Overview of Advanced RL in Finance:

Week 2: RL, Market Frictions, and Stock Dynamics

Igor Halperin

NYU Tandon School of Engineering

Brooklyn NY 2018

In these notes:

- **RL-inspired models for market dynamics**
- **Plan:**
  - From RL/IRL for individual investor to a market portfolio
  - RL-based market model
  - Role of frictions
  - Corporate defaults

Vapnik's principle

"One should avoid solving more difficult intermediate problems when solving a target problem"

V. Vapnik, *Statistical Learning Theory*, (1998)

Summary of Reinforcement Learning for stock trading

- A number of problems in stock trading (optimal execution, index tracking, portfolio optimization) amount to RL
- When rewards are unobservable, one can use methods of IRL.
- The IRL approach in the general case can be implemented using a variational EM algorithm.
- When applied to a particular trader, the model needs proprietary data.
- It can also be used for the market portfolio.
- A feedback loop (market impact) is critically important.

Inverse Reinforcement Learning for market modeling

The IRL model can be used in two settings:

- As a model of a particular trader - needs proprietary data
- As a model for the market portfolio - uses only public data
- As a model for the market portfolio with private signals, similar to the Black-Litterman model.

"Market-implied" optimal policy is:

$$\pi(\mathbf{a}_t|\mathbf{x}_t) = \mathcal{N}\left(\hat{\mathbf{A}}_0 + \hat{\mathbf{A}}_1\mathbf{x}_t, \Sigma_M\right)$$

Agent-based market models and and an 'Invisible Hand'

Two types of agent-based approaches to modeling market dynamics:

- A representative rational or bounded-rational investor (economics) - a 'mean' of all investors
- Multi-agent models (physics, computer science)
- To identify an agent whose optimal portfolio is a market portfolio, as in the BL model, we have to use an agent who is a 'sum' of all investors
- Such agent cannot be a rational agent, but should have a bounded rationality
- Embodies an 'Invisible Hand'-type market mechanism (Adam Smith, *etc.*)

Control question

Select all correct answers:
1. IRL can be applied to both an individual investor portfolio and a market portfolio.
2. Bounded rationality is a principle saying that rationality in scientific deduction methods has limits due to noise or quantum effects.
3. Multi-agent models produce identical results to single agent models if all agents are fully rational.
4. The notion of 'Invisible Hand' refers to the observation, first made by Adam Smith, that agents in multi-agent models are hard to identify, and therefore they should be modeled using hidden variables.
Correct answers: 1.

Geometric Brownian Motion

Geometric Brownian Motion model in Finance

The Geometric Brownian Motion (GBM) model (Samuelson, 1965), also known as the log-normal asset return model:

$$dX_t = (r_f + \mathbf{w}\mathbf{z}_t)X_t dt + \sigma X_t dW_t \qquad (1)$$

Here X_t is an asset price at time t, r_f is a risk-free rate, $\mathbf{z}_t$ are predictors ("alpha"-signals), $\mathbf{w}$ are weights, and W_t is a standard Brownian motion. The GBM model improved over the ABM (Arithmetic Brownian Motion) model of Bachelier (1900).
The GBM model can be viewed as a model with a linear drift $f(x) = (r_f + \mathbf{w}\mathbf{z}_t)x$. The ABM model has a constant drift and volatility.

The GBM model in discrete time

The Geometric Brownian Motion (GBM) model:

$$dX_t = (r_f + \mathbf{w}\mathbf{z}_t) X_t dt + \sigma X_t dW_t \qquad (2)$$

The GBM Eq.(2) is a continuous-time limit $\Delta t \to dt$ of a discrete-time dynamics

$$\Delta X_t = r_t X_t \Delta t, \quad r_t = r_f + \mathbf{w}\mathbf{z}_t + \frac{\sigma}{\sqrt{\Delta t}} \xi_t, \qquad (3)$$

where $\xi_t \sim \mathcal{N}(\cdot|0,1)$.
Equivalently can write

$$X_{t+\Delta t} = (1 + r_t \Delta t) X_t \qquad (4)$$

Uses and misuses of the GBM model

- Many models that use the GBM model include Capital Asset Pricing Model (CAMP) and the Black-Scholes option pricing model.
- The model does **not** incorporate the following:
  - Defaults and market crashes
  - Rare events of large market moves
  - Market frictions
  - Exchange of capital with an outside world
  - Volatility patterns

Traditional approaches to improving the GBM model

- Extend a set of predictions $\mathbf{z_t}$
- Include non-linear dependencies on predictors $\mathbf{z_t}$
- Include more complex state-dependent or/and stochastic noise coefficients
- All these approaches preserve linearity of dynamics in the state variable X_t
- (Beyond ML and RL?) We will see that including instead *non-linearities in* X_t may be more important!

Control question

Select all correct answers:
1. The Geometric Brownian Motion (GBM) is applied for a Brownian motion in non-trivial geometries, e.g. for a diffusion on a finite interval.
2. The GBM model overestimates probabilities of large market moves or defaults.
3. The GBM model is incompatible with defaults, because the boundary $X = 0$ in the GBM model is unattainable.
4. 'Non-linear' extensions of the GBM model may involve non-linearities in space or non-linearities in predictors.
Correct answers: 3,4.

The GBM model: an unbounded growth without defaults

Corporate defaults are beyond the GBM model

- Corporate defaults are similar to absorbing state: once a system gets there, it cannot escape.
- The zero level $X = 0$ could naturally serve as a default/absorbing boundary
- The problem is that in the GBM model, the zero level $X = 0$ is *unattainable*: defaults *cannot happen* in the GBM model at $X = 0$
- Defaults can be described as level crossing at some $\bar{X} > 0$ (e.g. the Merton model), but this approach has some issues too
- As it is hard to have defaults in the GBM model, we need other state variables such as credit spreads
- But credit spreads and stock prices are *not* independent - leading to highly complex joint dynamics of stock prices and spreads

Unbounded growth in the GBM model

$$\Delta X_t = r_t X_t \Delta t, \quad r_t = r_f + \mathbf{w} \mathbf{z}_t + \frac{\sigma}{\sqrt{\Delta t}} \xi_t, \tag{5}$$

This equation has a *linear drift* $f(x) = r_t X_t$. Taking averages on both sides, we obtain an equation for the mean $\bar{X}_t$:

$$d\bar{X}_t = r_f X_t dt \quad \Leftrightarrow \bar{X}_t = \bar{X}_0 e^{r_f t} \tag{6}$$

We got an **exponential growth** of the mean asset price! This is a consequence of the linearity of the drift $f(x) = r_t X_t$ and resulting *scale invariance* of Eq.(2) with respect to scale transformation $X_t \to \alpha X_t$.

Are unbounded returns reasonable?

In the GBM world, you can get **infinitely rich** (due to the *linear drift* $f(x) = r_t X_t$ and scale invariance):

$$d\bar{X}_t = r_f X_t dt \quad \Leftrightarrow \bar{X}_t = \bar{X}_0 e^{r_f t} \qquad (7)$$

But the market is typically considered a closed system without any exchange of capital with an outside world.
How can you get infinitely reach in such market?
A simple hypothesis is that there are some *saturation effects* for large values of X_t, so that you will not get *infinitely* rich at the end.

Control question

Select all correct answers:
1. The origin of unbounded returns in the GBM model is linearity of the drift and resulting scale invariance of the GBM model.
2. The origin of unbounded returns in the GBM model is a desire to make the model more attractive to investors.
3. Effects of saturations in the market, that are produced by interactions and a finite depth of the market, can change returns from unbounded to bounded.
4. Credit spreads should be independent from stock prices, because doing otherwise produces overly complex models.

Correct answers: 1,3.

Dynamics with saturation: the Verhulst model

Dynamics with saturation: the Verhulst model

The Verhulst model is popular in physics, biology and ecology as a model for the dynamics of a size of population x_t that competes for a limited resource such as food:

$$dx_t = \left(\theta x_t - \kappa x_t^2\right) dt = \kappa x_t \left(\frac{\theta}{\kappa} - x_t\right) dt \quad (8)$$

The Verhulst model for market dynamics?

The Verhulst model is popular in physics, biology and ecology as a model for the dynamics of a size of population x_t that competes for a limited resource such as food:

$$dx_t = \left(\theta x_t - \kappa x_t^2\right) dt = \kappa x_t \left(\frac{\theta}{\kappa} - x_t\right) dt \quad (9)$$

If x_t is used to model a stock price, this means that our model has *state-dependent* diminishing returns as an effect of competition for a limited resource (a market value):

$$\bar{r}_t = \bar{r}(x_t) = \frac{\theta}{\kappa} - x_t \quad (10)$$

This spells a boundedness of the total wealth, as we will see shortly.

Exponential growth as an initial "inflation"

The Verhulst model:

$$dx_t = \kappa x_t \left(\frac{\theta}{\kappa} - x_t \right) dt$$

Consider the 'normal' regime with $\kappa > 0$. For 'small fields" $x_t \ll \theta / \kappa$, we have an exponential growth:

$$dx_t \simeq \theta x_t dt \quad \Rightarrow x_t \simeq x_0 e^{\theta t} \tag{11}$$

But this 'inflationary' behavior is only approximate: it is valid only for short times (or small fields $x_t \ll \theta / \kappa$). In the long term, the system reaches an equilibrium at $\bar{x} = \frac{\theta}{\kappa}$.

The opposite limit $x_t \gg \frac{\theta}{\kappa}$

1. Let's neglect $\frac{\theta}{\kappa}$ in parenthesis. We obtain

$$dx_t \simeq -\kappa x_t^2 dt \quad \Leftrightarrow x_t = C + \frac{1}{\kappa t} \tag{12}$$

The solution approaches a constant C (which should be equal θ/κ) in the long run, the speed of convergence is controlled by κ.

The opposite limit $x_t \gg \frac{\theta}{\kappa}$

1. Let's neglect $\frac{\theta}{\kappa}$ in parenthesis. We obtain

$$dx_t \simeq -\kappa x_t^2 dt \quad \Leftrightarrow x_t = C + \frac{1}{\kappa t} \tag{13}$$

The solution approaches a constant C (which should be equal θ/κ) in the long run, the speed of convergence is controlled by κ.

2. To insure $x_t \gg \frac{\theta}{\kappa}$, we could set $x_t = \frac{\theta}{\kappa} + y_t$ where y_t is large. Substituting into the Verhulst model, we get an equation for y_t:

$$dy_t = \kappa y_t \left(-\frac{\theta}{\kappa} - y_t \right) dt \tag{14}$$

This is the same as the original Verhulst model but with a flipped sign of θ - an interesting symmetry of the model!

Control question

Select all correct answers:
1. In the limit $x_t \gg \frac{\theta}{\kappa}$, the Verhulst process grows exponentially
2. In the limit $x_t \ll \frac{\theta}{\kappa}$, the Verhulst process grows logarithmically
3. In the limit $x_t \ll \frac{\theta}{\kappa}$, the Verhulst process grows exponentially
4. In the limit $\to \infty$, the Verhulst process converges to a long-term mean $\frac{\theta}{\kappa}$
Correct answers: 3, 4.

The Singularity is Near!

Full time-dependent solution for the Verhulst model

The full solution is

$$X_t = X_0 \frac{e^{\theta t}}{1 + \frac{k}{\theta} X_0 \left(e^{\theta t} - 1 \right)} = \frac{\theta}{\kappa} \frac{1}{1 - \left(1 - \frac{\theta}{\kappa X_0} \right) e^{-\theta t}} \tag{15}$$

The only stable stationary solution for $\kappa, \theta > 0$ is $\bar{x} = \frac{\theta}{\kappa}$.

Full time-dependent solution: positive κ

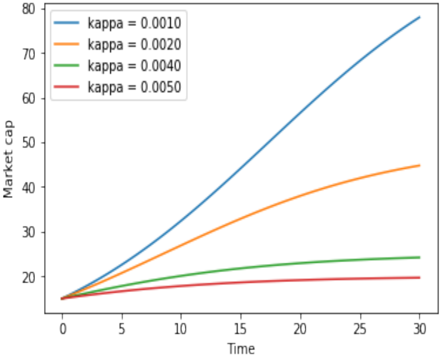

Figure: Solutions for positive values of κ.

Full time-dependent solution: negative κ

For $\kappa < 0$, the only stable solution is $\bar{x} = 0$
(requires $X_0 = 0!$)

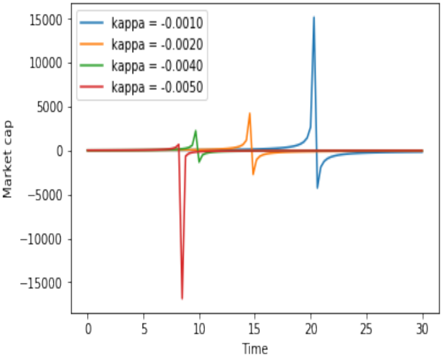

Figure: Solutions for negative values of κ.

Time asymmetry: The Singularity is near!

The solution is strongly asymmetric in time. For an arbitrarily $X_0 > 0$, the process (9) explodes to infinity in a finite time t_∞:

$$t_\infty = \frac{1}{\theta} \log \left(1 - \frac{\theta}{\kappa X_0} \right) \tag{16}$$

A *positive* singularity is in the **future** ($t_\infty > 0$) only for 'non-physical' choices $\theta > 0, \kappa X_0 \leq 0$ or $\theta < 0, \theta \leq \kappa X_0 \leq 0$.
The model should be 'regularized' to treat such 'non-physical' parameter values!
When $\kappa, \theta > 0$, the singularity is in the **past**: $t_\infty < 0$ ("emergence" from a negative singularity)

Control question

Select all correct answers:
1. The Verhulst model have a singularity either in the past, or in the future, depending on the parameters.
2. If $\frac{\theta}{\kappa X_0} > 1$, the solution becomes singular for complex-valued times.
3. It is only singularities for positive times that matter, others are just irrelevant mathematical details, especially for financial models.
4. Singularities of the Verhulst model point to the need to regularize the model.
Correct answers: 1,2, 4 .

Where are defaults?

Default as a crash to zero?

- When a firm defaults (more precisely, only when it goes bankrupt), its stock drops to zero, the company is closed.
- So let's just describes bankruptcies/defaults as a drop price to zero in the GBM model!
- Oops, sorry, we can't - the boundary $X = 0$ is unaccessible in the GBM model!
- A smart way out (the Merton 1974 model): let's model an unobservable firm value process, instead of the stock value process. The default boundary is at a non-zero level for the firm value process.

Defaults in the GBM model: the Merton model

The Merton 1974 model of corporate defaults as a level crossing phenomenon

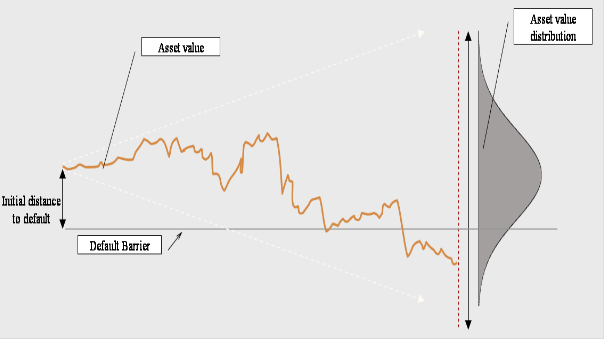

Figure: The Merton model of corporate defaults

Problems with the Merton default model

- The firm value process used by the model is unobservable
- The exact default position is unobservable too
- Noise in the default position can be brought simply by observational noise in the simplest model formulation
- Explicitly uncertain default barrier models can also be constructed
- As a result of noise in the barrier, the default event itself becomes uncertain - we can't say with certainty if a firm defaulted or not if we stay within the model
- Resembles the Schrodinger cat in quantum mechanics

Control question

Select all correct answers:
1. When a firm goes bankrupt, its stock price drops to zero.
2. A stock price cannot be negative because of limited liability of stockholders.
3. The Merton default model describes a default as a level crossing event for an unobservable firm value process.
4. If the default barrier position in the Merton model is not exactly known, at each moment in time we actually do not know if the default happened or not.

Correct answers: 1,2,3,4.

"Quantum Equilibrium-Disequilibrium"

"Quantum Equilibrium-Disequilibrium"

- Competitive market equilibrium models: markets near a state of a thermodynamic equilibrium, with zero exchange of money or information with an outside world
- Produce an unbounded growth of asset
- An alternative: an "equilibrium disequilibrium" in the market (Amihud et. a. 2005)
- "Quantum Equilibrium-Disequilibrium" - to emphasize the role of noise (the same as quantum effects)

IRL-inspired market dynamics model

Let X_t be a total capitalization of a firm at time t, rescaled to a dimensionless quantity $X_t \sim 1$. Discrete-time dynamics:

$$
\begin{aligned}
X_{t+\Delta t} &= (1 + r_t \Delta t)(X_t - cX_t \Delta t + u_t \Delta t) \\
r_t &= r_f + \mathbf{w}\mathbf{z}_t - \mu u_t + \frac{\sigma}{\sqrt{\Delta t}} \varepsilon_t
\end{aligned}
\tag{17}
$$

where μ is a market impact parameter, and c is the dividend rate.
Here $u_t \Delta t$ a new capital injected in the market by investors at the start of the interval $[t, t + \Delta t]$, after which the new capital $X_t - cX_t \Delta t + u_t \Delta t$ grows at rate r_t. When $u_t = 0, \ \forall t$ and $c = 0$, we recover the GBM model.

Capital supply function

In general, u_t should be a function of X_t. We consider a simple quadratic specification

$$u_t = u(X_t) = \phi X_t + \lambda X_t^2 \tag{18}$$

We assume that $0 < \lambda \ll 1$. Then in a parametrically wide region $|X_t| \ll |\phi/\lambda|$:

$$u(x) \simeq \phi x, \quad x \ll \phi/\lambda \tag{19}$$

$\phi > 0$: capital is injected ('growth')
$\phi < 0$: capital is withdrawn ('contraction').

"Quantum Equilibrium-Disequilibrium" (QED) model

Substituting Eq.(18) into Eqs.(17), neglecting term $(\Delta t)^2$ and taking the continuous time limit $\Delta t \to dt$ we obtain the "Quantum Equilibrium-Disequilibrium" (QED) model:

$$dX_t = \kappa X_t \left(\frac{\theta}{\kappa} - X_t - \frac{g}{\kappa} X_t^2 \right) dt + \sigma X_t \left(dW_t + \mathbf{wz}_t \right) \tag{20}$$

where we introduced parameters

$$\theta = r_f - c + \phi, \;\; \kappa = \mu\phi - \lambda, \;\; g = \mu\lambda \tag{21}$$

If we keep $\mu > 0$, the mean reversion parameter κ can be of either sign, depending on the sign of ϕ and the value of λ.

Control question

Select all correct answers:
1. The "QED" model is a model with an inflow/outflow of capital into the market.
2. The 'Q' in the name of the QED model stands for Q-Learning.
3. If we set $g = 0$, $\mathbf{w} = 0$ and $\sigma = 0$ in the QED model, we recover the Verhulst model.
4. If we set $\kappa = 0$ in the QED model, we recover the GBM model.
5. A steady non-equilibrium state is only possible for open systems that interact with an outside world.

Correct answers: 1,2,3,5.