Data Driven Methods in Finance and Insurance

Peter Forsyth

Cheriton School of Computer Science
University of Waterloo
paforsyt@uwaterloo.ca

Faculty of Mathematics
University of Waterloo

February 12, 2018
New York
“Everything in finance is basically an optimal stochastic control problem.” *Robert Ferstenberg*, (Global Head of Quantitative Research, ITG; previously Managing Director, electronic order flow, Morgan-Stanley).

This looks like an exaggeration, but

- In finance/insurance, we are faced with uncertainty (stochastic processes)
- Based on what we observe (which is noisy), we have to make decisions (controls)
- We try to make optimal decisions, based on some criteria (objective function)
Traditional Approach

Fit parametric model of stochastic process to data
- Different problems require different models
- Difficult optimization problem to determine parameters

Solve optimal control problem
- Stochastic dynamic programming
- Nonlinear Hamilton-Jacobi-Bellman equation
- Difficult for high-dimensional problems
Examples

Hedging/risk management of a large book of variable annuities
  - Traditional approach: nested Monte Carlo $\rightarrow$ computationally very demanding, tens of thousands of individual contracts in portfolio

Defined Contribution (DC) pension plans asset allocation glide path construction
  - $\simeq$ One trillion AUM in Target Date Funds (TDFs) in US
  - Almost all these TDFs use deterministic glide paths $\rightarrow$ Problem: deterministic strategies are provably sub-optimal
  - What’s needed:
    - Optimal adaptive (responds to current market conditions) and personalized (based on individual retirement goals/circumstances) strategy
Examples: II

Hospital admissions:
- Based on health insurance claim data
  - Predict if insured is readmitted to hospital, days of stay
  - Control: preventive therapy

Asset allocation for Defined Benefit (DB) pension plans
- Liability driven asset allocation
- Minimize volatility of funding ratio
Data-Driven Approach

Skip step of developing parametric model based on historical/market data.

Operate directly on the data (bootstrapping if necessary)
- Optimize objective function directly using resampled data
- Use machine learning techniques to represent optimal controls

Advantages:
- Adjusts automatically to new data ("Let the data speak for itself")
- Similar computational approach for many seemingly different problems
- Handles high dimensional problems, complex decision making
Does this method work?

Optimal glide path for DC pension plan
- From parametric model of historical data, generate *ground truth* in synthetic market
  - Determine provably optimal adaptive glide path asset allocation using traditional methods
- Generate sample data sets from synthetic market
  - Machine learning approach gets very close to the known optimal strategy, just by processing the data!
- Similar results on real historical data

Hedging options
- Based on historical samples of option prices and underlying asset prices
  - Learn hedging strategy
- Out of sample tests
  - Performs better than traditional approach, i.e. parametric model, solve option pricing equation, determine delta, gamma
Transformative technology

- Determine controls (decisions) by operating directly on the data (machine learning)
- Can develop automated, personalized approach for clients
- Optimal controls $\rightarrow$ to achieve financial/risk management goals