# Consistance Dynamique et Commande Optimale Stochastique

## Groupe de travail FiME

Pierre Girardeau[1,2,3]
Travail conjoint avec Pierre Carpentier[2], Jean-Philippe Chancelier[3], Guy Cohen[3] et Michel De Lara[3]

[1]OSIRIS, EDF R&D, France

[2]ENSTA-ParisTech, France

[3]École des Ponts-ParisTech, CERMICS, France

22 octobre 2010

# Plan

1. Motivation
   - Formulation du problème
   - Mesures de risque dynamiques
   - Notion de consistance
2. Programmation Dynamique et Consistance
   - Cas déterministe
   - Cas stochastique sans contrainte de risque
   - Cas stochastique avec contrainte de risque
3. Exemples

# Plan

# Introduction

## Commande optimale stochastique

- On considère un **système dynamique** influencé par des **bruits** exogènes.
- À chaque pas de temps :
  - des **observations** sont faites sur le système et conservées.
  - une **décision** est prise, et le système évolue.
- On cherche à **minimiser un certain objectif** fonction de l'évolution du système.

# Introduction
## Consistance dynamique

**La propriété de consistance dynamique concerne une suite de problèmes de décision.**

### Définition informelle de la consistance

- On peut souvent écrire un **nouveau problème** de commande optimale aux **instants ultérieurs** en "tronquant" le problème original.
- On se demande si la commande issue du problème original **reste optimale pour les problèmes de décision ultérieurs** [Ekeland and Lazrak, 2006].

# Introduction
## Consistance dynamique

**La propriété de consistance dynamique concerne une suite de problèmes de décision.**

### Définition informelle de la consistance

- On peut souvent écrire un **nouveau problème** de commande optimale aux **instants ultérieurs** en “tronquant” le problème original.
- On se demande si la commande issue du problème original **reste optimale pour les problèmes de décision ultérieurs** [Ekeland and Lazrak, 2006].

# Introduction
## Pour fixer les idées

$\boldsymbol{X}_t$ : stock[1], $\boldsymbol{U}_t$ : commande, $\boldsymbol{W}_t$ : bruit.

### Problème standard

$$\begin{aligned} \min_{\boldsymbol{X},\boldsymbol{U}} \quad & \mathbb{E}\left(\sum_{t=t_0}^{T-1} L_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right) + K\left(\boldsymbol{X}_T\right)\right), \\ \text{s.c.} \quad & \boldsymbol{X}_{t_0} \text{ est donné}, \end{aligned}$$

et, pour tout $t = t_0, \ldots, T-1$ :

$$\begin{aligned} & \boldsymbol{X}_{t+1} = f_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right), \\ & \boldsymbol{U}_t \text{ est } \sigma\{\boldsymbol{X}_{t_0}, \boldsymbol{W}_{t_1}, \ldots, \boldsymbol{W}_t\}\text{-mesurable.} \end{aligned}$$

1. On l'appellera état sous certaines hypothèses.

# Introduction
## Motivation(s)

**État (inspiré de [Whittle, 1982, Section 1.1])**

C'est une **variable** $x$ valant $x_t$ à l'instant $t$ et **suffisante** pour que :

- la **décision optimale à l'instant** $t$ ne dépende que de $x_t$ et de $t$ ;
- la variable $x_{t+1}$ soit calculable à partir de $x_t$, de la décision à l'instant $t$ et de la nouvelle observation à $t+1$.

Pour beaucoup de systèmes, on a :

- un **état "habituel"** (stock d'énergie, mémoire d'un processus) qui découle du principe de **programmation dynamique** ;
- **consistance dynamique** du fait du même principe.

# Introduction

## Motivation(s)

### État (inspiré de [Whittle, 1982, Section 1.1])

C'est une **variable** $x$ valant $x_t$ à l'instant $t$ et **suffisante** pour que :

- la **décision optimale à l'instant** $t$ ne dépende que de $x_t$ et de $t$ ;
- la variable $x_{t+1}$ soit calculable à partir de $x_t$, de la décision à l'instant $t$ et de la nouvelle observation à $t+1$.

Pour beaucoup de systèmes, on a :

- un **état "habituel"** (stock d'énergie, mémoire d'un processus) qui découle du principe de **programmation dynamique** ;
- **consistance dynamique** du fait du même principe.

# Introduction
## Motivation(s)

On peut alors chercher à :

- **changer le critère** en espérance à l'aide d'une autre mesure de risque,
- ajouter de **nouvelles contraintes**.

Parfois, la suite de problèmes de décision devient **inconsistante** [Haviv, 1996, Ruszczyński, 2009, Shapiro, 2009].

**Objectif de la présentation**

Étudier la **consistance dynamique** d'une suite de problèmes à la lumière du concept de **variable d'état**.

$\hookrightarrow$ Plus de précisions dans [Carpentier et al., 2010].

# Motivation 1

## Mesures de risque dynamiques

### Littérature des mesures de risque

- De nombreux travaux sur la cohérence et la consistance des mesures de risque [Detlefsen and Scandolo, 2005, Cheridito et al., 2006, Artzner et al., 2007].

$\hookrightarrow$ Ce ne sera pas notre propos ici.

### Littérature *Stochastic Programming*

- Plusieurs travaux [Ruszczyński, 2009, Shapiro, 2009] s'attachent à **introduire de l'aversion au risque** adns les problèmes d'optimisation stochastique.
- Ils cherchent alors des **classes de mesures de risque préservant une certaine structure** pour le problème.

$\hookrightarrow$ On propose ici un point de vue différent.

# Motivation 1

## Mesures de risque dynamiques

### Littérature des mesures de risque

- De nombreux travaux sur la cohérence et la consistance des mesures de risque [Detlefsen and Scandolo, 2005, Cheridito et al., 2006, Artzner et al., 2007].

$\hookrightarrow$ Ce ne sera pas notre propos ici.

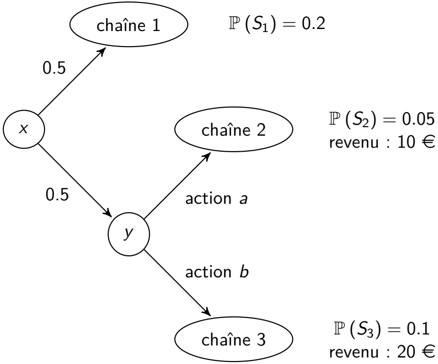
### Littérature *Stochastic Programming*

- Plusieurs travaux [Ruszczyński, 2009, Shapiro, 2009] s'attachent à **introduire de l'aversion au risque** adns les problèmes d'optimisation stochastique.
- Ils cherchent alors des **classes de mesures de risque préservant une certaine structure** pour le problème.

$\hookrightarrow$ On propose ici un point de vue différent.

# Motivation 2

## Exemple [Haviv, 1996]

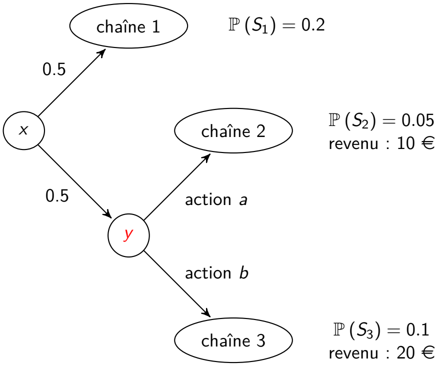



Maximiser le revenu espéré sous la contrainte :

$$\mathbb{P}(S_1 \cup S_2 \cup S_3) \leq 0.125.$$

# Motivation 2

Exemple [Haviv, 1996]

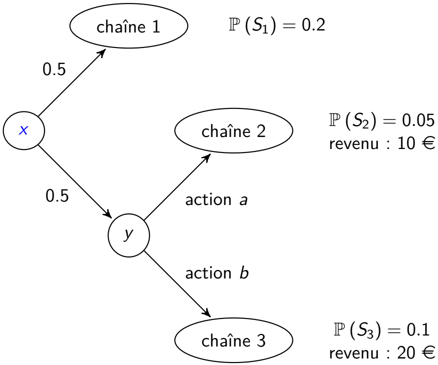



Vu de $y$, l'action $b$ est admissible et optimale :

$$0.1 < 0.125.$$

# Motivation 2

## Exemple [Haviv, 1996]

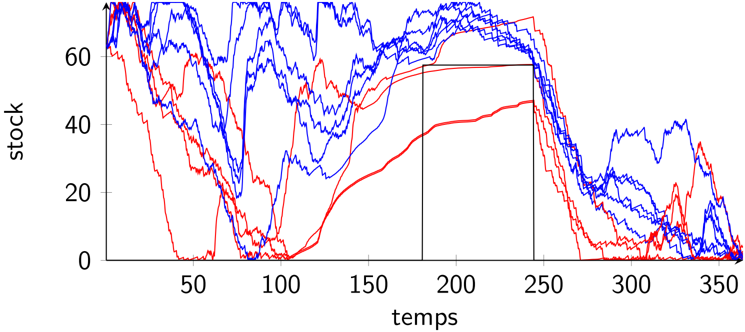



Vu de $x$, l'action $b$ n'est même pas admissible :

$$0.5 \times 0.2 + 0.5 \times 0.1 = 0.15 > 0.125.$$

# Motivation 3

## Problème EDF de la cote touristique




Certaines réglementations liées aux activités estivales près des réservoirs imposent :

$$\mathbb{P}\left(\boldsymbol{X}_t \geq x_{\text{ref}}, \forall t \in \mathbb{T}_{\text{été}}\right) \geq p.$$

# Plan

# Un premier problème...

## Problème partant de $t_0$

$$\begin{aligned} \min_{x,u} \quad & \sum_{t=t_0}^{T-1} L_t\left(x_t, u_t\right) + K\left(x_T\right), \\ \text{s.c.} \quad & x_{t_0} \text{ donné}, \\ & x_{t+1} = f_t\left(x_t, u_t\right), \qquad \forall t = t_0, \ldots, T-1. \end{aligned}$$

## Boucle ouverte

Une solution est une suite de décisions $u^\sharp_{t_0,t_0}, \ldots, u^\sharp_{t_0,T-1}$.

## Attention

La solution dépend de la condition initiale $x_{t_0}$.

# Un premier problème...

## Problème partant de $t_0$

$$\begin{aligned} \min_{x,u} \quad & \sum_{t=t_0}^{T-1} L_t\left(x_t, u_t\right) + K\left(x_T\right), \\ \text{s.c.} \quad & x_{t_0} \text{ donné}, \\ & x_{t+1} = f_t\left(x_t, u_t\right), \qquad \forall t = t_0, \ldots, T-1. \end{aligned}$$

## Boucle ouverte

Une solution est une suite de décisions $u^\sharp_{t_0,t_0}, \ldots, u^\sharp_{t_0,T-1}$.

## Attention

La solution dépend de la condition initiale $x_{t_0}$.

# Un premier problème...

## Problème partant de $t_0$

$$
\begin{aligned}
\min_{x,u} \quad & \sum_{t=t_0}^{T-1} L_t\left(x_t, u_t\right) + K\left(x_T\right), \\
\text{s.c.} \quad & x_{t_0} \text{ donné}, \\
& x_{t+1} = f_t\left(x_t, u_t\right), \qquad \forall t = t_0, \dots, T-1.
\end{aligned}
$$

## Boucle ouverte

Une solution est une suite de décisions $u^\sharp_{t_0,t_0}, \dots, u^\sharp_{t_0,T-1}$.

## Attention

La solution dépend de la condition initiale $x_{t_0}$.

# ... et la suite de problèmes qui en découle

## Problème partant de $t_i$

$$\begin{aligned} \min_{x,u} \quad & \sum_{t=t_i}^{T-1} L_t(x_t, u_t) + K(x_T), \\ \text{s.c.} \quad & x_{t_i} \text{ donné}, \\ & x_{t+1} = f_t(x_t, u_t), \qquad \forall t = t_i, \ldots, T-1. \end{aligned}$$

Soit $u^\sharp_{t_i,t_i}, \ldots, u^\sharp_{t_i,T-1}$ une solution.

## Attention

La solution dépend de la condition initiale $x_{t_i}$.

# ... et la suite de problèmes qui en découle

## Problème partant de $t_i$

$$\min_{x,u} \quad \sum_{t=t_i}^{T-1} L_t\left(x_t, u_t\right) + K\left(x_T\right),$$

$$\text{s.c.} \quad x_{t_i} \text{ donné},$$

$$x_{t+1} = f_t\left(x_t, u_t\right), \qquad \forall t = t_i, \ldots, T-1.$$

Soit $u^\sharp_{t_i,t_i}, \ldots, u^\sharp_{t_i,T-1}$ une solution.

## Attention

La solution dépend de la condition initiale $x_{t_i}$.

# Première observation

Retour

## Indépendance vis-à-vis de la condition initiale

Si les solutions des problèmes aux différents instants ne dépendent pas de la condition initiale, alors on a consistance dynamique.

Preuve : Application du principe de Bellman au problème à l'instant $t_0$.

## Exemple

Soient $l_t : \mathbb{U}_t \to \mathbb{R}$, $K$ scalaire, et $f_t : \mathbb{U}_t \to \mathbb{R}$ tels que $x_t > 0, \forall t$.

$$\begin{aligned} \min_{x,u} \quad & \sum_{t=t_0}^{T-1} l_t(u_t)\, x_t + K x_T, \\ \text{s.c.} \quad & x_{t_0} \text{ donné}, \\ & x_{t+1} = f_t(u_t)\, x_t, \qquad \forall t = t_0, \ldots, T-1. \end{aligned}$$

$x_{t_0}$ **n'influe** que sur la valeur du problème, **pas sur l'argmin**.

# Première observation

Retour

## Indépendance vis-à-vis de la condition initiale

Si les solutions des problèmes aux différents instants ne dépendent pas de la condition initiale, alors on a consistance dynamique.

Preuve : Application du principe de Bellman au problème à l'instant $t_0$.

## Exemple

Soient $l_t : \mathbb{U}_t \rightarrow \mathbb{R}$, $K$ scalaire, et $f_t : \mathbb{U}_t \rightarrow \mathbb{R}$ tels que $x_t > 0, \forall t$.

$$\begin{aligned} \min_{x,u} \quad & \sum_{t=t_0}^{T-1} l_t\left(u_t\right) x_t + K x_T, \\ \text{s.c.} \quad & x_{t_0} \text{ donné}, \\ & x_{t+1} = f_t\left(u_t\right) x_t, \qquad \forall t = t_0, \ldots, T-1. \end{aligned}$$

$x_{t_0}$ **n'influe** que sur la valeur du problème, **pas sur l'argmin**.

# Processus de décision

Le **processus de décision réel** est en fait soumis à des **perturbations** $\varepsilon_{t_1}, \ldots, \varepsilon_{T-1}$ **non modélisées** :

- À l'instant $t_0$ :
  - on écrit le problème à $t_0$ ;
  - on obtient une décision $u^\sharp_{t_0,t_0}$ pour $t_0$, et une suite de décisions $u^\sharp_{t_0,t_1}$, …, $u^\sharp_{t_0,T-1}$ pour le futur.
- À l'instant $t_1$ :
- Le même processus se poursuit jusqu'à la fin de l'horizon de temps.

# Processus de décision

Le **processus de décision réel** est en fait soumis à des **perturbations** $\varepsilon_{t_1}, \ldots, \varepsilon_{T-1}$ **non modélisées** :

- À l'instant $t_0$ :
  - on écrit le problème à $t_0$ ;
  - on obtient une décision $u^\sharp_{t_0,t_0}$ pour $t_0$, et une suite de décisions $u^\sharp_{t_0,t_1}$, $\ldots$, $u^\sharp_{t_0,T-1}$ pour le futur.
- À l'instant $t_1$ :
- Le même processus se poursuit jusqu'à la fin de l'horizon de temps.

# Processus de décision

Le **processus de décision réel** est en fait soumis à des **perturbations** $\varepsilon_{t_1}, \ldots, \varepsilon_{T-1}$ **non modélisées** :

- À l'instant $t_0$ :
  - on écrit le problème à $t_0$ ;
  - on obtient une décision $u^\sharp_{t_0,t_0}$ pour $t_0$, et une suite de décisions $u^\sharp_{t_0,t_1}$, ..., $u^\sharp_{t_0,T-1}$ pour le futur.
- À l'instant $t_1$ :
  - on observe le **stock perturbé** $f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on écrit un **nouveau problème** partant de $t_1$ en $x_{t_1} = f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on obtient une décision $u^\sharp_{t_1,t_1}$, **généralement différente** de $u^\sharp_{t_0,t_1}$.
- Le même processus se poursuit jusqu'à la fin de l'horizon de temps.

# Processus de décision

Le **processus de décision réel** est en fait soumis à des **perturbations** $\varepsilon_{t_1}, \ldots, \varepsilon_{T-1}$ **non modélisées** :

- À l'instant $t_0$ :
  - on écrit le problème à $t_0$ ;
  - on obtient une décision $u^\sharp_{t_0,t_0}$ pour $t_0$, et une suite de décisions $u^\sharp_{t_0,t_1}$, $\ldots$, $u^\sharp_{t_0,T-1}$ pour le futur.
- À l'instant $t_1$ :
  - on observe le **stock perturbé** $f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on écrit un **nouveau problème partant de** $t_1$ en $x_{t_1} = f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on obtient une **décision** $u^\sharp_{t_1,t_1}$, **généralement différente de** $u^\sharp_{t_0,t_1}$.
- Le même processus se poursuit jusqu'à la fin de l'horizon de temps.

# Processus de décision

Le **processus de décision réel** est en fait soumis à des **perturbations** $\varepsilon_{t_1}, \ldots, \varepsilon_{T-1}$ **non modélisées** :

- À l'instant $t_0$ :
  - on écrit le problème à $t_0$ ;
  - on obtient une décision $u^\sharp_{t_0,t_0}$ pour $t_0$, et une suite de décisions $u^\sharp_{t_0,t_1}$, $\ldots$, $u^\sharp_{t_0,T-1}$ pour le futur.
- À l'instant $t_1$ :
  - on observe le **stock perturbé** $f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on écrit un **nouveau problème partant de** $t_1$ en $x_{t_1} = f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on obtient une **décision** $u^\sharp_{t_1,t_1}$, **généralement différente de** $u^\sharp_{t_0,t_1}$.
- Le même processus se poursuit jusqu'à la fin de l'horizon de temps.

# Processus de décision

Le **processus de décision réel** est en fait soumis à des **perturbations** $\varepsilon_{t_1}, \ldots, \varepsilon_{T-1}$ **non modélisées** :

- À l'instant $t_0$ :
  - on écrit le problème à $t_0$ ;
  - on obtient une décision $u^\sharp_{t_0,t_0}$ pour $t_0$, et une suite de décisions $u^\sharp_{t_0,t_1}$, $\ldots$, $u^\sharp_{t_0,T-1}$ pour le futur.
- À l'instant $t_1$ :
  - on observe le **stock perturbé** $f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on écrit un **nouveau problème partant de** $t_1$ en $x_{t_1} = f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on obtient une **décision** $u^\sharp_{t_1,t_1}$, **généralement différente de** $u^\sharp_{t_0,t_1}$.
- Le même processus se poursuit jusqu'à la fin de l'horizon de temps.

# Processus de décision

Le **processus de décision réel** est en fait soumis à des **perturbations** $\varepsilon_{t_1}, \ldots, \varepsilon_{T-1}$ **non modélisées** :

- À l'instant $t_0$ :
  - on écrit le problème à $t_0$ ;
  - on obtient une décision $u^\sharp_{t_0,t_0}$ pour $t_0$, et une suite de décisions $u^\sharp_{t_0,t_1}$, $\ldots$, $u^\sharp_{t_0,T-1}$ pour le futur.
- À l'instant $t_1$ :
  - on observe le **stock perturbé** $f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on écrit un **nouveau problème partant de** $t_1$ en $x_{t_1} = f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on obtient une **décision** $u^\sharp_{t_1,t_1}$, **généralement différente de** $u^\sharp_{t_0,t_1}$.
- Le même processus se poursuit jusqu'à la fin de l'horizon de temps.

# Processus de décision

Le **processus de décision réel** est en fait soumis à des **perturbations** $\varepsilon_{t_1}, \dots, \varepsilon_{T-1}$ **non modélisées** :

- À l'instant $t_0$ :
  - on écrit le problème à $t_0$ ;
  - on obtient une décision $u^\sharp_{t_0,t_0}$ pour $t_0$, et une suite de décisions $u^\sharp_{t_0,t_1}$, $\dots$, $u^\sharp_{t_0,T-1}$ pour le futur.
- À l'instant $t_1$ :
  - on observe le **stock perturbé** $f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on écrit un **nouveau problème partant de** $t_1$ en $x_{t_1} = f_{t_0}(x_{t_0}, u^\sharp_{t_0,t_0}) + \varepsilon_{t_1}$ ;
  - on obtient une **décision** $u^\sharp_{t_1,t_1}$, **généralement différente de** $u^\sharp_{t_0,t_1}$.
- Le même processus se poursuit jusqu'à la fin de l'horizon de temps.

# Deuxième observation

Retour

## Monde parfaitement déterministe

Dans le cas où le **modèle est exact** (si les perturbations $\varepsilon_{t_i}$ sont nulles), les problèmes successifs avec comme conditions initiales :

$$x_{t_i} = x^{\sharp}_{t_i} := f_{t_i}\left(x^{\sharp}_{t_{i-1}}, u^{\sharp}_{t_0, t_{i-1}}\right),$$

sont **consistants dynamiquement**.

En général, cette observation est fausse !

# Deuxième observation

Retour

## Monde parfaitement déterministe

Dans le cas où le **modèle est exact** (si les perturbations $\varepsilon_{t_i}$ sont nulles), les problèmes successifs avec comme conditions initiales :

$$x_{t_i} = x^\sharp_{t_i} := f_{t_i}\left(x^\sharp_{t_{i-1}}, u^\sharp_{t_0,t_{i-1}}\right),$$

sont **consistants dynamiquement**.

En général, cette observation est fausse !

# Troisième observation

Retour

## Quantité suffisante d'information

- Supposons que le **problème original soit résolu par programmation dynamique**,
- obtenant ainsi une suite de stratégies optimales $\Phi^\sharp_{t_0,t_0}, \ldots, \Phi^\sharp_{t_0,T-1}$ en **feedback** sur la variable $x$,
- alors les problèmes successifs sont **consistants dynamiquement** pour $t = t_0, \ldots, T-1$.

# Un premier problème...

## Problème partant de $t_0$

$$\begin{aligned} \min_{\boldsymbol{X},\boldsymbol{U}} \quad & \mathbb{E}\left(\sum_{t=t_0}^{T-1} L_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right) + K\left(\boldsymbol{X}_T\right)\right), \\ \text{s.c.} \quad & \boldsymbol{X}_{t_0} \text{ donné,} \end{aligned}$$

et, pour tout $t = t_0, \ldots, T-1$ :

$$\begin{aligned} & \boldsymbol{X}_{t+1} = f_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right), \\ & \boldsymbol{U}_t \text{ est } \sigma\{\boldsymbol{X}_{t_0}, \boldsymbol{W}_{t_1}, \ldots, \boldsymbol{W}_t\}\text{-mesurable.} \end{aligned}$$

# Principe de Programmation Dynamique

## Hypothèse d'indépendance temporelle

Les variables aléatoires $\boldsymbol{X}_{t_0}, \boldsymbol{W}_{t_1}, \ldots, \boldsymbol{W}_T$ sont indépendantes entre elles.

## Principe de Programmation Dynamique

- Solution $\Phi_t \ : \mathbb{X}_t \to \mathbb{U}_t$ en **feedback** sur $\boldsymbol{X}_t$.
- Soit $V_t(x)$ le coût optimal partant de $(t, x)$, alors :

$$V_T(x) = K(x),$$
$$V_t(x) = \min_{u \in \mathbb{U}_t} \mathbb{E}\Big(L_t(x, u, \boldsymbol{W}_{t+1}) + V_{t+1}\big(f_t(x, u, \boldsymbol{W}_{t+1})\big)\Big).$$

**permet de calculer les stratégies optimales** $\Phi^\sharp_{t_0,t_0}, \ldots, \Phi^\sharp_{t_0,T-1}$.

# Consistance dynamique

- On est dans le cas le plus standard en commande optimale.
- La suite de stratégies $\Phi^\sharp_{t_0,t_0}, \dots, \Phi^\sharp_{t_0,T-1}$ reste optimale pour les problèmes ultérieurs :

## Problème partant de $t_i$

$$
\begin{aligned}
\min_{\boldsymbol{X},\boldsymbol{U}} \quad & \mathbb{E}\left(\sum_{t=t_i}^{T-1} L_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right) + K\left(\boldsymbol{X}_T\right)\right), \\
\text{s.c.} \quad & \boldsymbol{X}_{t_i} \text{ donné}, \\
& \boldsymbol{X}_{t+1} = f_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right), \qquad \forall t \geq t_i, \\
& \boldsymbol{U}_t \text{ est } \sigma\{\boldsymbol{X}_{t_i}, \boldsymbol{W}_{t_{i+1}}, \dots, \boldsymbol{W}_t\}\text{-mesurable}, \qquad \forall t \geq t_i.
\end{aligned}
$$

$\hookrightarrow$ Les problèmes successifs sont **dynamiquement consistants**.

# Consistance dynamique

- On est dans le cas le plus standard en commande optimale.
- La suite de stratégies $\Phi^\sharp_{t_0,t_0}, \dots, \Phi^\sharp_{t_0,T-1}$ reste optimale pour les problèmes ultérieurs :

**Problème partant de $t_i$**

$$\min_{\boldsymbol{X},\boldsymbol{U}} \quad \mathbb{E}\left(\sum_{t=t_i}^{T-1} L_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right) + K\left(\boldsymbol{X}_T\right)\right),$$

$$\text{s.c.} \quad \boldsymbol{X}_{t_i} \text{ donné},$$

$$\boldsymbol{X}_{t+1} = f_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right), \qquad \forall t \geq t_i,$$

$$\boldsymbol{U}_t \text{ est } \sigma\{\boldsymbol{X}_{t_i}, \boldsymbol{W}_{t_{i+1}}, \dots, \boldsymbol{W}_t\}\text{-mesurable}, \qquad \forall t \geq t_i.$$

$\hookrightarrow$ Les problèmes successifs sont **dynamiquement consistants**.

# Analogie avec le cas déterministe

## Question

Est-on dans le cas :

1. Indépendance vis-à-vis de la condition initiale ? ▸ 1re obs.
2. Monde parfaitement déterministe ? ▸ 2e obs.
3. Quantité suffisante d'information ? ▸ 3e obs.

# Réponse

## Indépendance vis-à-vis de la condition initiale

Étant donné un feedback $\Phi_t$ pour tout $t$, on définit :

- l'ensemble $\Psi_t$ des fonctions de $\mathbb{X}_t$ dans $\mathbb{R}$ ;
- le produit scalaire :

$$\langle \psi_t, \mu_t \rangle := \mathbb{E}\left(\psi_t(\boldsymbol{X}_t)\right), \qquad \text{avec } \boldsymbol{X}_t \sim \mu_t .$$

- l'opérateur linéaire $A_t^{\Phi_t} : \Psi_{t+1} \to \Psi_t$ :

$$\left(A_t^{\Phi_t} \psi_{t+1}\right)(\cdot) := \mathbb{E}\left(\psi_{t+1} \circ f_t\left(\cdot, \Phi_t\left(\cdot\right), \boldsymbol{W}_{t+1}\right)\right)$$

*(transport rétrograde du coût à $t+1$ si on applique le feedback $\Phi_t$)*

# Réponse
## Indépendance vis-à-vis de la condition initiale

Étant donné un feedback $\Phi_t$ pour tout $t$, on définit :

- l'ensemble $\Psi_t$ des fonctions de $\mathbb{X}_t$ dans $\mathbb{R}$ ;
- le produit scalaire :

$$\langle \psi_t, \mu_t \rangle := \mathbb{E}\left(\psi_t(\boldsymbol{X}_t)\right), \qquad \text{avec } \boldsymbol{X}_t \sim \mu_t.$$

- l'opérateur linéaire $A_t^{\Phi_t} : \Psi_{t+1} \rightarrow \Psi_t$ :

$$\left(A_t^{\Phi_t} \psi_{t+1}\right)(\cdot) := \mathbb{E}\left(\psi_{t+1} \circ f_t\left(\cdot, \Phi_t(\cdot), \boldsymbol{W}_{t+1}\right)\right)$$

*(transport rétrograde du coût à $t+1$ si on applique le feedback $\Phi_t$)*

# Réponse

## Indépendance vis-à-vis de la condition initiale

Étant donné un feedback $\Phi_t$ pour tout $t$, on définit :

- l'ensemble $\Psi_t$ des fonctions de $\mathbb{X}_t$ dans $\mathbb{R}$ ;
- le produit scalaire :

$$\langle \psi_t, \mu_t \rangle := \mathbb{E}\left(\psi_t(\boldsymbol{X}_t)\right), \qquad \text{avec } \boldsymbol{X}_t \sim \mu_t.$$

- l'opérateur linéaire $A_t^{\Phi_t}$ : $\Psi_{t+1} \to \Psi_t$ :

$$\left(A_t^{\Phi_t}\psi_{t+1}\right)(\cdot) := \mathbb{E}\left(\psi_{t+1} \circ f_t\left(\cdot, \Phi_t\left(\cdot\right), \boldsymbol{W}_{t+1}\right)\right)$$

*(transport rétrograde du coût à $t+1$ si on applique le feedback $\Phi_t$)*

# Réponse

## Indépendance vis-à-vis de la condition initiale

Enfin :

- l'opérateur adjoint $(A_t^{\Phi_t})^\star$ qui transporte les mesures $\mu_t$.
- un opérateur $\Lambda_t^{\Phi_t} : \mathcal{X}_t \to \mathbb{R}$ de la forme suivante.

$$\Lambda_t^{\Phi_t}(\cdot) := \mathbb{E}\left(L_t\left(\cdot, \Phi_t(\cdot), \boldsymbol{W}_{t+1}\right)\right),$$

*(coût instantané en $(t, \cdot)$ lorsque le feedback $\Phi_t$ est appliqué)*

# Réponse

## Indépendance vis-à-vis de la condition initiale

On aboutit à la **formulation distribuée** [Witsenhausen, 1973] **du problème original** :

$$\min_{\Phi,\mu} \quad \sum_{t=t_0}^{T-1} \left\langle \Lambda_t^{\Phi_t}, \mu_t \right\rangle + \left\langle K, \mu_T \right\rangle ,$$

$$\text{s.c.} \quad \mu_{t_0} \text{ donné},$$

$$\mu_{t+1} = \left(A_t^{\Phi_t}\right)^{\star} \mu_t, \qquad \forall t \geq t_0, \qquad \textit{(Fokker-Planck)}$$

↪ Contrôle optimal **déterministe en dimension infinie**.

↪ **Linéaire en les** $\mu_t$, $t = t_0, \ldots, T$, non-linéaire en les $\Phi_t$, $t = t_0, \ldots, T-1$.

# Réponse

## Indépendance vis-à-vis de la condition initiale

On aboutit à la **formulation distribuée** [Witsenhausen, 1973] **du problème original** :

$$\min_{\Phi,\mu} \quad \sum_{t=t_0}^{T-1} \left\langle \Lambda_t^{\Phi_t}, \mu_t \right\rangle + \left\langle K, \mu_T \right\rangle ,$$

$$\text{s.c.} \quad \mu_{t_0} \text{ donné},$$

$$\mu_{t+1} = \left(A_t^{\Phi_t}\right)^\star \mu_t, \qquad \forall t \geq t_0, \qquad \textit{(Fokker-Planck)}$$

↪ Contrôle optimal **déterministe en dimension infinie**.

↪ **Linéaire en les** $\mu_t$, $t = t_0, \ldots, T$, non-linéaire en les $\Phi_t$, $t = t_0, \ldots, T-1$.

# Réponse

## Indépendance vis-à-vis de la condition initiale

On aboutit à la **formulation distribuée** [Witsenhausen, 1973] **du problème original** :

$$\begin{aligned}
\min_{\Phi,\mu} \quad & \sum_{t=t_0}^{T-1} \left\langle \Lambda_t^{\Phi_t}, \mu_t \right\rangle + \left\langle K, \mu_T \right\rangle , \\
\text{s.c.} \quad & \mu_{t_0} \text{ donné}, \\
& \mu_{t+1} = \left( A_t^{\Phi_t} \right)^{\star} \mu_t , \qquad \forall t \geq t_0, \qquad \textit{(Fokker-Planck)}
\end{aligned}$$

$\hookrightarrow$ Contrôle optimal **déterministe en dimension infinie**.

$\hookrightarrow$ **Linéaire en les** $\mu_t$, $t = t_0, \ldots, T$, non-linéaire en les $\Phi_t$, $t = t_0, \ldots, T-1$.

# Réponse

## Indépendance vis-à-vis de la condition initiale

### Résolution du problème distribué par programmation dynamique

$$\mathcal{V}_T(\mu) = \langle K, \mu \rangle ,$$

$$\mathcal{V}_{T-1}(\mu) = \min_{\Phi} \langle \Lambda^{\Phi}_{Tt-1}, \mu \rangle + \mathcal{V}_T\left( \left(A^{\Phi}_{T-1}\right)^{\star} \mu \right) .$$

Le feedback optimal $\Gamma^{\sharp}_{T-1} : \mu \to \Phi^{\sharp}_{\mu}(\cdot)$ **dépend a priori de** $x$ **et de** $\mu$.

$$\mathcal{V}_{T-1}(\mu) = \min_{\Phi} \langle \Lambda^{\Phi}_{T-1} + A^{\Phi}_{T-1} K, \mu \rangle .$$

$$= \min_{\Phi(\cdot)} \int_{\mathbb{X}} \left(\Lambda^{\Phi}_{T-1} + A^{\Phi}_{T-1} K\right)(x)\, \mu(\mathrm{d}x) .$$

La minimisation peut se faire "$x$ par $x$".

# Réponse

## Indépendance vis-à-vis de la condition initiale

### Résolution du problème distribué par programmation dynamique

$$\mathcal{V}_T(\mu) = \langle K, \mu \rangle,$$
$$\mathcal{V}_{T-1}(\mu) = \min_{\Phi} \langle \Lambda^{\Phi}_{Tt-1}, \mu \rangle + \mathcal{V}_T\left(\left(A^{\Phi}_{T-1}\right)^{\star} \mu\right).$$

Le feedback optimal $\Gamma^{\sharp}_{T-1} : \mu \to \Phi^{\sharp}_{\mu}(\cdot)$ **dépend a priori de** $x$ **et de** $\mu$.

$$\mathcal{V}_{T-1}(\mu) = \min_{\Phi} \langle \Lambda^{\Phi}_{T-1} + A^{\Phi}_{T-1} K, \mu \rangle.$$
$$= \min_{\Phi(\cdot)} \int_{\mathbb{X}} \left(\Lambda^{\Phi}_{T-1} + A^{\Phi}_{T-1} K\right)(x)\, \mu(\mathrm{d}x).$$

La minimisation peut se faire "$x$ par $x$".

# Conclusion partielle

L'interversion de l'intégrale et de la minimisation conduit à :

- des **feedbacks optimaux ne dépendant plus de** $\mu$ : $\Gamma^{\sharp}_{T-1} \equiv \Phi^{\sharp}_{\mu}$
- $\mathcal{V}_{T-1}$ ne dépend de $\mu$ que de façon multiplicative.

$\rightarrow$ On est à nouveau dans le **cas très particulier de l'observation 1**.

La classe de problèmes considérée n'est pas si générale !

# Plusieurs types de contraintes en optimisation stochastique

On peut faire face à des contraintes :

- presque-sûres : $g(\boldsymbol{X}_T) \leq a$ ;
- en probabilité : $\mathbb{P}(g(\boldsymbol{X}_T) \leq a) \geq p$ ;
- en espérance : $\mathbb{E}(g(\boldsymbol{X}_T)) \leq a$ ;
- ...

Les contraintes en probabilité peuvent être mises sous la forme de contraintes en espérance :

$$\mathbb{P}(g(\boldsymbol{X}_T) \leq a) = \mathbb{E}\left(\mathbf{1}_{\mathbb{X}^{\mathrm{ad}}}(\boldsymbol{X}_T)\right),$$

avec $\mathbb{X}^{\mathrm{ad}} = \{x \in \mathbb{X}, g(x) \leq a\}$.

Contraintes en proba

# Plusieurs types de contraintes en optimisation stochastique

On peut faire face à des contraintes :

- presque-sûres : $g\left(\boldsymbol{X}_T\right) \leq a$ ;
- en probabilité : $\mathbb{P}\left(g\left(\boldsymbol{X}_T\right) \leq a\right) \geq p$ ;
- en espérance : $\mathbb{E}\left(g\left(\boldsymbol{X}_T\right)\right) \leq a$ ;
- ...

Les contraintes en probabilité peuvent être mises sous la forme de contraintes en espérance :

$$\mathbb{P}\left(g\left(\boldsymbol{X}_T\right) \leq a\right) = \mathbb{E}\left(\mathbf{1}_{\mathbb{X}^{\mathrm{ad}}}\left(\boldsymbol{X}_T\right)\right),$$

avec $\mathbb{X}^{\mathrm{ad}} = \{x \in \mathbb{X}, g(x) \leq a\}$.

Contraintes en proba

- Les contraintes en $\mathbb{P}$ apportent des difficultés supplémentaires [Prékopa, 1995].

# Plusieurs types de contraintes en optimisation stochastique

On peut faire face à des contraintes :

- presque-sûres : $g\left(\boldsymbol{X}_T\right) \leq a$ ;
- en probabilité : $\mathbb{P}\left(g\left(\boldsymbol{X}_T\right) \leq a\right) \geq p$ ;
- en espérance : $\mathbb{E}\left(g\left(\boldsymbol{X}_T\right)\right) \leq a$ ;
- ...

Les contraintes en probabilité peuvent être mises sous la forme de contraintes en espérance :

$$\mathbb{P}\left(g\left(\boldsymbol{X}_T\right) \leq a\right) = \mathbb{E}\left(\mathbf{1}_{\mathbb{X}^{\mathrm{ad}}}\left(\boldsymbol{X}_T\right)\right),$$

avec $\mathbb{X}^{\mathrm{ad}} = \{x \in \mathbb{X}, g(x) \leq a\}$.

# Plusieurs types de contraintes en optimisation stochastique

On peut faire face à des contraintes :

- presque-sûres : $g\left(\boldsymbol{X}_T\right) \leq a$ ;
- en probabilité : $\mathbb{P}\left(g\left(\boldsymbol{X}_T\right) \leq a\right) \geq p$ ;
- en espérance : $\mathbb{E}\left(g\left(\boldsymbol{X}_T\right)\right) \leq a$ ;
- ...

Les contraintes en probabilité peuvent être mises sous la forme de contraintes en espérance :

$$\mathbb{P}\left(g\left(\boldsymbol{X}_T\right) \leq a\right) = \mathbb{E}\left(\mathbf{1}_{\mathbb{X}^{\mathrm{ad}}}\left(\boldsymbol{X}_T\right)\right),$$

avec $\mathbb{X}^{\mathrm{ad}} = \{x \in \mathbb{X}, g(x) \leq a\}$.

Contraintes en proba

- Les contraintes en $\mathbb{P}$ apportent des difficultés supplémentaires [Prékopa, 1995].
- La difficulté qui nous intéresse est commune à $\mathbb{P}$ et $\mathbb{E}$.

$\longrightarrow$ On se concentre par la suite sur les contraintes en espérance.

# Plusieurs types de contraintes en optimisation stochastique

On peut faire face à des contraintes :

- presque-sûres : $g(\boldsymbol{X}_T) \leq a$ ;
- en probabilité : $\mathbb{P}(g(\boldsymbol{X}_T) \leq a) \geq p$ ;
- en espérance : $\mathbb{E}(g(\boldsymbol{X}_T)) \leq a$ ;
- ...

Les contraintes en probabilité peuvent être mises sous la forme de contraintes en espérance :

$$\mathbb{P}(g(\boldsymbol{X}_T) \leq a) = \mathbb{E}\left(\mathbf{1}_{\mathbb{X}^{\mathrm{ad}}}(\boldsymbol{X}_T)\right),$$

avec $\mathbb{X}^{\mathrm{ad}} = \{x \in \mathbb{X}, g(x) \leq a\}$.

Contraintes en proba

- Les contraintes en $\mathbb{P}$ apportent des difficultés supplémentaires [Prékopa, 1995].
- La difficulté qui nous intéresse est commune à $\mathbb{P}$ et $\mathbb{E}$.

$\hookrightarrow$ On se concentre par la suite sur les contraintes en espérance.

# Plusieurs types de contraintes en optimisation stochastique

On peut faire face à des contraintes :

- presque-sûres : $g\left(\boldsymbol{X}_T\right) \leq a$ ;
- en probabilité : $\mathbb{P}\left(g\left(\boldsymbol{X}_T\right) \leq a\right) \geq p$ ;
- en espérance : $\mathbb{E}\left(g\left(\boldsymbol{X}_T\right)\right) \leq a$ ;
- ...

Les contraintes en probabilité peuvent être mises sous la forme de contraintes en espérance :

$$\mathbb{P}\left(g\left(\boldsymbol{X}_T\right) \leq a\right) = \mathbb{E}\left(\mathbf{1}_{\mathbb{X}^{\mathrm{ad}}}\left(\boldsymbol{X}_T\right)\right),$$

avec $\mathbb{X}^{\mathrm{ad}} = \{x \in \mathbb{X}, g(x) \leq a\}$.

Contraintes en proba

- Les contraintes en $\mathbb{P}$ apportent des difficultés supplémentaires [Prékopa, 1995].
- La difficulté qui nous intéresse est commune à $\mathbb{P}$ et $\mathbb{E}$.

$\hookrightarrow$ On se concentre par la suite sur les contraintes en espérance.

# Plusieurs types de contraintes en optimisation stochastique

On peut faire face à des contraintes :

- presque-sûres : $g\left(\boldsymbol{X}_T\right) \leq a$ ;
- en probabilité : $\mathbb{P}\left(g\left(\boldsymbol{X}_T\right) \leq a\right) \geq p$ ;
- en espérance : $\mathbb{E}\left(g\left(\boldsymbol{X}_T\right)\right) \leq a$ ;
- ...

Les contraintes en probabilité peuvent être mises sous la forme de contraintes en espérance :

$$\mathbb{P}\left(g\left(\boldsymbol{X}_T\right) \leq a\right) = \mathbb{E}\left(\mathbf{1}_{\mathbb{X}^{\mathrm{ad}}}\left(\boldsymbol{X}_T\right)\right),$$

avec $\mathbb{X}^{\mathrm{ad}} = \{x \in \mathbb{X}, g(x) \leq a\}$.

Contraintes en proba

- Les contraintes en $\mathbb{P}$ apportent des difficultés supplémentaires [Prékopa, 1995].
- La difficulté qui nous intéresse est commune à $\mathbb{P}$ et $\mathbb{E}$.

$\hookrightarrow$ On se concentre par la suite sur les contraintes en espérance.

# Ajout d'une contrainte en espérance

## Problème partant de $t_0$

$$\min_{\boldsymbol{X},\boldsymbol{U}} \quad \mathbb{E}\left(\sum_{t=t_0}^{T-1} L_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right) + K\left(\boldsymbol{X}_T\right)\right), \tag{1a}$$

$$\text{s.c.} \quad \mathbb{E}\left(g\left(\boldsymbol{X}_T\right) \leq a\right), \tag{1b}$$

et, pour tout $t = t_0, \ldots, T-1$ :

$$\boldsymbol{X}_{t_0} \text{ donné}, \tag{1c}$$

$$\boldsymbol{X}_{t+1} = f_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right), \tag{1d}$$

$$\boldsymbol{U}_t \text{ est } \sigma\{\boldsymbol{X}_{t_0}, \boldsymbol{W}_{t_1}, \ldots, \boldsymbol{W}_t\}\text{-mesurable.} \tag{1e}$$

# Formulation distribuée

## Problème équivalent

$$\min_{\Phi,\mu} \quad \sum_{t=t_0}^{T-1} \left\langle \Lambda_t^{\Phi_t}, \mu_t \right\rangle + \left\langle K, \mu_T \right\rangle ,$$

$$\text{s.c.} \quad \langle g, \mu_T \rangle \leq a,$$

$$\mu_{t_0} \text{ donné},$$

$$\mu_{t+1} = \left(A_t^{\Phi_t}\right)^\star \mu_t, \qquad \forall t = t_0, \ldots, T-1.$$

On peut inclure la contrainte dans la fonction coût avec :

$$\chi_{\{\langle g, \mu_T \rangle \leq a\}}.$$

- La mesure $\mu_{t_0}$ n'intervient plus de manière linéaire.
- On perd la propriété d'indépendance à la condition initiale.

**Les feedbacks optimaux dépendent maintenant, en général, de $\mu_{t_0}$.**

# Formulation distribuée

## Problème équivalent

$$\min_{\Phi,\mu} \quad \sum_{t=t_0}^{T-1} \left\langle \Lambda_t^{\Phi_t}, \mu_t \right\rangle + \left\langle K, \mu_T \right\rangle ,$$

$$\text{s.c.} \quad \langle g, \mu_T \rangle \leq a,$$

$$\mu_{t_0} \text{ donné},$$

$$\mu_{t+1} = \left(A_t^{\Phi_t}\right)^{\star} \mu_t, \qquad \forall t = t_0, \ldots, T-1.$$

On peut inclure la contrainte dans la fonction coût avec :

$$\chi_{\{\langle g, \mu_T \rangle \leq a\}}.$$

- La mesure $\mu_{t_0}$ n'intervient plus de manière linéaire.
- On perd la propriété d'indépendance à la condition initiale.

Les feedbacks optimaux dépendent maintenant, en général, de $\mu_{t_0}$.

# Formulation distribuée

## Problème équivalent

$$
\begin{aligned}
\min_{\Phi,\mu} \quad & \sum_{t=t_0}^{T-1} \left\langle \Lambda_t^{\Phi_t}, \mu_t \right\rangle + \left\langle K, \mu_T \right\rangle , \\
\text{s.c.} \quad & \langle g, \mu_T \rangle \leq a, \\
& \mu_{t_0} \text{ donné}, \\
& \mu_{t+1} = \left(A_t^{\Phi_t}\right)^{\star} \mu_t, \qquad \forall t = t_0, \dots, T-1.
\end{aligned}
$$

On peut inclure la contrainte dans la fonction coût avec :

$$\chi_{\{\langle g, \mu_T \rangle \leq a\}}.$$

- La mesure $\mu_{t_0}$ n'intervient plus de manière linéaire.
- On perd la propriété d'indépendance à la condition initiale.

**Les feedbacks optimaux dépendent maintenant, en général, de $\mu_{t_0}$.**

# Retour à la consistance

Si on résout le problème **déterministe** précédent, on obtient une suite de feedbacks $(\Phi^\sharp_{t_0}, \ldots, \Phi^\sharp_{T-1})$.

- On a consistance dynamique si les $\mu$ suivent Fokker-Planck.
- Or, à l'instant $t$, la meilleure observation de la loi de $\boldsymbol{X}_t$ est le Dirac !

Résolvons le problème par programmation dynamique :

$$\mathcal{V}_T(\mu) = \langle K \,, \mu \rangle + \chi_{\{\langle g, \mu \rangle \leq a\}}(\mu),$$
$$\mathcal{V}_{T-1}(\mu) = \min_{\phi} \left\langle \Lambda^{\phi}_{T-1} \,, \mu \right\rangle + \mathcal{V}_T\Big( \big(A^{\phi}_{T-1}\big)^{\star} \mu \Big),$$

- On obtient des feedbacks $\Gamma^\sharp_t$ dépendant de $x$ et de $\mu$.
- On retrouve la consistance dynamique.
- Mais la variable d'état est de dimension infinie.

# Retour à la consistance

Si on résout le problème **déterministe** précédent, on obtient une suite de feedbacks $(\Phi^\sharp_{t_0}, \dots, \Phi^\sharp_{T-1})$.

- On a consistance dynamique si les $\mu$ suivent Fokker-Planck.
- Or, à l'instant $t$, la meilleure observation de la loi de $\boldsymbol{X}_t$ est le Dirac !

Résolvons le problème par programmation dynamique :

$$\mathcal{V}_T(\mu) = \langle K \,, \mu \rangle + \chi_{\{\langle g, \mu \rangle \le a\}}(\mu),$$

$$\mathcal{V}_{T-1}(\mu) = \min_{\phi} \left\langle \Lambda^{\phi}_{T-1} \,, \mu \right\rangle + \mathcal{V}_T\Big(\big(A^{\phi}_{T-1}\big)^{\star} \mu\Big),$$

- On obtient des feedbacks $\Gamma^\sharp_t$ dépendant de $x$ et de $\mu$.
- On retrouve la consistance dynamique.
- Mais la variable d'état est de dimension infinie.

# Retour à la consistance

Si on résout le problème **déterministe** précédent, on obtient une suite de feedbacks $(\Phi^{\sharp}_{t_0}, \dots, \Phi^{\sharp}_{T-1})$.

- On a consistance dynamique si les $\mu$ suivent Fokker-Planck.
- Or, à l'instant $t$, la meilleure observation de la loi de $\boldsymbol{X}_t$ est le Dirac !

Résolvons le problème par programmation dynamique :

$$\mathcal{V}_T(\mu) = \langle K , \mu \rangle + \chi_{\{\langle g, \mu \rangle \leq a\}}(\mu),$$
$$\mathcal{V}_{T-1}(\mu) = \min_{\phi} \left\langle \Lambda^{\phi}_{T-1} , \mu \right\rangle + \mathcal{V}_T\Big( \big(A^{\phi}_{T-1}\big)^{\star} \mu \Big),$$

- On obtient des feedbacks $\Gamma^{\sharp}_t$ dépendant de $x$ et de $\mu$.
- On retrouve la consistance dynamique.
- Mais la variable d'état est de dimension infinie.

# Retour à la consistance

Si on résout le problème **déterministe** précédent, on obtient une suite de feedbacks $(\Phi^{\sharp}_{t_0}, \ldots, \Phi^{\sharp}_{T-1})$.

- On a consistance dynamique si les $\mu$ suivent Fokker-Planck.
- Or, à l'instant $t$, la meilleure observation de la loi de $\boldsymbol{X}_t$ est le Dirac !

Résolvons le problème par programmation dynamique :

$$\mathcal{V}_T(\mu) = \langle K \,, \mu \rangle + \chi_{\{\langle g, \mu \rangle \leq a\}}(\mu),$$
$$\mathcal{V}_{T-1}(\mu) = \min_{\phi} \left\langle \Lambda^{\phi}_{T-1} \,, \mu \right\rangle + \mathcal{V}_T\Big(\big(A^{\phi}_{T-1}\big)^{\star} \mu\Big),$$

- On obtient des feedbacks $\Gamma^{\sharp}_t$ dépendant de $x$ et de $\mu$.
- On retrouve la consistance dynamique.
- Mais la variable d'état est de dimension infinie.

# Retour à la consistance

Si on résout le problème **déterministe** précédent, on obtient une suite de feedbacks $(\Phi^\sharp_{t_0}, \dots, \Phi^\sharp_{T-1})$.

- On a consistance dynamique si les $\mu$ suivent Fokker-Planck.
- Or, à l'instant $t$, la meilleure observation de la loi de $\boldsymbol{X}_t$ est le Dirac !

Résolvons le problème par programmation dynamique :

$$\mathcal{V}_T(\mu) = \langle K \, , \mu \rangle + \chi_{\{\langle g, \mu \rangle \leq a\}}(\mu),$$
$$\mathcal{V}_{T-1}(\mu) = \min_{\phi} \left\langle \Lambda^{\phi}_{T-1} \, , \mu \right\rangle + \mathcal{V}_T\Big(\big(A^{\phi}_{T-1}\big)^{\star} \mu\Big),$$

- On obtient des feedbacks $\Gamma^\sharp_t$ dépendant de $x$ et de $\mu$.
- On retrouve la consistance dynamique.
- Mais la variable d'état est de dimension infinie.

# Conclusion

- Pour une large classe de problèmes de commande optimale, **la notion de consistance est très liée à la définition d'un état**.
- En général, le **feedback doit dépendre de la loi** de la variable d'état "usuelle".
- Remarquons que nous avons choisi de ne pas modifier le niveau de contrainte :

$$\mathbb{E}\left(g\left(\boldsymbol{X}_T\right)\right) \leq a, \qquad \forall t_i \geq t_0.$$

- La dernière programmation dynamique introduite n'est **pas soluble en pratique** : les lois sont des objets de dimension infinie.

Retour à la dimension finie

# Conclusion

- Pour une large classe de problèmes de commande optimale, **la notion de consistance est très liée à la définition d'un état**.
- En général, le **feedback doit dépendre de la loi** de la variable d'état "usuelle".
- Remarquons que nous avons choisi de ne pas modifier le niveau de contrainte :

$$\mathbb{E}\left(g\left(\boldsymbol{X}_T\right)\right) \leq a, \qquad \forall t_i \geq t_0.$$

- La dernière programmation dynamique introduite n'est **pas soluble en pratique** : les lois sont des objets de dimension infinie.

Retour à la dimension finie

# Conclusion

- Pour une large classe de problèmes de commande optimale, **la notion de consistance est très liée à la définition d'un état**.
- En général, le **feedback doit dépendre de la loi** de la variable d'état "usuelle".
- Remarquons que nous avons choisi de ne pas modifier le niveau de contrainte :
$$\mathbb{E}\left(g\left(\boldsymbol{X}_T\right)\right) \leq a, \qquad \forall t_i \geq t_0.$$
- La dernière programmation dynamique introduite n'est **pas soluble en pratique** : les lois sont des objets de dimension infinie.

Retour à la dimension finie

# Conclusion

- Pour une large classe de problèmes de commande optimale, **la notion de consistance est très liée à la définition d'un état**.
- En général, le **feedback doit dépendre de la loi** de la variable d'état "usuelle".
- Remarquons que nous avons choisi de ne pas modifier le niveau de contrainte :

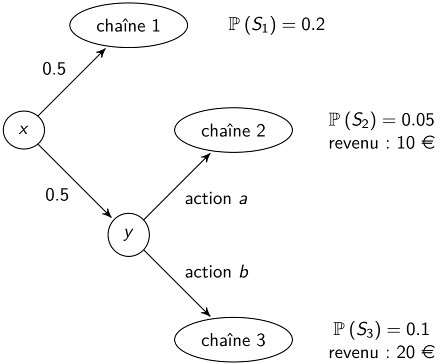
$$\mathbb{E}\left(g\left(\boldsymbol{X}_T\right)\right) \leq a, \qquad \forall t_i \geq t_0.$$

- La dernière programmation dynamique introduite n'est **pas soluble en pratique** : les lois sont des objets de dimension infinie.

Retour à la dimension finie

# Plan

# Retour sur le problème de Haviv

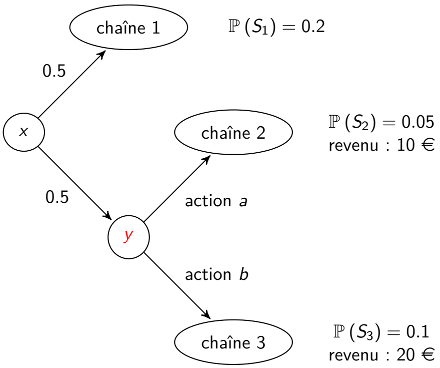



Il faut préciser comment réécrire la contrainte :

$$\mathbb{P}(S_1 \cup S_2 \cup S_3) \leq 0.125,$$

au second pas de temps.

# Retour sur le problème de Haviv



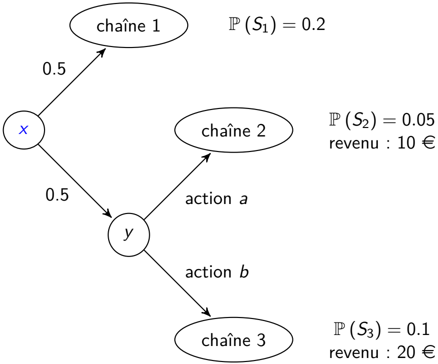

Une réécriture (naturelle ?) de la contrainte en $y$ est :

$$0.5 \times \mathbb{P}(S_2 \cup S_3) \leq \underbrace{0.125 - 0.5 \times \mathbb{P}(S_1)}_{0.025},$$

soit :

$$\mathbb{P}(S_2 \cup S_3) \leq 0.05.$$

# Retour sur le problème de Haviv




Vu de $x$ comme de $y$, l'action $b$ est alors non-admissible.

# Références I

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., and Ku, H. (2007).
Coherent multiperiod risk-adjusted values and Bellman's principle.
*Annals of Operations Research*, 152(1) :5–22.

Bouchard, B., Elie, R., and Touzi, N. (2009).
Stochastic target problems with controlled loss.
*SIAM Journal on Control and Optimization*, 48(5) :3123–3150.

Carpentier, P., Chancelier, J.-P., Cohen, G., De Lara, M., and Girardeau, P. (2010).
Dynamic consistency for stochastic optimal control problems.
arXiv :1005.3605.

Cheridito, P., Delbaen, F., and Kupper, M. (2006).
Dynamic monetary risk measures for bounded discrete-time processes.
*Electronic Journal of Probability*, 11(3) :57–106.

Detlefsen, K. and Scandolo, G. (2005).
Conditional and dynamic convex risk measures.
*Finance and Stochastics*, 9(4) :539–561.

Ekeland, I. and Lazrak, A. (2006).
Being serious about non-commitment : subgame perfect equilibrium in continuous time.
*arXiv*, math.OC 0604264.

# Références II


Haviv, M. (1996).
On constrained Markov decision processes.
*Operations Research Letters*, 19 :25–28.

Prékopa, A. (1995).
*Stochastic Programming*.
Kluwer, Dordrecht.

Ruszczyński, A. (2009).
Risk-averse dynamic programming for markov decision processes.
*Optimization Online*, to appear in Mathematical Programming.

Shapiro, A. (2009).
On a time consistency concept in risk averse multistage stochastic programming.
*Operations Research Letters*, 37(3) :143 – 147.

Whittle, P. (1982).
*Optimization over time*.
John Wiley & Sons.

Witsenhausen, H. S. (1973).
A standard form for sequential stochastic control.
*Mathematical Systems Theory*, 7(1) :5–11.

# Contraintes en probabilité

## Contraintes en probabilité jointes

Plutôt que $\mathbb{P}\left(h\left(\boldsymbol{X}_T\right) \geq b\right) \leq a$, on peut s'intéresser à :

$$\mathbb{P}\left(h_t\left(\boldsymbol{X}_t\right) \geq b_t, \forall t = t_1, \ldots, T\right) \leq a.$$

Via l'introduction d'une variable aléatoire supplémentaire à valeurs binaires :

$$\begin{aligned} \boldsymbol{Y}_{t_0} &= 1, \\ \boldsymbol{Y}_{t+1} &= \boldsymbol{Y}_t \times \mathbf{1}_{\{h_{t+1}(\boldsymbol{X}_{t+1}) \geq b_{t+1}\}}, \qquad \forall t = t_0, \ldots, T-1, \end{aligned}$$

elle s'écrit alors : $\mathbb{E}\left(\boldsymbol{Y}_T\right) \leq a$.

Retour

# Retour à la dimension finie I

## Problème équivalent

- Idée de [Bouchard et al., 2009] adaptée au temps discret.
- Consiste à ajouter une commande $\boldsymbol{V}_t$ et un état $\boldsymbol{Z}_t$.

$$\min_{\boldsymbol{X},\boldsymbol{Z},\boldsymbol{U},\boldsymbol{V}} \quad \mathbb{E}\left(\sum_{t=t_0}^{T-1} L_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right) + K\left(\boldsymbol{X}_T\right)\right), \tag{2a}$$

sous les contraintes dynamiques sur $\boldsymbol{X}$ :

$$\boldsymbol{X}_{t_0} = x_{t_0}, \tag{2b}$$

$$\boldsymbol{X}_{t+1} = f_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right), \qquad \forall t = t_0, \dots, T-1, \tag{2c}$$

les contraintes de non-anticipativité :

$$\boldsymbol{U}_t \preceq \mathcal{A}_t, \qquad \forall t = t_0, \dots, T-1, \tag{2d}$$

# Retour à la dimension finie II

## Problème équivalent

les contraintes dynamiques sur $\boldsymbol{Z}$ :

$$\boldsymbol{Z}_{t_0} = 0, \tag{2e}$$

$$\boldsymbol{Z}_{t+1} = \boldsymbol{Z}_t + \boldsymbol{V}_t, \qquad \forall t = t_0, \ldots, T-1, \tag{2f}$$

les contraintes de non-anticipativité :

$$\boldsymbol{V}_t \preceq \mathcal{A}_{t+1}, \qquad \forall t = t_0, \ldots, T-1, \tag{2g}$$

la contrainte presque-sûre portant sur l'instant final :

$$g\left(\boldsymbol{X}_T\right) - \boldsymbol{Z}_T \leq a, \qquad \mathbb{P}\text{-p.s.}, \tag{2h}$$

et les contraintes intermédiaires sur les nouvelles commandes :

$$\mathbb{E}\left(\boldsymbol{V}_t \mid \mathcal{A}_t\right) = 0, \qquad \forall t = t_0, \ldots, T-1. \tag{2i}$$

# Retour à la dimension finie III

## Équivalence avec le problème original : (1) $\subset$ (2)

Soient $(\boldsymbol{U}_{t_0}, \dots, \boldsymbol{U}_{T-1})$ et $(\boldsymbol{X}_{t_0}, \dots, \boldsymbol{X}_T)$ satisfaisant (1). On définit $\boldsymbol{Z}$ et $\boldsymbol{V}$ par :

$$\begin{aligned}
\boldsymbol{Z}_{t_0} &= 0, \\
\boldsymbol{Z}_{t+1} &= \boldsymbol{Z}_t + \boldsymbol{V}_t, \qquad \forall t \geq t_0, \\
\boldsymbol{V}_t &= \mathbb{E}\big(g(\boldsymbol{X}_T) \bigm| \mathcal{A}_{t+1}\big) - \mathbb{E}\big(g(\boldsymbol{X}_T) \bigm| \mathcal{A}_t\big), \qquad \forall t \geq t_0.
\end{aligned}$$

Par sommation, on obtient la relation :

$$\boldsymbol{Z}_T = g(\boldsymbol{X}_T) - \mathbb{E}\big(g(\boldsymbol{X}_T) \bigm| \mathcal{A}_{t_0}\big) = g(\boldsymbol{X}_T) - \mathbb{E}\big(g(\boldsymbol{X}_T)\big) \geq g(\boldsymbol{X}_T) - a.$$

Les processus $\boldsymbol{X}$, $\boldsymbol{Z}$, $\boldsymbol{U}$ et $\boldsymbol{V}$ vérifient donc bien les contraintes de (2).

# Retour à la dimension finie III

## Équivalence avec le problème original : (1) $\subset$ (2)

Soient $(\boldsymbol{U}_{t_0}, \dots, \boldsymbol{U}_{T-1})$ et $(\boldsymbol{X}_{t_0}, \dots, \boldsymbol{X}_T)$ satisfaisant (1). On définit $\boldsymbol{Z}$ et $\boldsymbol{V}$ par :

$$
\begin{aligned}
\boldsymbol{Z}_{t_0} &= 0, \\
\boldsymbol{Z}_{t+1} &= \boldsymbol{Z}_t + \boldsymbol{V}_t, \qquad \forall t \geq t_0, \\
\boldsymbol{V}_t &= \mathbb{E}\big(g(\boldsymbol{X}_T) \bigm| \mathcal{A}_{t+1}\big) - \mathbb{E}\big(g(\boldsymbol{X}_T) \bigm| \mathcal{A}_t\big), \qquad \forall t \geq t_0.
\end{aligned}
$$

Par sommation, on obtient la relation :

$$
\boldsymbol{Z}_T = g(\boldsymbol{X}_T) - \mathbb{E}\big(g(\boldsymbol{X}_T) \bigm| \mathcal{A}_{t_0}\big) = g(\boldsymbol{X}_T) - \mathbb{E}\big(g(\boldsymbol{X}_T)\big) \geq g(\boldsymbol{X}_T) - a.
$$

Les processus $\boldsymbol{X}$, $\boldsymbol{Z}$, $\boldsymbol{U}$ et $\boldsymbol{V}$ vérifient donc bien les contraintes de (2).

# Retour à la dimension finie III

## Équivalence avec le problème original : (1) $\subset$ (2)

Soient $(\boldsymbol{U}_{t_0}, \dots, \boldsymbol{U}_{T-1})$ et $(\boldsymbol{X}_{t_0}, \dots, \boldsymbol{X}_T)$ satisfaisant (1). On définit $\boldsymbol{Z}$ et $\boldsymbol{V}$ par :

$$\begin{aligned} \boldsymbol{Z}_{t_0} &= 0, \\ \boldsymbol{Z}_{t+1} &= \boldsymbol{Z}_t + \boldsymbol{V}_t, \qquad \forall t \geq t_0, \\ \boldsymbol{V}_t &= \mathbb{E}\big(g(\boldsymbol{X}_T) \bigm| \mathcal{A}_{t+1}\big) - \mathbb{E}\big(g(\boldsymbol{X}_T) \bigm| \mathcal{A}_t\big), \qquad \forall t \geq t_0. \end{aligned}$$

Par sommation, on obtient la relation :

$$\boldsymbol{Z}_T = g(\boldsymbol{X}_T) - \mathbb{E}\big(g(\boldsymbol{X}_T) \bigm| \mathcal{A}_{t_0}\big) = g(\boldsymbol{X}_T) - \mathbb{E}\big(g(\boldsymbol{X}_T)\big) \geq g(\boldsymbol{X}_T) - a.$$

Les processus $\boldsymbol{X}$, $\boldsymbol{Z}$, $\boldsymbol{U}$ et $\boldsymbol{V}$ vérifient donc bien les contraintes de (2).

# Retour à la dimension finie IV

## Équivalence avec le problème original : (2) $\subset$ (1)

Soient des processus $\boldsymbol{X}$, $\boldsymbol{Z}$, $\boldsymbol{U}$ et $\boldsymbol{V}$ vérifiant les contraintes de (2). On remarque que $\boldsymbol{Z}_T = \sum_{t=t_0}^{T-1} \boldsymbol{V}_t$, d'où :

$$\mathbb{E}\left(\boldsymbol{Z}_T\right) = \sum_{t=t_0}^{T-1} \mathbb{E}\big(\mathbb{E}\left(\boldsymbol{V}_t \mid \mathcal{A}_t\right)\big) = 0.$$

Comme, par hypothèse, on a que $g(\boldsymbol{X}_T) - \boldsymbol{Z}_T \leq a$, on en déduit que $\mathbb{E}\left(g(\boldsymbol{X}_T)\right) \leq a$, ce qui montre que les processus $\boldsymbol{X}$ et $\boldsymbol{U}$ vérifient les contraintes du problème (1).

Les deux problèmes ont même ensemble admissible et même critère, d'où l'équivalence.

# Retour à la dimension finie V

La contrainte $\mathbb{E}\left(\boldsymbol{V}_t \mid \mathcal{A}_t\right) = 0$ ne pose pas de problème à la programmation dynamique

## Cadre simplifié

On se donne :

- une variable $\boldsymbol{X}$ mesurable par rapport à $\mathcal{G} := \sigma(\boldsymbol{Y})$,
- une variable $\boldsymbol{W}$ indépendante de $\boldsymbol{Y}$,
- la tribu $\mathcal{F}$ engendrée par $(\boldsymbol{Y}, \boldsymbol{W})$.

On cherche à résoudre :

$$\min_{\boldsymbol{U} \preceq \mathcal{F}} \mathbb{E}\big(j(\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{W}) \bigm| \mathcal{G}\big) \quad \text{sous} \quad \mathbb{E}\big(\boldsymbol{U} \bigm| \mathcal{G}\big) = 0.$$

## Théorème

Il existe une solution $\boldsymbol{U}^\sharp$ du problème telle que $\boldsymbol{U}^\sharp$ soit mesurable par rapport au couple $(\boldsymbol{X}, \boldsymbol{W})$.

# Retour à la dimension finie VI

## Preuve dans le cas où $\boldsymbol{Y}$ est finie

- On note $y_1, \ldots, y_N$ les valeurs que prend $\boldsymbol{Y}$.
- Le problème est formé de $N$ sous-problèmes d'optimisation indépendants :

$$\min_{\boldsymbol{U} \preceq (\boldsymbol{Y}, \boldsymbol{W})} \mathbb{E}\Big(j(\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{W}) \;\Big|\; \boldsymbol{Y} = y_i\Big) \quad \text{sous} \quad \mathbb{E}\big(\boldsymbol{U} \;\big|\; \boldsymbol{Y} = y_i\big) = 0.$$

  $\boldsymbol{X}$ mesurable par rapport à $\boldsymbol{Y} \Rightarrow \exists \vartheta$ mesurable telle que $\boldsymbol{X} = \vartheta(\boldsymbol{Y})$ ; on note $x_i = \vartheta(y_i)$.

- $\boldsymbol{U}$ mesurable par rapport à $(\boldsymbol{Y}, \boldsymbol{W}) \Rightarrow$

$$\boldsymbol{U} = \sum_{i=1}^{N} \boldsymbol{U}_i \mathbf{1}_{\{\boldsymbol{Y} = y_i\}},$$

  où chaque $\boldsymbol{U}_i$ est mesurable par rapport à $\boldsymbol{W}$.

# Retour à la dimension finie VII

## Preuve dans le cas où $\boldsymbol{Y}$ est finie

- Le problème "$i$" se met sous la forme :

$$\min_{\boldsymbol{U}_i \preceq \boldsymbol{W}} \mathbb{E}\big(j(x_i, \boldsymbol{U}_i, \boldsymbol{W})\big) \quad \text{sous} \quad \mathbb{E}\big(\boldsymbol{U}_i\big) = 0.$$

La solution $\boldsymbol{U}_i^\sharp$ de ce dernier problème :

- est mesurable par rapport à $\boldsymbol{W}$,
- dépend paramétriquement de la valeur $x_i$ (plutôt que de $y_i$).

On en déduit que toute solution $\boldsymbol{U}^\sharp$ du problème de départ se met sous la forme :

$$\boldsymbol{U}^\sharp = \sum_{i=1}^{N} \boldsymbol{U}_i^\sharp \mathbf{1}_{\{\boldsymbol{X}=x_i\}},$$

et donc que $\boldsymbol{U}^\sharp$ est mesurable par rapport à $(\boldsymbol{X}, \boldsymbol{W})$.

Retour