Dynamic choice theory and temporal consistency

J.-Ph. Chancelier

Cermics, École des Ponts ParisTech.
SOWG

July 6 2011

Informal presentation

For a sequence of dynamic optimization problems, we aim at discussing a notion of consistency over time.

At the very first time step t_0, formulate an optimization problem that yields optimal decision rules for all the forthcoming time steps $t_0, t_1, \dots, T$;

At the next time step t_1, formulate a new optimization problem starting at time t_1 that yields a new sequence of optimal decision rules. This process can be continued until the final time T is reached.

A family of optimization problems formulated in this way is said to be dynamically consistent if the optimal strategies obtained when solving the original problem remain optimal for all subsequent problems.

A typical case : dynamic programing

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

$$\text{s.t.} \quad \mathbf{X}_{t_0} \sim \mu_{t_0}, \tag{1b}$$

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

$$\mathbf{U}_t \preceq \mathbf{X}_{t_0}, \mathbf{W}_{t_1}, \ldots, \mathbf{W}_t, \qquad \forall t = t_0, \ldots, T-1. \tag{1d}$$

Joint probability constraints

We add a chance constraint of the form :

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

Note that this constraint can be modeled through an expectation constraint by introducing a new binary state variable :

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

and considering constraint $\mathbb{E}\left[\mathbf{Y}_T\right] \geq a$. □

Questions

- How to recover dynamic consistency in the previous problem ?
- What happens if expected utility is replaced by risk measures ?
- How was this question formulated in economics when studying preferences ?
- How to get conditions on risks measures to obtain dynamic consistency.

Outline :

- Kreps and Porteus work on expected utility.
- Extensions to risks measures.

Preferences in a dynamic context

The aim of Kreps and Porteus work was to study preferences in a dynamic framework :
At time $t = 0$ do I prefer the sequence $(Z_0, \ldots, Z_T)$ of random variables or the sequence $(W_0, \ldots, W_T)$.
In a similar way we want to compare preferences for random variables at time t but in the light of past events.

- Compare $(z_0, \ldots, z_{t-1}, Z_t \ldots, Z_T)$ and $(z_0, \ldots, z_{t-1}, W_t \ldots, W_T)$.

Main question : Are my choices consistent between times.

Mathematical framework

First, consider a given sequence of spaces $(\mathcal{Z}_t)_{t=0,\dots,T}$. Then recursively define a sequence of spaces $(\mathcal{D}_t)_{t=0,\dots,T}$:

$$\mathcal{D}_T = \mathcal{P}(\mathcal{Z}_T)$$

Borel probability measures on $\mathcal{Z}_T$. Then for $t < T$:

$$\mathcal{D}_t = \mathcal{P}(\mathcal{Z}_t \times \mathcal{D}_{t+1}) .$$

We also define : $\mathcal{C}_t = \mathcal{Z}_t \times \mathcal{D}_{t+1}$ ($\mathcal{C}_T = \mathcal{Z}_T$) and information space at time t : $\mathcal{Y}_t = \mathcal{Z}_0 \times \cdots \times \mathcal{Z}_{t-1}$.

Actions

- An element $d_t \in D_t$ at time t is called an action.
- Note that, given T random variables $(Z_0, \ldots, Z_T)$ such that $Z_i \in \mathcal{Z}_i$ for $i = 0, \ldots, T$ there exists a strategy d_0 for which the probability induced by d_0 on the space $\prod_{i=0}^{T} \mathcal{Z}_i$ is the probability law of $(Z_1, \ldots, Z_T)$.
  When the probability is given by $d_0 \in \mathcal{D}_0$ it is obtained as $(\mathbb{P}_{Z_1}, \mathbb{P}_{Z_2}^{Z_1}, \ldots, \mathbb{P}_{Z_T}^{Z_1,\ldots,Z_{T-1}})$.
- We can consider two actions d_0 and d_0' in D_0 which are such that they give the same marginal law on $(Z_0, \ldots, Z_{t-1})$ which is the degenerated law $\delta_{z_0,\ldots,z_t}$. These two actions also induce a marginal law on $(Z_t, \ldots, Z_T)$ and can thus be considered as actions in D_t.

Utility functions

If the preferences follows three given axioms, then they are driven by a utility function [1, Lemma 3].

Lemma

[1, Lemma 3] Axioms 2.1, 2.2, and 2.3 are necessary and sufficient for there exists, for each t, a (bounded) continuous function $U_{y_t} : \mathcal{C}_t \to \mathbb{R}$ *such that for* $d, d' \in D_t$, $d \succcurlyeq_{y_t} d'$ *if and only if* $\mathbf{U}_{y_t}(d) \geq \mathbf{U}_{y_t}(d')$ *where* $\mathbf{U}$ *is defined by* $\mathbf{U}_{y_t}(d) \overset{\mathrm{def}}{=} \mathbb{E}_d\left[U_{y_t}\right]$.

In the previous Lemma, $y_t \in \mathcal{Y}_t$ is here to tell that the actions at time t are taken knowing a common past described by y_t.

Temporal Consistency

In order to tie together preferences at different times a temporal consistency axiom is introduced :

Axiom

[1, Axom 3.1 (Temporal consistency)] For all t, $y \in \mathcal{Y}_t$, $z \in \mathcal{Z}_t$ *and* x, $x' \in \mathcal{D}_{t+1}$, $\delta_{(z,x)} \succcurlyeq_y \delta_{(z,x')}$ *at time* t *if and only if* $x \succcurlyeq_{(y,z)} x'$ *at time* $t+1$.

Temporal Consistency for expected utility

When the preferences are given by expected utility :

Lemma

The if (resp. only if) part of Axiom 2 is equivalent to the following statement : For all t, $y \in \mathcal{Y}_t$, $z \in \mathcal{Z}_t$ and all x, $x' \in \mathcal{D}_{t+1}$, $U_y(z,x) \geq U_y(z,x')$ if (resp. only if) $\mathbf{U}_{(y,z)}(x) \geq \mathbf{U}_{(y,z)}(x')$.

Proof : The proof is straightforward using representation lemma. We have :

$$\mathbf{U}_{(y,z)}(x) \geq \mathbf{U}_{(y,z)}(x') \Leftrightarrow x \succcurlyeq_{(y,z)} x'$$

and

$$U_y(z,x) \geq U_y(z,x') \Leftrightarrow \mathbf{U}_y(\delta_{(z,x)}) \geq \mathbf{U}_y(\delta_{(z,x')}) \Leftrightarrow \delta_{(z,x)} \succcurlyeq_{(y)} \delta_{(z,x')} \,.$$

□

Utility at different times

Now we consider utility at different times. For that purpose, given $y_t \in \mathcal{Y}_t$ we define the multi-application $u_{y_t} : \mathcal{Z}_t \times \mathbb{R} \rightrightarrows \mathbb{R}$ as follows :

$$u_{y_t}(z,\gamma) \stackrel{\text{def}}{=} \left\{ U_{y_t}(z,x) : x \in \mathbf{U}^{-1}_{(y_t,z)}(\gamma) \right\} .$$

The right-hand side of the previous equality is reduced to the empty set if $\gamma \notin \text{im}(\mathbf{U}_{(y_t,z)})$.

We then have the following lemma :

Lemma

The if part of Axiom (2) *is satisfied at time* t *if and only if the multi-application* u_{y_t} *is in fact an application which is increasing in its second argument. Moreover, when the multi-application* u_{y_t} *is an application, it satisfies :*

$$U_{y_t}(z,x) = u_{y_t}(z, \mathbf{U}_{(y_t,z)}(x)) \forall (z,x) \mathcal{Z}_t \times \mathcal{D}_{t+1} .$$

Idea of the proof

We consider two functions $f : \mathcal{X} \to \mathbb{R}$ and $g : \mathcal{X} \to \mathbb{R}$ and we suppose that they satisfy the following property

$$(\mathbb{T}) \qquad \forall (x, x') \in \mathcal{X}^2 \quad f(x) \geq f(x') \Longrightarrow g(x) \geq g(x')$$

We associate to the pair (f, g) a multi-application $\phi : \mathbb{R} \rightrightarrows \mathbb{R}$ by :

$$\phi(\gamma) \stackrel{\text{def}}{=} \left\{ g(x) : x \in f^{-1}(\gamma) \right\} \ .$$

For $x \in \mathcal{X}$, $\phi(\gamma) \neq \emptyset$ for $\gamma = f(x)$ and we have :

$$\forall x \in \mathcal{X} \quad g(x) \in \phi(f(x)) \, .$$

This means that if ϕ is an application then it must satisfy $g(x) = \phi(f(x))$ for all $x \in \mathcal{X}$.

Idea of the proof

The next theorem give the link between property $(\mathbb{T})$ and properties of ϕ.

Theorem

The two functions f and g satisfy property $(\mathbb{T})$ if and only if the multi-application ϕ is in fact an application which is also increasing.

Note the similarity with the measurability condition.
Note that the the only if part just gives the strictly increasing property.

A simple two time steps example

Two random variables (Z_0, Z_1) and preferences given $\mathbb{E}\left[f(Z_0, Z_1)\right]$.
At time 0, $(Z_0^1, Z_1^1) \succcurlyeq (Z_0^2, Z_1^2)$ if and only if
$\mathbb{E}\left[f(Z_0^1, Z_1^1)\right] \geq \mathbb{E}\left[f(Z_0^2, Z_1^2)\right]$.
Using $\mathbb{P}_{Z_0}$ the probability law of Z_0 and $\mathbb{P}_{Z_1}^{Z_O}$ the conditional probability of Z_1 knowing Z_0 we have :

$$\mathbb{E}\left[f(Z_0, Z_1)\right] = \int_{\mathbb{R}} \mathbb{P}_{Z_0}(dz_0) \int_{\mathbb{R}} f(z_0, z_1) \mathbb{P}_{Z_1}^{Z_0 = z_0}(dz_1)$$

Introduce $U_{z_0}(z_1) \stackrel{\text{def}}{=} f(z_0, z_1)$ and for a given probability $p_1 \in \mathcal{P}(\mathbb{R})$ we have $\mathbf{U}_{z_0}(p_1) = \int_{\mathbb{R}} f(z_0, z_1) p_1(dz_1)$. Moreover, we can define $U(z_0, p_1) = \mathbf{U}_{z_0}(p_1)$ and a definition of $\mathbf{U}(p_0)$ is derived by expectation. Considering the probability p_0 which gives a probability $\mathbb{P}_{Z_0}(x)$ to the point $x \times \mathbb{P}_{Z_1}^{Z_0 = x}$ we obtain :

$$\mathbf{U}(p_0) = \int_{\mathbb{R}} U((z_0, \mathbb{P}_{Z_1}^{Z_0 = z_0})) = \int_{\mathbb{R}} p_0(dz_0) \mathbf{U}_{z_0}(\mathbb{P}_{Z_1}^{Z_0 = z_0}) = \mathbb{E}\left[f(Z_0, Z_1)\right]$$

A simple two time steps example

We specialize the previous example to produce a more involved u function. We suppose here that $f(z_0, z_1) = g(z_0) + h(z_0)j(z_0, z_1)$. Then, we can decide to use at time $t = 1$ the utility function based on function j and use the utility function $f((z_0, p_1)) = g(z_0) + h(z_0) \int_{\mathbb{R}} j(z_0, z_1) p_1(dz1)$ at time $t = 0$. The function $u(z_0, \gamma)$ which ties utility functions at time 0 and 1 is then given by :

$$u(z_0, \gamma) = g(z_0) + h(z_0)\gamma \,.$$

For this new choice of utility functions, we still have the fact that for p_0 which is associated to the law of (Z_0, Z_1), $\mathbf{U}(p_0) = \mathbb{E}\left[f(Z_0, Z_1)\right]$. The utility functions are time consistent if h is a strictly positive function.

The case of conditional risk measures

We suppose now that the preferences at time t are given by a risk measure $\rho_{t,T}$. The sequence $\rho_{t,T}$ for $t = 1, \dots, T$ is defined on $(\Omega, (\mathcal{F}_t)_{t=1,\dots,T}, \mathcal{F}, \mathbb{P})$

$$\rho_{t,T} : L_p^{(t)} \times \cdots \times L_p^{(T)} \to L_p^{(t)} , \quad L_p^{(t)} = L_p(\Omega, \mathcal{F}_t, \mathbb{P})$$

and satisfies

$$\rho_{t,T}(Z) \leq \rho_{t,T}(W) \quad \text{if} \quad Z \leq W .$$

The case of conditional risk measures

Copying what has been done so far on preferences with expected utility and in order to tie together risk measures at different times we introduce a definition of *t*ime-consistent dynamic risk measure. We use a definition similar to if part of Axiom (2) :

Definition

A dynamic risk measure $\{\rho_{t,T}\}_{t=0,\dots,T}$ *is called time-consistent if for all* $t \in \{0, \dots, T\}$ *and all sequences* Z, W *the conditions* $Z_t = W_t$ *and* $\rho_{t+1,T}(Z_{t+1}, \dots, Z_T) \leq \rho_{t+1,T}(W_{t+1}, \dots, W_T)$ *imply that* $\rho_{t,T}(Z_t, \dots, Z_T) \leq \rho_{t,T}(W_t, \dots, W_T)$.

Using this definition it is possible to obtain a representation theorem similar to the previous one.

Ruszczynski definition

Note that in [2] a different definition is used

Definition

A dynamic risk measure $\{\rho_{t,T}\}_{t=0,\dots,T}$ *is called time-consistent if for all* $0 \leq t < \theta \leq T$ *and all sequences* Z, W *the conditions* $Z_k = W_k$, $k = t, \dots, \theta - 1$ *and* $\rho_{\theta,T}(Z_\theta, \dots, Z_T) \leq \rho_{\theta,T}(W_\theta, \dots, W_T)$ *imply that* $\rho_{t,T}(Z_t, \dots, Z_T) \leq \rho_{t,T}(W_t, \dots, W_T)$.

It is easy to prove that the two definitions are equivalent.

A new theorem

Theorem

Let $\{\rho_{t,T}\}_{t=0,\dots,T}$ be a time-consistent dynamic risk measure, then there exist for each t, a measurable function $\phi_t : \mathcal{Z}_t \times \mathcal{D}_{t+1} \to \mathbb{R}$ such that :

$$\rho_{t,T}((X_t,\dots,X_T)) = \phi_t(X_t, \rho_{t+1,T}((X_{t+1},\dots,X_T)))\,.$$

The function ϕ_t is increasing in its second argument. The reverse statement is also true. If equation (20) is satisfied with a function ϕ_t increasing in its second argument then $\{\rho_{t,T}\}_{t=0,\dots,T}$ is time-consistent.

A new theorem 1/3

Proof : The proof is very similar to the one given for expected utility preferences. We consider the multi-application :

$$\phi_t(X_t, \Gamma_{t+1}) = \left\{ \rho_{t,T}((X_t, \dots, X_T)) : (X_{t+1}, \dots, X_T) \in \rho^{-1}_{t+1,T}(\Gamma_{t+1}) \right\} .$$

One easily check that the multi-application is in fact an application if the dynamic risk measure is time-consistent. The reverse statement is easily verified. □

A new theorem 2/3

Lemma

[2] Let $\{\rho_{t,T}\}_{t=0,\ldots,T}$ *be a dynamic risk measure and suppose that for all* $t \in (0, T)$ *we have :*

$$\rho_{t,T}(X_t, X_{t+1}, \ldots, X_T) = \rho_{t,T}(X_t, \rho_{t+1,T}(X_{t+1}, \ldots, X_T), 0, \ldots, 0)$$

then $\{\rho_{t,T}\}_{t=0,\ldots,T}$ *is a time-consistent dynamic risk measure.*

A new theorem 3/3

Lemma

Let $\{\rho_{t,T}\}_{t=0,\dots,T}$ be a time consistent dynamic risk measure such that $\rho_{t,T}(X_t, 0, \dots, 0) = X_t$ then we have

$$\rho_{t,T}(X_t, X_{t+1}, \dots, X_T) = \rho_{t,T}(X_t, \rho_{t+1,T}(X_{t+1}, \dots, X_T), 0, \dots, 0) \,.$$

Proof : There exist for each t, $\phi_t : \mathcal{Z}_t \times \mathcal{D}_{t+1} \to \mathbb{R}$:

$$\rho_{t,T}((X_t, \dots, X_T)) = \phi_t(X_t, \rho_{t+1,T}((X_{t+1}, \dots, X_T))) \,.$$

Thus we have

$$\rho_{t,T}(X_t, X_{t+1}, 0, \dots, 0) = \phi_t(X_t, \rho_{t+1,T}(X_{t+1}, 0, \dots, 0))) = \phi_t(X_t, X_{t+1})$$

Thus $\phi_t(X, Y) = \rho_{t,T}(X, Y, 0, \dots, 0)$ and we conclude. □

D. M. Kreps, E. L. Porteus, Temporal resolution of uncertainty and dynamic choice theory, Econometrica 46 (1) (1978) 185–200.

A. Ruszczyński, Risk-averse dynamic programming for markov decision processes, Mathematical Programming Series BTo appear.