VOL. 3, NO. 7

OMI Research Newsletter – May 2024



The Oxford-Man Institute of Quantitative Finance (OMI) is a world-leading centre for interdisciplinary research in financial markets. The OMI addresses fundamental problems in quantitative finance with a strong focus on machine learning and data driven models.

*The **OMI Research Newsletter** is a curated summary that covers the most recent research, insights, and tools related to financial decision making.*

Highlights

Research focus. Key contributions in this month's edition include works from the OMI, Imperial College London, King's College, University of Michigan, and Stevens Institute of Technology. Michael, Cucuringu and Howison (2024) extend HAR models to extract volatility estimators using option price data. Mc Greevy et al. (2024) study the dynamics of multivariate time series with unsupervised learning for pairs trading. Bayraktar, Cohen and Nellis (2024) study the equilibrium liquidity distribution in Uniswap v3 and investigate the effect of MEV bots. Finally, Sadeghi and Feinstein (2024) introduce a centrality measure of shock propagation in financial networks.

Machine Learning & Statistics

Options-driven Volatility Forecasting

N. MICHAEL, M. CUCURINGU, S. HOWISON
OXFORD-MAN INSTITUTE, UNIVERSITY OF OXFORD

We augment the Heterogeneous Autoregressive Regression model for forecasting realized volatility, using various measurements for the daily, weekly, and monthly volatilities, in addition to other predictive features. The main focus is on novel methods for extracting volatility estimators using option price data. Firstly, we provide a dimensionality reduction method for implied volatility surfaces built under the Black–Scholes model, whereby we combine simple row-wise and column-wise decompositions of the implied volatility surface with principal component analysis. Secondly, we provide a method for extracting the implied volatility under a Heston model. This is achieved by a calibration of the Heston model while assuming that some of the model parameters remain constant. We demonstrate that these augmentations result in improved daily forecasts for realized volatility in a selection of different stocks. These volatility forecasts, can be also be utilized to further increase predictive performance for the realized volatility of other instruments, and can be combined to provide accurate forecasts for VIX.

Detecting Multivariate Market Regimes Via Clustering Algorithms

J. MC GREEVY, A. MUGURUZA, Z. ISSA, C. SALVI, J. CHAN, Z. ZURIC
IMPERIAL COLLEGE LONDON, KING'S COLLEGE LONDON

In this paper we study the joint dynamics of multivariate time series using an unsupervised learning technique and demonstrate its use in pairs trading and portfolio design. We present a novel non-parametric market regime detection method for multidimensional data. The detection procedure is based on a k-means clustering algorithm which makes use of distances between distributions to discriminate between different market regimes, which are categorised by their mean, variance and correlation. In particular, we empirically investigate the performance of the algorithm endowed with either Wasserstein distances or Maximum Mean Discrepancies. We suggest a two-step approach to clustering multivariate data using this new method, which we show to be effective on both synthetic and real world data. We demonstrate how our new approach can be used to obtain approximations to the mean, variance and correlation between two assets at a given point in time. We further show that these values can be used in the context of Modern Portfolio Theory to

form profitable trading strategies when using two assets.

Conformal Prediction via Regression-as-Classification

E. Guha, S. Natarajan, T. Möllenhoff, M. E. Khan, E. Ndiaye

RIKEN Center for AI Project

Conformal prediction (CP) for regression can be challenging, especially when the output distribution is heteroscedastic, multimodal, or skewed. Some of the issues can be addressed by estimating a distribution over the output, but in reality, such approaches can be sensitive to estimation error and yield unstable intervals. Here, we circumvent the challenges by converting regression to a classification problem and then use CP for classification to obtain CP sets for regression. To preserve the ordering of the continuous-output space, we design a new loss function and make necessary modifications to the CP classification techniques. Empirical results on many benchmarks shows that this simple approach gives surprisingly good results on many practical problems.

Leave No Context Behind: Efficient Infinite Context Transformers with Infini-attention

T. Munkhdalai, M. Faruqui, S. Gopa

Google

This work introduces an efficient method to scale Transformer-based Large Language Models (LLMs) to infinitely long inputs with bounded memory and computation. A key component in our proposed approach is a new attention technique dubbed Infini-attention. The Infini-attention incorporates a compressive memory into the vanilla attention mechanism and builds in both masked local attention and long-term linear attention mechanisms in a single Transformer block. We demonstrate the effectiveness of our approach on long-context language modeling benchmarks, 1M sequence length passkey context block retrieval and 500K length book summarization tasks with 1B and 8B LLMs. Our approach introduces minimal bounded memory parameters and enables fast streaming inference for LLMs

Variational Bayesian Last Layers

J. Harrison, J. Willes, J. Snoek

Google DeepMind, Vector Institute

We introduce a deterministic variational formulation for training Bayesian last layer neural networks. This yields a sampling-free, single-pass model and loss that effectively improves uncertainty estimation. Our variational Bayesian last layer (VBLL) can be trained and evaluated with only quadratic complexity in last layer width, and is thus (nearly) computationally free to add to standard architectures. We experimentally investigate VBLLs, and show that they improve predictive accuracy, calibration, and out of distribution detection over baselines across both regression and classification. Finally, we investigate combining VBLL layers with variational Bayesian feature learning, yielding a lower variance collapsed variational inference method for Bayesian neural networks.

Theoretical Foundations of Deep Selective State-Space Models

N. M. Muca Cirone, A. Orvieto, B. Walker, C. Salvi, T. Lyons

Imperial College London

Structured state-space models (SSMs) such as S4, stemming from the seminal work of Gu et al., are gaining popularity as effective approaches for modeling sequential data. Deep SSMs demonstrate outstanding performance across a diverse set of domains, at a reduced training and inference cost compared to attention-based transformers. Recent developments show that if the linear recurrence powering SSMs allows for multiplicative interactions between inputs and hidden states (e.g. GateLoop, Mamba, GLA), then the resulting architecture can surpass in both in accuracy and efficiency attention-powered foundation models trained on text, at scales of billion parameters. In this paper, we give theoretical grounding to this recent finding using tools from Rough Path Theory: we show that when random linear recurrences are equipped with simple input-controlled transitions (selectivity mechanism), then the hidden state is provably a low-dimensional projection of a powerful mathematical object called the signature of the input – capturing non-linear interactions between tokens at distinct timescales. Our theory not only motivates the success of modern selective state-space models such as Mamba but also provides a solid framework to understand the expressive power of future SSM variants.

Finance and Economics

Reference Model Based Learning in Expectation Formation: Experimental Evidence

J. Pei

Nanyang Technological University

How do people form expectations about future prices in financial markets? One of the dominant learning rules that explains the forecasting behavior is the Adaptive Expectation Rule (ADA), which suggests that people adjust their predictions by adapting to the most recent prediction error at a constant weight. However, this rule also implies that they will continually learn and adapt until the prediction error is zero, which

contradicts recent experimental evidence showing that people usually stop learning long before reaching zero prediction error. A more recent learning rule, Reference Model Based Learning (RMBL), extends and generalizes ADA, hypothesizing that: i) People apply ADA but dynamically adjust the adaptive coefficient with regards to the auto-correlation of the prediction error in the most recent two periods; ii) Meanwhile, they also utilize a satisficing rule so that people would only adjust their adaptive coefficient when the prediction error is higher than their anticipation. This paper utilizes a rich set of experimental data with observations of 41,490 predictions from 801 subjects from the Learning-to-Forecast Experiments (LtFEs), i.e., the experiment that has been used to study expectation formation. Our results concludes that RMBL fits better than ADA in all the experiments.

Long run consequence of p-hacking

X. Wang
Dongbei University of Finance and Economics

We study the theoretical consequence of p-hacking on the accumulation of knowledge under the framework of mis-specified Bayesian learning. A sequence of researchers, in turn, choose projects that generate noisy information in a field. In choosing projects, researchers need to carefully balance as projects generates big information are less likely to succeed. In doing the project, a researcher p-hacks at intensity ε so that the success probability of a chosen project increases (unduly) by a constant ε. In interpreting previous results, researcher behaves as if there is no p-hacking because the intensity ε is unknown and presumably small. We show that over-incentivizing information provision leads to the failure of learning as long as $\varepsilon \neq 0$. If the incentives of information provision is properly provided, learning is correct almost surely as long as ε is small.

Belief Bias Identification

P. Gonzalez-Fernandez
Maastricht University

This paper proposes a unified theoretical model to identify and test a comprehensive set of probabilistic updating biases within a single framework. The model achieves separate identification by focusing on the updating of belief distributions, rather than classic point-belief measurements. Testing the model in a laboratory experiment reveals significant heterogeneity at the individual level: All tested biases are present, and each participant exhibits at least one identifiable bias. Notably, motivated-belief biases (optimism and pessimism) and sequence-related biases (gambler's fallacy and hot hand fallacy) are identified as key drivers of biased inference. Moreover, at the population level, base rate neglect emerges as a persistent influence. This study contributes to the belief-updating literature by providing a methodological toolkit for researchers examining links between different conflicting biases, or exploring connections between updating biases and other behavioural phenomena.

Strategic Informed Trading and the Value of Private Information

M. Anthropelos, S. Robertson
University of Piraeus

We consider a market of risky financial assets where the participants are an informed trader, a mass of uniformed traders and noisy liquidity providers. We prove the existence of a market-clearing equilibrium when the insider internalizes her power to impact prices. In the price-impact equilibrium the insider strategically reveals a noisier (compared to when the insider takes prices as given) signal, and prices are less reactive to the publicly available information. In contrast to the related literature, we show that in the price-impact equilibrium, the insider's ex-ante welfare monotonically increases in the signal precision. This clarifies when a trader with market power is motivated to both obtain and refine her private information. Furthermore, even though the uniformed traders act as price-takers, the effect of price impact is ex-ante welfare improving for them. By contrast, internalization of price impact may reduce insider ex-ante welfare. This happens provided the insider is sufficiently risk averse and the uninformed traders are sufficiently risk tolerant.

Mathematical Finance

DEX Specs: A Mean Field Approach to DeFi Currency Exchanges

E. Bayraktar, A. Cohen, A. Nellis
University of Michigan

We investigate the behavior of liquidity providers (LPs) by modeling a decentralized cryptocurrency exchange (DEX) based on Uniswap v3. LPs with heterogeneous characteristics choose optimal liquidity positions subject to uncertainty regarding the size of exogenous incoming transactions and the prices of assets in the wider market. They engage in a game among themselves, and the resulting liquidity distribution determines the exchange rate dynamics and potential arbitrage opportunities of the pool. We calibrate the distribution of LP characteristics based on Uniswap data and the equilibrium strategy resulting from this mean-field game produces pool exchange rate dynamics and liquidity evolution consistent with observed pool behavior. We subsequently introduce Maximal Extractable Value (MEV) bots who

perform Just-In-Time (JIT) liquidity attacks, and develop a Stackelberg game between LPs and bots. This addition results in more accurate simulated pool exchange rate dynamics and stronger predictive power regarding the evolution of the pool liquidity distribution.

Statistical Validation of Contagion Centrality in Financial Networks

A. Sadeghi, Z. Feinstein
Stevens Institute of Technology

In this paper, we introduce a novel centrality measure to evaluate shock propagation on financial networks capturing a notion of contagion and systemic risk contributions. In comparison to many popular centrality metrics (e.g., eigenvector centrality) which provide only a relative centrality between nodes, our proposed measure is in an absolute scale permitting comparisons of contagion risk over time. In addition, we provide a statistical validation method when the network is estimated from data, as is done in practice. This statistical test allows us to reliably assess the computed centrality values. We validate our methodology on simulated data and conduct empirical case studies using financial data. We find that our proposed centrality measure increases significantly during times of financial distress and is able to provide insights in to the (market implied) risk-levels of different firms and sectors.

Acknowledgement

The OMI Research Newsletter is curated by Álvaro Cartea, Director of the Oxford-Man Institute, and Fayçal Drissi, postdoctoral researcher at the institute. Paper suggestions are from fellow members at the institute: Álvaro Arroyo, Patrick Chang, Gerardo Durán-Martín, Marcello Monga, and Harrison Waldon.

References

Anthropelos, Michail, and Scott Robertson. 2024. "Strategic Informed Trading and the Value of Private Information." *arXiv preprint arXiv:2404.08757*.

Bayraktar, Erhan, Asaf Cohen, and April Nellis. 2024. "DEX Specs: A Mean Field Approach to DeFi Currency Exchanges." *arXiv preprint arXiv:2404.09090*.

Cirone, Nicola Muca, Antonio Orvieto, Benjamin Walker, Cristopher Salvi, and Terry Lyons. 2024. "Theoretical Foundations of Deep Selective State-Space Models." *arXiv preprint arXiv:2402.19047*.

Gonzalez-Fernandez, Pedro. 2024. "Belief Bias Identification." *arXiv preprint arXiv:2404.09297*.

Guha, Etash, Shlok Natarajan, Thomas Möllenhoff, Mohammad Emtiyaz Khan, and Eugene Ndiaye. 2024. "Conformal Prediction via Regression-as-Classification." *arXiv preprint arXiv:2404.08168*.

Harrison, James, John Willes, and Jasper Snoek. 2024. "Variational Bayesian last layers." *arXiv preprint arXiv:2404.11599*.

Mc Greevy, James, Aitor Muguruza, Zach Issa, Cristopher Salvi, Jonathan Chan, and Zan Zuric. 2024. "Detecting Multivariate Market Regimes Via Clustering Algorithms." *Available at SSRN 4758243*.

Michael, Nikolas, Mihai Cucuringu, and Sam Howison. 2024. "Options-driven Volatility Forecasting." *Available at SSRN 4790644*.

Munkhdalai, Tsendsuren, Manaal Faruqui, and Siddharth Gopal. 2024. "Leave No Context Behind: Efficient Infinite Context Transformers with Infini-attention." *arXiv preprint arXiv:2404.07143*.

Pei, Jiaoying. 2024. "Reference Model Based Learning in Expectation Formation: Experimental Evidence." *arXiv preprint arXiv:2404.08908*.

Sadeghi, Agathe, and Zachary Feinstein. 2024. "Statistical Validation of Contagion Centrality in Financial Networks." *arXiv preprint arXiv:2404.14337*.

Wang, Xuanye. 2024. "Long run consequence of p-hacking." *arXiv preprint arXiv:2404.08984*.