Five (US) Perspectives on Big Data and Mathematical Finance

Knut Sølna, University of California at Irvine

1 Big Data as a Market∼Technology Feature

- Spectrum of electronic order-driven markets[†]:

Regime	Time scale	Issues
Ultra-high ferquency (UHF)	$\sim 10^{-3} - .1s$	Microstructure Latency
High Frequency (HF)	$\sim 1 - 100s$	Trade execution
"Daily"	$\sim 10^3 - 10^4 s$	Trading strategies Option hedging

$\rightarrow$ Calls for modeling of price interaction & impact via the order (flow) book, as well as frequency coupling and multiscale analysis.

- Computational power enables the use of collateral data and statistical learning techniques.

[†] Cont, 2011, "Statistical Modeling of High-Frequency Financial Data".

2 Big Data as Viewed Through the SIAM Meeting Lens

SIAM: Society of Industrial and Applied Mathematics.

- Activity Group on Financial Mathematics and Engineering organizes every second year:

$\hookrightarrow$ Meeting on Financial Mathematics and Engineering.

Big Data at the Origin of Time: Boston Meeting July 2006

• One Minisymposium, out of 22, on Market Models with one talk on: "Application of Machine Learning to Short-Term Equity Return Prediction", Yan et al.

$\rightarrow$ Use "kernel regression" approach to predict tomorrow's return based on previous returns and volumes.

$\hookrightarrow$ Approximately 1 out of 90 talks ...

Big Data Today: Austin Meeting Nov 2016

$\rightarrow$ A total of 62 Minisymposia

- 3 minisymposia on Machine Learning

Tools as also used in imaging, speech recognition, for instance, "learning patterns from big data rather than programming patterns"

- 2 minisymposia on Algorithmic and High Frequency Trading

Algorithms and (microstructure) models for electronic markets

- 1 minisymposium on Optimal Trading and Limit Order Books on price impact and optimal execution using optimization tools

- 1 minisymposium on Computational Optimization and Data Mining for Finance, Data mining (machine learning) for risk management (insurance) and portfolio optimization

- 1 minisymposium on High-Frequency Financial Statistics, Statistics of high-frequency data, in particular returns, volatility and coupling, that is "leverage"

$\hookrightarrow$ Approximately 30 out of 250 talks ...

What is “big” anyway?

$\rightarrow$ Other problems where dimensionality plays an important role (typically $N \sim 100$):

- Portfolio optimization, 28 talks.
- Credit portfolios and default risk (Collateral Debt Obligation), 6 talks.
- Financial networks and (systemic) risk, 17 talks.

$\hookrightarrow$ Coupling/correlation make such problems challenging.

$\hookrightarrow$ Approximately a net of 80 out of 250 talks in total ...

3 Big Data and Machine Learning: Further Perspectives

$\rightarrow$ Samples of talks:

- Returns Modeling:

Ex: “Sparse Signals in the Cross-Section of Returns”, Chinco et al.

- Limit Order Book (High-frequency Trading):

Ex: “Deep Learning for Limit Order Books”, Sirignano et al.

- Mortgage Risk:

Ex: “Deep Learning for Mortgage Risk”, Giesecke et al.

- Volatility:

Ex: “When Arma Meets High-Frequency Data”, Xiu et al.

- Regulators:

Ex: “Do U.S. Financial Regulators Listen to the Public?”, Kirilenko et al.

4 Big Data and the Firm as Experienced by my Students

- Sam Luxenberg - intership in start-up "I know first".

$\rightarrow$ The start-up considers: stocks, indices, currencies, commodities in more than 200 markets and how these markets interact. The company uses machine learning tools (as well as wavelets) for forecasting and identification of "market inefficiency..."

- Fan Wang - current "data" project at JP Morgan:

Computation of leverage, or risk, for (hedge) funds:
Systemic and idiosyncratic risks computed via (un)supervised learning approaches rather than a scenario simulation-based approach.
$\rightarrow$ Trader learning increasingly supplemented by statistical learning.
$\rightarrow$ Quantification of risk increasingly data driven.

5 Final Remarks

• Machine learning tools will likely have an impact on how you drive your car, but also how your retirement savings are being invested.

• Artificial intelligence is not so artificial in a big data age.

• New challenges and possibilities in research and jobs: where is the math?

• Big data in finance involves a number and the integration of different tools – stochastics, statistics, numerics, optimization, computer science & financial modeling.
– How can they borrow strength from each other rather than one being merely a tool in the other?