Stochastics in high dimensions

S2F2 Hauptseminar Stochastische Prozesse, S4F3 Graduate Seminar Applied Probability (WS 2024/25, Thursdays 14 ct)

This seminar will consider topics in Probability theory and Statistics in high dimension with applications to data science. It requires only a background in measure-theoretic probability. A prior knowledge of statistics is not necessary. There will be both more basic and more advanced topics available, so that the seminar is suitable both for third year Bachelor and for Master students.

Main references:

- Vershynin: Four lectures on probabilistic methods in data science. https://arxiv.org/abs/1612.06661
- Rigollet, Hütter: High dimensional statistics. https://arxiv.org/abs/2310.19244
- Misiakiewisz, Montanari: Six lectures on linearized neural networks. https://arxiv.org/abs/2308.13431
- Tropp: High dimensional probability. https://tropp.caltech.edu/notes/Tro21-Probability-High-LN-corr.pdf

Additional references are the books “High-Dimensional Statistics” by M.J. Wainwright

https://people.eecs.berkeley.edu/~wainwrig/BibPapers/Wai19.pdf

and “High-Dimensional Probability: An introduction with applications to data science” by R. Vershynin

https://www.math.uci.edu/~rvershyn/papers/HDP-book/HDP-book.pdf

		Thema	Name	E-Mail	Datum
		High dimensional probability (Vershynin)			
		Concentration of sums of independent r.v., Johnson-Lindenstrauss			
		Concentration of sums of indep. random matrices, community recovery			
		Covariance estimation and matrix completion			
		Matrix deviation inequality			
		*Poincaré and log Sobolev inequalities			
		High dimensionsal statistics (Rigollet, Hütter)			
		Linear regression model			
		High dimensional linear regression			
		Misspecified linear models			
		Minimax lower bounds			
		Multivariate regression models			
		Principal component analysis			
		*Graphical models			
		Linearized neural networks (Misiakiewisz, Montanari)			
		Linear regression under feature concentration assumptions			
		Kernel ridge regression in high dimension			
		Random features			
		Neural tangent features			
		Why stop being lazy (and how)			