Time Series and Machine Learning Reading Group 2023 Autumn/Winter

We will read some of the following references on deep nerual networks theories and relevant techniques, casual inference, and time series:

References

Schmidt-Hieber, J. (2020). Nonparametric regression using deep neural networks with ReLU activation function. *The Annals of Statistics*, **48**, 1875–1897.

Bauer, B., and Kohler, M. (2019). On deep learning as a remedy for the curse of dimensionality in nonparametric regression. *The Annals of Statistics*, **47**, 2261-2285.

Kohler, M., and Langer, S. (2021). On the rate of convergence of fully connected deep neural network regression estimates. *The Annals of Statistics*, **49**, 2231–2249.

Song, S., Wang, T., Shen, G., Lin, Y., and Huang, J. (2023). Wasserstein generative regression. *arXiv preprint*, arXiv:2306.15163.

He, X., Pan, X., Tan, K. M., and Zhou, W. X. (2021). Smoothed quantile regression with large-scale inference. *Journal of Econometrics.*, **232**, 367–388.

Koltchinskii, V. (2006). Local Rademacher complexities and oracle inequalities in risk minimization. *The Annals of Statistics*, **34**, 2593–2656.

Koltchinskii, V., and Panchenko, D. (2000). Rademacher processes and bounding the risk of function learning. In *High Dimensional Probability II*, Springer, 443–457.

Mendelson, S. (2018). Learning without concentration for general loss functions. *Probability Theory and Related Fields*, **171**, 459–502.

Duan, Y., and Wang, K. (2023). Adaptive and robust multi-task learning. *The Annals of Statistics*, forthcoming.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal*, C1–C68.

Athey, S., and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, **113**, 7353–7360.

Semenova, V., Goldman, M., Chernozhukov, V., and Taddy, M. (2023). Inference on heterogeneous treatment effects in high-dimensional dynamic panels under weak dependence. *Quantitative Economics*, **14**, 471–510.

We will also read some of the following non-technical papers on state-of-the-art deep learning methodology and time series forecasting:

References

Zeng, A., Chen, M., Zhang, L., and Xu, Q. (2023). Are transformers effective for time series forecasting? In *Proceedings of the AAAI conference on artificial intelligence*, 2023.

Elsayed, S., Thyssens, D., Rashed, A., Jomaa, H. S., and Schmidt-Thieme, L. (2021). Do we really need deep learning models for time series forecasting? *arXiv preprint*, arXiv:2101.02118.

Li, L., Yan, J., Wang, H., and Jin, Y (2020). Anomaly detection of time series with smoothness-inducing sequential variational auto-encoder. *IEEE transactions on neural networks and learning systems*, **32**, 1177–1191.

Ataee Tarzanagh, D., Li, Y., Thrampoulidis, C., and Oymak, S. (2023). Transformers as Support Vector Machines. *arXiv preprint*, arXiv:2308.16898.

Hartford, J., Lewis, G., Leyton-Brown, K., and Taddy, M. (2017). Deep IV: A flexible approach for counterfactual prediction. In *International Conference on Machine Learning.*