universität wien
Fakultät für Mathematik

VORTRÄGE

Tenure Track „Stochastic Methods for Data Science“

Dienstag, 12.1.2021, Zoom-Meeting

Nikita Zhivotovskiy
(Google Research Zürich)

https://zoom.us/j/91200487889?pwd=Ny9YMytJbW5lTUhZakhSZjNVbVYwZz09

14:00 Uhr – 14:20 Uhr: Didaktischer Vortrag

„Chernoff bound, Hoeffding’s inequality and moments of the sum of independent random variables“

In this lecture, we are discussing several basic results relating moment generating functions, exponential tail bounds, and moments of the sums of independent random variables. First, we introduce the moment generating function and prove a version of Hoeffding’s lemma for bounded random variables. Then, we discuss the Chernoff method and show how our bound on the moment generating function implies a concentration inequality for the sum of independent, bounded random variables. Finally, we show how similar tail inequalities imply the bound on the moments of this sum.

14:50 Uhr – 15:35 Uhr: Wissenschaftlicher Vortrag

„Quantization of probability distributions under minimal moment assumptions“

In this talk we discuss the problem of quantization of a probability distribution, sometimes referred to as the k-means clustering problem, where based on n independent observations sampled according to the underlying distribution, one is aiming to find k centers minimizing the distortion. We only assume the existence of the two moments of the underlying distribution and work in general separable Hilbert spaces. Since we consider potentially heavy-tailed distributions, our estimation methods should be robust to the presence of outliers in the set of observations. We begin our talk with a short survey of the field. In the context of k-means clustering, we present the median of means based estimators together with corresponding non-asymptotic distortion bounds. Our results extend the renowned asymptotic result of David Pollard who showed that the existence of two moments is necessary and sufficient for strong consistency of an empirically optimal quantizer in R^d. In a special case of clustering in R^d, under two bounded moments, we show matching non-asymptotic upper and lower bounds on the distortion, which depend on the probability mass of the lightest cluster of an optimal quantizer.

Mittwoch, 13.1.2021, Zoom-Meeting

Felix Krahmer
(TU München)

https://zoom.us/j/93618316488?pwd=Y0VHV1JaR09HdVNBOTEwWEVBTzJjQT09

14:00 Uhr – 14:20 Uhr: Didaktischer Vortrag

„The covering argument“

In this short lecture we prove that the norm of a matrix can be estimated just by its action on a finite set of vectors. This estimate is of great importance for random matrix theory and many problems in data science.

14:50 Uhr – 15:35 Uhr: Wissenschaftlicher Vortrag

„Structure and Randomness in Signal Processing“

In this talk, we will discuss various examples how randomized measurement designs for signal processing applications yield improved performance with provable recovery guarantees. We will show how these scenarios give rise to problems of independent interest at the interface of high-dimensional probability and statistical learning theory, whose solutions then help to advance the understanding of the measurement systems.
Firstly, motivated by imaging applications, we discuss the problem of sparse recovery from subsampled random convolutions. We advance techniques related to the theory of empirical processes to establish near-optimal recovery guarantees.
Secondly, we present recent results about the geometry of random polytopes generated by heavy-tailed random vectors and discuss their implications for noise-blind compressed sensing.
Lastly, motivated by applications in wireless communication, we establish local instabilities arising for convex regularizations of the randomized blind deconvolution problem.
These are joint works with the speaker's PhD students Christian Kümmerle and Dominik Stöger, well as with Olivier Guédon (Université de Paris Est), Shahar Mendelson (Australian National University) and Holger Rauhut (RWTH Aachen University).

Donnerstag, 14.1.2021, Zoom-Meeting

Yura Malitsky
(EPFL)

https://zoom.us/j/94381358170?pwd=RHhhU3dnL0lLb2R4bkpjc1dqbzE3QT09

14:00 Uhr – 14:20 Uhr: Didaktischer Vortrag

„Stochastic gradient descent“

In this short lecture, we will try to understand what stochastic gradient descent is, why do we need it, and what open questions still exist in this area. Prerequisites require calculus and basic numerical methods.

14:50 Uhr – 15:35 Uhr: Wissenschaftlicher Vortrag

„Variance reduction for saddle point problems“

Variance reduction was transformative for stochastic methods in the last decade. However, most results have been obtained only for finite sum minimization and not much has been done for saddle point problems. In this talk, we will present our recent progress in this direction. We will show how to achieve the same improvement with variance reduction as in the pure minimization case.

Dienstag, 19.1.2021, Zoom-Meeting

Lukas Gonon
(LMU München)

https://zoom.us/j/92444681556?pwd=ZWNETnA0OGVTRUdjMUVya1UxS25lUT09

14:00 Uhr – 14:20 Uhr: Didaktischer Vortrag

„Introduction to Rademacher complexities“

This lecture introduces the concept of Rademacher complexity, which is frequently used in machine learning as a means to measure the richness or capacity of function classes in a given learning problem. By proving bounds on Rademacher complexities of affine function classes and by outlining the relevance of these bounds for classes of neural networks, we show how Rademacher complexities are employed and why they are useful.

14:50 Uhr – 15:35 Uhr: Wissenschaftlicher Vortrag

„Dynamic learning for stochastic processes: neural networks, reservoir computing systems and applications to mathematical finance“

In this talk I present recent mathematical results regarding different dynamic learning techniques for stochastic processes. I will also elaborate on the central role that Rademacher complexities play in the proofs of these results.
The talk begins with a discussion of our joint research on deep neural network expression rates for option prices in high-dimensional exponential Lévy models. We show that under mild growth conditions on the Lévy triplets deep neural networks are able to approximate option prices without the curse of dimensionality. The talk then continues with results from our joint research on approximation and learning based on random recurrent neural networks. We obtain high-probability bounds on the approximation error in terms of the network parameters and generalization error bounds for weakly dependent input data. Notably, these results imply a universal approximation theorem for random recurrent neural networks. In closing, I will elaborate on further research directions.

Mittwoch, 20.1.2021, Zoom-Meeting

Elizaveta Rebrova
(UCLA)

https://zoom.us/j/99861337985?pwd=bGhiSEZGREgrNjByUytURnU5M205dz09

17:00 Uhr – 17:20 Uhr: Didaktischer Vortrag

„Solving linear systems: optimization perspective“

Suppose we aim to solve a system Ax = b where the matrix A is square and full rank. Then, as we know from linear algebra, it has a unique solution given by x* = A^{-1}b. However, as soon as we allow rectangular or rank deficient matrices in the left hand side, the problem becomes more interesting. For example, when the system is overdetermined and exact solution does not exist, we could search for "the best" approximation for the solution by solving an optimization (so-called, least-squares) problem, and if the system is underdetermined and there are multiple solutions, we could search for "the best" solution among them. In this class, we will formalize what could "the best" mean in the previous sentence, and discuss several (direct and iterative) optimization-based methods for solving linear systems.

17:50 Uhr – 18:35 Uhr: Wissenschaftlicher Vortrag

„Random matrices, low-rank tensors and beyond: using high-dimensional probability to study complex data“

It is not a secret that probabilistic view in general and random matrix theory in particular present amazing tools to understand, process and learn from large high-dimensional data. However, in many cases, one has to go beyond "simple" matrix models to correctly represent and treat the data. For example, inherently multimodal data is better represented with a tensor, that is, higher-order generalization of a matrix. Transition to more advanced data structures sometimes can survive re-using old algorithms, however, the development of the special tools that honor the full structure within the data pays off by making the algorithms both much more efficient and better interpretable. Simultaneously, it presents many interesting and challenging non-trivialities from the math point of view. In this talk, I will focus on our new provable methods for modewise (that is, structure preserving) tensor dimension reduction. I will also discuss the connections to interpretable learning from multi-modal data through tensor decompositions, and to our new randomized algorithms for solving linear systems with corrupted equations.

Donnerstag, 21.1.2021, Zoom-Meeting

Lyudmilla Grigoryeva
(Universität Konstanz)

https://zoom.us/j/99888853039?pwd=TjFWMFBGNDRiQkNkeUxCdFhWWnpRUT09

14:00 Uhr – 14:20 Uhr: Didaktischer Vortrag

„Prediction with Experts Advice“

In this lecture we consider a particular example of online learning methods which gain increasing attention in machine learning applications. I will introduce the so-called online prediction with expert advice, a method which consists in constructing the online forecasting strategy for the player/learner combining the predictions from different experts. The online data processing makes this method computationally cheap and easy to implement. Moreover, the available theoretical guarantees make it a method of choice for reliable learning.

14:50 Uhr – 15:35 Uhr: Wissenschaftlicher Vortrag

„Learning Dynamic Processes with Reservoir Computing“

Many dynamical problems in engineering (including financial), control theory, signal processing, time series analysis and forecasting can be described using input/output (IO) systems. State-space systems are known to provide a parsimonious and computationally efficient way to model the relation between time evolving explanatory variables (the input) and a collection of dependent or explained variables of interest (the output). Whenever a true functional IO relation cannot be derived from first principles, various classes of state-space systems can be used as universal approximants. We shall show that particular families of such state-space systems, the so-called Reservoir Computing (RC) systems, with extremely simple and easy to implement architectures, enjoy universal approximation properties which have been proved in different setups. The defining feature of RC systems is the fact that some their components (usually the state map) are randomly generated and the observation equation has an easily tractable form. From the machine learning perspective, RC systems can be seen as recurrent neural networks with randomly generated and non-trainable weights and a simple-to-train readout layer (often a linear map). RC systems serve as efficient, randomized, online computational tools for dynamic processes and enjoy generalization properties which can be explicitly

derived. We will make a general introduction into up-to-date theoretical developments, will discuss connections with research contributions in other fields, and will address details of applications of RC systems for data processing.