Machine Learning and Finance

Yue Wu RCC Apr 26 2012

Computational and Biological Learning
Finance Flowchart

• Data – prices, volumes, indicators, news events, analyst opinions, etc
• Extract signal and make predictions
• Optimize policy from predictions
• Execution/Trading
• Market impact
• “It’s like a GPS system: you can stick it in a boat, a plane or a car,” he explained. “We’re asking Zoubin to build us a GPS.” - Institutional Investor (Nov 2009)
Outline

• Efficient Market Hypothesis
• Prediction Tasks (numerical data)
  – Returns size and direction
  – Return direction
  – Volatilities
  – Covariances
• Optimization Tasks
  – Mean-Variance Optimization
  – Alternate Objective Functions
  – Risk constraints
• Execution Tasks
• Market Impact
• Prediction Tasks (text + numerical data)
  – Expert Opinions
  – Market Sentiment
  – News events and market trigger words
Efficient Market Hypothesis

- Louis Bachelier in *The Theory of Speculation* (1900)
- Fama (1965)
- Cannot consistently achieve excess returns on risk-adjusted basis
- Weak form: asset prices reflect all past publicly available information
- Semi-strong form: prices instantly update to reflect new information
Financial Factors

• Fama-French 3 Factor Model:

\[ r = R_f + \beta(R_m - R_f) + b_s \cdot SMB + b_v \cdot HML + \alpha \]

• Foreign Exchange Factors:
  – Carry
  – Risk

• Identifying Factors
  – Manually constructed based on fundamentals
  – Focus on explaining data with intuitive factors
Dynamic Factor Models

• Systematic methods for finding factors: Factor models and PCA

• Factors evolve over time

\[ X_t = \Lambda_t F_t + e_t \]

• Forni and Lippi (2001), Stock and Watson (2002)

Predicting Future Returns

• Predict future returns $Y$, $n \times p$, based on past data $X$, $n \times q$:
  \[ Y_{t+h} = \beta X_t + \epsilon_{t+h} \]

• Data problems:
  – How to design $X$? Past returns, market data, risk data, etc
  – Dimensionality issues, $q >> p$
  – Extremely noisy

• Model concerns:
  – Linear assumption
  – Robustness
  – Time-varying relationships

• Time frame of predictions
Predicting Return Direction

- Technical Analysis
  - Rule based
  - Subjective
  - Some evidence (Lo et al., 2000)

- Logistic Regression

- Neural Networks (Trippi & Turban 1992)
Predicting Volatility

• Returns size and direction hard to predict
• Volatility has structure:
  – Time-varying
  – Clustering of high vol and low periods
  – Spikes in vol.
• Volatility predictions can be used to trade options and variance swaps
Volatility Models

• **GARCH:**
  - Can overfit with p, q
  - Usually p=1, q=1
  - Symmetric effect of negative and positive returns
  - Linear relation assumed.
  - Past variances not observed.

• **Stochastic Volatility:**
  \[ \log \sigma_t^2 \sim N(\mu - \alpha(\mu - \log \sigma_{t-1}^2), \tau^2) \]
Black-Scholes Options Model

• Assumptions:
  – Markov
  – Gaussian Innovations
  – Continuous time
  – Stationary volatility

• Call Option Price, \( C(S,T) \) depends on:
  – Time to Maturity, \( T \)
  – \( S \), spot price
  – \( K \), strike price
  – \( r \), risk free rate
  – \( \sigma \), vol. estimate. Constant for all prices, \( S \) and times, \( T \)

\[
\frac{\delta S}{S} \sim N(\mu \delta t, \sigma^2 \delta t)
\]

\[
\log S_T - \log S_0 \sim N \left( \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma^2 T \right)
\]

\[
C(S,T) = \Phi(d_1)S - \Phi(d_2)Ke^{-rT}
\]

\[
d_1 = \frac{\log(S/K) + (r + \sigma^2/2)T}{\sigma \sqrt{T}}
\]

\[
d_2 = d_1 - \sigma \sqrt{T}
\]
Volatility Smile

- Implied volatility not constant
- Volatility expectation dependent on price, and sometimes on events.
Predicting Covariances

• Multivariate-GARCH/BEKK class:
  – Eg. BEKK(1,1)
  \[ y_t \sim N(0, \Sigma_t) \]
  \[ \Sigma_t = Ay_{t-1}y_{t-1}^T A^T + B\Sigma_{t-1}B^T + C^T C \]

• Problems:
  – Highly parametrised, \(O(d^2)\)
  – Can overfit with MLE
  – Susceptible to local optima
  – Computationally expensive

• Multivariate Stochastic Volatility Models:
Policy Optimization

• Modern portfolio theory:
  – Expected return \( r_p = E_y(w^T y_t) \)
  – Expected variance \( \sigma_p^2 = E(w^T \Sigma_t w) \)

• Efficient Frontier

• Mean-Variance Optimization:
  – Maximize Sharpe Ratio
  – Subject to transaction costs, \( C_t \)

\[
\arg \max_w \frac{r_p - r_f}{\sigma_p} - |w - w_{t-1}| \cdot C_t
\]

\[
\sum_i w_i = 1
\]
Alternate Optimization Objectives

• Risk Appetite
  – Risk Neutral: Mean-Variance optimization assumes perfectly rational, risk neutral agent.
  – Risk Aversion: behavioral finance studies suggest hyperbolic utility functions
  – Risk Friendly: Gamblers
Value at Risk

• Loss-averse
• VaR is a risk measurement of losses

\[ V(r_p) = -\inf\{r_p \in \mathbb{R} : P(r_p > \tau) \leq 1 - \alpha\} \]

• Threshold \( \tau \) and significance level \( \alpha \)
• VaR can be estimated through modeling the return distribution or bootstrap simulation of portfolio holdings on historical data.
• VaR can be additional constraint to mean-variance optimization
Execution

• Execution objectives:
  – High frequency arbitrage: Ultra-fast communication, co-location, loss-less communication
  – Effective trade execution on optimized predictions or client orders.

• Effective trade execution
  – Execute within signal time frame
  – Minimize transaction costs
  – Minimize market impact
  – Benchmark: Volume Weighted Average Price (VWAP)
  – Often implemented with automated, algorithmic trading programs.
Execution Problem

• Partially observe order book (bids, asks and respective volumes)
  – Observe recent transacted prices
  – Observe recent transaction volumes on some exchanges
  – How do you place order?
  – What level? Best bid/offer or at offset?
  – How much to bid at each level?

• Equity Dark Pools:
  – Do not observe order book
  – Buyers and sellers submit volume of trade.
  – Orders queued, and matched first in first out.
  – Ganchev, Kearns et al. (2010)
Market Impact & Feedback

• Market Impact:
  – Signals fade as your action changes market conditions
  – Other market participants executing similar strategy

• Feedback:
  – Signals gets stronger and stronger, do you increase positions?
‘Expert’ Opinions

• TV analysts and business columnists have short-term effect

• Combining expert opinion:
  – Marshall Wace LLP ranks 750k analyst ideas, and pays commissions if idea is traded.
Market Sentiment Caricature
Market Sentiment

• Existing sentiment indices
  – Consumer Confidence Index
  – Purchasing Managers Index
  – Manufacturing Industry Index

• Twitter sentiment (Bollen, Mao, Zeng 2011)
News Events

• Words with significant impact:
  – Merger: X buys Y for $Z per share -> X and Y become correlated and Y trades in a band around $Z
  – Rights offering: shareholder can buy at discount
  – Buyback: company buys back at premium

• Can we learn key words that cause stock price movement?

• Is word significance related to document type?

• Mittermayer, (2004)
• Luss and d'Aspremont (2008)
• Schumaker & Chen (2009)
Summary

• Finance –
  – Data rich
  – Model poor (possibly)
• Many areas open to machine learning methods and models:
  – Prediction
  – Optimization
  – Execution
  – Market impact
Bibliography


• Trippi and Turban, Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real World Performance, McGraw-Hill, 1992

Bibliography (2)

  Volume 24, Issue 3
• Wilson and Ghahramani, "Generalised Wishart Processes“, UAI 2011
• Schumaker and Chen (2009), “Textual analysis of stock market prediction using breaking financial news: The AZFin text system”, ACM Transactions on Information Systems
• Luss and d'Aspremont (2008), Predicting Abnormal Returns From News Using Text Classification, Arxiv
• Bollen, Mao, Zeng, “Twitter mood predicts the stock market”, Journal of Computational Science, Volume 2, Issue 1, March 2011, Pages 1–8