MA668: Algorithmic and High Frequency Trading
Lecture 26

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

Combined Diffusion and Jumps

1. There are many situations in which the agent is exposed to more than one source of uncertainty.
2. Typically, an agent is faced with control problems where both diffusive and jump uncertainty appear, and she/he may be able to control all or only parts of the system.
3. Accordingly, we simply state the main results for a fairly general class of models.
4. Let:
   - A. $\mathbf{N}^{\mathbf{u}} = (\mathbf{N}^{\mathbf{u}}_t)_{\{0 \le t \le T\}}$: Collection of counting processes (of dim p).
   - B. $\lambda^{\mathbf{u}} = (\lambda^{\mathbf{u}}_t)_{\{0 \le t \le T\}}$: Controlled intensities.
   - C. $\mathbf{u} = (\mathbf{u}_t)_{\{0 \le t \le T\}}$: Control processes (of dim m).
5. We next assume that the controlled processes $\mathbf{X}^{\mathbf{u}} = (\mathbf{X}^{\mathbf{u}}_t)_{\{0 \le t \le T\}}$ satisfy the SDEs:

$$d\mathbf{X}^{\mathbf{u}}_t = \mu^{\mathbf{u}}_t dt + \sigma^{\mathbf{u}}_t d\mathbf{W}_t + \gamma^{\mathbf{u}}_t d\mathbf{N}^{\mathbf{u}}_t, \tag{1}$$

   with $\mu^{\mathbf{u}}_t := \mu(t, \mathbf{X}^{\mathbf{u}}_t, \mathbf{u}_t)$ ($m \times 1$ vector), $\sigma^{\mathbf{u}}_t := \sigma(t, \mathbf{X}^{\mathbf{u}}_t, \mathbf{u}_t)$ ($m \times m$ matrix) and $\gamma^{\mathbf{u}}_t := \gamma(t, \mathbf{X}^{\mathbf{u}}_t, \mathbf{u}_t)$ ($m \times p$ matrix).

Combined Diffusion and Jumps (Contd ...)

1. We assume that the controlled intensity takes the form $\lambda_t^{\mathbf{u}} := \lambda(t, \mathbf{X}_t^{\mathbf{u}}, \mathbf{u}_t)$.
2. The modelling approach above implies that the agent can control, in general, the drift, volatility, jump size and jump arrivals.
3. The agent's performance criteria is given by:

$$H^{\mathbf{u}}(t, \mathbf{x}) = \mathbb{E}_{t,\mathbf{x}} \left[G\left(\mathbf{X}_T^{\mathbf{u}}\right) + \int_t^T F\left(s, \mathbf{X}_s^{\mathbf{u}}, \mathbf{u}_s\right) ds \right]. \tag{2}$$

4. The value function is then given by the usual expression:

$$H(t, \mathbf{x}) = \sup_{\mathbf{u} \in \mathcal{A}_{[t,T]}} H^{\mathbf{u}}(t, \mathbf{x}). \tag{3}$$

DPP for Jump-Diffusions

The value function satisfies the DPP:

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

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

DPE/HJB for Jump-Diffusions

$$\partial_t H(t, x) + \sup_{\mathbf{u} \in \mathcal{A}} \left(\mathcal{L}_t^{\mathbf{u}} H(t, \mathbf{x}) + F(t, \mathbf{x}, \mathbf{u}) \right) = 0, \;\; H(T, \mathbf{x}) = G(\mathbf{x}), \tag{5}$$

where:

$$\begin{aligned} \mathcal{L}_t^{\mathbf{u}} H(t, \mathbf{x}) \quad &= \quad \mu(t, \mathbf{x}, \mathbf{u}) \cdot \mathcal{D}_x H(t, \mathbf{x}) + \frac{1}{2} \sigma(t, \mathbf{x}, \mathbf{u}) \sigma(t, \mathbf{x}, \mathbf{u})' \mathcal{D}_x^2 H(t, \mathbf{x}) \\ &+ \quad \sum_{j=1}^{p} \lambda_j(t, \mathbf{x}, \mathbf{u}) \left[H(t, \mathbf{x} + \gamma_{\cdot,j}(t, \mathbf{x}, \mathbf{u})) - H(t, \mathbf{x}) \right]. \end{aligned}$$

Combined Diffusion and Jumps (Contd ...)

1. The various terms in the generator can be easily interpreted.
2. In the first line, the first term represents the (controlled) drift of **X** and the second term represents the (controlled) volatility of **X**.
3. In the second line, the intensity for each counting process is shown separately and each term in the sum has the (controlled) rate of arrival of that jump component through λ_j.
4. The difference terms that appear are due to the jump arriving and causing that component of **N** to increment and simultaneously cause (potentially all components of) **X** to jump by $\gamma_{\cdot,j}$.

Optimal Stopping

1. In many circumstances the agent wishes to find the best time at which to enter or exit a given strategy.
2. The classical finance example of this is the "American put option" in which an agent who owns the option has the right, at any point in time up to and including the maturity date T, to exercise the option by receiving the amount of cash K in exchange for the underlying asset.
3. The net cash value of this transaction at the exercise date τ is $(K - S_\tau)$.
4. Naturally, the agent will not exercise when $S_\tau > K$ and hence the effective payoff is $(K - S_\tau)_+$ where $(\cdot)_+ = \max(\cdot, 0)$.
5. This simple observation provides only a trivial part of the strategy.
6. In order to determine the full strategy, the agent seeks the stopping time τ, which maximizes the discounted value of the payoff, that is, she/he searches for the stopping time which attains the supremum (if possible):

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}\left[e^{-r\tau}(K - S_\tau)_+\right],$$

where $\mathcal{T}$ are the $\mathcal{F}$-stopping times bounded by T.

Optimal Stopping (Contd ...)

1. This is just one of many such problems, and in general problems which seek optimal stopping times are called optimal stopping problems.
2. Similar to optimal control problems, optimal stopping problems admit a DPP and have an infinitesimal version, that is, a DPE.
3. Accordingly, we provide a concise outline of the DPP and DPE for optimal stopping problems.
4. Rather than first developing the diffusion case, then the jump, and then the jump-diffusion, here we begin immediately with a fairly general jump-diffusion model.
5. Let $\mathbf{X} = (\mathbf{X}_t)_{\{0\leq t\leq T\}}$ denoted a vector-valued processes of dim m satisfying the SDEs:

$$d\mathbf{X}_t = \mu(t, \mathbf{X}_t)dt + \sigma(t, \mathbf{X}_t)d\mathbf{W}_t + \gamma(t, \mathbf{X}_t)d\mathbf{N}_t,$$

where $\mu : [0, T] \times \mathbb{R}^m \to \mathbb{R}^m$, $\sigma : [0, T] \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$ and $\gamma : [0, T] \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$. Further, $\mathbf{W}$ is an m-dimensional Brownian motion with independent components, and $\mathbf{N}$ is an m-dimensional counting process with intensities $\lambda(t, \mathbf{X}_t)$.

Optimal Stopping (Contd ...)

1. The filtration $\mathcal{F}$ is the natural one generated by $\mathbf{X}$, and the generator of the process, $\mathcal{L}_t$ acts on twice differentiable functions as follows:

$$\begin{aligned} \mathcal{L}_t h(t,\mathbf{x}) &= \mu(t,\mathbf{x}) \cdot \mathcal{D}_x h(t,\mathbf{x}) + \frac{1}{2} \mathrm{Tr}\, \sigma(t,\mathbf{x})\sigma(t,\mathbf{x},)' \mathcal{D}_x^2 h(t,\mathbf{x}), \\ &+ \sum_{j=1}^{p} \lambda_j(t,\mathbf{x}) \left[h(t,\mathbf{x} + \gamma_{\cdot,j}(t,\mathbf{x})) - h(t,\mathbf{x}) \right]. \end{aligned}$$

2. The agent's problem is to find the value function:

$$H(t,\mathbf{x}) = \sup_{\tau \in [t,T]} H^\tau(t,\mathbf{x}), \text{ where } H^\tau(t,\mathbf{x}) = \mathbb{E}_{t,\mathbf{x}} \left[G\left(\mathbf{X}_\tau \right) \right], \qquad (6)$$

and the stopping time τ which attains the supremum if it exists.

Optimal Stopping (Contd ...)

1. At first glance, it appears that we have omitted the running reward/penalty $\int_t^\tau F(s, \mathbf{X}_s)ds$ that we included when studying optimal control.
2. Such terms can, however, be cast into the above form by choosing one of the components of $\mathbf{X}$, say X^1, to satisfy,

$$dX_t^1 = F(t, X_t^2, \ldots, X_t^m)dt,$$

   and writing $G(x) = x^1 + \widetilde{G}(x^2, \ldots, x^m)$.
3. It is therefore no loss of generality to consider only terminal rewards.
4. In the stochastic control case, we kept the explicit running reward/penalty since it traditionally appears there.

Optimal Stopping (Contd ...)

1. From the posing of the stopping problem in (6), intuitively, we see that the agent is attempting to decide between:
   - (A) "Stopping now" and receiving the reward G OR
   - (B) "Continuing" to hold off in hopes of receiving a larger reward in the future.
2. However, on closer examination it seems clear that the agent should continue to wait at the point $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^m$ as long as the value function has not attained a value of $G(\mathbf{x})$ there.
3. This motivates the definition of the stopping region, $\mathcal{S}$, which we define as:
$$\mathcal{S} = \{(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^m : H(t, \mathbf{x}) = G(\mathbf{x})\} .$$
4. It should be evident that whenever $(t, \mathbf{x}) \in \mathcal{S}$, that is, if the state of the system lies in $\mathcal{S}$, it is optimal to stop immediately, since the agent cannot improve beyond G, by the definition of the value function in (6).
5. The complement of this region, $\mathcal{S}^{\mathrm{c}}$, is known as the continuation region, where the agent can still improve her/his value and therefore continues to wait.

Figure 5.3

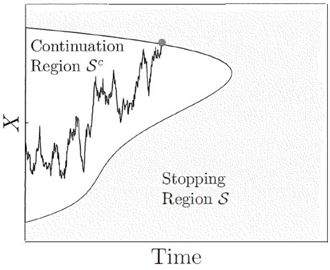

Figure 5.3 A generic depiction of the continuation and stopping regions for the optimal stopping problem (5.45). A sample path is also shown and the red dot corresponds to the optimal time to stop.

Figure: Figure 5.3

Optimal Stopping (Contd ...)

1. The difficulties in optimal stopping problems arise because the region $\mathcal{S}$ (or equivalently its boundary $\partial\mathcal{S}$) must be solved simultaneously with the value function itself.
2. Hence, the DPEs which arise in this context are free-boundary problems, also called obstacle problems, or variational inequalities.
3. Such PDEs are more difficult to solve, and even very simple examples typically do not admit explicit solutions (for example the American put option when the underlying asset is a geometric Brownian motion).
4. Nonetheless, a non-linear PDE is often enough to characterize the solution and many numerical schemes exist for solving them.

DPP for Stopping Problems

The value function satisfies the DPP:

$$H(t,\mathbf{x}) = \sup_{\tau\in\mathcal{T}_{[t,T]}} \mathbb{E}_{t,\mathbf{x}}\left[G(\mathbf{X}_\tau)\mathbb{1}_{\tau<\theta} + H(\theta,\mathbf{X}_\theta)\mathbb{1}_{\tau\geq\theta}\right], \tag{7}$$

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

DPE/HJB for Stopping Problems

Assume that the value function $H(t,\mathbf{x})$ is once differentiable in t and all second order derivatives in $\mathbf{x}$ exist, that is, $H \in C^{1,2}([0,T],\mathbb{R}^m)$, and that $G:\mathbb{R}^m\to\mathbb{R}$ is continuous. Then H solves the "variational inequality", also known as an obstacle problem, or free-boundary, problem:

$$\max\left\{\partial_t H + \mathcal{L}_t H, G - H\right\} = 0, \text{ on } \mathcal{D}, \tag{8}$$

where $\mathcal{D} = [0,T]\times\mathbb{R}^m$

Variational Inequality: An Interpretation

1. We briefly exploring the interpretation of the variational inequality:

$$\max\{\partial_t H + \mathcal{L}_t H, G - H\} = 0, \text{ on } \mathcal{D}.$$

2. This really represents two possibilities:
    1. $\partial_t H + \mathcal{L}_t H = 0$ and $H > G$.
    2. $H = G$ and $\partial_t H + \mathcal{L}_t H < 0$.
3. The first of these possibilities corresponds to the value function H being higher than the reward G, and in this region the value function satisfies a linear PDE.
4. The second of these possibilities corresponds to the value function H equaling the reward G, and hence occurs in the stopping region $\mathcal{S}$.

DPP for Optimal Stopping and Control

The value function satisfies the DPP:

$$H(t, \mathbf{x}) = \sup_{\tau \in \mathcal{T}_{[t,T]}} \sup_{\mathbf{u} \in \mathcal{A}_{[t,T]}} \mathbb{E}_{t,\mathbf{x}} \left[G(\mathbf{X}_\tau^{\mathbf{u}}) \mathbb{1}_{\tau < \theta} + H(\theta, \mathbf{X}_\theta^{\mathbf{u}}) \mathbb{1}_{\tau \geq \theta} \right], \tag{9}$$

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

DPE/HJB for Stopping Problems

Assume that the value function $H(t, \mathbf{x})$ is once differentiable in t and all second order derivatives in $\mathbf{x}$ exist, that is, $H \in C^{1,2}([0, T], \mathbb{R}^m)$, and that $G : \mathbb{R}^m \to \mathbb{R}$ is continuous. Then H solves the "quasi-variational inequality":

$$\max \left\{ \partial_t H + \sup_{u \in \mathcal{A}} \mathcal{L}_t^{\mathbf{u}} H, G - H \right\} = 0, \text{ on } \mathcal{D}, \tag{10}$$

where $\mathcal{D} = [0, T] \times \mathbb{R}^m$.

A Prelude

1. A classical problem in finance is how an agent can sell or buy a large amount of shares and yet minimize adverse price movements, which are a consequence of her/his own trades.
2. Here, the term “large” means that the amount the agent is interested in buying or selling is too big to execute in one trade.
3. One way to think about a trade being too large is to compare it to the size of an average trade or to the volume posted on the Limit Order Book (LOB) at the best bid or best offer.
4. Clearly, if the number of shares that the agent seeks to execute is significantly larger than the average size of a trade, then it is probably not a good idea to try to execute all the shares in one trade.
5. Investors who regularly come to the market with large orders (orders that are a significant fraction of average daily volume) are institutional traders such as pension funds, hedge funds, mutual funds and sovereign funds.

A Prelude (Contd ...)

1. These investors often delegate their trades to an agency broker (the agent) who acts on their behalf.
2. The agent will slice the "parent order" into smaller parts (sometimes called "child orders") and try to execute each one of these child orders over a period of time, taking into account the balance between price impact (trade quicker) and price risk (take longer to complete all trades).
3. What we mean by this trade-off is the following:
   - A. Imagine the situation in which the agent is selling shares.
   - B. If she/he trades quickly, then her/his orders will walk through the buy side of the LOB and she/he will obtain progressively worse prices for her/his orders.
   - C. Even if she/he breaks up each order into small bits (so that each one does not walk the book) and sends them quickly to the market, then other traders will notice an excess of sell orders and reshuffle their quotes inducing again a negative price impact.
   - D. If on the other hand, she/he trades slowly, so as to avoid this price impact, then she/he will be exposed to the uncertainty of what precisely the future prices will be.
4. Hence, she/he must attempt to balance these two factors.

A Prelude (Contd ...)

1. The time the agent takes to space out and execute the smaller orders is crucial.
2. Short time horizons will lead to faster trading (and hence more price impact) and less price uncertainty, but there are also many reasons why a long trading horizon might not be desirable.
3. For instance, it might be that it is decided to sell a large chunk of shares because the price is convenient, but by the time the agent executes all child orders, the share price could have dropped to a less desirable level.
4. Another reason which constrains the time needed to sell all the shares, is that this particular operation is part of a bigger one, which is the result of a portfolio re-balance that also requires the purchase of a large number of shares in another firm, and both operations need to be completed over approximately the same time period.

A Prelude (Contd ...)

1. Hence the agent must formulate a model to help her/him decide how to optimally liquidate or acquire shares, where the aim is to minimize the cost of executing her/his trade(s) and balance it against price risk.
2. Execution costs are measured as the difference between a benchmark price and the actual price (measured as the average price per share) at which the trade was completed.
3. Our convention is that when the sign of the execution cost is positive, it means that there is loss of value in the operation because the actual price of the trade was worse than the benchmark price.
4. The benchmark price represents a perfectly executed price in a market with no frictions.
5. It is customary to use the mid-price of the asset at the time when the order is given to execute the trade.
6. This benchmark is known as the arrival price which is generally taken to be the average of the best bid and best ask, that is, the mid-price.
7. Moreover, when the arrival price is the benchmark, the execution cost is known as the implementation shortfall or slippage.

The Model

To pose the optimal execution problem we require notation to describe the number of shares the agent is holding (inventory), the dynamics of the mid-price, and how the agent's market orders (MOs) affect the mid-price. Accordingly, the key stochastic processes are:

A. $\nu = (\nu_t)_{\{0 \le t \le T\}}$: The trading rate or the speed at which the agent is liquidating or acquiring shares (it is also the variable the agent controls in the optimization problem).
B. $Q^\nu = (Q^\nu_t)_{\{0 \le t \le T\}}$: Agent's inventory, which is clearly affected by how fast she/he trades.
C. $S^\nu = (S^\nu_t)_{\{0 \le t \le T\}}$: Mid-price process which is also affected, in principle, by the speed of her/his trading.
D. $\widehat{S}^\nu = (\widehat{S}^\nu_t)_{\{0 \le t \le T\}}$: Corresponds to the price process at which the agent can sell or purchase the asset, that is, the execution price, by walking the LOB.
E. $X^\nu = (X^\nu_t)_{\{0 \le t \le T\}}$: Agent's cash process resulting from the agent's execution strategy.

The Model (Contd ...)

1. Whether liquidating or acquiring, the agent's controlled inventory process is given in terms of her trading rate as follows:

$$dQ_t^\nu = \pm \nu_t dt, \ Q_0^\nu = q. \tag{11}$$

2. The mid-price is assumed to satisfy the SDE:

$$dS_t^\nu = \pm g(\nu_t)dt + \sigma dW_t, \ S_0^\nu = S, \tag{12}$$

where $W = (W_t)_{\{0 \leq t \leq T\}}$ is the standard Brownian motion and $g : \mathbb{R}_+ \to \mathbb{R}_+$ denotes the "permanent price impact" that the agent's trading action has on the mid-price.

3. The execution price satisfies the SDE:

$$\widehat{S}_t^\nu = S_t^\nu \pm \left(\frac{\Delta}{2} + f(\nu_t) \right), \ \widehat{S}_0^\nu = \widehat{S}, \tag{13}$$

where $f : \mathbb{R}_+ \to \mathbb{R}_+$ denotes the "temporary price impact" that the agent's trading action has on the price (they can execute the trade at) and $\Delta \geq 0$ is the bid-ask spread, which is assumed here to be a constant.

The Model (Contd ...)

1. Equations (11), (12) and (13) apply to both liquidation and acquisition problems, where the sign $\pm$ changes depending on whether the problem is that of liquidating $(-)$ or acquiring $(+)$ shares.
2. In equity markets, the "fundamental price" of the asset (also known in the literature as the efficient price or true price of the asset) refers to the share price that reflects fundamental information about the value of the firm and this is impounded in the price of the share.
3. In this discussion, we assume that during the optimal execution trading period, the fundamental price is the same as the mid-price of the asset.
4. Thus, as new information about the actual and expected performance of the firm is revealed to the market, the mid-price changes.
5. This is partly captured in the model by the increments of the Brownian motion W in (12).

The Model (Contd ...)

1. A key element of the model is how the agent's trades affect the mid-price.
2. Here the agent's market orders affect the mid-price in two ways:
    1. Through $f(\nu_t)$ in (13).
    2. Through $g(\nu_t)$ in (12).

    These functions capture two different ways in which the agent's trades affect the mid-price.
3. At any one time, the number of shares displayed and available in the market at the quoted bid/ask prices $S_t^\nu \pm \dfrac{\Delta}{2}$ is limited.
4. A large MO will walk the book, so that the average price per share obtained will be worse than the current bid/ask price.
5. This is captured in the model, as an order of size νdt will obtain an execution price per share of $S_t^\nu \pm \left(\dfrac{\Delta}{2} + f(\nu)\right)$ with $f(\nu) \geq 0$.
6. Note, however, that the impact of the order as captured by $f(\nu_t)$ is limited to the execution price and does not affect the mid-price of the asset.

Figure 6.1

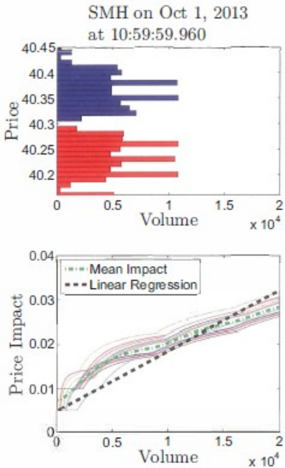

Figure 6.1 An illustration of how the temporary impact may be estimated from snapshots of the LOB using SMH on Oct 1, 2013. The top two panels are at 11:00am. The bottom left from 11:00am to 11:01am and the bottom right the entire day.

Figure 6.1 (Contd ...)

1. Figure 6.1: We show a snapshot of the LOB (top left panel) for SMH on Oct 1, 2013 at 11:00 AM, together with the price impact per share (top right panel) that an MO of various volumes would face as it walks through the buy side of the LOB.
2. The bottom left panel shows the impact every second from 11:00 AM to 11:01 AM as well as the average of those curves over the minute (the dash-dotted line).
3. We also include a linear regression (the dashed line) with intercept set to the half-spread: This would correspond to a linear impact function $f(\nu_t)$, which is the simple model we adopt in this discussion and which is also widely used.
4. Figure 6.1: Illustrates that the function $f(\nu_t)$ seems better described by a power law and the model can be extended to incorporated this.

Figure 6.1 (Contd ...)

1. Notice that the impact function fluctuates within the minute and with it the impact that trades of different size have.
2. The linear regression provides an approximation of the temporary impact during that one minute. The bottom right panel shows how the slope of this linear impact model fluctuates throughout the entire day.
3. We see that the largest impact tends to occur in the morning, then this impact flattens and stays flat throughout the day and towards the end of the day it lessens.
4. Such a pattern is seen in a number of assets.