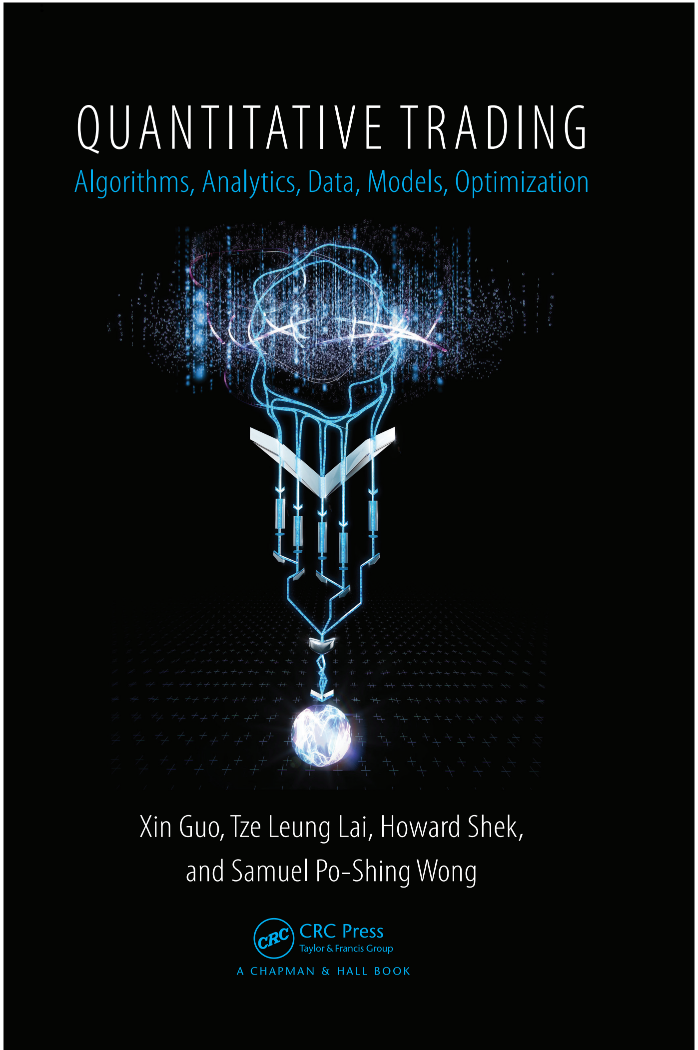
Quantitative Trading, Financial Technology, and Predictive Analytic

Tze Leung LAI

Department of Statistics

Stanford University



- **Data** cloud is first processed by an analytics machine.
- **Analytics** refers to both analysis of the data and the development of data-driven trading strategies which naturally make use of **Optimization**.
- **Models** provide the connection between the data and the trading strategies.
- **Algorithms** are step-by-step procedures for computing the solutions of not only optimization but also other mathematical and data analysis problems.

- Fundamentally motivated quant (FMQ): Buying a stock if it is undervalued with respect to fundamentals (such as earning quality, value of the firm, and investor sentiment). Since the fundamentals are based on quarterly earning reports and forecasts, the FMQ has a quarterly time-scale.
- Macro strategies:
  - Use macroeconomic analysis of market events and trends to identify opportunities for investment. They include discretionary and systematic strategies. Discretionary ones carried out by investment managers who identify and select the investments. Systematic macro strategies are model-based and executed by software with limited human involvement.
  - Example: Buy US dollars if the macroeconomic analysis suggests a rising trend for the dollar. Since the Federal Reserve Bank has to hold collaterals mainly in US Treasury debts and since the shortest maturity of Treasury bills is 28 days (about a month), such macro strategy has a monthly time scale.

- Convergence or relative value trades and other statistical arbitrage (StatArb) strategies:
  - Convergence trades: Trading in assets that are expected to converge in value.
  - Relative value strategies take simultaneously a short position in an overvalued asset and a long position in an undervalued asset, with the expectation that their spread will decrease over time from the current spread.
  - Assets include stocks, bonds and derivatives, and the time-scales range from minutes to months.
- High-frequency trading (HFT): The time scale of HFT is in milliseconds and the holding period of the traded securities is usually less than one second.

- m assets' return with mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_m)^\top$ and covariance matrix $\boldsymbol{\Sigma}$.
- Portfolio weight vector: $\mathbf{w} = (w_1, \ldots, w_m)^\top$ with $\mathbf{w}^\top \mathbf{1} = 1$.
- Markowitz's efficient portfolio for target mean return μ_* (short-selling is allowed):

$$\mathbf{w}_{\text{eff}} = \arg\min_{\mathbf{w}} \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w} \qquad \text{subject to } \mathbf{w}^\top \boldsymbol{\mu} = \mu_*, \ \mathbf{w}^\top \mathbf{1} = 1.$$

- $\mathbf{w}_{\text{eff}} = \left\{ B\boldsymbol{\Sigma}^{-1}\mathbf{1} - A\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \mu_* \left(C\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - A\boldsymbol{\Sigma}^{-1}\mathbf{1} \right) \right\} / D$ where $A = \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}$, $B = \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$, $C = \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}$, and $D = BC - A^2$.
- *Efficient frontier* is the collection of all possible $(\mu_*, \sqrt{\mathbf{w}_{\text{eff}}^\top \boldsymbol{\Sigma} \mathbf{w}_{\text{eff}}})$
- $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are actually unknown – Plug-in frontier: Replacing them by the sample mean vector $\widehat{\boldsymbol{\mu}}$ and covariance matrix $\widehat{\boldsymbol{\Sigma}}$ of a training sample of historical returns $\mathbf{r}_t = (r_{1t}, \ldots, r_{mt})^\top$, $1 \le t \le n$.
- However, this plug-in frontier is no longer optimal because $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$ actually differ from $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, and portfolios associated with the plug-in frontier can perform worse than an equally weighted portfolio that is highly inefficient.

- Dimension reduction in estimating $\mathbf{\Sigma}$ via multifactor models, i.e., relating the ith asset returns r_i to k factors $f_1, \ldots, f_k$ by $r_i = \alpha_i + (f_1, \ldots, f_k)^\top \boldsymbol{\beta}_i + \epsilon_i$, where α_i and $\boldsymbol{\beta}_i$ are unknown regression parameters and ϵ_i is an unobserved random disturbance that has mean 0 and is uncorrelated. Examples: CAPM, APT and Fama-French three-factor model.

- Use shrinkage estimates of $\mathbf{\Sigma}$ in the form of $\widehat{\mathbf{\Sigma}} = \widehat{\delta}\widehat{\mathbf{F}} + (1 - \widehat{\delta})\mathbf{S}$ where $\widehat{\delta}$ is an estimator of the *optimal shrinkage constant* and $\mathbf{S} = n^{-1} \sum_{i=1}^{n} (\mathbf{r}_i - \bar{\mathbf{r}})(\mathbf{r}_i - \bar{\mathbf{r}})^\top$. $\widehat{\mathbf{F}}$ is given by the mean of the prior distribution or a structured covariance matrix $\mathbf{F}$ with much fewer parameters than $m(m+1)/2$; see Ledoit and Wolf (2003, 2004). The estimate of $\boldsymbol{\mu}$ can also be handled by shrinkage similarly.

- To correct for the bias of $\widehat{\mathbf{w}}_{\text{eff}}$, use the average of the bootstrap weight vectors $\bar{\mathbf{w}} = B^{-1} \sum_{b=1}^{B} \widehat{\mathbf{w}}_b^*$, where $\widehat{\mathbf{w}}_b^*$ is the estimated optimal portfolio weight vector based on the bth bootstrap sample $\{\mathbf{r}_{b1}^*, \ldots, \mathbf{r}_{bn}^*\}$ drawn with replacement from the observed sample $\{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$; see Michaud (1989).

$$\max \left\{ E(\mathbf{w}^\top \mathbf{r}_{n+1}) - \lambda \mathrm{Var}(\mathbf{w}^\top \mathbf{r}_{n+1}) \right\} \tag{1}$$

Lai, Xing, and Chen (2011) solve (1) by rewriting it as the following maximization problem over η:

$$\max_{\eta} \left\{ E\left[\mathbf{w}^\top(\eta)\mathbf{r}_{n+1}\right] - \lambda \mathrm{Var}\left[\mathbf{w}^\top(\eta)\mathbf{r}_{n+1}\right] \right\}, \tag{2}$$

where $\mathbf{w}(\eta)$ is the solution of the stochastic optimization problem

$$\mathbf{w}(\eta) = \arg\min_{\mathbf{w}} \left\{ \lambda E\left[(\mathbf{w}^\top \mathbf{r}_{n+1})^2\right] - \eta E(\mathbf{w}^\top \mathbf{r}_{n+1}) \right\}. \tag{3}$$

Let P_t be the price of a stock (or a more general asset) at time t. Assuming no dividend over the period from time $t-1$ to time t, the *logarithmic return* is $r_t = \log(P_t/P_{t-1})$. The stylized facts of low-frequency (daily and weekly) r_t are:

- Non-normality in r_t (with kurtosis much greater than 3).
- Small autocorrelations of r_t.
- Volatility clustering: Strong (many lags) autocorrelations of the r_t^2.
- Leverage effect: The volatility response to a large positive return is considerably smaller than that of a negative return of the same magnitude.
- Marked changes of volatility in response to exogenous (e.g., macroeconomic) variables and external events (such as scheduled earnings announcements).

- Working model of Lai, Xing, and Chen (2011):

$$r_{it} = \boldsymbol{\beta}_i^\top \mathbf{x}_{i,t-1} + \epsilon_{it}, \tag{4}$$

where the components of $\mathbf{x}_{i,t-1}$ include 1, factor variables such as the return of a market portfolio like S&P 500 at time $t-1$, and lagged variables $r_{i,t-1}, r_{i,t-2}, \ldots$. Also, $\epsilon_{it} = s_{i,t-1}(\boldsymbol{\gamma}_i) z_{it}$, where z_{it} are i.i.d. with mean 0 and variance 1, $\boldsymbol{\gamma}_i$ is a parameter vector which can be estimated by maximum likelihood or generalized method of moments, and $s_{i,t-1}$ is a given function that depends on $r_{i,t-1}, r_{i,t-2}, \ldots$.

- A well-known example is the GARCH$(1,1)$ model

$$\epsilon_{it} = s_{i,t-1} z_{it}, \qquad s_{i,t-1}^2 = \omega_i + a_i s_{i,t-2}^2 + b_i r_{i,t-1}^2, \tag{5}$$

for which $\boldsymbol{\gamma}_i = (\omega_i, a_i, b_i)$.

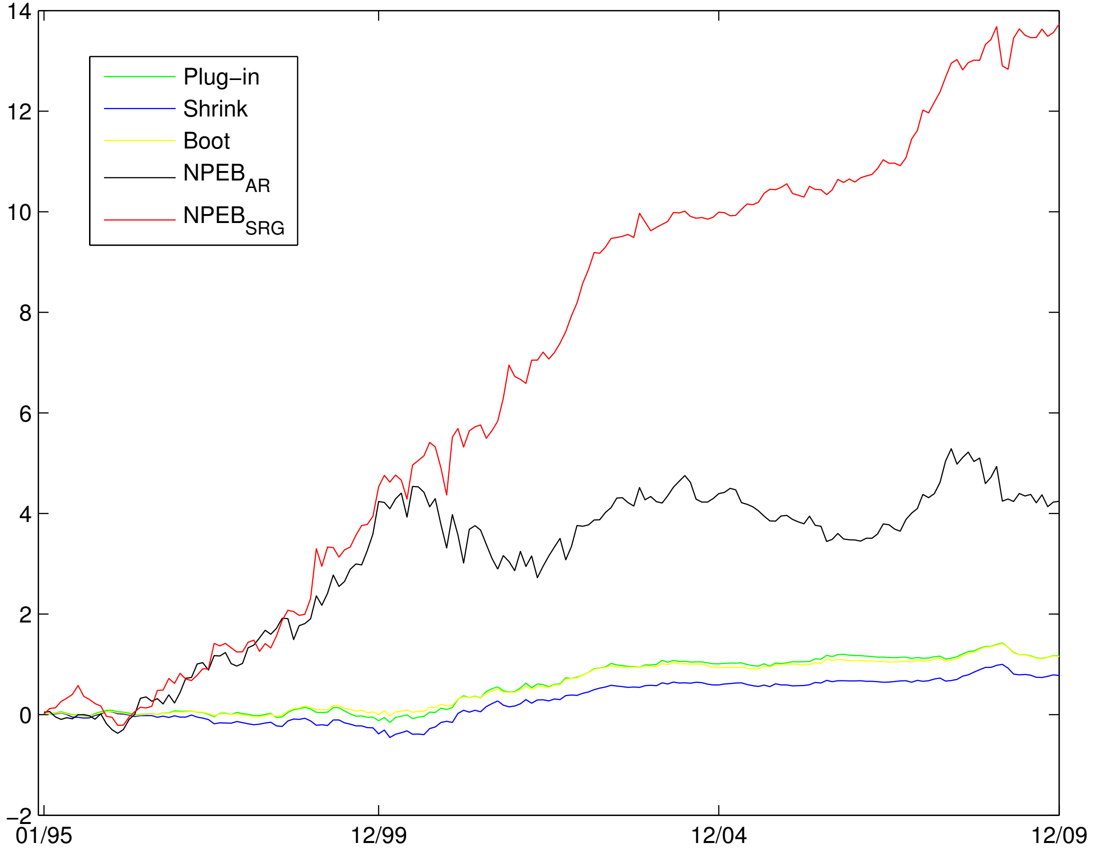
- Note that (4)–(5) models the asset returns separately, instead of jointly in a multivariate regression or multivariate GARCH model which has too many parameters to estimate. While the vectors $\mathbf{z}_t$ are assumed to be i.i.d., (5) does not assume their components to be uncorrelated since it treats the components separately rather than jointly.

- The conditional cross-sectional covariance between the returns of assets i and j given $\mathcal{R}_n = \{\mathbf{r}_1, \dots, \mathbf{r}_n\}$ is given by

$$\mathrm{Cov}(r_{i,n+1}, r_{j,n+1} | \mathcal{R}_n) = s_{i,n}(\boldsymbol{\gamma}_i) s_{j,n}(\boldsymbol{\gamma}_j) \mathrm{Cov}(z_{i,n+1}, z_{j,n+1} | \mathcal{R}_n),$$

  for the model (4)–(5). Thus, the estimator of $E(\mathbf{r}_{n+1} | \mathcal{R}_n)$ and $E(\mathbf{r}_{n+1}\mathbf{r}_{n+1}^\top | \mathcal{R}_n)$ are $\boldsymbol{\mu}_n = (\widehat{\boldsymbol{\beta}}_1^\top \mathbf{x}_{1,n}, \dots, \widehat{\boldsymbol{\beta}}_m^\top \mathbf{x}_{m,n})^\top$ and $\mathbf{V}_n = \boldsymbol{\mu}_n \boldsymbol{\mu}_n^\top + (\widehat{s}_{i,n}\widehat{s}_{j,n}\widehat{\sigma}_{ij})_{1 \le i,j \le n}$,, respectively, in which $\widehat{\boldsymbol{\beta}}_i$ is the least squares estimate of $\boldsymbol{\beta}_i$, and $\widehat{s}_{l,n}$ and $\widehat{\sigma}_{ij}$ are the usual estimates of $s_{l,n}$ and $\mathrm{Cov}(z_{i,1}, z_{j,1})$ based on $\mathcal{R}_n$.

- Let r_0 be the return of the benchmark investment. Take λ as a tuning parameter and choose it to maximize (over a grid of possible λ) the bootstrap estimate of the information ratio $E_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{w}_\lambda \mathbf{r} - r_0) \Big/ \sqrt{\mathrm{Var}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{w}_\lambda^\top \mathbf{r} - r_0)}$.



Realized cumulative excess returns over the S&P 500 Index.

- The cornerstones of quantitative portfolio management are prediction of asset returns from a large pool of investment possibilities, risk estimation, and portfolio optimization.
- There are two main styles of portfolio management – passive and active.
- Passive portfolio management constructs and administers portfolios that tracks some given index. Rationale: Index tracking incurs low cost as it does not require much information gathering on individual stocks. By reducing investment costs, the net return improves. Moreover, relatively infrequent trading results in fewer capital gains and therefore lower taxes.
- Goal of active portfolio management: Construct portfolios that aim to outperform some index or benchmark. The additional return that a portfolio generates relative to the benchmark is commonly known as the *alpha* of the portfolio. Performance is measured by the information ratio that expresses mean excess return in units of its standard deviation.

- Two main sources of alpha are (a) superior information, and (b) efficient information processing.
- Active portfolio management outperforms the passive approach in the absence of transaction costs, but the advantage may be outweighed by the transaction costs.
- Let $\boldsymbol{r} \in \mathbb{R}^m$ be the return vector of m assets. The returns of the portfolio and benchmark are $r_P = \boldsymbol{w}_P^\top \boldsymbol{r}$ and $r_B = \boldsymbol{w}_B^\top \boldsymbol{r}$ where $\boldsymbol{w}_P$ and $\boldsymbol{w}_B$ are the corresponding portfolio weights. Consider the one-factor model $r_P - r_f = \alpha + \beta(r_B - r_f) + \epsilon$ where r_f is the risk-free rate, subtracting $r_B - r_f$ from both sides of the equation defines the "active return" of the portfolio by $r_P - r_B = \alpha + \beta_P (r_B - r_f) + \epsilon$, in which the "active" α denotes the additional return of the portfolio over that of the benchmark, $\beta_P = \beta - 1$ is known as the active beta of the portfolio.
- Since the benchmark portfolio has beta equal to 1 (and therefore zero active beta), a portfolio with positive alpha and small $|\beta| < 1$ can have a high information ratio. Such a portfolio is said to have an "exotic beta".

- Samuelson-Merton (1969) theory of "lifetime portfolio selection": *Dynamic programming* for expected utilities of consumption and investment.
- Merton's continuous-time framework

$$
\begin{aligned}
dS_t &= S_t(\alpha\, dt + \sigma\, dB_t), \\
dX_t &= (rX_t - C_t)\, dt - dL_t + dM_t, \\
dY_t &= \alpha Y_t\, dt + \sigma Y_t\, dB_t + dL_t - dM_t.
\end{aligned}
$$

with stock price S_t, X_t = dollar value of investment in bond, $Y_t = y_t S_t$ and y_t = number of shares held in stock. Maximize the expected utility $J(t,x,y) = E\left[\int_t^\top e^{-\beta(s-t)} U_1(C_s)\, ds + e^{-\beta(T-t)} U_2(Z_T) \middle| X_t = x, Y_t = y\right]$, with the disount factor $\beta > 0$.

- U_1 and U_2 are CRRA :$-xU''(x)/U'(x)$ is constant.
- Optimal strategy is to devote a constant proportion (the Merton proportion) p of the investment to the stock and to consume at a rate proportional to wealth.
- *Extensions to the case with proportional transaction costs.*

- Gârleanu and Pedersen (2013): Quadratic transaction costs and reduction to LQG control problem
- Markowitz and van Dijk (2003): Linear transaction costs and heuristics to approximate the dynamic programming solution.
- Approximate dynamic programming (ADP)
  - Reinforcement learning
  - Monte Carlo tree search and deep learning
- Lai and Gao (2016): New approach that combines ADP, singular stochastic control and adaptive filtering in nonlinear state-space models.

- Transaction prices are in discrete units or ticks.
- Price clustering: The tendency for transaction prices to cluster around certain values (such as integers).
- Intraday seasonality of trading: Transactions tend to be heaviest near the beginning and close of trading hours and lightest around the middle of a trading day. It is equivalent to U-shaped (or diurnal) daily pattern in the durations between transactions.
- Negative lag 1 autocorrelation in (log) price changes from one transaction to the next.

- Roll's model of bid-ask bounce to explain negative lag 1 autocorrelations.
- General market microstructure model with additive noise $Y_{t_i} = X_{t_i} + \varepsilon_i$, where $E(\varepsilon_i) = 0$, X_t is the logarithm of the efficient price at time t, and $t_1 < \cdots < t_n$ is the set of transaction times belonging to $[0, T]$, and Y_{t_i} is the logarithm of the transaction price at time t_i.
  - Roll's model is a special case: X_t as the mid price, Y_t as the transaction price and $\varepsilon_i = \pm 1/2$ with equal probability.
- Methods to estimate integrated variance $[X]_T = \underset{\text{mesh}(\Pi)\to 0}{\text{p-lim}} \sum_{i=1}^{n} (X_{t(i)} - X_{t(i-1)})^2$, in which Π denotes a partition $0 = t_0 < \cdots < t_n = T$ of $[0, T]$ and $\text{mesh}(\Pi) = \max_{1\le i\le n}(t_i - t_{i-1})$. Extension to multiple assets.
- Autoregressive conditional duration (ACD) models of inter-transaction times. Self-exciting point process models.
- Joint modeling of point process and its marks.
- Realized GARCH and other models relating low-frequency to high-frequency volatilities.

- Limit order book (LOB) consists of all untransacted limit bid and ask orders for a specific asset.
- Example: Marketable limit order to buy 150 shares, GTC (stands for good till cancelled) with limit price of $114.50, resulting in 100 shares of the limit order traded at $114.50, and with the remaining 50 shares of the original order now becoming the best bid in the LOB.
- Such matching rule is known as the price-time priority rule.

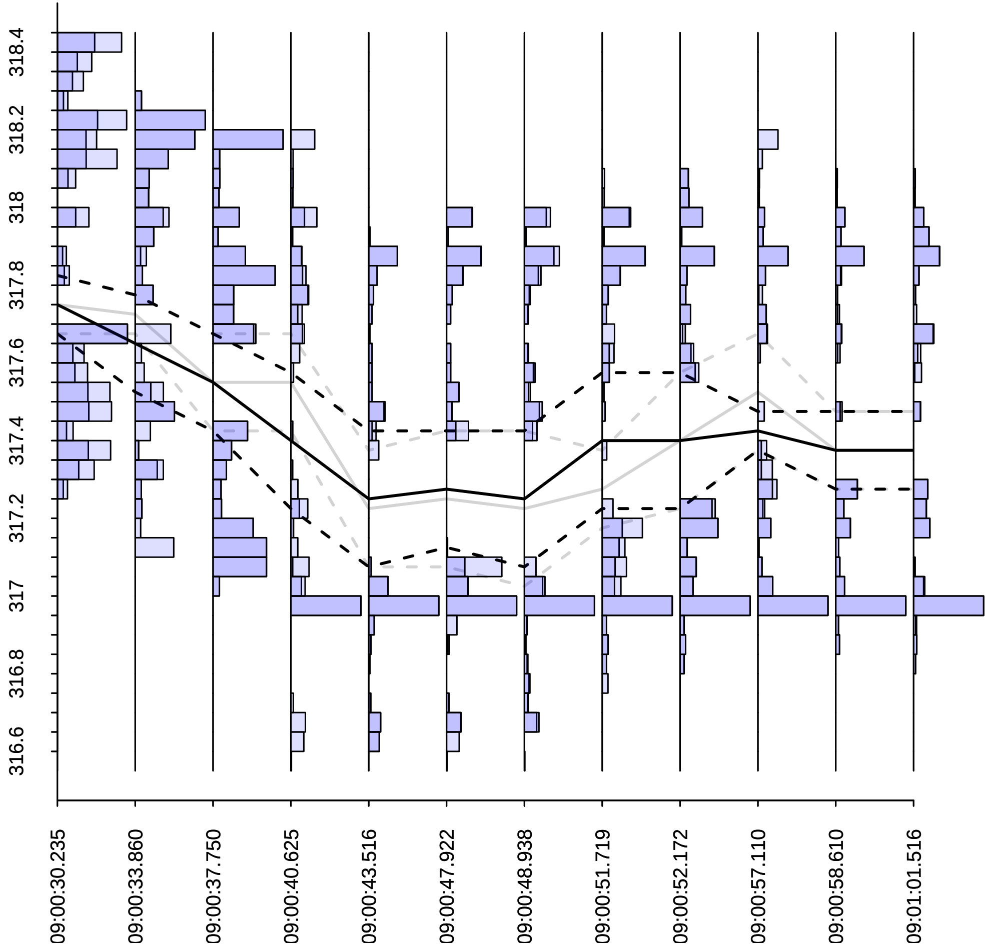
		114.52	145
		114.51	51
		114.50	100
80	114.49		
150	114.48		

Buy 150 shares at $114.50

⇓

		114.53	350
		114.52	145
		114.51	51
50	114.50		
80	114.49		

Snapshot of LOB for British Petroleum on June 25, 2010



- Let w_{1i} and w_{2j} be the trade sizes for the bid and ask market orders at times t_i and t_j, respectively. The intensity process can be modeled by

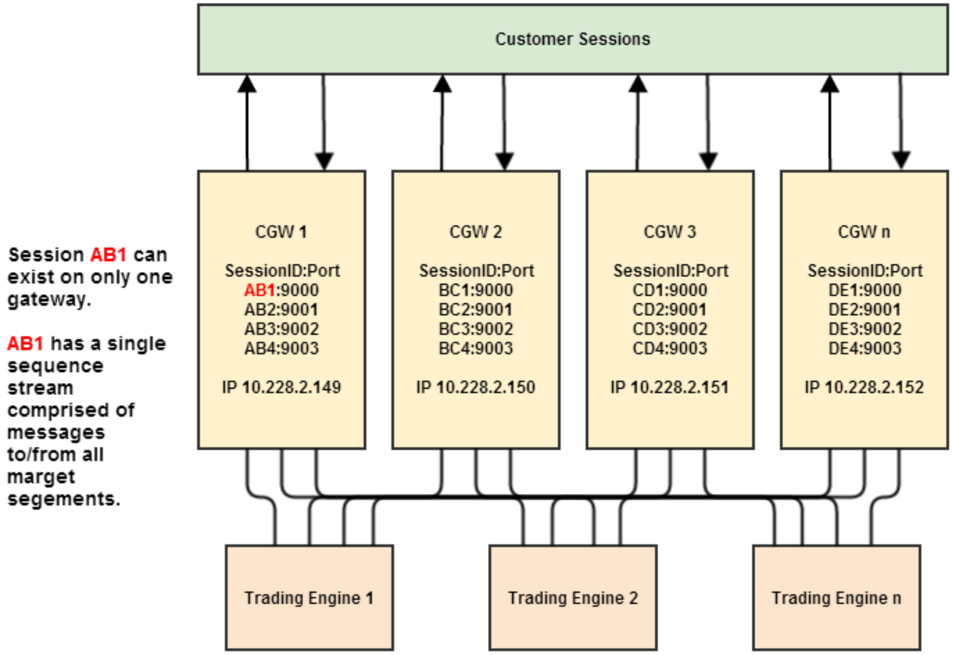
$$\lambda_t^{(1)} = \mu^{(1)}\bar{v}_{2t} + \frac{1}{\bar{w}_1}\sum_{t_i<t}\alpha_{11}w_{1i}e^{-\beta_{11}(t-t_i)} + \frac{1}{\bar{w}_2}\sum_{t_j<t}\alpha_{12}w_{2j}e^{-\beta_{12}(t-t_j)},$$

$$\lambda_t^{(2)} = \mu^{(2)}\bar{v}_{1t} + \frac{1}{\bar{w}_2}\sum_{t_j<t}\alpha_{22}w_{2j}e^{-\beta_{22}(t-t_j)} + \frac{1}{\bar{w}_1}\sum_{t_i<t}\alpha_{21}w_{1i}e^{-\beta_{21}(t-t_i)},$$

  where $\bar{w}_1$ (or $\bar{w}_2$) is the average of the trade sizes w_{1i} (or w_{2i}) in the period $[0, t)$ and $\bar{v}_{1t}$ is the probability weighted volume for bid orders and $\bar{v}_{2t}$ is that for ask orders.

- Use maximum likelihood to estimate the parameters.

- A typical exchange platform consists of the order gateway, the matching engine and the market data dissemination system.
- The matching engine takes orders routed by the order gateway and performs the matching of buy and sell orders based on some predetermined algorithms (such as price-time priority).
- The exchange market data dissemination system broadcasts public order book events, such as the cancellation and execution of an existing order or the addition of new (unhidden) limit orders. These market data, which show the real-time LOB dynamics, are important input of modern quantitative trading.
- To facilitate high-frequency trading activities (such as market making), exchanges provide colocation services to participants who want to place their trading servers at the exchange data center to reduce latency. The colocation fee depends on the space and power consumption that client servers require.



CME CGW iLink overview showing four order gateways and three matching engines. Note that for CME, matching engine is also known as trading engine. Four gateways nodes: CGW1, CGW2,CGW3, CGW4; three matching engines Trading Engine 1, Trading Engine 2, Trading Engine 3; CGW1 handles four sessions: AB1:9000, AB2:90001, . . . , AB4:9003, where 9000 to 9003 are the ports to which the sessions are connected. Used with permission of CME.

- *Limit order* and *market order* are the two most common order types for buying or selling financial securities.
- A *limit order* is an order to buy or sell a specific amount of shares of a stock at a specific price.
- A *market order* is an order to buy or sell a stock at the prevailing market price.
- The advantage of a market order is that it is almost always guaranteed to be executed. However, it could execute at a significantly worse price than the best bid and ask prices if there is insufficient liquidity to fill the original order at those prices. Thus, most exchanges offer various features (such as marketable limit order) to protect market orders from executing beyond some predetermined price band.
- *A marketable limit order* is a buy order with a price limit at or above the lowest offer in the market or a sell order with a price limit at or below the highest bid in the market.

Order qualifiers

- Midpoint-peg (abbreviated as MP) orders are hidden orders pegged to the midpoint of the National Best Bid and Offer (NBBO). Only the execution of these orders will reveal their presence. For odd spreads, the pegged orders can execute in half-penny increments, effectively reducing the cost to less than the minimum tick size. The key benefits of MP order are price improvement and anonymity.
- The qualifier of iceberg order specifies the display quantify of the order. Upon execution of the displayed quantity, the order is replenished automatically, until the full amount is filled. Note that the replenishment order is treated as a new order in queue priority.
- A post-only order only executes if the price crosses the market by an amount that is economically advantageous for the order to execute, otherwise the order is repriced to one tick away from the best bid and ask prices. Market makers, who use the rebate to control trading costs, can use this qualifier to reduce the risk of taking liquidity and paying a fee when the snapshot of the market maker's order book is delayed. Post-only orders must be marked as display and limit, or else are rejected by the exchange.

Operational Risk

- Software error risk
  - e.g. undefined behavior, memory leak.
  - potential remedy: Unit tests, regression and non-regression testing.
- Order transmission protocol risk
  - e.g. incorrect contract multiplier, incorrect FIX de/encoding
  - potential remedy: exchange certification (testing) environment.
- Network transmission error risk
  - e.g. dropped packets, dropped connection
  - potential remedy: heart-beat checks, recovery procedure.

Strategy Risk

- Model error
  - e.g. ill conditioned matrix, model logic error
  - potential remedy: condition number check, unit test, simulation.
- Trading exposure
  - e.g. portfolio exposure (delta, vega, scenario limits), order rate
  - potential remedy: position limits, P&L limits, order rate limits (cap the number of orders that the strategy can submit within a second).

- Order Protection Rule (Rule 611)
  - Target fragmented nature of equities trading
  - Ensure transaction price is the best price available: National-Best-Bid-and-Offer (NBBO)
  - 1975 Securities Amendments by SEC
- Circuit Breakers
  - Reduce volatility and panic sell-off, e.g. Black Monday 1987
  - Multiple levels of price drops (e.g. NYSE: 7% (Level 1), 13% (Level 2), 20% (Level 3)). Trading halts for specific period of time (e.g. 15 min. for NYSE) upon price breach.
  - Rolling benchmark (the transaction price in the preceding five minutes) is implemented after Flash Crash (June 11, 2010). (The old benchmark is determined at a quarterly basis after 1987 Black Monday.)
  - Limit-Up-Limit-Down mechanism was implemented in May 2012. Market Wide Circuit Breakers was also implemented with the prior days closing S&P 500 index as benchmark.

- Optimal execution is to find a trading strategy that buys and sells some large amount of shares within a short period of time under some assumptions on how the stock price is impact by these trading activities.
- Classical result: Bertsimas and Lo (1998) assume the price impact is linear under the random walk model (for share price) and prove that the optimal trading strategy is a volume-average type deterministic procedure.
- By assuming a Geometric Brownian Motion model with a multiplicative price impact, Guo and Zervos (2015) applies singular control theory to show that the optimal execution strategy is stochastic and Markovian, i.e., it depends on the state variable associated with the efficient price process.
- Obizhaeva and Wang (2013) and Alfonsi, Fruth, and Schied (2010) solve the optimal execution problem a LOB-based market impact model. In particular, Obizhaeva and Wang assume a block-shaped LOB which leads to a closed-form optimal execution strategy.

- Traders need to decide on (a) whether to use market orders, limit orders, or both, (b) the number of orders to place at different price levels, and (c) the optimal sequence of orders in a given time frame to execute each child order (as instructed by the optimal execution strategy).
- In terms of cost when using limit orders, traders do not need to pay the spread and can even get a *rebate*.
- When using market orders, one has to pay both the spread between the limit and the market orders and the fee in exchange for a guaranteed immediate execution.
- Goal: To balance between paying the spread and fees when placing market orders against execution/inventory risks when placing limit orders.
- Assuming a continuous-time Markov chain model for the LOB, Hult and Kiessling (2010) solves the optimal placement problem by using the potential theory.
- Guo et al. (2013) assume a Markov random walk model with mean reversion for the share price and obtain some more general optimal placement strategies.

- After the tumultuous period marked by the 2007-2008 Financial Crisis and the Great Recession of 2009, the financial industry has entered a new era.The onset of this era is marked by two "revolutions" that have transformed modern life and business.

- One is technological, dubbed "the FinTech revolution" for financial services by the May 9, 2015, issue of *The Economist* which says: "In the years since the crash of 2007-08, policymakers have concentrated on making finance safer.... Away from the regulator spotlight, another revolution is under way.... From payments to wealth management, from peer-to-peer lending to crowdfunding, a new generation of startups is taking aim at the heart of the industry – and a pot of revenues that Goldman Sachs estimates is worth $4.7 trillion. ...fintech firms are growing fast." The other is called "big data revolution".

- Quantitative trading in electronic markets epitomizes Dynamic Optimization, Financial Technology, and Risk Control in the aforementioned new era of the financial industry, covering all aspects ranging from portfolio/wealth management to order placement and routing.

References

A. Alfonsi, A. Fruth, and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance*, 10(2):143–157, 2010.

D. Bertsimas and A. W. Lo. Optimal control of execution costs. *J. Financial Markets*, 1(1):1–50, 1998.

N. Gârleanu and L. H. Pedersen. Dynamic trading with predictable returns and transaction costs. *J. Finance*, 68(6):2309–2340, 2013.

X. Guo and M. Zervos. Optimal execution with multiplicative price impact. *SIAM J. Finan. Math.*, 6(1):281–306, 2015.

X. Guo, A. de Larrard, and Z. Ruan. Optimal placement in a limit order book: An analytical approach. Preprint, 2013.

H. Hult and J. Kiessling. *Algorithmic trading with Markov chains*. PhD thesis, Stockholm University, Sweden, 2010.

T. L. Lai, H. Xing, and Z. Chen. Mean-variance portfolio optimization when means and covariances are unknown. *Ann. Appl. Statist.*, 5(2A):798–823, 2011.

O. Ledoit and M. Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empirical Finance*, 10(5):603–621, 2003.

O. Ledoit and M. Wolf. Honey, I shrunk the sample covariance matrix. *J. Portfolio Management*, 30(4):110–119, 2004.

H. M. Markowitz and E. L. van Dijk. Single-period mean-variance analysis in a changing world (corrected). *Finan. Analysts J.*, 59(2):30–44, 2003.

R. O. Michaud. *Efficient Asset Management.* Harvard Business School Press, 1989.

A. A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *J. Financial Markets*, 16(1):1–32, 2013.