow an optimal trajectory $(x_t^{(k)})_t$ of the approximated problem, and add a so-called "cut" to improve each Bellman function

Deterministic SDDP



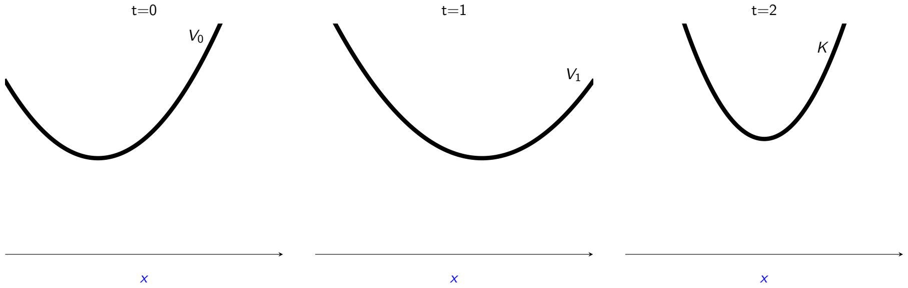
Final Cost $V_2 = K$

Deterministic SDDP



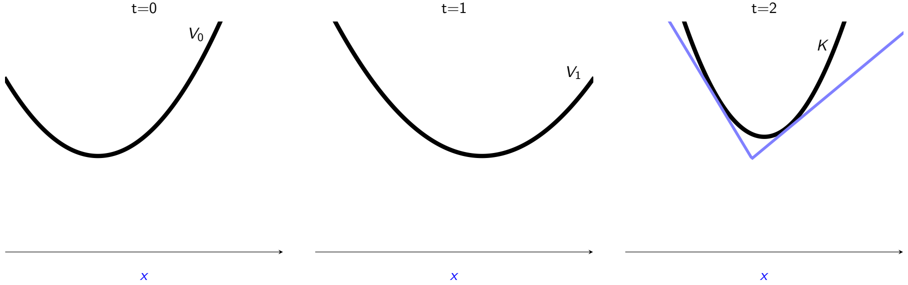
Real Bellman function $V_1 = T_1(V_2)$

Deterministic SDDP



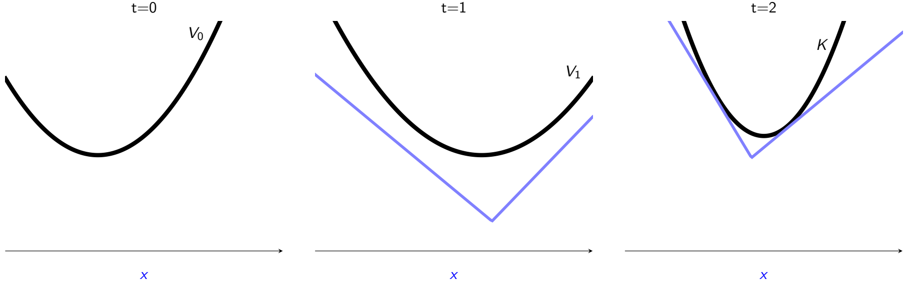
Real Bellman function $V_0 = T_0(V_1)$

Deterministic SDDP



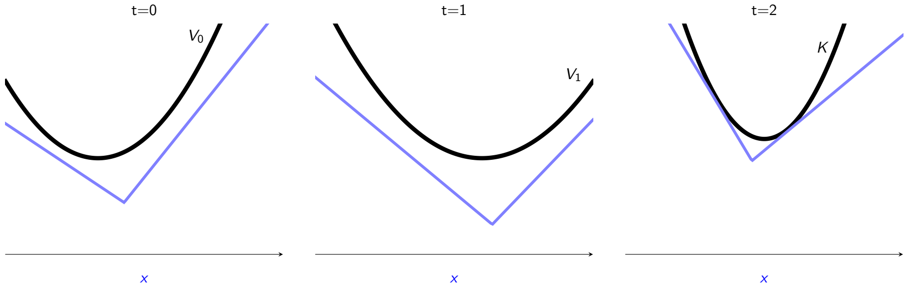
Lower polyhedral approximation $\underline{K}$ of K

Deterministic SDDP



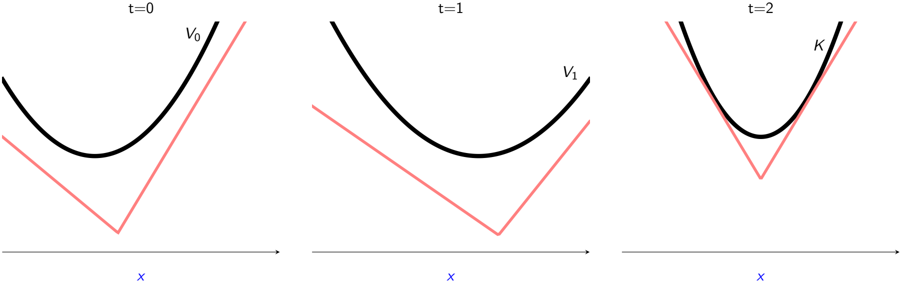
Lower polyhedral approximation $\underline{V}_1 = \mathcal{T}_t(\underline{K})$ of V_1

Deterministic SDDP



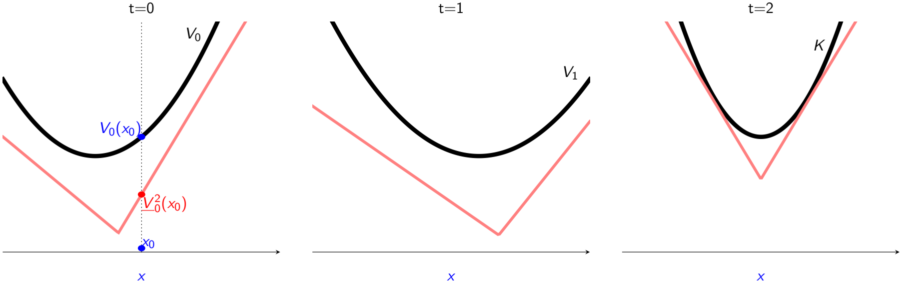
Lower polyhedral approximation $\underline{V}_0 = \mathcal{T}_t(\underline{V}_1)$ of V_0

Deterministic SDDP



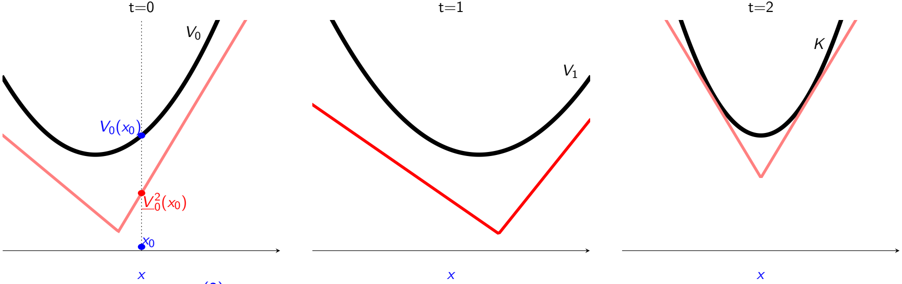
Assume that we have lower polyhedral approximations of V_t

Deterministic SDDP



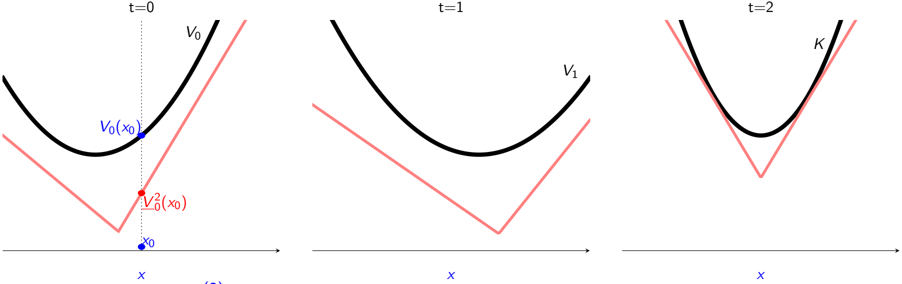
Thus we have a lower bound on the value of our problem

Deterministic SDDP



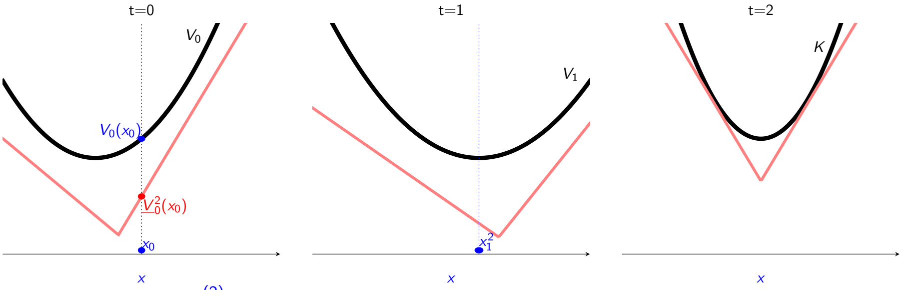
We apply $\pi_0^{\underline{V}_1^{(2)}}$ to x_0 and obtain $x_1^{(2)}$

Deterministic SDDP



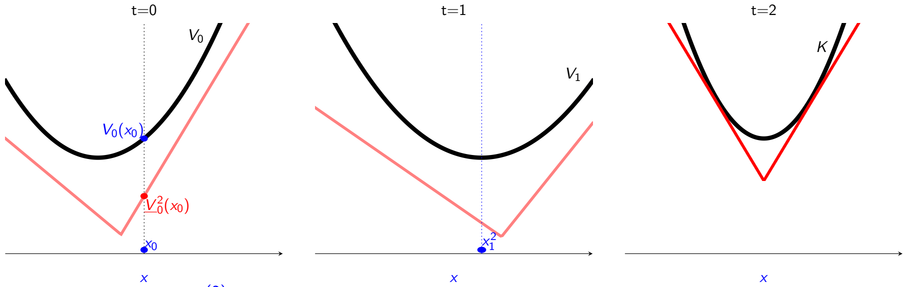
We apply $\pi_0^{\underline{V}_1^{(2)}}$ to x_0 and obtain $x_1^{(2)}$

Deterministic SDDP



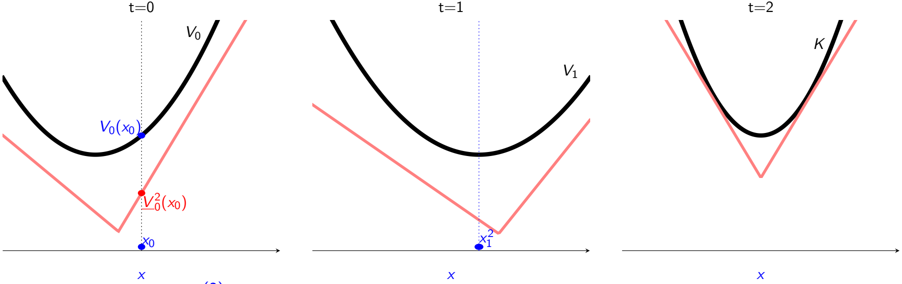
We apply $\pi_0^{\underline{V}_1^{(2)}}$ to x_0 and obtain $x_1^{(2)}$

Deterministic SDDP



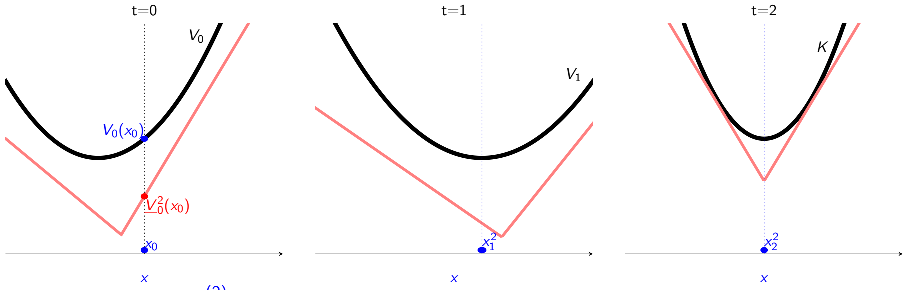
We apply $\pi_1^{\underline{V}_1^{(2)}}$ to $x_1^{(2)}$ and obtain $x_2^{(2)}$

Deterministic SDDP



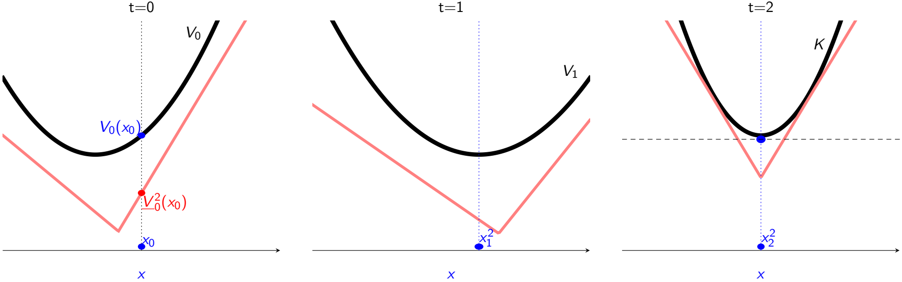
We apply $\pi_1^{\underline{V}_1^{(2)}}$ to $x_1^{(2)}$ and obtain $x_2^{(2)}$

Deterministic SDDP



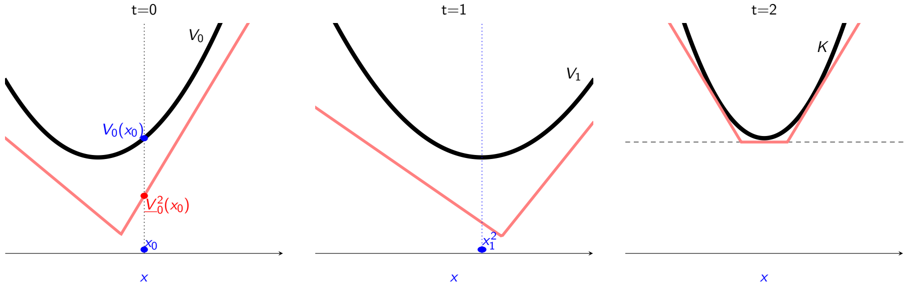
We apply $\pi_1^{\underline{V}_1^{(2)}}$ to $x_1^{(2)}$ and obtain $x_2^{(2)}$

Deterministic SDDP



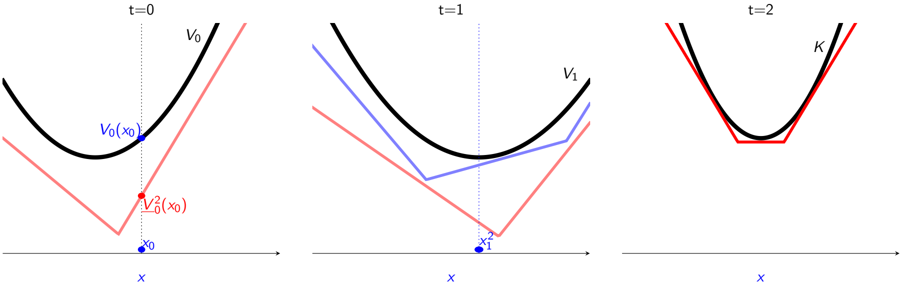
Compute a cut for K at $x_2^{(2)}$

Deterministic SDDP



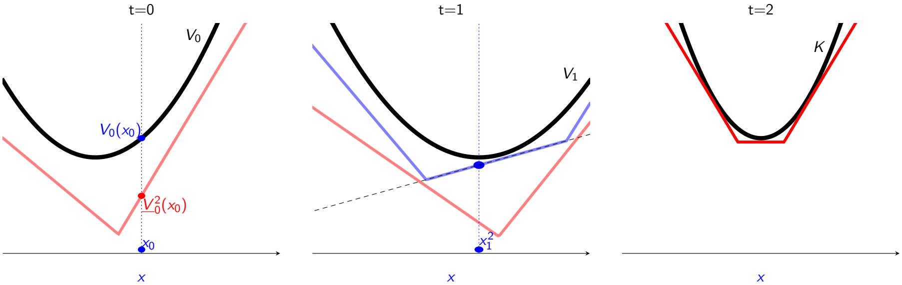
Add the cut to $\underline{V}_2^{(2)}$ which gives $\underline{V}_2^{(3)}$

Deterministic SDDP



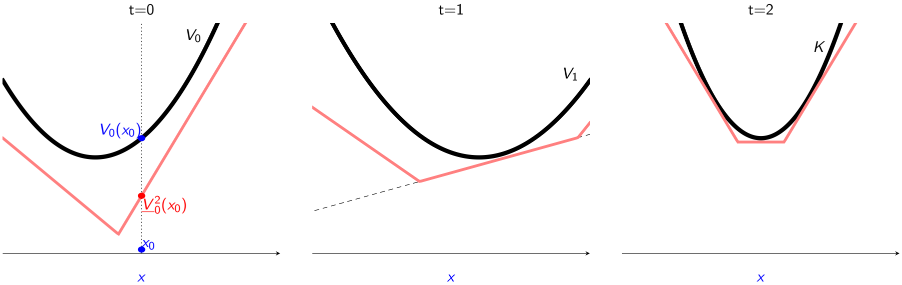
A new lower approximation of V_1 is $T_1(\underline{V}_2^{(3)})$

Deterministic SDDP



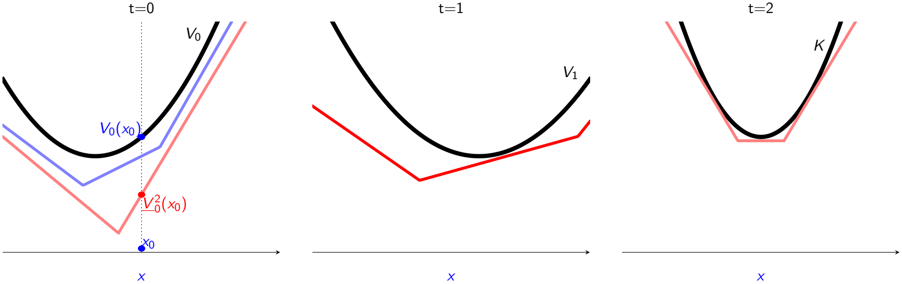
We only compute the face active at $x_1^{(2)}$

Deterministic SDDP



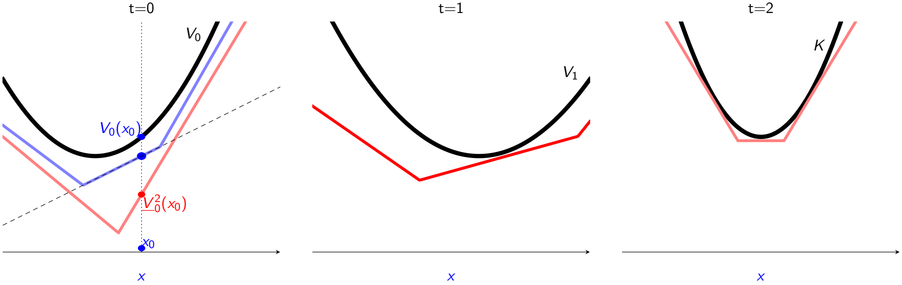
Add the cut to $\underline{V}_1^{(2)}$ which gives $\underline{V}_1^{(3)}$

Deterministic SDDP



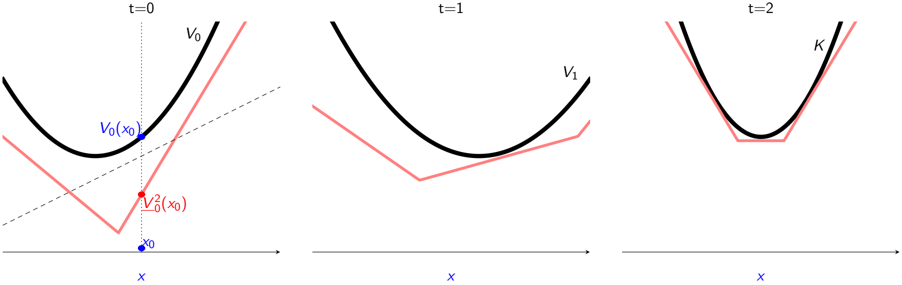
A new lower approximation of V_0 is $T_0(\underline{V}_1^{(3)})$

Deterministic SDDP



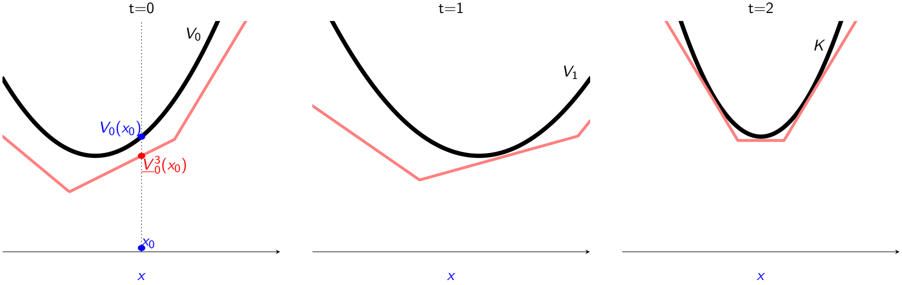
We only compute the face active at x_0

Deterministic SDDP



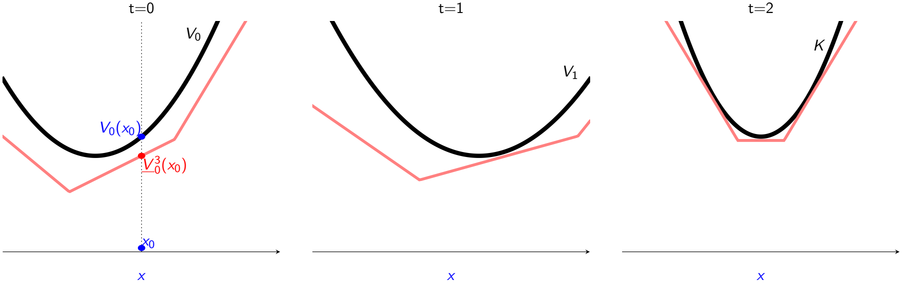
We only compute the face active at x_0

Deterministic SDDP



We obtain a new lower bound

Deterministic SDDP



We obtain a new lower bound

Stage k of SDDP description (1/2)

- Begin a "forward in time" loop by setting $t = 0$ and $x_t^{(k)} = x_0$
- Solve

$$\min_{x,u} \quad L_t(x,u) + \underline{V}_{t+1}^{(k)} \circ f_t(x,u)$$
$$x = x_t^{(k)} \qquad [\lambda_t^{(k+1)}]$$

where we call
  - $\beta_t^{(k+1)}$ the value of the problem
  - $\lambda_t^{(k+1)}$ a multiplier of the constraint $x = x_t^{(k)}$
  - $u_t^{(k)}$ an optimal control
- By construction, we have that

$$\beta_t^{(k+1)} = \mathcal{T}_t\left(\underline{V}_{t+1}^{(k)}\right)\left(x_t^{(k)}\right),$$
$$\lambda_t^{(k+1)} \in \partial\mathcal{T}_t\left(\underline{V}_{t+1}^{(k)}\right)\left(x_t^{(k)}\right).$$

Stage k of SDDP description (1/2)

- Begin a "forward in time" loop by setting $t = 0$ and $x_t^{(k)} = x_0$
- Solve

$$\min_{x,u} \quad L_t(x,u) + \underline{V}_{t+1}^{(k)} \circ f_t(x,u)$$
$$x = x_t^{(k)} \qquad [\lambda_t^{(k+1)}]$$

where we call
  - $\beta_t^{(k+1)}$ the value of the problem
  - $\lambda_t^{(k+1)}$ a multiplier of the constraint $x = x_t^{(k)}$
  - $u_t^{(k)}$ an optimal control
- By construction, we have that

$$\beta_t^{(k+1)} = \mathcal{T}_t\left(\underline{V}_{t+1}^{(k)}\right)\left(x_t^{(k)}\right),$$
$$\lambda_t^{(k+1)} \in \partial\mathcal{T}_t\left(\underline{V}_{t+1}^{(k)}\right)\left(x_t^{(k)}\right).$$

Stage k of SDDP description (2/2)

- We deduce that

$$\beta_t^{(k+1)} + \langle \lambda_t^{(k+1)}, \cdot - x_t^{(k)} \rangle \leq \mathcal{T}_t\left(\underline{V}_{t+1}^{(k)}\right) \leq \mathcal{T}_t\left(V_{t+1}\right) = V_t$$

- Thus $\mathcal{C}_t^{(k+1)} \ : \ x \mapsto \beta_t^{(k+1)} + \left\langle \lambda_t^{(k+1)}, x - x_t^{(k)} \right\rangle$ is a cut
- We update our approximation $\underline{V}_t^{(k)}$ of V_t by defining

$$\underline{V}_t^{(k+1)} = \max\left\{\underline{V}_t^{(k)}, \mathcal{C}_t^{(k+1)}\right\} = \max_{\kappa \leq k+1} \left\{\mathcal{C}_t^{(\kappa)}\right\}$$

  so that $\underline{V}_t^{(k+1)}$ is convex and is lower than V_t
- We set

$$x_{t+1}^{(k)} = f_t\left(x_t^{(k)}, u_t^{(k)}\right)$$

- Upon reaching time $t = T$ we have completed iteration k of the algorithm.

Stage k of SDDP description (2/2)

- We deduce that

$$\beta_t^{(k+1)} + \langle \lambda_t^{(k+1)}, \cdot - x_t^{(k)} \rangle \leq \mathcal{T}_t\left(\underline{V}_{t+1}^{(k)}\right) \leq \mathcal{T}_t\left(V_{t+1}\right) = V_t$$

- Thus $\mathcal{C}_t^{(k+1)} \ : \ x \mapsto \beta_t^{(k+1)} + \left\langle \lambda_t^{(k+1)}, x - x_t^{(k)} \right\rangle$ is a cut
- We update our approximation $\underline{V}_t^{(k)}$ of V_t by defining

$$\underline{V}_t^{(k+1)} = \max\left\{\underline{V}_t^{(k)}, \mathcal{C}_t^{(k+1)}\right\} = \max_{\kappa \leq k+1} \left\{\mathcal{C}_t^{(\kappa)}\right\}$$

  so that $\underline{V}_t^{(k+1)}$ is convex and is lower than V_t
- We set

$$x_{t+1}^{(k)} = f_t\left(x_t^{(k)}, u_t^{(k)}\right)$$

- Upon reaching time $t = T$ we have completed iteration k of the algorithm.

Stage k of SDDP description (2/2)

- We deduce that

$$\beta_t^{(k+1)} + \langle \lambda_t^{(k+1)}, \cdot - x_t^{(k)} \rangle \leq \mathcal{T}_t \left(\underline{V}_{t+1}^{(k)} \right) \leq \mathcal{T}_t \left(V_{t+1} \right) = V_t$$

- Thus $\mathcal{C}_t^{(k+1)} \ : \ x \mapsto \beta_t^{(k+1)} + \left\langle \lambda_t^{(k+1)}, x - x_t^{(k)} \right\rangle$ is a cut
- We update our approximation $\underline{V}_t^{(k)}$ of V_t by defining

$$\underline{V}_t^{(k+1)} = \max\left\{ \underline{V}_t^{(k)}, \mathcal{C}_t^{(k+1)} \right\} = \max_{\kappa \leq k+1} \left\{ \mathcal{C}_t^{(\kappa)} \right\}$$

  so that $\underline{V}_t^{(k+1)}$ is convex and is lower than V_t
- We set

$$x_{t+1}^{(k)} = f_t \left(x_t^{(k)}, u_t^{(k)} \right)$$

- Upon reaching time $t = T$ we have completed iteration k of the algorithm.

Contents

Initialization and stopping rule

- To initialize the algorithm, we need a lower bound $\underline{V}_t^{(0)}$ for each value function V_t. This lower bound can be computed backward by arbitrarily choosing a point x_t and using the standard cut computation.
- At any step k we have an admissible, non optimal solution $(u^{(k)})_t$, with
  - an upper bound

$$\sum_{t=0}^{T-1} L_t\left(x_t^{(k)}, u_t^{(k)}\right) + K\left(x_T^{(k)}\right)$$

  - a lower bound $\underline{V}_0^{(k)}(x_0)$
- A reasonable stopping rule for the algorithm is given by checking that the (relative) difference between the upper and lower bounds is small enough

Contents

What's new ?

Now we introduce random variables ξ_t in our problem, which complexifies the algorithm in different ways:

- we need some probabilistic assumptions
- for each stage k we need to do a forward phase, for each sequence of realizations of the random variables, that yields a trajectory $(x_t^{(k)})_t$, and a backward phase that gives a new cut
- we cannot compute an exact upper bound for the problem value

Problem statement

We consider the optimization problem

$$\begin{aligned} \min_{\pi} \quad & \mathbb{E}\left(\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{\xi}_t) + K(\boldsymbol{X}_T)\right) \\ s.t. \quad & \boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{\xi}_t) \quad \boldsymbol{X}_0 = x_0 \\ & \boldsymbol{U}_t = \pi_t(\boldsymbol{X}_t, \boldsymbol{\xi}_t) \in U_t(x, \boldsymbol{\xi}_t) \end{aligned}$$

under the crucial assumption that $(\boldsymbol{\xi}_t)_{t\in\{1,\cdots,T\}}$ is a white noise

$\rightsquigarrow$ we are in an hazard-decision framework.

Problem statement

We consider the optimization problem

$$\begin{aligned} \min_{\pi} \quad & \mathbb{E}\left(\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{\xi}_t) + K(\boldsymbol{X}_T)\right) \\ s.t. \quad & \boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{\xi}_t) \quad \boldsymbol{X}_0 = x_0 \\ & \boldsymbol{U}_t = \pi_t(\boldsymbol{X}_t, \boldsymbol{\xi}_t) \in U_t(x, \boldsymbol{\xi}_t) \end{aligned}$$

under the crucial assumption that $(\boldsymbol{\xi}_t)_{t\in\{1,\cdots,T\}}$ is a white noise

$\leadsto$ we are in an hazard-decision framework.

Stochastic Dynamic Programming

By the white noise assumption, this problem can be solved by dynamic programming, where the Bellman functions satisfy

$$\begin{cases} V_T(x) & = & K(x) \\ \hat{V}_t(x,\xi) & = & \min\limits_{u_t \in U_t(x,\xi)} L_t(x,u_t,\xi) + V_{t+1} \circ f_t(x,u_t,\xi) \\ V_t(x) & = & \mathbb{E}\Big(\hat{V}_t(x,\boldsymbol{\xi}_t)\Big) \end{cases}$$

Indeed, an optimal policy for this problem is given by

$$\pi_t(x,\xi) \in \mathop{\arg\min}_{u_t \in U_t(x,\xi)} \Big\{ L_t(x,u_t,\xi) + V_{t+1} \circ f_t(x,u_t,\xi) \Big\}$$

Bellman operator

For any time t, and any function A mapping the set of states and noises $\mathbb{X} \times \Xi$ into $\mathbb{R}$, we define

$$\begin{cases} \hat{\mathcal{T}}_t(A)(x,\xi) & := \min_{u_t \in U_t(x,\xi)} L_t(x,u_t,\xi) + A \circ f_t(x,u_t,\xi) \\ \mathcal{T}_t(V_{t+1}) & := \mathbb{E}\left(\hat{\mathcal{T}}_t(V_{t+1})(x,\boldsymbol{\xi}_t)\right) \end{cases}$$

Thus the Bellman equation simply reads

$$\begin{cases} V_T & = & K \\ V_t & = & \mathcal{T}_t(V_{t+1}) \end{cases}$$

The Bellman operators have the same properties as in the deterministic case

Contents

Duality theory (1/2)

Suppose that we have $\underline{V}_{t+1}^{(k+1)} \leq V_{t+1}$

$$\hat{\beta}_t^{(k+1)}(\xi) = \min_{x,u} \quad L_t(x,u,\xi) + \underline{V}_{t+1}^{(k+1)} \circ f_t(x,u,\xi)$$
$$s.t \quad x = x_t^{(k)} \qquad [\hat{\lambda}_t^{(k+1)}(\xi)]$$

This can also be written as

$$\hat{\beta}_t^{(k+1)}(\xi) = \hat{\mathcal{T}}_t\left(\underline{V}_{t+1}^{(k+1)}\right)(x,\xi)$$
$$\hat{\lambda}_t^{(k+1)}(\xi) \in \partial_x \hat{\mathcal{T}}_t\left(\underline{V}_{t+1}^{(k+1)}\right)(x,\xi)$$

Thus, for all ξ, $\hat{\mathcal{C}}_t^{(k+1),\xi} : x \mapsto \hat{\beta}_t^{(k+1)}(\xi) + \left\langle \hat{\lambda}_t^{(k+1)}(\xi), x - x_t^{(k)} \right\rangle$ satisfy

$$\hat{\mathcal{C}}_t^{(k+1),\xi}(x) \leq \hat{\mathcal{T}}_t\left(\underline{V}_{t+1}^{(k+1)}\right)(x,\xi) \leq \hat{\mathcal{T}}_t\left(V_{t+1}\right)(x,\xi) = \hat{V}_t(x,\xi)$$

Duality theory (1/2)

Suppose that we have $\underline{V}_{t+1}^{(k+1)} \leq V_{t+1}$

$$\begin{aligned} \hat{\beta}_t^{(k+1)}(\xi) = \min_{x,u} \quad & L_t(x,u,\xi) + \underline{V}_{t+1}^{(k+1)} \circ f_t(x,u,\xi) \\ s.t \quad & x = x_t^{(k)} \qquad [\hat{\lambda}_t^{(k+1)}(\xi)] \end{aligned}$$

This can also be written as

$$\hat{\beta}_t^{(k+1)}(\xi) = \hat{\mathcal{T}}_t\left(\underline{V}_{t+1}^{(k+1)}\right)(x,\xi)$$

$$\hat{\lambda}_t^{(k+1)}(\xi) \in \partial_x \hat{\mathcal{T}}_t\left(\underline{V}_{t+1}^{(k+1)}\right)(x,\xi)$$

Thus, for all ξ, $\hat{\mathcal{C}}_t^{(k+1),\xi} : x \mapsto \hat{\beta}_t^{(k+1)}(\xi) + \left\langle \hat{\lambda}_t^{(k+1)}(\xi), x - x_t^{(k)} \right\rangle$ satisfy

$$\hat{\mathcal{C}}_t^{(k+1),\xi}(x) \leq \hat{\mathcal{T}}_t\left(\underline{V}_{t+1}^{(k+1)}\right)(x,\xi) \leq \hat{\mathcal{T}}_t\left(V_{t+1}\right)(x,\xi) = \hat{V}_t(x,\xi)$$

Duality theory (1/2)

Suppose that we have $\underline{V}_{t+1}^{(k+1)} \leq V_{t+1}$

$$\hat{\beta}_t^{(k+1)}(\xi) = \min_{x,u} \quad L_t(x,u,\xi) + \underline{V}_{t+1}^{(k+1)} \circ f_t(x,u,\xi)$$
$$s.t \quad x = x_t^{(k)} \qquad [\hat{\lambda}_t^{(k+1)}(\xi)]$$

This can also be written as

$$\hat{\beta}_t^{(k+1)}(\xi) = \hat{\mathcal{T}}_t\left(\underline{V}_{t+1}^{(k+1)}\right)(x,\xi)$$
$$\hat{\lambda}_t^{(k+1)}(\xi) \in \partial_x \hat{\mathcal{T}}_t\left(\underline{V}_{t+1}^{(k+1)}\right)(x,\xi)$$

Thus, for all ξ, $\hat{\mathcal{C}}_t^{(k+1),\xi} : x \mapsto \hat{\beta}_t^{(k+1)}(\xi) + \left\langle \hat{\lambda}_t^{(k+1)}(\xi), x - x_t^{(k)} \right\rangle$ satisfy

$$\hat{\mathcal{C}}_t^{(k+1),\xi}(x) \leq \hat{\mathcal{T}}_t\left(\underline{V}_{t+1}^{(k+1)}\right)(x,\xi) \leq \hat{\mathcal{T}}_t\left(V_{t+1}\right)(x,\xi) = \hat{V}_t(x,\xi)$$

Duality theory (2/2)

Thus, we have an affine minorant of $\hat{V}_t(x, \boldsymbol{\xi}_t)$ for each realization of $\boldsymbol{\xi}_t$

Replacing ξ by the random variable $\boldsymbol{\xi}_t$ and taking the expectation yields the following affine minorant

$$\beta_t^{(k+1)} + \left\langle \lambda_t^{(k+1)}, \cdot - x_t^{(k)} \right\rangle \leq V_t$$

where

$$\begin{cases} \beta_t^{(k+1)} & := & \mathbb{E}\left(\hat{\beta}_t^{(k+1)}(\boldsymbol{\xi}_t)\right) = \mathcal{T}_t\left(\underline{V}_{t+1}^{(k)}\right)(x) \\ \lambda_t^{(k+1)} & := & \mathbb{E}\left(\hat{\lambda}_t^{(k+1)}(\boldsymbol{\xi}_t)\right) \in \partial\mathcal{T}_t\left(\underline{V}_{t+1}^{(k)}\right)(x) \end{cases}$$

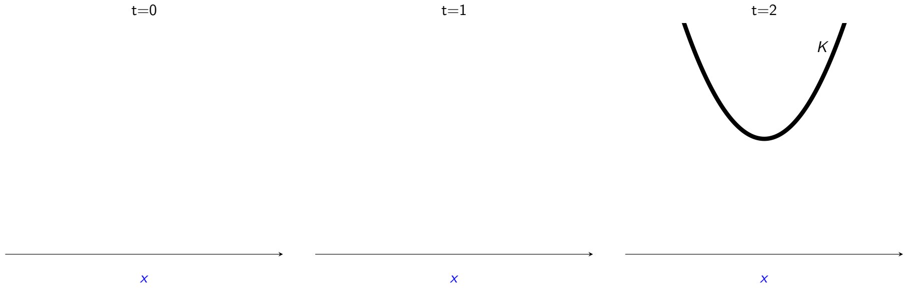
Contents

At the beginning of step k

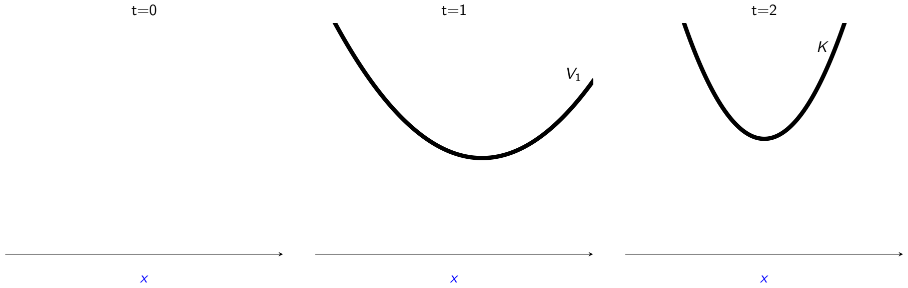
At the beginning of step k, we suppose that we have, for each stage t, an approximation $\underline{V}_t^{(k)}$ of V_t satisfying

- $\underline{V}_t^{(k)} \leq V_t$
- $\underline{V}_T^{(k)} = K$
- $\underline{V}_t^{(k)}$ is convex

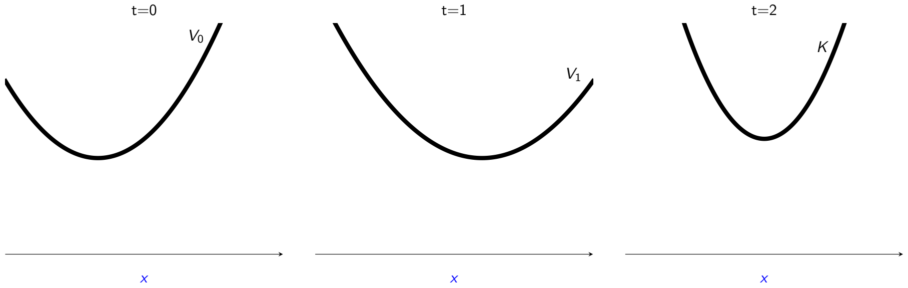
Abstract SDDP



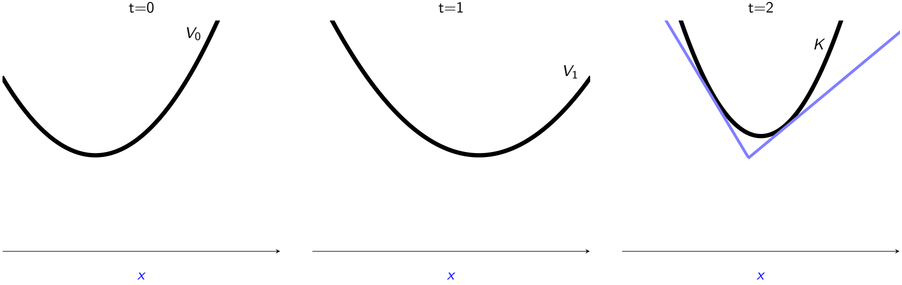
Abstract SDDP



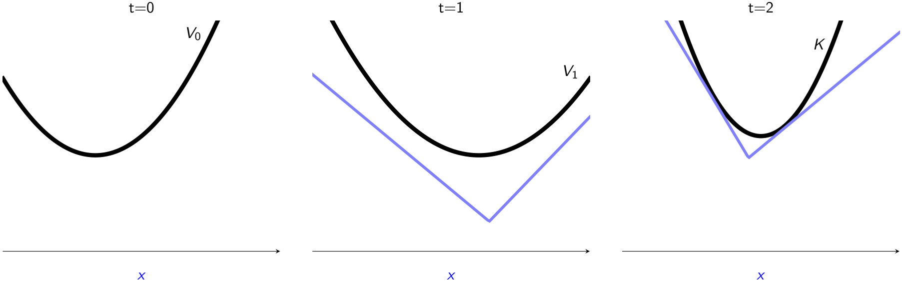
Abstract SDDP



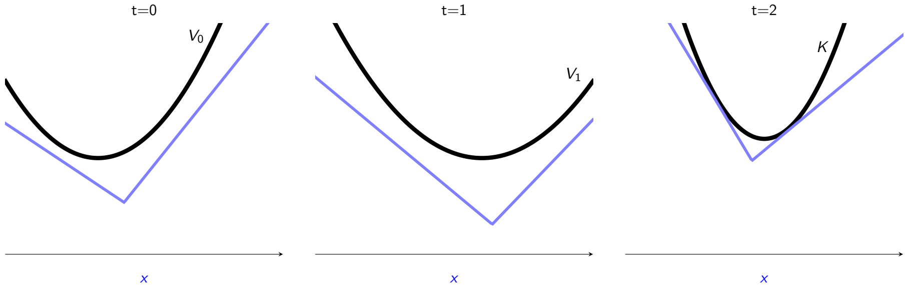
Abstract SDDP



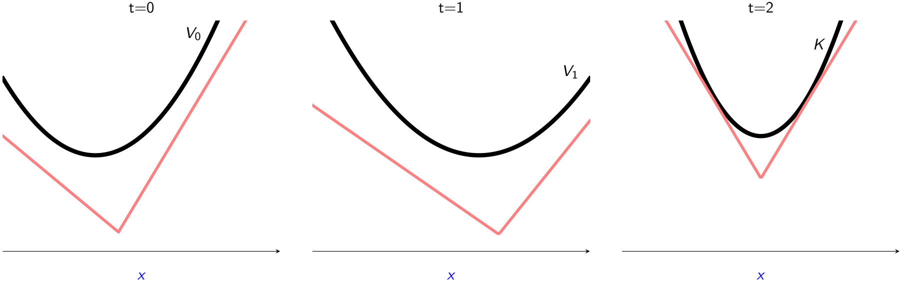
Abstract SDDP



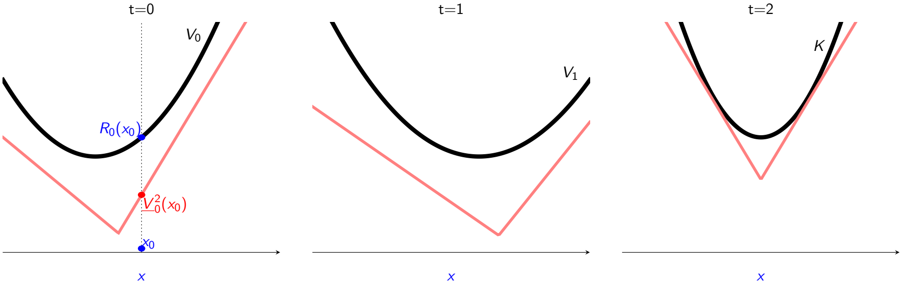
Abstract SDDP



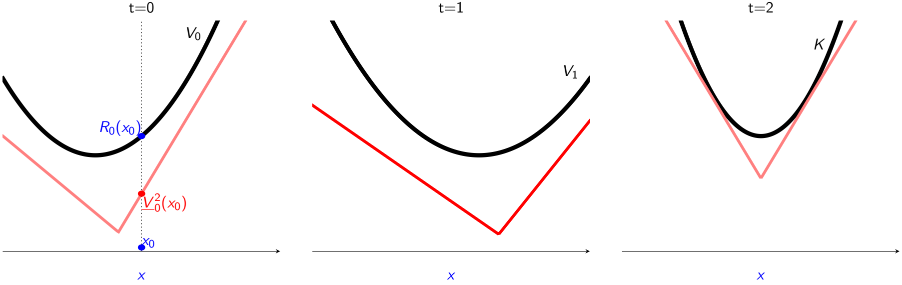
Abstract SDDP



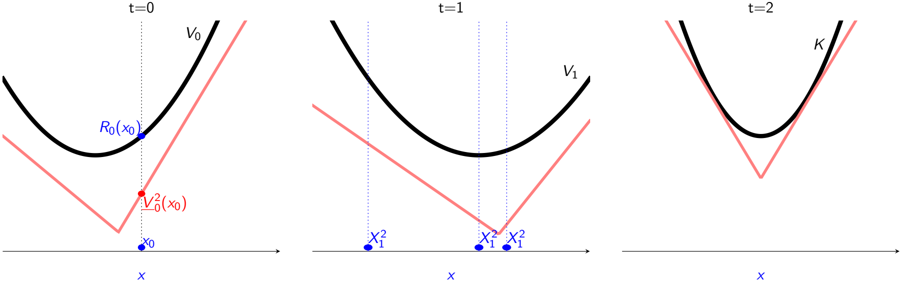
Abstract SDDP



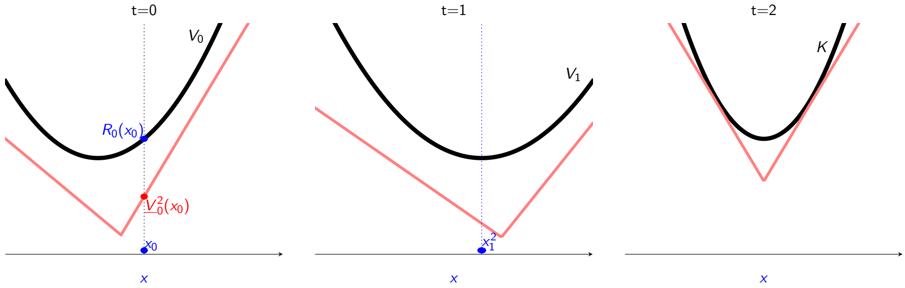
Abstract SDDP



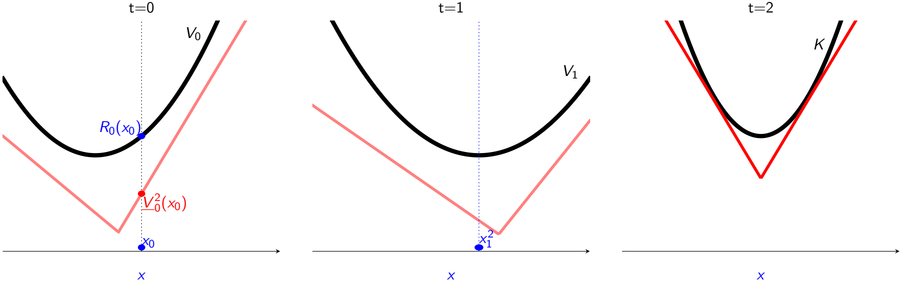
Abstract SDDP



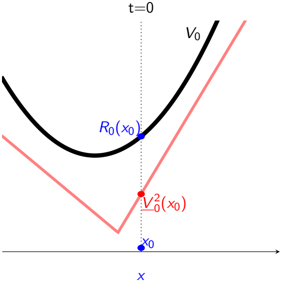
Abstract SDDP



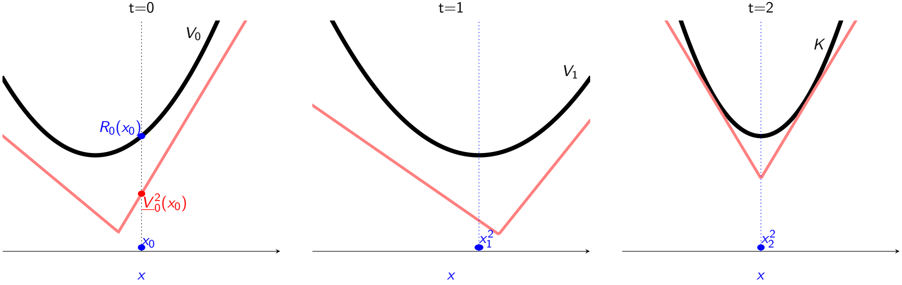
Abstract SDDP



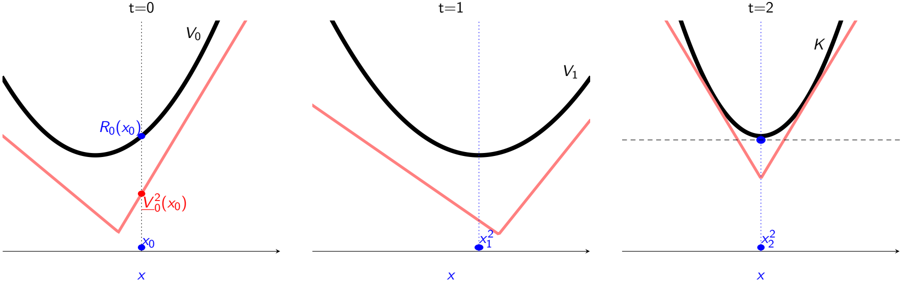
Abstract SDDP



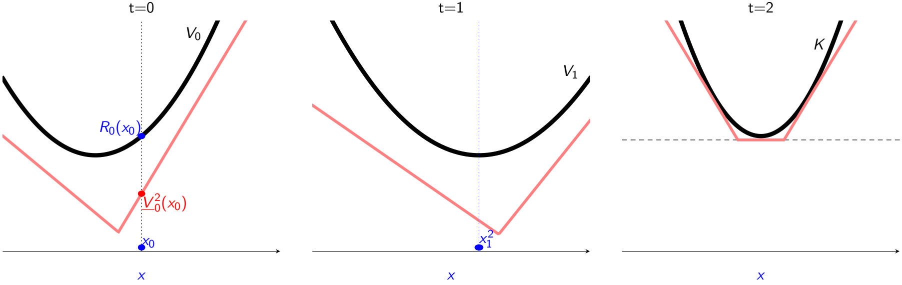
Abstract SDDP



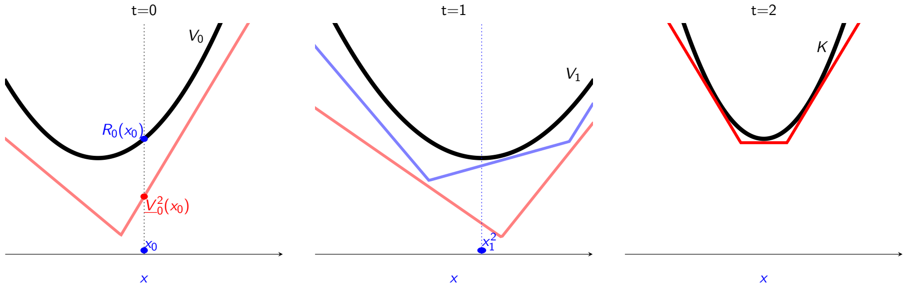
Abstract SDDP



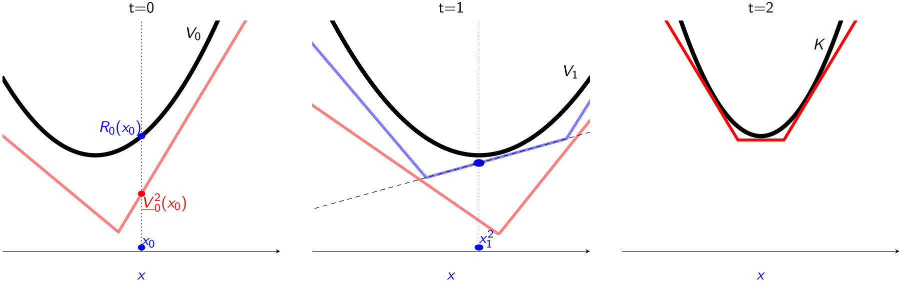
Abstract SDDP



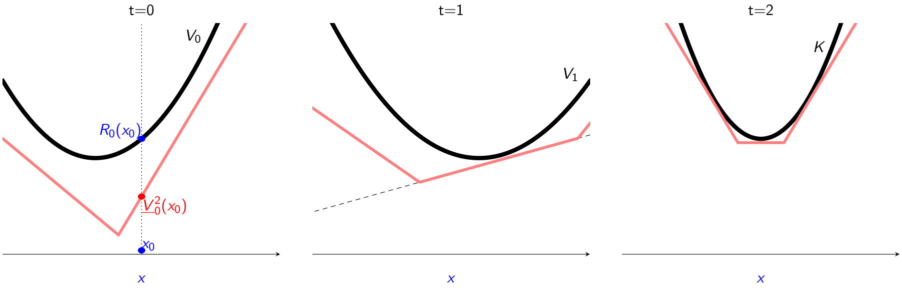
Abstract SDDP



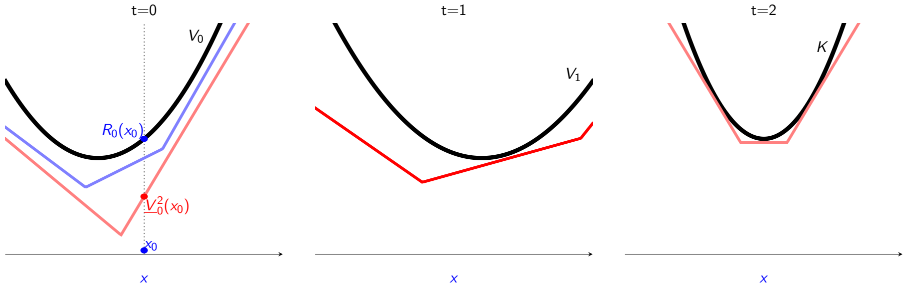
Abstract SDDP



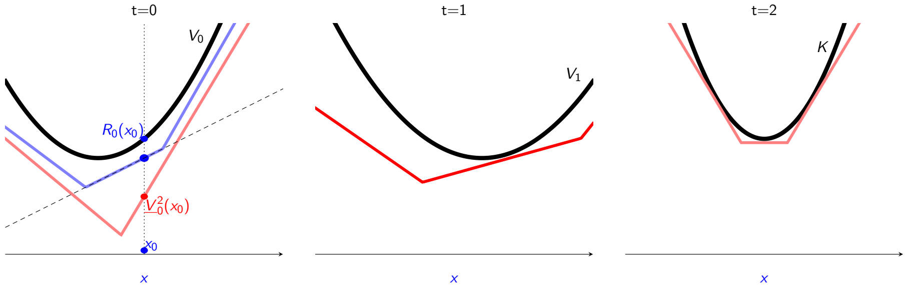
Abstract SDDP



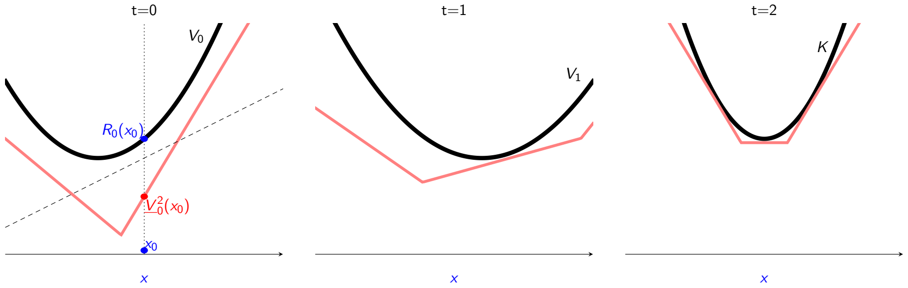
Abstract SDDP



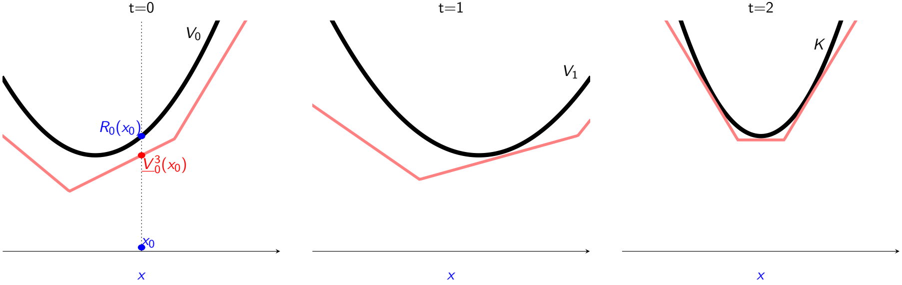
Abstract SDDP



Abstract SDDP



Abstract SDDP



Forward path: define a trajectory

- Randomly select a scenario $(\xi_0, \dots, \xi_{T-1}) \in \Xi^T$
- Define a trajectory $(x_t^{(k)})_{t=0,\dots,T}$ by

$$x_{t+1}^{(k)} = f_t(x_t^{(k)}, u_t^{(k)}, \xi_t)$$

where $u_t^{(k)} = \pi_t^{V_t^{(k)}}(x_t^{(k)})$ is an optimal solution of

$$\min_{u \in U_t(x,\xi)} L_t\left(x_t^{(k)}, u, \xi_t\right) + V_{t+1}^{(k)} \circ f_t\left(x_t^{(k)}, u, \xi_t\right)$$

Backward path: add cuts

- For any t we want to add a cut to $\underline{V}_t^{(k)}$ of V_t
- At time t solve, for any possible ξ,

$$\begin{aligned} \hat{\beta}_t^{(k+1)}(\xi) = \min_{x,u} \quad & L_t(x,u,\xi) + \underline{V}_{t+1}^{(k+1)} \circ f_t(x,u,\xi), \\ s.t \quad & x = x_t^{(k)} \qquad [\hat{\lambda}_t^{(k+1)}(\xi)] \end{aligned}$$

- Def $\lambda_t^{(k+1)} := \mathbb{E}\left(\lambda_t^{(k+1)}(\boldsymbol{\xi}_t)\right)$ and $\beta_t^{(k+1)} := \mathbb{E}\left(\beta_t^{(k+1)}(\boldsymbol{\xi}_t)\right)$
- Add a cut

$$V_t^{(k+1)}(x) = \max\left\{\underline{V}_t^{(k)}(x), \beta_t^{(k+1)} + \left\langle \lambda_t^{(k+1)}, x - x_t^{(k)} \right\rangle\right\}$$

- Go one step back in time: $t \leftarrow t-1$. Upon reaching $t=0$, we have completed step k of the algorithm

Recall on CLT

- Let $\{C_i\}_{i\in\mathbb{N}}$ be a sequence of identically distributed random variables with finite variance.
- Then the Central Limit Theorem ensures that

$$\sqrt{n}\Big(\frac{\sum_{i=1}^n \boldsymbol{C}_i}{n} - \mathbb{E}[\boldsymbol{C}_1]\Big) \implies G \sim \mathcal{N}(0, Var[\boldsymbol{C}_1]) ,$$

  where the convergence is in law.
- In practice it is often used in the following way. Asymptotically,

$$\mathbb{P}\Big(\mathbb{E}\big[C_1\big] \in \Big[\bar{\boldsymbol{C}}_n - \frac{1.96\boldsymbol{\sigma}_n}{\sqrt{n}}, \bar{\boldsymbol{C}}_n + \frac{1.96\boldsymbol{\sigma}_n}{\sqrt{n}}\Big]\Big) \simeq 95\% ,$$

  where $\bar{\boldsymbol{C}}_n = \frac{\sum_{i=1}^n \boldsymbol{C}_i}{n}$ is the empirical mean and $\boldsymbol{\sigma}_n = \sqrt{\frac{\sum_{i=1}^n (\boldsymbol{C}_i - \bar{\boldsymbol{C}}_n)^2}{n-1}}$ the empirical standard semi-deviation.

Bounds

- Exact lower bound on the value of the problem: $\underline{V}_0^{(k)}(x_0)$.
- Exact upper bound on the value of the problem:

$$\mathbb{E}\left(\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t^{(k)}, \boldsymbol{U}_t^{(k)}, \boldsymbol{\xi}_t) + K(\boldsymbol{X}_T)\right)$$

  where $\boldsymbol{X}_t^{(k)}$ and $\boldsymbol{U}_t^{(k)}$ are the trajectories induced by $\underline{V}_t^{(k)}$.
- This bound cannot be computed exactly, but can be estimated by Monte-Carlo method as follows
  - Draw N scenarios $\{\xi_1^n, \dots, \xi_t^n\}$.
  - Simulate the corresponding N trajectories $X_t^{(k),n}, \boldsymbol{U}_t^{(k),n}$, and the total cost for each trajectory $C^{(k),n}$.
  - Compute the empirical mean $\bar{C}^{(k),N}$ and standard dev. $\sigma^{(k),N}$.
  - Then, with confidence 95% the upper bound on the problem is

$$\Big[\bar{C}^{(k),N} - \frac{1.96\sigma^{(k),N}}{\sqrt{N}}, \underbrace{\bar{C}^{(k),N} + \frac{1.96\sigma^{(k),N}}{\sqrt{N}}}_{UB_k}\Big]$$

Stopping rule

- One stopping test consist in fixing an a priori relative gap ε, and stopping if

$$\frac{UB_k - V_0^{(k)}(x_0)}{V_0^{(k)}(x_0)} \leq \varepsilon$$

  in which case we know that the solution is ε-optimal with probability 97.5%.
- It is not necessary to evaluate the gap at each iteration.
- To alleviate the computational load, we can estimate the upper bound by using the trajectories of the recent forward phases.
- Another more practical stopping rule consists in stopping after a given number of iterations or fixed computation time.

Contents

Non-independent inflows

- In most cases the stagewise independence assumption is not realistic.
- One classical way of modelling dependencies consists in considering that the inflows I_t follow an AR-k process

$$I_t = \alpha_1 I_{t-1} + \cdots + \alpha_k I_{t-k} + \beta_t + \xi_t$$

where ξ_t is the residual, forming an independent sequence.
- The state of the system is now $(X_t, I_{t-1}, \dots, I_{(t-k)})$.

A few other implementations

- We presented DOASA: select one scenario (one realization of $(\xi_1, \dots, \xi_{T-1})$) to do a forward and backward path
- Classical SDDP: select a number N of scenarios to do the forward path (computation can be parallelized); then during the backward path we add N cuts to V_t before computing the cuts on V_{t-1}.
- CUPPS algorithm suggests to use $\underline{V}_{t+1}^{(k)}$ instead of $\underline{V}_{t+1}^{(k+1)}$ in the computation of the cuts. In practice:
  - select randomly a scenario $(\xi_t)_{t=0,\dots,T-1}$
  - at time t we have a state $x_t^{(k)}$, we compute the new cut for V_t
  - choose the optimal control corresponding to the realization $W_t = w_t$ in order to compute the state $x_{t+1}^{(k)}$ where the cut for V_{t+1} will be computed, and go to the next step

Numerical tricks

- We can compute some cuts before starting the algorithm. For example by bypassing the forward phase by properly choosing the trajectory $(x_t^{(k)})_{t=0,\dots,T}$.
- With iterations the number of cuts can become exceedingly large and pruning (i.e. eliminate some cuts) can be numerically efficient.
- Eliminate some non-convexity through Lagrange dualization of the non-convex constraint.
- The number of simulations in the forward phase can vary throughout the algorithm, leading to better numerical results.

Cut Selection methods

- Let $\underline{V}_t^{(k)}$ be defined as $\max_{\ell \leq k} \left\{ \beta_t^{(\ell)} + \left\langle \lambda_t^{(\ell)}, \cdot - x_t^{(l-1)} \right\rangle \right\}$.
- For $j \leq k$, if

$$\begin{aligned} \min_{x,\alpha} \quad & \alpha - \left(\beta_t^{(j)} + \left\langle \lambda_t^{(j)}), x - x_t^{(j-1)} \right\rangle \right) \\ s.t. \quad & \alpha \geq \beta_t^{(\ell)} + \left\langle \lambda_t^{(\ell)}, x - x_t^{(\ell-1)} \right\rangle \qquad \forall \ell \neq j \end{aligned}$$

  is non-negative, then cut j can be discarded without modifying $\underline{V}_t^{(k)}$
- this technique is exact but time-consuming.

Cut Selection methods II

- Instead of comparing a cut everywhere, we can choose to compare it only on the already visited points.
- The Level-1 cut method goes as follow:
  - keep a list of all visited points $x_t^{(\ell)}$ for $\ell \leq k$.
  - for ℓ from 1 to k, tag each cut that is active at $x_t^{(\ell)}$.
  - Discard all non-tagged cut.

Contents

Coherent Risk Measure I

To take into account some risk aversion we can replace the expectation by a *risk measure*. A risk measure is a function giving to a random cost $\boldsymbol{X}$ a determinitic equivalent $\rho(\boldsymbol{X})$ A Coherent Risk Measure $\rho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is a functionnal satisfying

- Monotonicity: if $\boldsymbol{X} \geq \boldsymbol{Y}$ then $\rho(\boldsymbol{X}) \geq \rho(\boldsymbol{Y})$,
- Translation equivariance: for $c \in \mathbb{R}$ we have $\rho(\boldsymbol{X} + c) = \rho(\boldsymbol{X}) + c$,
- Convexity: for $t \in [0, 1]$, we have

$$\rho(t\boldsymbol{X} + (1-t)\boldsymbol{Y}) \leq t\rho(\boldsymbol{X}) + (1-t)\rho(\boldsymbol{Y}),$$

- Positive homogeneity: for $\lambda \in \mathbb{R}^+$, we have $\rho(\lambda\boldsymbol{X}) = \lambda\rho(\boldsymbol{X})$.

Coherent Risk Measure II

From convex analysis we obtain the main theorem over coherent risk measure.

Theorem

Let ρ be a coherent risk measure, then there exists a (convex) set of probability $\mathcal{P}$ such that

$$\forall \boldsymbol{X}, \qquad \rho(\boldsymbol{X}) = \sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[\boldsymbol{X}].$$

Average Value at Risk

One of the most practical and used coherent risk measure is the Average Value at Risk at level α. Roughly, it is the expectation of the cost over the α-worst cases. For a random variable $\boldsymbol{X}$ admitting a density, we define de value at risk of level α, as the quantile of level α, that is

$$VaR_\alpha(\boldsymbol{X}) = \inf \Big\{ t \in \mathbb{R} \quad | \quad \mathbb{P}(\boldsymbol{X} \geq t) \leq \alpha \Big\}.$$

And the average value at risk is

$$AVaR_\alpha(\boldsymbol{X}) = \mathbb{E}\big[\boldsymbol{X} \mid \boldsymbol{X} \geq VaR_\alpha(\boldsymbol{X})\big]$$

Average Value at Risk II

One of the best aspect of the AVaR, is the following formula

$$AVaR_\alpha(\boldsymbol{X}) = \min_{t\in\mathbb{R}} \Big\{ t + \frac{\mathbb{E}\big[X - t\big]^+}{\alpha} \Big\}.$$

Indeed it allow to linearize the AVaR.

SDDP and risk

- The problem studied was risk neutral
- However a lot of works has been done recently about how to solve risk averse problems
- Most of them are using AVAR, or a mix between AVAR and expectation either as objective or constraint
- Indeed AVAR can be used in a linear framework by adding other variables
- Another easy way is to use "composed risk measures"
- Finally a convergence proof with convex costs (instead of linear costs) exists, although it requires to solve non-linear problems

Contents

Assumptions

- Noises are time-independent, with finite support.
- Decision and state constraint sets are compact convex subsets of finite dimensional spaces.
- Dynamic is linear, costs are convex and lower semicontinuous.
- There is a strict relatively complete recourse assumption.

Remark, if we take the tree-view of the algorithm

- stage-independence of noise is not required to have theoretical convergence
- node-selection process should be admissible (e.g. independent, SDDP, CUPPS...)

Convergence result

Theorem

With the preceding assumption, we have that the upper and lower bound are almost surely converging toward the optimal value, and we can obtain an $\varepsilon-$optimal strategy for any $\varepsilon > 0$.
More precisely, if we call $\underline{V}_t^{(k)}$ the outer approximation of the Bellman function V_t at step k of the algorithm, and $\pi_t^{(k)}$ the corresponding strategy, we have

$$\underline{V}_0^{(k)}(x_0) \to_k V_0(x_0)$$

and

$$\mathbb{E}\Big[L_t\big(\boldsymbol{x}_t^{(k)}, \pi_t^{(k)}(\boldsymbol{x}_t^{(k)}), \boldsymbol{\xi}_t\big) + \underline{V}_{t+1}^{(k)}(\boldsymbol{x}_{t+1}^{(k)})\Big] \to_k V_t(\boldsymbol{x}_t^{(k)}).$$

Contents

Conclusion

SDDP is an algorithm, more precisely a class of algorithms, that

- exploits convexity of the value functions (from convexity of costs...)
- does not require state discretization
- constructs outer approximations of V_t, those approximations being precise only "in the right places"
- gives bounds:
  - "true" lower bound $\underline{V}_0^{(k)}(x_0)$
  - estimated (by Monte-Carlo) upper bound
- constructs linear-convex approximations, thus enabling to use linear solver like CPLEX
- can be shown to display asymptotic convergence

Bibliography

R. Van Slyke and R. Wets (1969).
L-shaped linear programs with applications to optimal control and stochastic programming.
SIAM Journal on Applied Mathematics

M. Pereira, L.Pinto (1991).
Multi-stage stochastic optimization applied to energy planning
Mathematical Programming

A. Shapiro (2011).
Analysis of stochastic dual dynamic programming method.
European Journal of Operational Research.

P.Girardeau, V.Leclère, A. Philpott (2014).
On the convergence of decomposition methods for multi-stage stochastic convex programs.
Mathematics of Operations Research.