Riccardo Grazzi

riccardo.grazzi.4@gmail.com
prolearner.github.io/riccardograzzi.github.io
Riccardo.Grazzi
Riccardo Grazzi
prolearner

Curriculum Vitae

Profile: Post Doc at MSR Cambridge. I like to study and design efficient and scalable machine learning methods. Currently Interested in how to improve foundation models making them more efficient and expressive. My recent work has been focused on studying how to improve Linear RNNs.

Work Experience

2025 - now **Post Doc**, *Microsoft Research Cambridge*, Cambridge, UK
Developing quantization techniques

2023 - 2024 **Post Doc**, *Computational Statistics and Machine Learning, IIT*, Genova, Italy

2018 - 2023 **PhD fellow**, *Computational Statistics and Machine Learning, IIT*, Genova, Italy

2020 **Reasearch Intern**, *Amazon Web Services (AWS)*, Berlin, Germany
Project : "Meta-learning for time-series forecasting". Manager : Matthias Seeger

Education

2018 – 2023 **PhD in Computer Science**, *University College of London, UK*
Thesis : "Principled and Efficient Bilevel Optimization for Machine Learning"
Supervisor: Massimiliano Pontil. Mentor: Saverio Salzo.

2015 – 2017 **Master Degree in Computer Engineering**, *University of Florence*, Italy
Thesis : "Truncated Hyper-gradient for Hyperparameter Optimization", *Advisor : Paolo frasconi.*
Final grade: 110/110 cum laude. GPA 30/30

2012 – 2015 **Bachelor Degree in Computer Engineering**, *University of Florence*, Italy
Final grade: 110/110 cum laude. GPA 29.2/30

Selected Papers

1. J. Siems*, T. Carstensen*, A. Zela, F. Hutter, M Pontil, R. Grazzi* DeltaProduct: Improving State-Tracking in Linear RNNs via Householder Products. Preprint.
2. R. Grazzi*, J. Siems*, A. Zela, J. Franke, F. Hutter, M. Pontil Unlocking state-tracking in linear rnns through negative eigenvalues. (Oral) ICLR 2025.
3. R. Grazzi, M. Pontil, S. Salzo. Nonsmooth Implicit Differentiation: Deterministic and Stochastic Convergence Rates. ICML 2024.
4. R. Grazzi*, J. Siems*, S. Schrodi, T. Brox, H Frank. Is Mamba Capable of In-Context Learning?. AutoML 2024.
5. V. Kostic, P. Novelli, R. Grazzi, K. Lounici, M. Pontil Learning invariant representations of time-homogeneous stochastic dynamical systems. ICLR 2024.
6. R. Grazzi, M. Pontil, S. Salzo. Bilevel Optimization with a Lower-level Contraction: Optimal Sample Complexity without Warm-Start. JMLR 2023.
7. R. Grazzi, A. Akhavan, J. Falk, L. Cella, M. Pontil. Group Meritocratic Fairness in Linear Contextual Bandits. NeurIPS 2022.
8. R. Grazzi, M. Pontil, S. Salzo. Convergence Properties of Stochastic Hypergradients. AISTATS 2021.
9. R. Grazzi, L. Franceschi, M. Pontil and S. Salzo. On the Iteration Complexity of Hypergradient Computation. ICML 2020.
10. G. Denevi, C. Ciliberto, R. Grazzi, M. Pontil. Learning-to-Learn Stochastic Gradient Descent with Biased Regularization. ICML 2019.
11. L. Franceschi, P. Frasconi, S. Salzo, R. Grazzi and M. Pontil. Bilevel programming for hyperparameter optimization and meta-learning. ICML 2018.

Main Programming Skills

Proficient Python, PyTorch, Git, Sci-kit learn.
Notions Triton, Jax, C/C++.

Selected Open Source Repos

1. hypertorch (⋆ 100+) Lightweight flexible research-oriented package to compute hypergradients in PyTorch.
2. hyper-representation Official repo for the experiments in the paper "Bilevel Programming for Hyperparameter Optimization and Meta-Learning".

Awards

2022 **NeurIPS Top Reviewer**.

2021 **NeurIPS Outstanding Reviewer Award**. Top 8% of reviewers according to AC/Authors.

Research Activities

Teaching

2023, 2024 **University College of London**, *London, UK*

TA for the convex optimization part of the Master course "Advanced Topics in Machine Learning"

August 2023 **Bilevel Summer School**, *Southampton, UK*

Lecture (3h): "Principled and Efficient Bilevel Optimization for Machine Learning"

Talks and Presentations

2025 **ASAP Seminars**, *Online*

Talk: "State Tracking in Scalable Linear RNNs", (YouTube link)

2025 **AutoML/FlaNN Seminars**, *Online*

Talk: "Unlocking State-tracking in Linear RNNs with Negative Eigenvalues", (AutoML) (FlaNN) (ASAP)

April 2024 **AutoML Seminars**, *Online*

Talk: "Is Mamba Capable of In-Context Learning?", (YouTube link)

December 2023 **NeurIPS Conference**, *New Orleans, USA*

Poster: "Bilevel Optimization with a Lower-level Contraction: Optimal Sample Complexity without Warm-start"

July 2023 **Mind Team at Inria-Saclay**, *Paris, France*

Talk: "Principled and Efficient Bilevel Optimization for Machine Learning"

March 2023 **ML Crash Course**, *Genoa, IT*

Talk: "Principled and Efficient Bilevel Optimization for Machine Learning"

November 2022 **NeurIPS Conference**, *New Orleans, USA*

Video and Poster Presentation: "Group Meritocratic Fairness in Stochastic Linear Bandits"

September 2021 **IFIP TC7**, *Online*

Talk: "On the Iteration Complexity of Hypergradient Computation"

April 2020 **AISTATS Conference**, *Online*

Video Presentation: "Convergence Properties of Stochastic Hypergradients"

July 2020 **ICML Conference**, *Online*

Video Presentation: "On the Iteration Complexity of Hypergradient Computation"

June 2019 **ICML Conference**, *Long Beach, USA*

Poster:"Learning-to-Learn Stochastic Gradient Descent with Biased Regularization"

September 2018 **RIKEN & IIT workshop**, *Genoa, Italy*

Talk: "Bilevel programming for hyperparmeter optimization and meta-learning"

July 2018 **AutoML workshop, ICML**, *Stocholm, Sweden*

Poster and short talk: "Far-HO: A Bilevel Programming Package for Hyperparameter Optimization and Meta-Learning"

Reviewer

Conferences **NeurIPS 2021, 2022, 2025. ICML 2022, 2024. ICLR 2025.**

Journals **JMLR 2021, 2025. TPAMI 2021, SIMODS 2024**

Workshops **AutoML workshop, ICML 2019**