OMI Research Newsletter – February 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, University of Southern California, The University of Texas at Austin, Columbia University, Université Paris Dauphine-PSL, University of Warwick, University of York, Macquarie University, National University of Singapore, and University of Southampton. Leng et al. (2024) use a functional principal factor approach to estimate high-dimensional covariance matrices of random functions. Xu, Zariphopoulou and Zhang (2023) study a decision-making problem with information acquisition and adaptive information sources framed as a sequential optimal stopping problem. Bergault and Sánchez-Betancourt (2024) study the equilibrium between a broker and informed traders in a stochastic game. Finally, Luo, Ma and Cucuringu (2024) use a novel Graph Neural Network, designed for better spatial-temporal processing of data, to predict stock price movements.

Machine Learning & Statistics

Generating Synthetic Data for Neural Operators

E. HASANI, R. A. WARD
UNIVERSITY OF TEXAS AT AUSTIN

Numerous developments in the recent literature show the promising potential of deep learning in obtaining numerical solutions to partial differential equations (PDEs) beyond the reach of current numerical solvers. However, data-driven neural operators all suffer from the same problem: the data needed to train a network depends on classical numerical solvers such as finite difference or finite element, among others. In this paper, we propose a new approach to generating synthetic functional training data that does not require solving a PDE numerically. The way we do this is simple: we draw a large number N of independent and identically distributed 'random functions' u_j from the underlying solution space (e.g., $H_0^1(\Omega)$) in which we know the solution lies according to classical theory. We then plug each such random candidate solution into the equation and get a corresponding right-hand side function f_j for the equation, and consider $(f_j, u_j)_{j=1}^N$ as supervised training data for learning the underlying inverse problem $f \to u$. This 'backwards' approach to generating training data only requires derivative computations, in contrast to standard 'forward' approaches, which require a numerical PDE solver, enabling us to generate a large number of such data points quickly and efficiently. While the idea is simple, we hope that this method will expand the potential for developing neural PDE solvers that do not depend on classical numerical solvers.

Spatial-Temporal Stock Movement Prediction and Portfolio Selection based on the Semantic Company Relationship Graph

C. LUOA, T. MAA, M. CUCURINGU
OXFORD-MAN INSTITUTE, UNIVERSITY OF EDINBURGH, UNIVERSITY OF SOUTHAMPTON

In this paper, we approach stock price movements as a spatial-temporal prediction task, advancing beyond the traditional view of stocks as standalone entities. We first represent companies as vector embeddings based on company name co-occurrence statistics within a large financial news corpus, and then construct a Semantic Company Relationship Graph (SCRG) using

cosine similarities between vectors to define the mutual relationships. To tackle the financial prediction task, we introduce a novel Non-Independent and Identically Distributed Spatial-Temporal Graph Neural Network (NIST-GNN), specifically designed to propagate features from both neighboring companies and internal historical data while effectively handling the inherent temporal non-IIDness in stock sequences. This innovative aspect of our NIST-GNN allows for a more nuanced understanding and processing of temporal data, setting it apart from traditional spatial-temporal approaches. Our experimental results demonstrate that this methodology significantly outperforms benchmark models, yielding superior profitability and enhancing the Sharpe Ratio by 0.40 compared to the best-performing baseline, with statistical significance. Importantly, our findings also provide valuable theoretical insights into the effect of information diffusion within the US market, revealing that public information from cross-correlated companies typically experiences a minimum one-day lag before diffusion, challenging conventional perceptions of market efficiency.

Covariance Function Estimation for High-Dimensional Functional Time Series with Dual Factor Structures

C. Leng, D. Li, H. Shang, Y. Xia

University of Warwick, University of York, Macquarie University, National University of Singapore

We propose a flexible dual functional factor model for modelling high-dimensional functional time series. In this model, a high-dimensional fully functional factor parametrisation is imposed on the observed functional processes, whereas a low-dimensional version (via series approximation) is assumed for the latent functional factors. We extend the classic principal component analysis technique for the estimation of a low-rank structure to the estimation of a large covariance matrix of random functions that satisfies a notion of (approximate) functional "low-rank plus sparse" structure; and generalise the matrix shrinkage method to functional shrinkage in order to estimate the sparse structure of functional idiosyncratic components. Under appropriate regularity conditions, we derive the large sample theory of the developed estimators, including the consistency of the estimated factors and functional factor loadings and the convergence rates of the estimated matrices of covariance functions measured by various (functional) matrix norms. Consistent selection of the number of factors and a data-driven rule to choose the shrinkage parameter are discussed. Simulation and empirical studies are provided to demonstrate the finite-sample performance of the developed model and estimation methodology.

Neural Hawkes: Non-Parametric Estimation in High Dimension and Causality Analysis in Cryptocurrency Markets

T. Fabre, I. Muni Toke

Centrale Supélec, Université Paris-Saclay

We propose a novel approach to marked Hawkes kernel inference which we name the moment-based neural Hawkes estimation method. Hawkes processes are fully characterized by their first and second order statistics through a Fredholm integral equation of the second kind. Using recent advances in solving partial differential equations with physics-informed neural networks, we provide a numerical procedure to solve this integral equation in high dimension. Together with an adapted training pipeline, we give a generic set of hyperparameters that produces robust results across a wide range of kernel shapes. We conduct an extensive numerical validation on simulated data. We finally propose two applications of the method to the analysis of the microstructure of cryptocurrency markets. In a first application we extract the influence of volume on the arrival rate of BTC-USD trades and in a second application we analyze the causality relationships and their directions amongst a universe of 15 cryptocurrency pairs in a centralized exchange.

Towards Transparent Time Series Forecasting

Anonymous

Transparent machine learning (ML) models are essential for ensuring interpretability and trustworthiness in decision-making systems, particularly in high-stakes domains such as healthcare, finance, and criminal justice. While transparent machine learning models have been proposed for classification and regression, time series forecasting presents some unique challenges for ensuring transparency. In particular, currently used bottom-up approaches that focus on the values of the time series at specific time points (usually regularly spaced) do not provide a holistic understanding of the entire time series. This limits the applicability of ML in many critical areas. To open up these domains for ML, we propose a top-down framework of bi-level transparency, which involves understanding the higher-level trends and the lower-level properties of the predicted time series. Applying this framework, we develop TIMEVIEW, a transparent ML model for time series forecasting based on static features, complemented with an interactive visualization tool. Through a series of experiments, we demonstrate the efficacy and interpretability of our approach, paving the way for more transparent and reliable applications of ML in various domains.

MEDUSA: Simple LLM Inference Acceleration Framework with Multiple Decoding Heads

T. Cai, Y. Li, Z. Geng, H. Peng, J. D. Lee, D. Chen, T. Dao

Princeton University, University of Illinois, Carnegie Mellon University, University of Connecticut

The inference process in Large Language Models (LLMs) is often limited due to the absence of parallelism in the auto-regressive decoding process, resulting in most operations being restricted by the memory bandwidth of accelerators. While methods such as speculative decoding have been suggested to address this issue, their implementation is impeded by the challenges associated with acquiring and maintaining a separate draft model. In this paper, we present Medusa, an efficient method that augments LLM inference by adding extra decoding heads to predict multiple subsequent tokens in parallel. Using a tree-based attention mechanism, Medusa constructs multiple candidate continuations and verifies them simultaneously in each decoding step. By leveraging parallel processing, Medusa introduces only minimal overhead in terms of single-step latency while substantially reducing the number of decoding steps required. We present two levels of fine-tuning procedures for Medusa to meet the needs of different use cases: Medusa-1: Medusa is directly fine-tuned on top of a frozen backbone LLM, enabling lossless inference acceleration. Medusa-2: Medusa is fine-tuned together with the backbone LLM, enabling better prediction accuracy of Medusa heads and higher speedup but needing a special training recipe that preserves the backbone model's capabilities. Moreover, we propose several extensions that improve or expand the utility of Medusa, including a self-distillation to handle situations where no training data is available and a typical acceptance scheme to boost the acceptance rate while maintaining generation quality. We evaluate Medusa on models of various sizes and training procedures. Our experiments demonstrate that Medusa-1 can achieve over 2.2x speedup without compromising generation quality, while Medusa-2 further improves the speedup to 2.3-3.6x.

TSPP: A Unified Benchmarking Tool for Time-series Forecasting

J. Bączek, D. Zhylko, G. Titericz, S. Darabi, J-F. Puget, I. Putterman, D. Majchrowski, A. Gupta, K. Kranen, P. Morkisz

NVIDIA

While machine learning has witnessed significant advancements, the emphasis has largely been on data acquisition and model creation. However, achieving a comprehensive assessment of machine learning solutions in real-world settings necessitates standardization throughout the entire pipeline. This need is particularly acute in time series forecasting, where diverse settings impede meaningful comparisons between various methods. To bridge this gap, we propose a unified benchmarking framework that exposes the crucial modelling and machine learning decisions involved in developing time series forecasting models. This framework fosters seamless integration of models and datasets, aiding both practitioners and researchers in their development efforts. We benchmark recently proposed models within this framework, demonstrating that carefully implemented deep learning models with minimal effort can rival gradient-boosting decision trees requiring extensive feature engineering and expert knowledge.

Finance and Economics

Learning about a changing state

B. Davies

Stanford University

A long-lived Bayesian agent observes costly signals of a time-varying state. He chooses the signals' precisions sequentially, balancing their costs and marginal informativeness. I compare the optimal myopic and forward-looking precisions when the state follows a Brownian motion. I also compare the myopic precisions induced by other Gaussian processes.

On the use of artificial intelligence in financial regulations and the impact on financial stability

J. Danielsson, A. Uthemann

London School of Economics

Artificial intelligence (AI) can undermine financial stability because of malicious use, misaligned AI engines and since financial crises are infrequent and unique, frustrating machine learning. Even if the authorities prefer a conservative approach to AI adoption, it will likely become widely used by stealth, taking over increasingly high-level functions, driven by significant cost efficiencies and its superior performance on specific tasks. We propose six criteria against which to judge the suitability of AI use by the private sector for financial regulation and crisis resolution and identify the primary channels through which AI can destabilise the system.

Equity auction dynamics: latent liquidity models with activity acceleration

Mohammed Salek, Damien Challet, Ioane Muni Toke

Centrale Supélec, Université Paris-Saclay

Equity auctions display several distinctive characteristics in contrast to continuous trading. As the auction time approaches, the rate of events accelerates causing a substantial liquidity buildup around the indicative price. This, in turn, results in a reduced price impact and decreased volatility of the indicative price. In this study, we adapt the latent/revealed order book framework to the specifics of equity auctions. We provide precise measurements of the model parameters, including order submissions, cancellations, and diffusion rates. Our setup allows us to describe the full dynamics of the average order book during closing auctions in Euronext Paris. These findings support the relevance of the latent liquidity framework in describing limit order book dynamics. Lastly, we analyze the factors contributing to a sub-diffusive indicative price and demonstrate the absence of indicative price predictability.

Mathematical Finance

Decision Making under Costly Sequential Information Acquisition: the Paradigm of Reversible and Irreversible Decisions

R. Xu, T. Zariphopoulou, L. Zhang
University of Southern California, The University of Texas at Austin, Columbia University

Decision making in modern stochastic systems, including e-commerce platforms, financial markets, and healthcare systems, has evolved into a multifaceted process that involves information acquisition and adaptive information sources. This paper initiates a study on this integrated process, where these elements are not only fundamental but also interact in a complex and dynamically intertwined manner.

We introduce a relatively simple model, which, however, captures the novel elements we consider. Specifically, a decision maker (DM) can choose between an established product A with a known value and a new product B with an unknown value. The DM can observe signals about the unknown value of product B and can also opt to exchange it for product A if B is initially chosen. Mathematically, the model gives rise to a sequential optimal stopping problem with two different informational regimes (before and after buying product B), differentiated by the initial, coarser signal and the subsequent, finer one. We analyze the underlying problems using predominantly viscosity solution techniques, differing from the existing literature on information acquisition which is based on traditional optimal stopping techniques. Additionally, our modeling approach offers a novel framework for developing more complex interactions among decisions, information sources, and information costs through a sequence of nested obstacles.

Mean field games with unbounded controlled common noise in portfolio management with relative performance criteria

P. E. Souganidis, T. Zariphopoulou
University of Chicago, The University of Texas at Austin

Motivated by optimal allocation models with relative performance criteria, we introduce a mean field game in which the terminal expected utility of the representative agent depends on her own state as well as the average of her peers. We derive the master equation, which, in view of the presence of controls in the volatility, needs to be coupled with a compatibility condition for the mean field optimal feedback control. We concentrate on the class of separable payoffs under both general utilities and couplings. We derive a solution to the master equation and find the associated optimal feedback control expressed via the value function in the absence of competition and a dynamic coupling function solving a non-local quasilinear equation. In turn, we construct the related optimal state and control processes, and give representative examples. Projecting the mean field solutions on finite dimensions, we recover the solution of the N-game for linear couplings and arbitrary utilities, and, we study the proximity of these approximations to their N-player game counterparts.

A Mean Field Game between Informed Traders and a Broker

P. Bergault, L. Sánchez-Betancourt
Université Paris Dauphine-PSL, Oxford-Man Institute

We find closed-form solutions to the stochastic game between a broker and a mean-field of informed traders. In the finite player game, the informed traders observe a common signal and a private signal. The broker, on the other hand, observes the trading speed of each of his clients and provides liquidity to the informed traders. Each player in the game optimises wealth adjusted by inventory penalties. In the mean field version of the game, using a Gâteaux derivative approach, we characterise the solution to the game with a system of forward-backward stochastic differential equations that we solve explicitly. We find that the optimal trading strategy of the broker is linear on his own inventory, on the average inventory among informed traders, and on the common signal or the average trading speed of the informed traders. The Nash equilibrium we find helps informed traders decide how to use private information, and helps brokers decide how much of the order flow they should externalise or internalise when facing a large number of clients.

Nash Equilibria in Greenhouse Gas Offset Credit Markets

L. Welsh, S. Jaimungal
University of Toronto, Oxford-Man Institute

In response to the global climate crisis, governments worldwide are introducing legislation to reduce greenhouse gas (GHG) emissions to help mitigate environmental catastrophes. One method to encourage emission reductions is to incentivize carbon capturing and carbon reducing projects while simultaneously penalising excess GHG output. Firms that invest in carbon capturing projects or reduce their emissions can receive offset credits (OCs) in return. These OCs can be used for regulatory purposes to offset their excess emissions in a compliance period. OCs may also be traded between firms. Thus, firms have the choice between investing in projects to generate OCs or to trade OCs. In this work, we present a novel market framework and characterise the optimal behaviour of GHG OC market participants in both single-player and two-player settings. We analyse both a single-period and multi-period setting. As the market model does not elicit a closed form solution, we develop a numerical methodology to estimate players' optimal behaviours in accordance to the Nash equilibria. Our findings indicate the actions players take are dependent on the scale of their project opportunities as well as their fellow market participants. We demonstrate the importance of behaving optimally via simulations in order to offset emission penalties and the importance of investing in GHG reducing or capturing projects from a financial perspective.

Volatility models in practice: Rough, Path-dependent or Markovian?

E. Abi Jaber, S. Li
Ecole Polytechnique, Université Paris 1 Panthéon-Sorbonne

An extensive empirical study of the class of Volterra Bergomi models using SPX options data between 2011 and 2022 reveals the following fact-check on two fundamental claims echoed in the rough volatility literature:

Do rough volatility models with Hurst index $H \in (0, 1/2)$ really capture well SPX implied volatility surface with very few parameters? No, rough volatility models are inconsistent with the global shape of SPX smiles. They suffer from severe structural limitations imposed by the roughness component, with the Hurst parameter $H \in (0, 1/2)$ controlling the smile in a poor way. In particular, the SPX at-the-money skew is incompatible with the power-law shape generated by rough volatility models. The skew of rough volatility models increases too fast on the short end, and decays too slow on the longer end where "negative" H is sometimes needed.

Do rough volatility models really outperform consistently their classical Markovian counterparts? No, for short maturities they underperform their one-factor Markovian counterpart with the same number of parameters. For longer maturities, they do not systematically outperform the one-factor model and significantly underperform when compared to an under-parameterized two-factor Markovian model with only one additional calibratable parameter.

On the positive side: our study identifies a (non-rough) path-dependent Bergomi model and an under-parameterized two-factor Markovian Bergomi model that consistently outperform their rough counterpart in capturing SPX smiles between one week and three years with only 3 to 4 calibratable parameters.

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 from Harrison Waldon.

References

Abi Jaber, Eduardo, and Shaun Xiaoyuan Li. 2024. "Volatility models in practice: Rough, Path-dependent or Markovian?"

Bączek, Jan, Dmytro Zhylko, Gilberto Titericz, Sajad Darabi, Jean-Francois Puget, Izzy Putterman, Dawid Majchrowski, Anmol Gupta, Kyle Kranen, and Pawel Morkisz. 2023. "TSPP: A Unified Benchmarking Tool for Time-series Forecasting." *arXiv preprint arXiv:2312.17100*.

Bergault, Philippe, and Leandro Sánchez-Betancourt. 2024. "A Mean Field Game between Informed Traders and a Broker." *arXiv preprint arXiv:2401.05257*.

Cai, Tianle, Yuhong Li, Zhengyang Geng, Hongwu Peng, Jason D Lee, Deming Chen, and Tri Dao. 2024. "Medusa: Simple LLM Inference Acceleration Framework with Multiple Decoding Heads." *arXiv preprint arXiv:2401.10774*.

Danielsson, Jon, and Andreas Uthemann. 2023. "On the use of artificial intelligence in financial regulations and the impact on financial stability." *arXiv preprint arXiv:2310.11293*.

Davies, Benjamin. 2024. "Learning about a changing state." *arXiv preprint arXiv:2401.03607*.

Fabre, Timothée, and Ioane Muni Toke. 2024. "Neural Hawkes: Non-Parametric Estimation in High Dimension and Causality Analysis in Cryptocurrency Markets." *arXiv preprint arXiv:2401.09361*.

Hasani, Erisa, and Rachel A Ward. 2024. "Generating synthetic data for neural operators." *arXiv preprint arXiv:2401.02398*.

Leng, Chenlei, Degui Li, Hamlin Shang, and Yingcun Xia. 2024. "Covariance Function Estimation for High-Dimensional Functional Time Series with Dual Factor Structures." *arXiv preprint arXiv:2401.05784*.

Luo, Chang, Tiejun Ma, and Mihai Cucuringu. 2024. "Spatial-Temporal Stock Movement Prediction and Portfolio Selection based on the Semantic Company Relationship Graph." *Available at SSRN*.

Salek, Mohammed, Damien Challet, and Ioane Muni Toke. 2024. "Equity auction dynamics: latent liquidity models with activity acceleration." *arXiv preprint arXiv:2401.06724*.

Souganidis, Panagiotis E, and Thaleia Zariphopoulou. 2023. "Mean field games with unbounded controlled common noise in portfolio management with relative performance criteria."

Welsh, Liam, and Sebastian Jaimungal. 2024. "Nash Equilibria in Greenhouse Gas Offset Credit Markets." *arXiv preprint arXiv:2401.01427*.

Xu, Renyuan, Thaleia Zariphopoulou, and Luhao Zhang. 2023. "Decision Making under Costly Sequential Information Acquisition: the Paradigm of Reversible and Irreversible Decisions." *arXiv preprint arXiv:2401.00569*.