MA668: Algorithmic and High Frequency Trading
Lecture 23

Prof. Siddhartha Pratim Chakrabarty
Department of Mathematics
Indian Institute of Technology Guwahati

The Dynamic Programming Principle (Contd ...)

1. Note that on the right-hand side of the above, the arbitrary control $\mathbf{u}$ only acts over the interval $[t, T]$ and the optimal one is implicitly incorporated in the value function $H(\tau, \mathbf{X}^{\mathbf{u}}_{\tau})$ but starting at the point to which the arbitrary control $\mathbf{u}$ caused the process $\mathbf{x}$ to flow, namely, $\mathbf{x}^{\mathbf{u}}_{\tau}$.
2. Taking supremum over admissible strategies on the left-hand side, so that the left-hand side also reduces to the value function, we have that:

$$H(t, \mathbf{x}) \leq \sup_{\mathbf{u} \in \mathcal{A}} \mathbb{E}_{t,\mathbf{x}} \left[H(\tau, \mathbf{X}^{\mathbf{u}}_{\tau}) + \int_t^{\tau} F\left(s, \mathbf{X}^{\mathbf{u}}_s, \mathbf{u}_s\right) ds \right]. \tag{1}$$

3. Next, we aim to show that the inequality above can be reversed. Take an arbitrary admissible control $\mathbf{u} \in \mathcal{A}$ and consider what is known as an ϵ-optimal control denoted by $\mathbf{v}^{\epsilon} \in \mathcal{A}$ and defined as a control which is better than $H(t, \mathbf{x}) - \epsilon$, but of course not as good as $H(t, \mathbf{x})$ i.e., a control such that

$$H(t, \mathbf{x}) \geq H^{\mathbf{v}^{\epsilon}}(t, \mathbf{x}) \geq H(t, \mathbf{x}) - \epsilon. \tag{2}$$

The Dynamic Programming Principle (Contd ...)

1. Such a control exists, assuming that the value function is continuous in the space of controls.
2. Consider next the modification of the ϵ-optimal control.

$$\widetilde{\mathbf{v}}^{\epsilon} = \mathbf{u}_t \mathbb{1}_{t \leq \tau} + \mathbf{v}^{\epsilon} 1_{t > \tau}, \tag{3}$$

i.e., the modification is ϵ-optimal after the stopping time T, but potentially sub-optimal on the interval $[t, T]$. Then we have:

$$\begin{aligned}
H(t, \mathbf{x}) &\geq H^{\widetilde{\mathbf{v}}^{\epsilon}}(t, \mathbf{x}) \\
&= \mathbb{E}_{t,\mathbf{x}} \left[H^{\widetilde{\mathbf{v}}^{\epsilon}}(\tau, \mathbf{X}_{\tau}^{\widetilde{\mathbf{v}}^{\epsilon}}) + \int_t^{\tau} F\left(s, \mathbf{X}_s^{\widetilde{\mathbf{v}}^{\epsilon}}, \widetilde{\mathbf{v}}_s^{\epsilon}\right) ds \right], \\
&= \mathbb{E}_{t,\mathbf{x}} \left[H^{\widetilde{\mathbf{v}}^{\epsilon}}(\tau, \mathbf{X}_{\tau}^{\mathbf{u}}) + \int_t^{\tau} F\left(s, \mathbf{X}_s^{\mathbf{u}}, \mathbf{u}_s\right) ds \right], \text{ (using (3))} \\
&\geq \mathbb{E}_{t,\mathbf{x}} \left[H(\tau, \mathbf{X}_{\tau}^{\mathbf{u}}) + \int_t^{\tau} F\left(s, \mathbf{X}_s^{\mathbf{u}}, \mathbf{u}_s\right) ds \right]. \text{ (by (2))}
\end{aligned}$$

The Dynamic Programming Principle (Contd ...)

1. Taking limit as $\epsilon \downarrow 0$, we have,

$$H(t, \mathbf{x}) \geq \mathbb{E}_{t,\mathbf{x}} \left[H(\tau, \mathbf{X}^{\mathbf{u}}_{\tau}) + \int_t^{\tau} F\left(s, \mathbf{X}^{\mathbf{u}}_s, \mathbf{u}_s\right) ds \right].$$

2. Moreover, since the above holds true for every $\mathbf{u} \in \mathcal{A}$, we have that:

$$H(t, \mathbf{x}) \geq \sup_{\mathbf{u} \in \mathcal{A}} \mathbb{E}_{t,\mathbf{x}} \left[H(\tau, \mathbf{X}^{\mathbf{u}}_{\tau}) + \int_t^{\tau} F\left(s, \mathbf{X}^{\mathbf{u}}_s, \mathbf{u}_s\right) ds \right]. \tag{4}$$

3. The upper bound (1) and lower bound (4) form the dynamic programming inequalities. Putting them together, we obtain the following Theorem.

The Dynamic Programming Principle (Contd ...)

Theorem

Dynamic Programming Principle for Diffusions. The value function satisfies the DPP:

$$H(t, \mathbf{x}) = \sup_{\mathbf{u} \in \mathcal{A}} \mathbb{E}_{t,\mathbf{x}} \left[H(\tau, \mathbf{X}^{\mathbf{u}}_{\tau}) + \int_t^{\tau} F\left(s, \mathbf{X}^{\mathbf{u}}_s, \mathbf{u}_s\right) ds \right], \tag{5}$$

for all $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^n$ *and all stopping times* $\tau \leq T$*.*

1. This equation is really a sequence of equations that tie the value function to its future expected value, plus the running reward/penalty.
2. Since it is a sequence of equations, an even more powerful equation can be found by looking at its infinitesimal version, the so-called DPE.

DPE/HJB Equation

The DPE is an infinitesimal version of the dynamic programming principle (DPP). There are two key ideas involved:

Idea 1

Setting the stopping time τ in the DPP to be the minimum between:

- (a) The time it takes for the process $\mathbf{X}_t^{\mathbf{u}}$ to exit a ball of size ϵ around its starting point AND
- (b) A fixed (small) time h: all while keeping it bounded by T.

This can be viewed in Figure 5.2 and can be stated precisely as:

$$\tau = T \wedge \inf \left\{ s > t : (s - t, |\mathbf{X}_s^{\mathbf{u}} - \mathbf{x}|) \notin [0, h) \times [0, \epsilon) \right\}.$$

Notice that as $h \downarrow 0$, $\tau \downarrow t$, a.s. and that $\tau = t + h$ whenever h is sufficiently small: since as the time span h shrinks, it is less and less likely that $\mathbf{X}$ will exit the ball first.

Figure 5.2

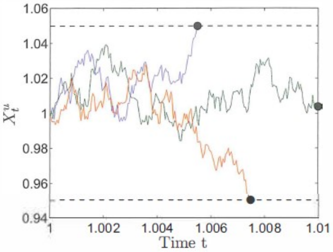

Figure 5.2 The DPE is an infinitesimal version of the DPP where the stopping time τ is the first exit time from a ball of size ϵ or a small time h, whichever occurs first. This sample plot of three paths, and the corresponding value X_τ^u and stopping time τ indicated by the black circles for $\epsilon = 0.05$ and $h = 0.01$.

Figure: Figure 5.2

Idea 2

1. Writing the value function (for an arbitrary admissible control $\mathbf{u}$) at the stopping time τ in terms of the value function at t using Ito's lemma. Specifically, assuming enough regularity of the value function, we can write:

$$H(\tau, \mathbf{X}_\tau^{\mathbf{u}}) = H(t, \mathbf{x}) + \int_t^\tau \left(\partial_s + \mathcal{L}_s^{\mathbf{u}}\right) H(s, \mathbf{X}_s^{\mathbf{u}}) ds + \int_t^\tau \mathcal{D}_x H(s, \mathbf{X}_s^{\mathbf{u}})' \sigma_s^{\mathbf{u}} d\mathbf{W}_s, \tag{6}$$

where $\sigma_t^{\mathbf{u}} := \sigma\left(t, \mathbf{X}_t^{\mathbf{u}}, \mathbf{u}_t\right)$, $\mathcal{L}_t^{\mathbf{u}}$ represents the infinitesimal generator of $\mathbf{X}_t^{\mathbf{u}}$ and $\mathcal{D}_x H(\cdot)$ denotes the vector of partial derivatives with components $[\mathcal{D}_x H(\cdot)]_i = \partial_x^i H(\cdot)$.

2. For example, in the one-dimensional case:

$$\mathcal{L}_t^u = \mu_t^u \partial_x + \frac{1}{2} \left(\sigma_t^u\right)^2 \partial_{xx} = \mu(t, x, u)\partial_x + \frac{1}{2}\sigma^2(t, u, x)\partial_{xx}.$$

DPE/HJB Equation (Contd ...)

1. As before, we derive the DPE in two stages by obtaining two inequalities. First, taking $\mathbf{v} \in \mathcal{A}$ to be constant over the interval $[t, \tau]$, applying the lower bound and substituting (6) into the right-hand side implies that:

$$
\begin{aligned}
H(t, \mathbf{x}) &\geq \sup_{\mathbf{u} \in \mathcal{A}} \mathbb{E}_{t,\mathbf{x}} \left[H(\tau, \mathbf{X}^{\mathbf{u}}_{\tau}) + \int_t^{\tau} F\left(s, \mathbf{X}^{\mathbf{u}}_s, \mathbf{u}_s\right) ds \right], \\
&\geq \mathbb{E}_{t,\mathbf{x}} \left[H(\tau, \mathbf{X}^{\mathbf{v}}_{\tau}) + \int_t^{\tau} F\left(s, \mathbf{X}^{\mathbf{v}}_s, \mathbf{v}_s\right) ds \right], \\
&= \mathbb{E}_{t,\mathbf{x}} \left[H(t, \mathbf{x}) + \int_t^{\tau} \left(\partial_s + \mathcal{L}^{\mathbf{v}}_s\right) H(s, \mathbf{X}^{\mathbf{v}}_s) ds \right. \\
&+ \left. \int_t^{\tau} \mathcal{D}_x H(s, \mathbf{X}^{\mathbf{v}}_s)' \sigma^{\mathbf{v}}_s d\mathbf{W}_s + \int_t^{\tau} F\left(s, \mathbf{X}^{\mathbf{v}}_s, \mathbf{v}_s\right) ds \right].
\end{aligned}
$$

DPE/HJB Equation (Contd ...)

1. The integrand in the stochastic integral above, $\mathcal{D}_x H(s, \mathbf{X}_s^{\mathbf{v}})' \sigma_s^{\mathbf{v}}$, is bounded on the interval $[t, \tau]$, since we have ensured that $|\mathbf{X}_t^{\mathbf{v}} - \mathbf{x}| \leq \epsilon$ on the interval.
2. Hence, this stochastic integral is the increment of a martingale and we can be assured that its expectation is zero.
3. Therefore:

$$H(t, \mathbf{x}) \geq \mathbb{E}_{t,\mathbf{x}} \left[H(t, \mathbf{x}) + \int_t^\tau \left\{ (\partial_s + \mathcal{L}_s^{\mathbf{v}}) H(s, \mathbf{X}_s^{\mathbf{v}}) + \int_t^\tau F(s, \mathbf{X}_s^{\mathbf{v}}, \mathbf{v}) \right\} ds \right],$$

and recall that $\tau = t + h$.

DPE/HJB Equation (Contd ...)

1. Moving the $H(t, \mathbf{x})$ on the left-hand side over to the right-hand side, dividing by h and taking the limit as $h \downarrow 0$ yields:

$$
\begin{aligned}
0 \quad &\geq \quad \lim_{h \downarrow 0} \mathbb{E}_{t,\mathbf{x}} \left[\frac{1}{h} \int_t^\tau \left\{ \left(\partial_s + \mathcal{L}_s^{\mathbf{v}} \right) H(s, \mathbf{X}_s^{\mathbf{v}}) + F\left(s, \mathbf{X}_s^{\mathbf{v}}, \mathbf{v}\right) \right\} ds \right], \\
&= \quad \left(\partial_t + \mathcal{L}_t^{v} \right) H(t, \mathbf{x}) + F(t, \mathbf{x}, \mathbf{v}).
\end{aligned}
$$

2. The second line follows from:
   - (i) As $h \downarrow 0$, $\tau = t + h$ a.s. since the process will not hit the barrier of ϵ in extremely short periods of time,
   - (ii) The condition that $|\mathbf{X}_\tau^{\mathbf{u}} - \mathbf{x}| \leq \epsilon$, which implies that if the process does hit the barrier it is bounded,
   - (iii) The Mean-Value Theorem allows us to write $\lim_{h \downarrow 0} \dfrac{1}{h} \displaystyle\int_t^{t+h} \omega_s ds = \omega_t$, and
   - (iv) The process starts at $\mathbf{X}_t^{\mathbf{v}} = \mathbf{x}$.