Cong Ma

Address: Jones 313, 5747 South Ellis Avenue, Chicago, IL 60637

Homepage: https://congma1028.github.io/

Email: congm@uchicago.edu

Employment

Assistant professor *July 2021 - present*

- Department of Statistics and the College, University of Chicago

Postdoctoral scholar *June 2020 - June 2021*

- Department of Statistics and Department of Electrical Engineering and Computer Sciences, University of California, Berkeley
- Advisor: Professor *Martin J. Wainwright*

Education

Ph.D. in **Operations Research and Financial Engineering** *May 2020*

- Princeton University, advised by Professor *Yuxin Chen* and Professor *Jianqing Fan*

B.E. in **Electrical Engineering** *July 2015*

- Tsinghua University

Exchange student *Aug. 2013 - Dec. 2013*

- School of Electrical and Computer Engineering, Georgia Institute of Technology

Research Interests

Mathematics of data science, nonconvex optimization, transfer learning, reinforcement learning, high-dimensional statistics, inference and uncertainty quantification, and their applications to neuroscience and signal processing.

Current External Research Support

1. National Science Foundation grant DMS-2311127 (no overlap with the proposed project)

Preprints

[P1] R. Pathak, **C. Ma**, "On the Design-Dependent Suboptimality of the Lasso," 2024.

[P2] Y. Gui, **C. Ma**, Y. Zhong, "Unraveling Projection Heads in Contrastive Learning: Insights from Expansion and Shrinkage," 2023.

[P3] G. Li, W. Wu, Y. Chi, **C. Ma**, A. Rinaldo, Y. Wei, "Sharp High-Probability Sample Complexities for Policy Evaluation with Linear Function Approximation," 2023.

Journal articles

[J1] **C. Ma**, R. Pathak, M. J. Wainwright, "Optimally Tackling Covariate Shift in RKHS-based Nonparametric Regression," *Annals of Statistics*, vol. 51, no. 2, pp. 738-761, 2023.

[J2] Y. Yang, **C. Ma**, "Optimal Tuning-Free Convex Relaxation for Noisy Matrix Completion," *IEEE Transactions on Information Theory*, vol. 69, no. 10, pp. 6571-6585, Oct. 2023.

[J3] H. Dong, T. Tong, **C. Ma**, Y. Chi, "Fast and Provable Tensor Robust Principal Component Analysis via Scaled Gradient Descent," *Information and Inference: A Journal of the IMA*, , vol. 12, no. 3, pp. 1716-1758, 2023.

[J4] P. Rashidinejad, B. Zhu, **C. Ma**, J. Jiao, S. Russel, "Bridging Offline Reinforcement Learning and Imitation Learning: A Tale of Pessimism," *IEEE Transactions on Information Theory*, vol. 68, no. 12, pp. 8156-8196, June 2022.

[J5] **C. Ma**, B. Zhu, J. Jiao, M. J. Wainwright, "Minimax Off-Policy Evaluation for Multi-Armed Bandits," *IEEE Transactions on Information Theory*, vol. 68, no. 8, pp. 5314-5339, Mar. 2022.

[J6] T. Tong, **C. Ma**, A. Prater-Bennette, E. Tripp, Y. Chi, "Scaling and Scalability: Provable Nonconvex Low-Rank Tensor Estimation from Incomplete Measurements," *Journal of Machine Learning Research*, vol. 23, no. 1, pp. 7312-7388, 2022.

[J7] Y. Chen, Y. Chi, J. Fan, **C. Ma** (**alphabetical order**), "Spectral Methods for Data Science: A Statistical Perspective," *Foundations and Trends in Machine Learning*, 2021.

[J8] Y. Chen, J. Fan, **C. Ma**, Y. Yan (**alphabetical order**), "Bridging Convex and Nonconvex Optimization in Robust PCA: Noise, Outliers, and Missing Data," *Annals of Statistics*, vol. 49, no. 5, pp. 2948-2971, 2021.

[J9] Y. Chen, **C. Ma**, H. V. Poor, Y. Chen, "Learning Mixtures of Low-Rank Models," *IEEE Transactions on Information Theory*, vol. 67, no. 7, pp. 4613-4636, July 2021.

[J10] T. Tong, **C. Ma**, Y. Chi, "Low-Rank Matrix Recovery with Scaled Subgradient Methods: Fast and Robust Convergence Without the Condition Number," *IEEE Transactions on Signal Processing*, vol. 69, no. 3, pp. 2396-2409, 2021.

[J11] J. Fan, **C. Ma**, Y .Zhong (**alphabetical order**), "A Selective Overview of Deep Learning," *Statistical Science*, vol. 36, no. 2, pp. 264-290, May 2021 (**invited overview article**).

[J12] T. Tong, **C. Ma**, Y. Chi, "Accelerating Ill-Conditioned Low-Rank Matrix Estimation via Scaled Gradient Descent," *Journal of Machine Learning Reseasrch*, vol. 22, no. 150, pp. 1-63, May 2021.

[J13] Y. Li, **C. Ma**, Y. Chen, Y. Chi, "Nonconvex Matrix Factorization from Rank-One Measurements," *IEEE Transactions on Information Theory*, vol. 67, no. 3, pp. 1928-1950, March 2021.

[J14] **C. Ma**, Y. Li, Y. Chi, "Beyond Procrustes: Balancing-Free Gradient Descent for Asymmetric Low-Rank Matrix Sensing," *IEEE Transactions on Signal Processing*, vol. 69, pp. 867-877, Jan. 2021.

[J15] **C. Ma**, J. Lu, H. Liu, "Inter-Subject Analysis: A Partial Gaussian Graphical Model Approach," *Journal of the American Statistical Association*, vol. 116, no. 534, pp. 746-755, 2021.

[J16] Y. Chen, Y. Chi, J. Fan, **C. Ma**, Y. Yan (**alphabetical order**), "Noisy Matrix Completion: Understanding Statistical Guarantees for Convex Relaxation via Nonconvex Optimization," *SIAM Journal on Optimization*, vol. 30, no. 4, pp. 3098–3121, Oct. 2020.

[J17] **C. Ma**, K. Wang, Y. Chi, Y. Chen, "Implicit Regularization in Nonconvex Statistical Estimation: Gradient Descent Converges Linearly for Phase Retrieval, Matrix Completion, and Blind Deconvolution," *Foundations of Computational Mathematics*, vol. 20, no. 3, pp. 451-632, June 2020 (**2024 SIAM Activity Group on Imaging Science Best Paper Prize**).

[J18] Y. Chen, Y. Chi, J. Fan, **C. Ma** (**alphabetical order**), "Gradient Descent with Random Initialization: Fast Global Convergence for Nonconvex Phase Retrieval," *Mathematical Programming*, vol. 176, no. 1-2, pp. 5-37, July 2019.

[J19] Y. Chen, J. Fan, **C. Ma**, Y. Yan (**alphabetical order**), "Inference and Uncertainty Quantification for Noisy Matrix Completion," *Proceedings of the National Academy of Sciences (PNAS)* vol. 116, no. 46, pp. 22931-22937, Nov. 2019 (direct submission).

[J20] Y. Chen, J. Fan, **C. Ma**, K. Wang (**alphabetical order**), "Spectral Method and Regularized MLE Are Both Optimal for Top-K Ranking," *Annals of Statistics*, vol. 47, no. 4, pp. 2204-2235, Aug. 2019.

[J21] J. Zhang, J. Tang, **C. Ma**, H. Tong, Y. Jing, J. Li, W. Luyten, M-F. Moens, "Fast and Flexible Top-k Similarity Search on Large Networks," *ACM Transactions on Information Systems*, vol. 36, no. 2, pp. 13:1-13:30, Sept. 2017.

Conference papers

[C1] J. Ge, S. Tang, J. Fan, **C. Ma**, C. Jin, "Maximum Likelihood Estimation is All You Need for Well-Specified Covariate Shift," *International Conference on Learning Representations (ICLR)*, 2024.

[C2] Y. Gui, R. F. Barber, **C. Ma**, "Conformalized Matrix Completion," *Conference on Neural Information Processing Systems (Neurips)*, 2023.

[C3] X. Xu, Y. Shen, Y. Chi, **C. Ma**, "The Power of Preconditioning in Overparameterized Low-Rank Matrix Sensing," *International Conference on Machine Learning (ICML)*, 2023.

[C4] I. Uchendu, T. Xiao, Y. Lu, B. Zhu, M. Yan, J. Simon, M. Bennice, C. Fu, **C. Ma**, J. Jiao, S. Levine, K. Hausman, "Jump-Start Reinforcement Learning," *International Conference on Machine Learning (ICML)*, 2023.

[C5] Y. Yang, **C. Ma**, "$O(T^{-1})$ Convergence of Optimistic-Follow-the-Regularized-Leader in Two-Player Zero-Sum Markov Games," *International Conference on Learning Representations (ICLR)*, 2023.

[C6] R. Pathak, C. Ma, M. J. Wainwright, "A New Similarity Measure for Covariate Shift with Applications to Nonparametric Regression", *International Conference on Machine Learning (ICML)*, 2022 (long presentation).

[C7] G. Li, **C. Ma**, N. Srebro (**alphabetical order**), "Pessimism for Offline Linear Contextual Bandits using ℓ_p Confidence Sets, " *Conference on Neural Information Processing Systems (Neurips)*, 2022.

[C8] P. Rashidinejad, B. Zhu, **C. Ma**, J. Jiao, S. Russel, "Bridging Offline Reinforcement Learning and Imitation Learning: A Tale of Pessimism," *Conference on Neural Information Processing Systems (Neurips)*, 2021.

[C9] **C. Ma**, Y. Li, Y. Chi, "Beyond Procrustes: Balancing-Free Gradient Descent for Asymmetric Low-Rank Matrix Sensing," *Asilomar Conference on Signals, Systems and Computers*, Nov. 2019.

[C10] Y. Li, **C. Ma**, Y. Chen, Y. Chi, "Nonconvex Matrix Factorization from Rank-One Measurements," *International Conference on Artificial Intelligence and Statistics (AISTATS)*, Apr. 2019.

[C11] **C. Ma**, K. Wang, Y. Chi, Y. Chen, "Implicit Regularization in Nonconvex Statistical Estimation: Gradient Descent Converges Linearly for Phase Retrieval and Matrix Completion," *International Conference on Machine Learning (ICML)*, July 2018.

[C12] J. Zhang, J. Tang, **C. Ma**, H. Tong, Y. Jing, J. Li, "Panther: Fast Top-k Similarity Search on Large Networks," *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, Aug. 2015.

TUTORIALS

[Tut1] "Overview of classical ML: classification methods and decision trees," Data-Driven Materials Informatics Statistical Methods and Mathematical Analysis, Institute for Mathematical and Statistical Innovation (IMSI), 2024.

[Tut2] "Nonconvex Optimization for High-Dimensional Signal Estimation: Spectral and Iterative Methods," European Signal Processing Conference (EUSIPCO) 2020, together with Y. Chen and Y. Chi.

INVITED TALKS

[T1] University of Minnesota Twin Cities, Statistics Seminar, Feb. 2024.

[T2] Purdue University, Quantitative Methods Seminar, Dec. 2023.

[T3] UChicago, Big Data and Machine Learning in Econometrics, Finance, and Statistics, Oct. 2023.

[T4] Allerton Conference on Communication, Control, and Computing, Sept. 2023.

[T5] IDEAL Fall 2023 Kick-Off, Sept. 2023.

[T6] IDEAL Annual Meeting, June 2023.

[T7] University of Bristol, Statistics, Seminar, Nov. 2022.

[T8] SIAM Conference on Mathematics of Data Science, Sept. 2022.

[T9] Northwestern University, Industrial Engineering & Management Sciences Seminars, Feb. 2022.

[T10] UIC, Statistics and Data Science Seminar, Dec. 2021.

[T11] TTIC, Machine Learning Seminar, Oct. 2021.

[T12] Informs Annual Meeting, Oct. 2021.

[T13] UC Berkeley, FODSI Seminar, Apr. 2021.

[T14] UC Berkeley, BLISS Seminar, Nov. 2020.

[T15] ETH, Young Data Science Researcher Seminar, June 2020.

[T16] Georgia Tech, ISyE Seminar, School of Industrial & Systems Engineering, Feb. 2020.

[T17] Columbia University, IEOR Department Seminar, Feb. 2020.

[T18] UPenn, Department of Statistics, Feb. 2020.

[T19] UC Irvine, Department of Statistics, Feb. 2020.

[T20] UC Davis, Department of Statistics, Feb. 2020.

[T21] Rutgers University, Department of Statistics, Jan. 2020.

[T22] University of Chicago, Statistics Colloquium, Jan. 2020.

[T23] USC, Department of Data Sciences and Operations, Jan. 2020.

[T24] Purdue University, School of Industrial Engineering, Dec. 2019.

[T25] University of Maryland, College Park, Statistics Seminar, Dec. 2019.

[T26] AMS Fall Eastern Sectional Meeting, Oct. 2019.

[T27] Cornell University, Annual Young Researchers Workshop, Oct. 2019.

[T28] NYU, Mathematics, Information and Computation (MIC) Seminar, Apr. 2019.

[T29] Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, Oct. 2018.

[T30] CMU, Energy and Information Systems Seminar, July 2018.

[T31] ICML Workshop on Modern Trends in Nonconvex Optimization for Machine Learning, July 2018.

[T32] International Conference on Machine Learning (ICML), Stockholm, July 2018.

[T33] International Symposium on Mathematical Programming, Bordeaux, July 2018.

[T34] Tsinghua University, Department of Electrical Engineering, Jan. 2018.

[T35] International Conference on Data Science, Shanghai, Dec. 2017.

[T36] ICSA Applied Statistics Symposium, Chicago, June 2017.

Professional Service

1. Co-organizer of *Big Data and Machine Learning in Econometrics, Finance, and Statistics*, Stevanovich Center for Financial Mathematics, University of Chicago Oct. 2023
2. Co-organizer of *Statistics Colloquium*, Department of Statistics, University of Chicago Sept. 2022 – present
3. Co-organizer of *Wilks statistics seminar*, ORFE, Princeton University July 2018 – May 2019
4. Co-organizer of 6^{th} *Princeton Day of Statistics*, ORFE, Princeton University
5. Member of Technical Program Committee for *52nd Annual Conference on Information Sciences and Systems* (*CISS 2018*)
6. Reviewer for the following journals: *Annals of Statistics*, *Journal of the American Statistical Association*, *Statistical Science*, *Bernoulli Journal*, *Biometrika*, *Electronic Journal of Statistics*, *Journal of Computational and Graphical Statistics*, *Operations Research*, *Mathematics of Operations Research*, *Proceedings of the IEEE*, *Journal of Machine Learning Research*, *IEEE Transactions on Signal Processing*, *IEEE Transactions on Information Theory*, *Journal of Business & Economic Statistics*, *Transactions on Knowledge and Data Engineering*, *SIAM Journal on Mathematics of Data Science*, *Mathematical Programming*
7. Reviewer for the following conferences: *ACM Symposium on Theory of Computing (STOC)*, *Conference on Neural Information Processing Systems* (*NeurIPS*), *International Conference on Artificial Intelligence and Statistics*, *IEEE International Symposium on Information Theory*, *Annual Conference on Information Sciences and Systems*, *International Conference on Acoustics, Speech, and Signal Processing* (*ICASSP*)

Teaching

1. Topics in Learning Under Distribution Shifts, Winter 2023
2. Introduction to Probabilistic Models, Winter 2022, 2023, 2024
3. Topics in Mathematical Data Science: Spectral Methods and Nonconvex Optimization, Fall 2021, Winter 2024

Awards and Honors

1. 2024 SIAM Activity Group on Imaging Science Best Paper Prize *2024*
   This prize is awarded every two years to the author(s) of the most outstanding paper on mathematical and computational aspects of imaging published within the four calendar years preceding the year prior to the award year. Our paper on "Implicit Regularization in Nonconvex Statistical Estimation: Gradient Descent Converges Linearly for Phase Retrieval, Matrix Completion, and Blind Deconvolution" published on Foundations of Computational Mathematics is recognized "for the deep understanding and analysis of the interaction between optimization algorithms and the geometry of the landscape in which they operate."
2. Hannan Graduate Student Travel Award, Institute of Mathematical Statistics *2020*
3. School of Engineering and Applied Science Award for Excellence, Princeton University *2019*
   This award is given to SEAS advanced graduate students who have performed at the highest level as scholars and researchers.
4. AI Labs Fellowship, Hudson River Trading *2019*
5. Best Poster Award, Princeton Day of Optimization *2018*
6. ICSA Student Paper Award, International Chinese Statistical Association *2017*
7. First Year Engineering Fellowship, Princeton University *2015*
8. Outstanding Academic Performance Scholarship, Tsinghua University *2014*
9. National Scholarship, Tsinghua University (Highest honor, 2 out of 118 in department) *2013*
10. CNPC Scholarship, Tsinghua University (Highest honor, only recipient in department) *2012*