VOL. 3, NO. 11

OMI Research Newsletter – September 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 Imperial College London, University of Berlin, University of Munich, University of Bern, Peking University, University of Zurich, Princeton University, Boston University, and Harvard University. Zhu et al. (2024) show that deep neural networks trained on a large dataset of human decisions predicts people's choices better than classical theories of strategic behavior, indicating that some components of human decision are not explained by these theories. Knochenhauer, Merkel and Zhang (2024) study continuous-time decision making with costly acquisition of information. Finally, Benkert, Liu and Netzer (2024) study the informational content in the response times of agents.

Machine Learning & Statistics

Capturing the Complexity of Human Strategic Decision-Making with Machine Learning

J-Q. ZHU, J. C. PETERSON, B. ENKE, T. L. GRIFFITHS
PRINCETON UNIVERSITY, BOSTON UNIVERSITY, HARVARD UNIVERSITY

Understanding how people behave in strategic settings–where they make decisions based on their expectations about the behavior of others–is a long-standing problem in the behavioral sciences. We conduct the largest study to date of strategic decision-making in the context of initial play in two-player matrix games, analyzing over 90,000 human decisions across more than 2,400 procedurally generated games that span a much wider space than previous datasets. We show that a deep neural network trained on these data predicts people's choices better than leading theories of strategic behavior, indicating that there is systematic variation that is not explained by those theories. We then modify the network to produce a new, interpretable behavioral model, revealing what the original network learned about people: their ability to optimally respond and their capacity to reason about others are dependent on the complexity of individual games. This context-dependence is critical in explaining deviations from the rational Nash equilibrium, response times, and uncertainty in strategic decisions. More broadly, our results demonstrate how machine learning can be applied beyond prediction to further help generate novel explanations of complex human behavior.

Reinforcement Learning in High-frequency Market Making

Y. ZHENG, Z. DING
PRINCETON UNIVERSITY

This paper establishes a new and comprehensive theoretical analysis for the application of reinforcement learning (RL) in high-frequency market making. We bridge the modern RL theory and the continuous-time statistical models in high-frequency financial economics. Different with most existing literature on methodological research about developing various RL methods for market making problem, our work is a pilot to provide the theoretical analysis. We target the effects of sampling frequency, and find an interesting tradeoff between error and complexity of RL algorithm when tweaking the values of the time increment Δ — as Δ becomes smaller, the error will be smaller but the complexity will be larger. We also study the two-player case under the general-sum game framework and establish the

convergence of Nash equilibrium to the continuous-time game equilibrium as $\Delta \to 0$. The Nash Q-learning algorithm, which is an online multi-agent RL method, is applied to solve the equilibrium. Our theories are not only useful for practitioners to choose the sampling frequency, but also very general and applicable to other high-frequency financial decision making problems, e.g., optimal executions, as long as the time-discretization of a continuous-time markov decision process is adopted. Monte Carlo simulation evidence support all of our theories.

Strong denoising of financial time-series

M. J. Feiler
University of Zurich

In this paper we introduce a method for significantly improving the signal to noise ratio in financial data. The approach relies on combining a target variable with different context variables and use auto-encoders (AEs) to learn reconstructions of the combined inputs. The objective is to obtain agreement among pairs of AEs which are trained on related but different inputs and for which they are forced to find common ground. The training process is set up as a "conversation" where the models take turns at producing a prediction (speaking) and reconciling own predictions with the output of the other AE (listening), until an agreement is reached. This leads to a new way of constraining the complexity of the data representation generated by the AE. Unlike standard regularization whose strength needs to be decided by the designer, the proposed mutual regularization uses the partner network to detect and amend the lack of generality of the learned representation of the data. The integration of alternative perspectives enhances the denoising capacity of a single AE and allows us to discover new regularities in financial time-series which can be converted into profitable trading strategies.

Graph Language Models

Moritz Plenz, Anette Frank
Heidelberg University

While Language Models (LMs) are the workhorses of NLP, their interplay with structured knowledge graphs (KGs) is still actively researched. Current methods for encoding such graphs typically either (i) linearize them for embedding with LMs – which underutilize structural information, or (ii) use Graph Neural Networks (GNNs) to preserve the graph structure – but GNNs cannot represent text features as well as pretrained LMs. In our work we introduce a novel LM type, the Graph Language Model (GLM), that integrates the strengths of both approaches and mitigates their weaknesses. The GLM parameters are initialized from a pre- trained LM to enhance understanding of in- dividual graph concepts and triplets. Simul- taneously, we design the GLM's architecture to incorporate graph biases, thereby promot- ing effective knowledge distribution within the graph. This enables GLMs to process graphs, texts, and interleaved inputs of both. Empirical evaluations on relation classification tasks show that GLM embeddings surpass both LM- and GNN-based baselines in supervised and zero- shot setting, demonstrating their versatility.

Finance and Economics

Revealed Invariant Preference

P. Caradonna, C. P. Chambers
Georgetown University

We consider the problem of rationalizing choice data by a preference satisfying an arbitrary collection of invariance axioms. Examples of such axioms include quasilinearity, homotheticity, independence-type axioms for mixture spaces, constant relative/absolute risk and ambiguity aversion axioms, stationarity for dated rewards or consumption streams, separability, and many others. We provide necessary and sufficient conditions for invariant rationalizability via a novel approach which relies on tools from the theoretical computer science literature on automated theorem proving. We also establish a generalization of the Dushnik-Miller theorem, which we use to give a complete description of the out-of-sample predictions generated by the data under any such collection of axioms.

Deviations from the Nash equilibrium and emergence of tacit collusion in a two-player optimal execution game with reinforcement learning

F. Lillo, A. Macrì
Scuola Normale Superiore, University of Bologna

The use of reinforcement learning algorithms in financial trading is becoming increasingly prevalent. However, the autonomous nature of these algorithms can lead to unexpected outcomes that deviate from traditional game-theoretical predictions and may even destabilize markets. In this study, we examine a scenario in which two autonomous agents, modeled with Double Deep Q-Learning, learn to liquidate the same asset optimally in the presence of market impact, using the Almgren-Chriss (2000) framework. Our results show that the strategies learned by the agents deviate significantly from the Nash equilibrium of the corresponding market impact game. Notably, the learned strategies exhibit tacit collusion, closely aligning with the Pareto-optimal solution. We further explore how different levels of market volatility influence the agents' performance and the equilibria

they discover, including scenarios where volatility differs between the training and testing phases.

Time is Knowledge: What Response Times Reveal

J-M. Benkert, S. Liu, N. Netzer

University of Bern, Peking University, University of Zurich

Response times contain information about economically relevant but unobserved variables like willingness to pay, preference intensity, quality, or happiness. Here, we provide a general characterization of the properties of latent variables that can be detected using response time data. Our characterization generalizes various results in the literature, helps to solve identification problems of binary response models, and paves the way for many new applications. We apply the result to test the hypothesis that marginal happiness is decreasing in income, a principle that is commonly accepted but so far not established empirically.

Mathematical Finance

Continuous-Time Dynamic Decision Making with Costly Information

C. Knochenhauer, A. Merkel, Y. Zhang

Imperial College London, University of Berlin, University of Munich

We consider a continuous-time linear-quadratic Gaussian control problem with partial observations and costly information acquisition. More precisely, we assume the drift of the state process to be governed by an unobservable Ornstein–Uhlenbeck process. The decision maker can additionally acquire information on the hidden state by conducting costly tests, thereby augmenting the available information. Combining the Kalman–Bucy filter with a dynamic programming approach, we show that the problem can be reduced to a deterministic control problem for the conditional variance of the unobservable state. Optimal controls and value functions are derived in a semi-explicit form, and we present an extensive study of the qualitative properties of the model. We demonstrate that both the optimal cost and the marginal cost increase with model uncertainty. We identify a critical threshold: below this level, it is optimal not to acquire additional information, whereas above this threshold, continuous information acquisition is optimal, with the rate increasing as uncertainty grows. For quadratic information costs, we derive the precise asymptotic behavior of the acquisition rate as uncertainty approaches zero and infinity.

Pricing and hedging of decentralised lending contracts

L. Szpruch, M. Sabaté Vidales, T. Treetanthiploet, Y. Zhang

Imperial College London, University of Edinburgh, Infinitas by Krungthai

We study the loan contracts offered by decentralised loan protocols (DLPs) through the lens of financial derivatives. DLPs, which effectively are clearinghouses, facilitate transactions between option buyers (i.e. borrowers) and option sellers (i.e. lenders). The loan-to-value at which the contract is initiated determines the option premium borrowers pay for entering the contract, and this can be deduced from the non-arbitrage pricing theory. We show that when there are no market frictions, and there is no spread between lending and borrowing rates, it is optimal to never enter the lending contract. Next, by accounting for the spread between rates and transactional costs, we develop a deep neural network-based algorithm for learning trading strategies on the external markets that allow us to replicate the payoff of the lending contracts that are not necessarily optimally exercised. This allows hedge the risk lenders carry by issuing options sold to the borrowers, which can complement (or even replace) the liquidations mechanism used to protect lenders' capital. Our approach can also be used to exploit (statistical) arbitrage opportunities that may arise when DLP allow users to enter lending contracts with loan-to-value, which is not appropriately calibrated to market conditions or/and when different markets price risk differently. We present thorough simulation experiments using historical data and simulations to validate our approach.

Logarithmic regret in the ergodic Avellaneda–Stoikov market making model

J. Cao, D. Šiška, L. Szpruch

University of Edinburgh, Infinitas by Krungthai

We analyse the regret arising from learning the price sensitivity parameter κ of liquidity takers in the ergodic version of the Avellaneda–Stoikov market making model. We show that a learning algorithm based on a regularised maximum-likelihood estimator for the parameter achieves the regret upper bound of order $\ln^2 T$ in expectation. To obtain the result we need two key ingredients. The first are tight upper bounds on the derivative of the ergodic constant in the Hamilton–Jacobi–Bellman (HJB) equation with respect to κ. The second is the learning rate of the maximum-likelihood estimator which is obtained from concentration inequalities for Bernoulli signals. Numerical experiment confirms the convergence and the robustness of the proposed algorithm.

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

Benkert, Jean-Michel, Shuo Liu, and Nick Netzer. 2024. "Time is knowledge: what response times reveal." *arXiv preprint arXiv:2408.14872*.

Cao, Jialun, David Šiška, Lukasz Szpruch, and Tanut Treetanthiploet. 2024. "Logarithmic regret in the ergodic Avellaneda-Stoikov market making model." *arXiv preprint arXiv:2409.02025*.

Caradonna, Peter, and Christopher P Chambers. 2024. "Revealed Invariant Preference." *arXiv preprint arXiv:2408.04573*.

Feiler, Matthias J. 2024. "Strong denoising of financial time-series." *arXiv preprint arXiv:2408.05690*.

Knochenhauer, Christoph, Alexander Merkel, and Yufei Zhang. 2024. "Continuous-Time Dynamic Decision Making with Costly Information." *arXiv preprint arXiv:2408.09693*.

Lillo, Fabrizio, and Andrea Macrì. 2024. "Deviations from the Nash equilibrium and emergence of tacit collusion in a two-player optimal execution game with reinforcement learning." *arXiv preprint arXiv:2408.11773*.

Plenz, Moritz, and Anette Frank. 2024. "Graph Language Models." *arXiv preprint arXiv:2401.07105*.

Szpruch, Lukasz, Marc Sabaté Vidales, Tanut Treetanthiploet, and Yufei Zhang. 2024. "Pricing and hedging of decentralised lending contracts." *arXiv preprint arXiv:2409.04233*.

Zheng, Yuheng, and Zihan Ding. 2024. "Reinforcement Learning in High-frequency Market Making." *arXiv preprint arXiv:2407.21025*.

Zhu, Jian-Qiao, Joshua C Peterson, Benjamin Enke, and Thomas L Griffiths. 2024. "Capturing the Complexity of Human Strategic Decision-Making with Machine Learning." *arXiv preprint arXiv:2408.07865*.