A general Dynamic Programming Principle

V. Leclère (ENPC), M. De Lara (ENPC).

July 3, 2016

Contents

Contents

Classical optimization problem

$$\begin{aligned} \min_{s} \quad & \mathbb{E}\Big[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t) + K(\mathbf{X}_T)\Big] \\ \text{s.t.} \quad & \mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t) \\ & \mathbf{U}_t = s_t(\mathbf{X}_t) \end{aligned}$$

- $\mathbf{X}_t$: state,
- $\mathbf{U}_t$: control,
- $\mathbf{W}_t$: uncertainty,
- $s = (s_1, \dots, s_{T-1})$: strategy.

Time consistency of a sequence of problem

$$(\mathcal{P}_{t_0}) \qquad \min_{s} \quad \mathbb{E}\Big[\sum_{t=t_0}^{T} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t) + K(\mathbf{X}_T)\Big]$$

$$\text{s.t.} \quad \mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t)$$

$$\mathbf{U}_t = s_t(\mathbf{X}_t)$$

The sequence of problems $(\mathcal{P}_t)_{t\in[\![0,T-1]\!]}$ is said to be *time consistent* iff an optimal strategy $(s_t^{t_0})_{t\in[\![t_0,T-1]\!]}$ of Problem $(\mathcal{P}_{t_0})$ is such that its restriction to $[\![t_1, T-1]\!]$, $(s_t^{t_0})_{t\in[\![t_1,T-1]\!]}$ is optimal for $(\mathcal{P}_{t_1})$, $(\forall t_1 > t_0)$.

Dynamic Programming

The sequence of Bellman functions $(V_t)_{t\in[\![0,T]\!]}$ defined by

$$
\begin{aligned}
V_{t_0}(x) \quad = \quad & \min_{s} \quad \mathbb{E}\Big[\sum_{t=t_0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t) + K(\mathbf{X}_T)\Big], \\
\text{s.t.} \quad & \mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t), \\
& \mathbf{U}_t = s_t(\mathbf{X}_t), \\
& \mathbf{X}_{t_0} = x,
\end{aligned}
$$

satisfies the Bellman equation

$$
\begin{cases}
V_{T+1}(x) & = \quad K(x), \\
V_t(x) & = \quad \min_u \mathbb{E}\Big[L_t(x, u, \mathbf{W}_t) + V_{t+1} \circ f_t(x, u, \mathbf{W}_t)\Big].
\end{cases}
$$

Contents

Aggregation in time

$$\min_{s} \quad \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t) + K(\mathbf{X}_T)\right],$$

$$\text{s.t.} \quad \mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t),$$

$$\mathbf{U}_t = s_t(\mathbf{X}_t).$$

Here we consider the sum over time. Other choices :

- discounted sum $\sum_{t=0}^{T-1} \rho^t L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t) + \rho^T K(\mathbf{X}_T)$,
- maximin (Rawls) $\max_{t \in [\![0, T-1]\!]} \{L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t), K(\mathbf{X}_T)\}$,
- product, for example for a discounted final cost with a controlled interest rate, $\left(\prod_{t=0}^{T-1} \rho(\mathbf{X}_t)\right) \times K(\mathbf{X}_T)$.

Aggregation in uncertainty

$$\begin{aligned} \min_{s} \quad & \mathbb{E}\left(\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t) + K(x_T) \right), \\ \text{s.t.} \quad & \mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t), \\ & \mathbf{U}_t = s_t(\mathbf{X}_t). \end{aligned}$$

Here we consider the expectation over time. Other choices :

- expectation over another probability $\mathbb{E}_{\mathbb{Q}}\big(\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t) + K(\mathbf{X}_T)\big)$,
- worst case aggregator $\max_w \Big(\sum_{t=0}^{T-1} L_t(X_t, U_t, w_t) + K(X_T) \Big)$,
- coherent risk measures $\sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}\big(\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_t) + K(\mathbf{X}_T)\big)$.

Contents

Dynamic system and state

- Sets of states : $\mathbb{X}_t$, and $\mathbb{X} = \mathbb{X}_1 \times \cdots \times \mathbb{X}_T$.
- Sets of uncertainties : $\mathbb{W}_t$, and $\mathbb{W} = \mathbb{W}_1 \times \cdots \times \mathbb{W}_T$.
- Sets of controls $\mathbb{U}_t$, and $\mathbb{U} = \mathbb{U}_1 \times \cdots \times \mathbb{U}_T$.
- Dynamics $f_t : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_t \to \mathbb{X}_{t+1}$.
- Stage-costs $L_t : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_t \to \mathbb{R} \cup \{+\infty\}$,
- Control constraints $U_t : \mathbb{X}_t \rightrightarrows \mathbb{U}_t$.

A *strategy* $s = (s_t)_{t \in [\![1,T]\!]}$ is a sequence of functions such that for all $t \in [\![1, T]\!]$, s_t maps the set of states $\mathbb{X}_t$ into the set of control $\mathbb{U}_t$.

A *trajectory* $x^s(w)$ follows

$$\forall t \in [\![1, T]\!], \qquad x^s_{t+1}(w_t) = f_t(x_t, s_t(x_t), w_t) \ .$$

Stream of costs

For any admissible strategy s, we have a corresponding stream of costs

$$\Big(L_1\big(x_1^s, s_1(x_1^s), W_1\big), \cdots L_T\big(x_T^s, s_T(x_T^s), W_T\big)\Big)$$

which is a stochastic process. To choose the "best" strategy we first need to determine a way of comparing streams of costs.
Note that:

- there are T costs,
- each of them is uncertain.

They need to be aggregated into a single number to be compared.

Time and uncertainty aggregators

- A *time-aggregator* is given as $\Psi : \bar{\mathbb{R}}^{T+1} \to \bar{\mathbb{R}}$, (for example the sum discounted or not, the product, ...)
- An *uncertainty-aggregator* is given as $\mathbb{G} : \mathcal{F}(\mathbb{W}, \bar{\mathbb{R}}) \to \bar{\mathbb{R}}$ (for example an expectation, a worst case,...).

Thus the problem considered is

$$\min_{s \in S} \quad \mathbb{G}\Big[w \mapsto \Psi\Big\{L_1\big(x_1, u_1, w_1\big), \dots, L_T\big(x_T, u_T, w_T\big), K\big(x_{T+1}\big)\Big\}\Big],$$

$$x_{t+1} = f_t\big(x_t, u_t, w_t\big),$$

$$u_t = s_t(x_t),$$

$$u_t \in U_t(x_t).$$

Other problems

We denote $L_t^s = \big(x_t^s, s_1(x_t^s), W_t\big)$. We considered the problem of minimizing

$$(TU) \qquad \mathbb{G}\Big[\Psi\Big\{L_1^s, L_2^s, \dots, L_T^s\Big\}\Big]$$

But we could also have considered

$$(UT) \qquad \Psi\Big\{\mathbb{G}_1\big[L_1^s\big], \mathbb{G}_2\big[L_2^s\big] \dots, \mathbb{G}_T\big[L_T^s\big]\Big\}$$

or

$$(NTU) \qquad \mathbb{G}_1\Big[\Psi_1\Big\{L_1^s, \mathbb{G}_2\Big[\Psi_2\Big\{L_2^s, \cdots \Psi_{T-1}\big\{L_{T-1}^s, \mathbb{G}_T[L_T^s]\big\}\Big\}\Big]\Big\}\Big]$$

or inversing the order of aggregators we could write an (NUT) problem.

Contents

Composition of time aggregator

We consider $\Psi_t : \bar{\mathbb{R}} \times \bar{\mathbb{R}} \to \bar{\mathbb{R}}$.
We define the composition $\Psi_1 \odot \Psi_2 : \bar{\mathbb{R}}^3 \to \bar{\mathbb{R}}$ by

$$\Psi_1 \odot \Psi_2\{C_1, C_2, C_3\} := \Psi_1\Big\{C_1, \Psi_2\{C_2, C_3\}\Big\}.$$

More generally $\Psi = \Psi_1 \odot \cdots \odot \Psi_t : \bar{\mathbb{R}}^t \to \bar{\mathbb{R}}$ is recursively defined by, for all $c \in \bar{\mathbb{R}}^{t+1}$,

$$\Psi_{t'} \odot \cdots \odot \Psi_t\{c\} := \Psi_{t'}\Big\{c_{t'}, \Psi_{t'+1} \odot \cdots \odot \Psi_t\{c_t, \dots, c_{t+1}\}\Big\} .$$

Composition of uncertainty aggregators

We consider $\mathbb{G}_t : \mathcal{L}\big(\mathbb{W}_t; \bar{\mathbb{R}}\big) \to \bar{\mathbb{R}}$.
We define $\mathbb{G}_1 \boxdot \mathbb{G}_2 : \mathcal{F}\big(\mathbb{W}_1 \times \mathbb{W}_2, \bar{\mathbb{R}}\big) \to \bar{\mathbb{R}}$ by

$$\mathbb{G}_1 \boxdot \mathbb{G}_2\big[A\big] := \mathbb{G}_1\Big[w_1 \mapsto \mathbb{G}_2\big[w_2 \mapsto A(w_1, w_2)\big]\Big].$$

Their composition $\mathbb{G} = \mathbb{G}_1 \boxdot \cdots \boxdot \mathbb{G}_t : \mathcal{L}\big(\mathbb{W}; \bar{\mathbb{R}}\big) \to \bar{\mathbb{R}}$ is recursively defined by

$$\begin{aligned} &\mathbb{G}_{t'} \boxdot \cdots \boxdot \mathbb{G}_t\big[A\big] := \\ &\quad \mathbb{G}_{t'}\Big[w_{t'} \mapsto \mathbb{G}_{t'+1} \boxdot \cdots \boxdot \mathbb{G}_t \\ &\qquad\qquad \big[(w_{t'+1}, \dots, w_t) \mapsto A(w_{t'}, w_{t'+1}, \dots, w_t)\big]\Big] . \end{aligned}$$

Examples

If

$$\Psi\{c_1, c_2\} = c_1 + c_2, \quad \Psi'\{c_1, c_2\} = c_1 c_2,$$

then we have

$$\begin{aligned} \Psi \odot \Psi\{c_1, c_2, c_3\} &= c_1 + c_2 + c_3, \\ \Psi' \odot \Psi'\{c_1, c_2, c_3\} &= c_1 c_2 c_3, \\ \Psi \odot \Psi'\{c_1, c_2, c_3\} &= c_1 + c_2 c_3, \\ \Psi' \odot \Psi\{c_1, c_2, c_3\} &= c_1(c_2 + c_3). \end{aligned}$$

Moreover

$$\mathbb{G}_1 := \mathbb{E}_{\mathbb{P}_1}, \quad \mathbb{G}_2 := \mathbb{E}_{\mathbb{P}_2}, \qquad \mathbb{G}_1 \boxdot \mathbb{G}_2 = \mathbb{E}_{\mathbb{P}_1 \otimes \mathbb{P}_2}.$$

Commutation

A sequence of time aggregators $(\Psi_1, \dots, \Psi_t)$ is said to *commute* with a sequence of noise aggregators $(\mathbb{G}_1, \dots, \mathbb{G}_t)$ if for all $C_t \in \mathcal{L}(\mathbb{W}_t, \bar{\mathbb{R}})$,

$$\left(\mathbb{G}_1 \boxdot \cdots \boxdot \mathbb{G}_t\right)\left[\left(\Psi_1 \odot \cdots \odot \Psi_t\right)(C_1, \dots, C_{t+1})\right] = \mathbb{G}_1\left[\Psi_1\left(C_1, \mathbb{G}_2\left[\dots \mathbb{G}_t\left[\Psi_t\{C_t, C_{t+1}\}\right]\right]\right)\right] .$$

or, loosely speaking,

$$\left(\mathbb{G}_1 \boxdot \cdots \boxdot \mathbb{G}_t\right) \circ \left(\Psi_1 \odot \cdots \odot \Psi_t\right) = \left(\mathbb{G}_1 \circ \Psi_1\right) \boxdot \cdots \boxdot \left(\mathbb{G}_t \circ \Psi_t\right) .$$

Strong commutation

A noise aggregator $\mathbb{G}$ is said to strongly commute with a time aggregator Ψ if for any $A \in \mathcal{L}(\mathbb{W}; \bar{\mathbb{R}})$, and any $c \in \bar{\mathbb{R}}$,

$$\mathbb{G}\big[w \mapsto \Psi\{c, A(w)\}\big] = \Psi\Big\{c, \mathbb{G}\big[w \mapsto A(w)\big]\Big\}$$

Theorem

If for any $t \in [\![1, T-1]\!]$, $\mathbb{G}_{t+1}$ strongly commute with Ψ_t, then the sequence $(\mathbb{G}_1, \dots, \mathbb{G}_T)$ commutes with the sequence $(\Psi_1, \dots, \Psi_T)$.

Example : If $\mathbb{W}_2$ is a probability space $(\mathbb{W}_2, \mathcal{F}, \mathbb{P})$ and $\Psi_1(c_1, c_2) = c_1 + c_2$, then the extended expectation $\bar{\mathbb{E}}_{\mathbb{P}}$ is strongly commuting with Ψ_1.

General Dynamic Programming Principle

Theorem

Assume that $(\mathbb{G}_1, \dots, \mathbb{G}_T)$ *commute with* $(\Psi_1, \dots, \Psi_T)$*, and* $\Psi_t\{C_1, \cdot\}$ *is non decreasing and* $\mathbb{G}_t$ *is monotonous. If* $\mathbb{G} = \mathbb{G}_1 \boxdot \cdots \boxdot \mathbb{G}_T$*, and* $\Psi = \Psi_1 \odot \cdots \odot \Psi_T$*, then we define*

$$V_{T+1}(x) := K(x),$$

$$V_t(x) := \inf_{u \in \mathbb{U}_t(x)} \mathbb{G}_t\Big[\Psi_t\Big\{L_t(x,u,\cdot), V_{t+1} \circ f_t(x,u,\cdot)\Big\}\Big],$$

$$s_t^*(x) \in \arg\min_{u \in U_t(x)} \mathbb{G}_t\Big[\Psi_t\Big\{L_t(x,u,\cdot), V_{t+1} \circ f_t(x,u,\cdot)\Big\}\Big].$$

If s^* *exists and is admissible then* s^* *is an optimal strategy, and its value is* $V_0(x)$.

Contents

Maximum and infimum

Theorem

Let Ψ be a time aggregator non-decreasing in its second argument. If $(\mathbb{G}^i)_{i\in I}$ is a family of monotonous random aggregators strongly commuting with Ψ. Moreover if we have either one of the two following conditions

1. *for all $C' \in \mathcal{L}(\mathbb{W}, \bar{\mathbb{R}})$, $\sup_{i\in I} \mathbb{G}^i[C']$ is attained (always verified for I finite);*
2. *for all $C \in \bar{\mathbb{R}}$ $\Psi\{C, \cdot\}$ is lower-semicontinuous (resp. upper-semicontinuous);*

then the aggregator $\bar{\mathbb{G}} := \sup_{i\in I} \mathbb{G}^i$ (resp. $\underline{\mathbb{G}} := \inf_i \mathbb{G}^i$) is monotonous and strongly commute with Ψ.

Change of variable

Theorem

Let ν be an increasing bijection from $I \subset \bar{\mathbb{R}}$ onto $J \subset \bar{\mathbb{R}}$. Assume that we have an aggregator such that the image of $\mathbb{G} \circ \nu$ is contained in I. We define $\tilde{\mathbb{G}}_t := \nu^{-1} \circ \mathbb{G} \circ \nu$.
In a similar fashion, we define $\tilde{\Psi}_t$ given by, for all $C \in \bar{\mathbb{R}}^2$,

$$\tilde{\Psi}_t : (C_1, C_2) \mapsto \nu^{-1}\Big(\Psi_t\big\{C_1, \nu(C_2)\big\}\Big) .$$

If the sequence $(\mathbb{G}_1, \dots, \mathbb{G}_t)$ commute with $(\Psi_1, \dots, \Psi_t)$, then the sequence $(\tilde{\mathbb{G}}_1, \dots, \tilde{\mathbb{G}}_t)$ commute with $(\tilde{\Psi}_1, \dots, \tilde{\Psi}_t)$. Finally if $\mathbb{G}_t$ is monotonous then so is $\tilde{\mathbb{G}}_t$, and if $\Psi_t\big\{C, \cdot\big\}$ is monotonous so is $\tilde{\Psi}_t\big\{C, \cdot\big\}$.

Contents

Coherent risk measures 1/2

$$\min_{s \in S^{ad}} \sup_{\mathbb{P} \in \mathcal{P}} \bar{\mathbb{E}}_{\mathbb{P}} \left[\sum_{t=1}^{T} L_t(x_t, s_t(x_t), W_t) + K(x_{T+1}) \right].$$

$$s.t.\ x_{t+1} = f_t(x_t, s_t(x_t), W_t)$$

Where $\mathcal{P} = \mathcal{P}_1 \otimes \cdots \otimes \mathcal{P}_T$, $\mathcal{P}_t$ is a set of probability, and $\beta_t > 0$.
This problem can be solved by using the following DP Principle

$$\begin{aligned} V_{T+1}(x) &= K(x) \,, \\ V_t(x) &= \min_{u \in U_t(x)} \sup_{\mathbb{P}_t \in \mathcal{P}_t} \left\{ \bar{\mathbb{E}}_{\mathbb{P}_t} \big(L_t(x, u, W_t) + V_{t+1} \circ f_t(x, u, W_t) \big) \right\} . \end{aligned}$$

Coherent risk measures 2/2

$$\min_{s \in S^{ad}} \sup_{\mathbb{P} \in \mathcal{P}} \bar{\mathbb{E}}_{\mathbb{P}} \Bigg[\sum_{t=1}^{T} \Big\{ \alpha_t \big(L_t(x_t, s_t(x_t), W_t) \big) \prod_{t' < t} \beta_{t'} \big(L_{t'}(x_{t'}, s_{t'}(x_{t'}), W_{t'}) \big) \Big\} + \prod_{t=1}^{T} \beta_t \big(L_t(x_t, s_t(x_t), W_t) \big) K(x_{T+1}) \Bigg].$$

$$s.t.\ x_{t+1} = f_t(x_t, s_t(x_t), W_t)$$

Where $\mathcal{P} = \mathcal{P}_1 \otimes \cdots \otimes \mathcal{P}_T$, $\mathcal{P}_t$ is a set of probability. This problem can be solved by using the following DP Principle

$$V_{T+1}(x) = K(x) \, ,$$

$$V_t(x) = \min_{u \in U_t(x)} \sup_{\mathbb{P}_t \in \mathcal{P}_t} \Big\{ \bar{\mathbb{E}}_{\mathbb{P}_t} \Big[\alpha_t \big(J_t(x, u, W_t) \big) + \beta_t \big(L_t(x, u, W_t) \big) \cdot V_{t+1} \circ f_t(x, u, W_t) \Big]$$

Proof

This problem fit the framework with

- $\Psi = \Psi_1 \odot \cdots \odot \Psi_T$,
- $\mathbb{G} = \mathbb{G}_1 \odot \cdots \odot \mathbb{G}_T$,
- $\Psi_t\{C_1, C_2\} = \alpha_t(C_1) + \beta_t(C_1)C_2$,
- $\mathbb{G}_t[A] = \sup_{\mathbb{P}_t \in \mathcal{P}_t} \bar{\mathbb{E}}_{\mathbb{P}_t}[A]$.

We show that if β_t is non-negative, $\mathbb{G}_{t+1}$ strongly-commute with Ψ_t.

Convex risk measures

$$\min_{s \in S^{ad}} \quad \sup_{\mathbb{P} \in \mathcal{P}} \bar{\mathbb{E}}_{\mathbb{P}} \left[\sum_{t=1}^{T} \left\{ \alpha_t \big(L_t(x_t, s_t(x_t), W_t) \big) + K(x_{T+1}) \right\} - g\left(\mathbb{P}\right) \right],$$

$$s.t. \quad x_{t+1} = f_t(x_t, s_t(x_t), W_t)$$

where $g\left(\mathbb{P}\right) = \sum_{t=1}^{T} g_t(\mathbb{P}_t)$.
Can be solved by using the following Dynamic Programming Principle

$$V_{T+1}(x) = K(x) \,,$$

$$V_t(x) = \min_{u \in U_t(x)} \sup_{\mathbb{P}_t \in \mathcal{P}_t} \left\{ \bar{\mathbb{E}}_{\mathbb{P}_t} \left[\alpha_t \big(L_t(x, u, W_t) \big) + V_{t+1} \circ f_t(x, u, W_t) \right] \right\}$$

Contents

Fear operator

Let Ψ_t be non decreasing in its second argument, and $\tilde{\mathbb{W}} = \tilde{\mathbb{W}}_1 \times \cdots \times \tilde{\mathbb{W}}_t$. Then

$$\min_{s \in S^{ad}} \sup_{w \in \tilde{\mathbb{W}}} \Psi_1 \Big\{ L_1\big(x_1, s_1(x_1), w_1\big), \Psi_2\Big\{ \cdots, \Psi_T\Big\{ L_T(), K(x_{T+1}(w_T))\Big\}\Big\}\Big\}$$

$$s.t. x_{t+1} = f_t(x_t, s_t(x_t), W_t),$$

can be solved by using the following DP Principle

$$V_{T+1}(x) = K(x) ,$$

$$V_t(x) = \min_{u \in U_t(x)} \sup_{w_t \in \tilde{\mathbb{W}}_t} \Big\{ \Psi_t\big\{ L_t\big(x, u, w_t\big), V_{t+1} \circ f_t\big(x, u, w_t\big)\big\}\Big\} .$$

Risk sensitive approach

$$\inf_{s \in S^{ad}} -\frac{1}{\gamma} \ln \left(\bar{\mathbb{E}}\Big[\sum_{t=1}^{T} \beta^{t-1} e^{-\gamma L_t(x_t, s_t(x_t), W_t)} + \beta^T e^{-\gamma K(x_T)} \Big] \right) .$$

Can be solved by using the following Dynamic Programming Principle

$$\begin{aligned} V_{T+1}(x) &= K(x) , \\ V_t(x) &= \min_{u \in U_t(x)} -\frac{1}{\gamma} \ln \left(\bar{\mathbb{E}}\Big[e^{-\gamma L_t(x_t, s_t(x_t), W_t)} + \beta e^{-\gamma V_{t+1} \circ f_t(x_t, s_t(x_t), W_t)} \Big] \right. \end{aligned}$$

Conclusion

- Dynamic programming principle yield time consistency result.
- Dynamic programming can be used with aggregators that differs from the sum and the expectation.
- We mainly require :
  - that the aggregators are monotonous,
  - that the global aggregator is given as a composition of instantaneous ones,
  - and that they commutes.
- This framework can be extended to the conditionnal case and linked to time-consistency results known in the risk measure litterature.

The end

De Lara, Michel and Leclère, Vincent (2016).
Building-up Time-Consistency for Risk Measures and Dynamic Optimization.
European Journal of Operations Research

Thank you for your attention !