A Theory of Time Inconsistent Optimal Control

Tomas Björk   Agatha Murgoci

Department of Finance,
Stockholm School of Economics,

KTH, June 2008

Contents

- Recap of DynP.
- Problem formulation.
- Discrete time.
- Continuous time.
- Examples.

Standard problem

We are standing at time $t = 0$ in state $X_0 = x_0$.

$$\max_u \quad E\left[\int_0^T h(s, X_s, u_s)dt + F(X_T)\right]$$

$$dX_t = \mu(t, X_t, u_t)dt + \sigma(t, X_t, u_t)dW_t$$

For simplicty we assume that

- X is scalar.
- The adapted control u_t is scalar with no restrictions.

We denote this problem by $\mathcal{P}$

We restrict ourselves to **feedback controls** of the form

$$u_t = u(t, X_t).$$

Dynamic Programming

We embed the problem $\mathcal{P}$ in a family of problems $\mathcal{P}_{tx}$

$\mathcal{P}_{tx}$:

$$\max_{u} \quad E_{t,x}\left[\int_t^T h(s, X_s, u_s)dt + F(X_T)\right]$$

$$\begin{aligned} dX_s &= \mu(t, X_s, u_s)ds + \sigma(s, X_s, u_s)dW_s, \\ X_t &= x \end{aligned}$$

The original problem corresponds to $\mathcal{P}_{0,x_0}$.

Def:
For $\mathcal{P}_{tx}$, we denote the **optimal value function** by $V(t, x)$ and the **optimal control law** by $\hat{u}(s, y)$.

In principle, the optimal control law for $\mathcal{P}_{tx}$ should be denoted $\hat{u}_{t,x}(s, y)$, but:

Bellman

We now have the Bellman optimality principle, which says that the family $\{\mathcal{P}_{t,x};\ t \geq 0, x \in R\}$ are **time consistent**.

More precisely: If $\hat{u}$ is optimal on the time interval $[t, T]$, then it is also optimal on the sub-interval $[s, T]$ for every s with $t \leq s \leq T$.

We also have the Hamilton-Jacobi-Bellman equation

HJB:

$$
\begin{aligned}
V_t(t,x) + \sup_u \left\{ h(t,x,u) + \mu(t,x,u)V_x(t,x) + \frac{1}{2}\sigma^2(t,x,u)V_{xx}(t,x) \right\} &= 0, \\
V(T,x) &= F(x)
\end{aligned}
$$

Three Disturbing Examples

Hyperbolic discounting

$$\max_u \quad E_{t,x}\left[\int_t^T \varphi(T-t)h(X_s,u_s)dt + F(X_T)\right]$$

Mean variance utility

$$\max_u \quad E_{t,x}\left[X_T\right] - \frac{\gamma}{2}Var_{t,x}\left(X_T\right)$$

Endogenous habit formation

$$\max_u E_{t,x}\left[\ln\left(X_T - x\right)\right]$$

$$dX_t = [rX_t + (\alpha - r)u_t]dt + \sigma u_t dW_t$$

Moral

- These types of problems are **not** time consistent.
- We cannot use DynP.
- In fact, in these cases it is unclear what we mean by "optimality".

Possible ways out:

- **Easy way:** Dismiss the problem as being silly.
- **Pre-commitment:** Solve (somehow) the problem $\mathcal{P}_{0,x_0}$ and ignore the fact that later on, your "optimal" control will no longer be viewed as optimal.
- **Game theory:** Take the time inconsistency seriously. View the problems as a game and look for a Nash equilibrium point.

We use the game theoretic approach.

Ekeland-Lazrak-Pirvu

Maximize expected utility of investment/consumption with hyperbolic discounting

$$\max_u \quad E_{t,x}\left[\int_t^T \varphi(T-t)h(X_s,u_s)dt + F(X_T)\right]$$

Portfolio dynamics

$$dX_t = [rX_t + (\mu - r)u_t]dt + \sigma u_t dW_t$$

Results:

- Very precise problem statement and analysis.
- Verification theorem proved.
- Explicit solution when

$$\varphi(T-t) = \alpha e^{a(T-t)} + \beta e^{-b(T-t)}$$

Basak-Chabakauri:

Mean variance optimal investment.

$$\max_u \quad E_{t,x}\left[X_T\right] - \frac{\gamma}{2} Var_{t,x}\left(X_T\right)$$

$$dX_t = [rX_t + (\alpha - r)u_t]dt + \sigma u_t dW_t$$

Results:

- Very nice explicit solution, including hidden Markov model.
- The formal equilibrium problem is never given a precise definition.
- No verification theorem.
- Relies heavily on "total variance formula", so the method is hard to generalize from mean-variance.
- The arguments are not completely precise, but more heuristic.

Contributions of present paper

Present paper:

- We study a considerably more general problem than in previous papers.
- We derive a system of PDEs, extending the standard HJB equation from DynP.
- Earlier results included as special cases.
- Precise definition of equilibrium given (inspired by Ekeland *et al*).
- Verification theorem proved (inspired by Ekeland *et al*).
- We prove that for every time inconsistent problem there is an equivalent **consistent** problem with the same optimal strategy.
- Particular cases explicitly solved.

Our Basic Problem

$$\max_u \quad E_{t,x}\left[\int_t^T C(t,x,X_s,u_s)ds + F(t,x,X_T)\right] + G\left(t,x,E_{t,x}\left[X_T\right]\right)$$

$$\begin{aligned} dX_s &= \mu(t,X_s,u_s)ds + \sigma(s,X_s,u_s)dW_s, \\ X_t &= x \end{aligned}$$

This can be extended considerably.

For simplicity we will consider the easier problem

$$\max_u \quad E_{t,x}\left[F(X_T)\right] + G\left(E_{t,x}\left[X_T\right]\right)$$

The Game Theoretic Approach

- We view this as a game where there is one player for each t.

- Player No t chooses the control function $u(t, \cdot)$ at time t, and applies the control $u(t, X_t)$

- The value, to player No t, if all players use the control law u is

$$J(t, x; u) = E_{t,x}\left[F\left(X_T^u\right)\right] + G\left(E_{t,x}\left[X_T^u\right]\right)$$

Def: The strategy $\hat{u}$ is a **Nash subgame perfect equilibrium** if the following hold for all t.

- Assume that all players No s with $s > t$ use the control $\hat{u}(s, X_s)$.

- Then it is optimal for player No t also to use $\hat{u}(t, X_t)$.

- This is a bit delicate to formalize in continuous time.
- Thus we turn to discrete time, and then go to the limit.

Discrete Time

Given: A controlled Markov process $\{X_n : n = 0, 1, \ldots T\}$

Def:

- For each n and each fixed real number $u \in R$ we have the transition probabilities

$$p_n^u(x, dz) = P\left(X_{n+1} \in dz \,|X_n = x,\ u_n = u\right)$$

- The operator $\mathbf{P}^u$ is defined for a function sequence $\{f_n(x)\}$, where $f_n : R \to R$ by

$$\left(\mathbf{P}^u f\right)_n(x) = \int_R f_{n+1}(z) p_n^u(x, dz)$$

$$\left(\mathbf{P}^u f\right)_n(x) = E\left[f_{n+1}\left(X_{n+1}\right) \middle| X_n = x,\ u_n = u\right]$$

- The "infinitesimal operator" $\mathbf{A}^u$ is defined by

$$\mathbf{A}^u = \mathbf{P}^u - \mathbf{I}$$

Equilibrium

Def:

- The value function is defined by

$$J_n(x, \bar{u}) = E_{n,x}\left[F(X_T^{\bar{u}})\right] + G\left(E_{n,x}\left[X_T^{\bar{u}}\right]\right)$$

- The control law $\hat{u}$ is an **equilibrium strategy** if the following hold for each fixed n.
  - Assume that all players No k for $k = n+1, \ldots, T-1$ use $\hat{u}_k(\cdot)$.
  - Then it is optimal for player No n to use $\hat{u}_n(\cdot)$.

- The equilibrium value function is defined by

$$V_n(x) = J_n(x, \hat{u})$$

Important Idea

It turns out that a fundmental role is plyed by the function sequence f_n defined by

$$f_n(x) = E_{n,x}\left[X_T^{\hat{u}}\right]$$

where $\hat{u}$ is the equilibrium strategy.

The process $f_n(X_n)$ is of course a martingale under the equilibrium control $\hat{u}$ so we have

$$\begin{aligned} \mathbf{A}^{\hat{u}} f_n(x) &= 0, \\ f_T(x) &= x. \end{aligned}$$

Extending HJB

Proposition: The equilibrium value function satisfies the system

$$
\begin{aligned}
\sup_u \left\{ \mathbf{A}^u V_n(x) - \mathbf{A}^u (G \circ f)_n (x) + (\mathbf{H}^u f)_n (x) \right\} &= 0, \\
V_T(x) &= F(x) + G(x) \\
\mathbf{A}^{\hat{u}} f_n(x) &= 0, \\
f_T(x) &= x.
\end{aligned}
$$

$$
(\mathbf{H}^u f)_n (x) = G\left(\mathbf{P}^u f_n(x)\right) - G\left(f_n(x)\right), \quad f_n(x) = E_{n,x}\left[X_T^{\hat{u}}\right]
$$

Note the fixed point character of the problem.

Continuous Time

The discrete time results extend immediately to continuous time.

- Now X is a controlled continuous time Markov process with controlled infinitesimal generator

$$\mathbf{A}^u g(t,x) = \lim_{h \to 0} \frac{1}{h} \left\{ E_{t,x} \left[g(t+h, X_{t+h}^u) \right] - g(t,x) \right\}$$

- The extended HJB is now an equation with time step $[t, t+h]$.

- Divide the discrete time HJB equations by h and let $h \to 0$.

Extended HJB Continuous Time

Proposition: The optimal value function satisfies the system

$$
\begin{aligned}
\sup_u \left\{ \mathbf{A}^u V(t,x) - \mathbf{A}^u \left(G \circ f\right)(t,x) + \left(\mathbf{H}^u f\right)(t,x) \right\} &= 0, \\
\mathbf{A}^{\hat{u}} f(t,x) &= 0, \\
V(T,x) &= F(x) + G(x) \\
f(T,x) &= x.
\end{aligned}
$$

$$
\left(\mathbf{H}^u f\right)(t,x) = \lim_{h \to 0} \frac{1}{h} \left\{ G\left(E_{t,x}\left[f(t+h, X^u_{t+h}\right]\right) - G\left(f(t,x)\right) \right\}
$$

Note the fixed point character of the extended HJB.

The operator $\mathbf{H}^u$

$$\mathbf{H}^u f(t,x) = \lim_{h\to 0} \frac{1}{h} \left\{ G\left(E_{t,x}\left[f(t+h, X^u_{t+h}] \right) - G\left(f(t,x)\right) \right\}$$

We have, to first order,

$$E_{t,x}\left[f(t+h, X^u_{t+h}) \right] = f(t,x) + \mathbf{A}^u f(t,x) h$$

Thus, to first order,

$$\begin{aligned} & G\left(E_{t,x}\left[f(t+h, X^u_{t+h}] \right) \\ & = G\left(f(t,x)\right) + G'\left(f(t,x)\right) \cdot \mathbf{A}^u f(t,x) h \end{aligned}$$

Thus

$$\mathbf{H}^u f(t,x) = G'\left(f(t,x)\right) \cdot \mathbf{A}^u f(t,x)$$

Extended HJB Continuous Time

Proposition: The optimal value function satisfies the system

$$\begin{aligned}
\sup_u \left\{ \mathbf{A}^u V(t,x) - \mathbf{A}^u (G \circ f)(t,x) + G'(f(t,x)) \cdot \mathbf{A}^u f(t,x) \right\} &= 0, \\
\mathbf{A}^{\hat{u}} f(t,x) &= 0, \\
V(T,x) &= F(x) + G(x) \\
f(T,x) &= x.
\end{aligned}$$

Diffusion Case

If X is a scalar SDE of the form

$$dX_t = \mu(X_t, u_t)dt + \sigma(X_t, u_t)dW_t$$

then the extended HJB takes the form

$$\begin{aligned}
\sup_u \left\{ \mathbf{A}^u V(t,x) - \frac{1}{2}\sigma^2(x,u)G''\left(f(t,x)\right) f_x^2(t,x) \right\} &= 0, \\
\mathbf{A}^{\hat{u}} f(t,x) &= 0, \\
V(T,x) &= F(x) + G(x) \\
f(T,x) &= x.
\end{aligned}$$

Optimal for what?

In continuous time, it is not immediately clear how to define an equilibrium strategy. We follow Ekeland *et al*.

- Consider a fixed control law $\hat{u}$.
- Fix (t, x) and a "small" time increment h.
- Choose an arbitrary real number u.
- Consider the control law $\bar{u}_h(t, x)$ defined by

$$\bar{u}_h(s, y) = \begin{cases} \hat{u}(s, y) & \text{for} \quad t + h \leq s \leq T \\ u & \text{for} \quad t \leq s \leq t + h \end{cases}$$

Def: The control law $\hat{u}$ is an **equilibrium control** if

$$\lim_{h \to 0} \frac{J(t, x, \hat{u}) - J(t, x, \bar{u}_h)}{h} \geq 0$$

for all choices of t, x, h, u.

Verification Theorem

Theorem: Assume that V, f and $\hat{u}$ satisfies the extended HJB system. Then V is the equilibrium value function and $\hat{u}$ is the equilibrium control.

Connection to Standard Problems

- Assume that we **know** the equilibrium strategy $\hat{u}$.
- Then we can compute f.
- Now **define** the function $h(t,x,u)$ by

$$h(t,x,u) = (\mathbf{H}^u f)(t,x) - \mathbf{A}^u (G \circ f)(t,x)$$

The extended HJB takes the form

$$\begin{aligned} \sup_u \{\mathbf{A}^u V(t,x) + h(t,x,u)\} &= 0, \\ V(T,x) &= F(x) + G(x) \end{aligned}$$

This is the HJB for the **time consistent** problem

$$\max_u \quad E_{t,x}\left[\int_t^T h(s,X_s,u_s)dt + F(X_T) + G(X_T)\right]$$

Practical handling of the theory

- Make a parameterized Ansatz for V.
- Make a parameterized Ansatz for f.
- Plug everything into the extended HJB system and hope to obtain a system of ODEs for the parameters in the Ansatz.

Basak's Example (in a simple version)

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dB_t &= r B_t dt \end{aligned}$$

$$\begin{aligned} X_t &= \text{portfolio value process} \\ u &= \text{amount of money invested in risky asset} \end{aligned}$$

Problem:

$$\max_u \quad E_{t,x}\left[X_T\right] - \frac{\gamma}{2} Var_{t,x}\left(X_T\right)$$

$$dX_t = [rX_t + (\alpha - r)u_t]dt + \sigma u_t dW_t$$

This corresponds to our standard problem with

$$F(x) = x - \frac{\gamma}{2}x^2, \quad G(x) = \frac{\gamma}{2}x^2$$

Extended HJB

$$
\begin{aligned}
V_t + \sup_u \left\{ [rX_t + (\alpha - r)u]V_x + \frac{1}{2}\sigma^2 u^2 V_{xx} - \frac{\gamma}{2}\sigma^2 u^2 f_x^2 \right\} &= 0 \\
V(T, x) &= x \\
\mathcal{A}^{\hat{u}} f &= 0 \\
f(T, x) &= x
\end{aligned}
$$

Ansatz:

$$
\begin{aligned}
V(t, x) &= g(t)x + h(t) \\
f(t, x) &= A(t)x + B(t)
\end{aligned}
$$

Extended HJB

HJB equation becomes:

$$\begin{aligned} g_t x + h_t + \sup_u \left\{ [rx + (\alpha - r)u] g(t) - \frac{\gamma}{2} \sigma^2 u^2 A^2 \right\} &= 0 \\ g(T) &= 1 \\ h(T) &= 0 \end{aligned}$$

- Embedded static problem:

$$\max_u \left\{ (\alpha - r) g(t) u - \frac{\gamma}{2} \sigma^2 u^2 A^2 \right\}$$

- Optimal control

$$u = \frac{1}{\gamma} \frac{\alpha - r}{\sigma^2} \frac{g(t)}{A^2}$$

Plug back into HJB.

HJB equation becomes:

$$\begin{aligned} g_t x + h_t + grx + \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2}\frac{g(t)^2}{A^2} &= 0 \\ g(T) &= 1 \\ h(T) &= 0 \end{aligned}$$

Separation of variables gives us

$$\begin{aligned} g_t + gr &= 0 \\ g(T) &= 1 \end{aligned}$$

We obtain $g(t) = e^{r(T-t)}$.

Furthermore

$$\begin{aligned} h_t + \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2}\frac{e^{2r(T-t)}}{A^2} &= 0 \\ h(T) &= 0 \end{aligned}$$

We need to solve the PDE for the function f:

$$\begin{aligned}\mathcal{A}^{\hat{u}} f(t,x) &= 0 \\ f(T,x) &= x\end{aligned}$$

The PDE becomes:

$$\begin{aligned}A_t x + B_t + rxA + \frac{1}{\gamma}\frac{(\alpha - r)^2}{\sigma^2}\frac{e^{r(T-t)}}{A} &= 0 \\ A(T) &= 1 \\ B(T) &= 0\end{aligned}$$

Separation of variables gives us

$$\begin{aligned}A_t + Ar &= 0 \\ A(T) &= 1\end{aligned}$$

We obtain

$$A(t) = e^{r(T-t)}$$

Separation also gives us

$$\begin{aligned} B_t &= \frac{1}{\gamma}\frac{(\alpha - r)^2}{\sigma^2} \\ B(T) &= 0 \end{aligned}$$

with solution

$$B(t) = \frac{1}{\gamma}\frac{(\alpha - r)^2}{\sigma^2}(T - t)$$

We go back to the equation for h:

$$\begin{aligned} h_t + \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2} &= 0 \\ h(T) &= 0 \end{aligned}$$

We obtain

$$h(t) = \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2}(T - t)$$

Result

The equilibrium value function and strategy are given by

$$
\begin{aligned}
V(t,x) &= e^{r(T-t)}x + \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2}(T-t) \\
\hat{u}(t,x) &= \frac{1}{\gamma}\frac{\alpha - r}{\sigma^2}e^{-r(T-t)} \\
f(t,x) &= e^{r(T-t)}x + \frac{1}{\gamma}\frac{(\alpha - r)^2}{\sigma^2}(T-t)
\end{aligned}
$$

Equivalent Standard Problem

The Basak problem has the same optimal control as the **time consistent** problem

$$\max_u E_{t,x}\left[X_T - \frac{\gamma\sigma^2}{2}\int_t^T e^{2r(T-s)}u_s^2 ds\right]$$

$$dX_t = [rX_t + (\alpha - r)u_t]dt + \sigma u_t dW_t$$

We note in passing that

$$\sigma^2 u_t^2 dt = d\langle X\rangle_t$$