Winter school
"New frontiers in high-dimensional probability and statistics 2"

The school will be held at HSE University (20 Myasnitskaya street, Moscow) from February, 22 to February, 23, 2019 in the room 101.

Preliminary program (v. 5.02.2019)

February 22

	Speaker	Title
9:50-10:00	Welcome	
10:00-11:10	Cristina Butucea	Mini-course: Local differential privacy and adaptive estimation
11:10-11:40	**Coffee break**	
11:40-12:50	Zuoqiang Shi	Mini-course: PDE-based methods in interpolation on high-dimensional point cloud
12:50-14:10	**Lunch**	
14:10-14:30	Vladimir Spokoiny, Maxim Panov	SLT Master Program presentation
14:30-15:40	Karim Lounici	Mini-course: PCA: comparison of performances in the offline and online setting
15:40-16:00	**Tea break**	
16:00-16:25	Max Kaledin	Approximate Dynamic Programming with Approximation of Transition Density
16:30-16:55	Leonid Iosipoi	MCMC estimation for distributions with known characteristic function
16:55-17:25	**Coffee break**	
17:25-17:50	Sergey Samsonov	Concentration inequalities for functionals of Markov Chains with applications to variance reduction
17:55-18:20	Nikita Gladkov	Variants of the Monge-Kantorovich problem for the cost function xyz
18:25-18:50	Alexey Kroshnin	TBA
19:00 –	**Welcome party at Professor's club**	

February 23

	Speaker	Title
10:00-11:10	Zuoqiang Shi	Mini-course: PDE-based methods in interpolation on high-dimensional point cloud
11:10-11:40	**Coffee break**	
11:40-12:50	Cristina Butucea	Mini-course: Local differential privacy and adaptive estimation
12:50-14:10	**Lunch**	
14:10-15:20	Karim Lounici	Mini-course: PCA: comparison of performances in the offline and online setting
15:20-15:40	**Tea break**	
15:40-16:05	Valentina Shumovskaia	Hypothesis Testing for Random Graphs with a Low-Rank Structure
16:10-16:35	Franz Besold	Minimax clustering with adaptive weights
16:35-17:00	**Coffee break**	
17:00-17:25	Nikita Puchkin	Manifold estimation from noisy observations
17:30-17:55	Darina Dvinskikh	Complexity bounds for optimal distributed primal and dual methods for finite sum minimization problems
18:00 –	**The end**	

Abstracts

Cristina Butucea (Universite Paris-Est Marne-la-Vallee),
e-mail: cristina.butucea@univ-mlv.fr

Local differential privacy and adaptive estimation

In a privacy context, the data cannot be released as observed but need to be privatized in some way and only the modified data are available for later inference. We address the problem of nonparametric density estimation under the additional constraint of α-local differential privacy.

After discussing the recent statistical literature, we give the rates of convergence over Besov spaces under mean integrated L^r-risk. We show an optimal way to privatize data and an optimal estimator of the probability density - both adaptive of the smoothness. The rates of convergence in this case are deteriorated compared to the standard setup without privacy but reveal an analogous elbow effect (change of regime).

Karim Lounici (Ecole Polytechnique),
e-mail: karim.lounici@polytechnique.edu

PCA: a comparison of performances in the offline and online setting.

Principal Component Analysis (PCA) is a popular technique of reduction of dimension used in virtually everything of sciences that consists essentially in the analysis of the covariance structure of a random process. Let X be a zero mean random vector with unknown covariance operator Σ. In the offline setting, the problem is to recover the spectral decomposition of Σ (eigenvalues and corresponding eigenspaces) based on the sample $X_1, \ldots, X_n$ of n i.i.d. copies of X. In the online setting, the observations are not available at once but rather arrive in a stream and the goal is to update our estimators any time a new observation is obtained. This presentation will concentrate on standard PCA for the offline setting and the Krasulina scheme for the online setting. We provide new non-asymptotic bounds for the performances of these two methods. The analysis of these two methods is completely different. Standard PCA bounds rely on the combination of perturbation theory with concentration inequalities. Krasulina's scheme is analyzed through martingales theory.

This presentation is based on joint works with Vladimir Koltchinskii and Jiangning Chen.

Zuoqiang Shi (Yau Mathematical Sciences Center),
e-mail: zqshi@tsinghua.edu.cn

PDE-based methods in interpolation on high-dimensional point cloud.

Point cloud is very powerful to model high dimensional data. Interpolation on point cloud is a fundamental mathematical problem in machine learning and data analysis. In this lecture, I will introduce several PDE-based models to do interpolation on high dimensional point cloud and show some applications in image processing and data analysis. I will reveal the close connections between PDEs and some deep neural networks. Theoretical analysis and numerical simulations show that PDEs provide a powerful tool to understand high dimensional data.

Franz Besold (Humboldt-Universitat zu Berlin),
e-mail: franz.besold@googlemail.com

Minimax clustering with adaptive weights

We discuss a nonparametric clustering algorithm called Adaptive Weights Clustering (AWC) suggested by Kirill Efimov, Larisa Adamyan and Vladimir Spokoiny (2017). The main idea of the approach is to use a likelihood ratio test of "no gap" to dynamically update local weights, where the scale of locality increases during the procedure. Numerical results show a state-of-the-art performance of the new method and adaptivity to unknown features of the data, such as the number of clusters and clusters with different shape, size or density. Theoretical results ensure local propagation of homogeneous regions as well as optimal sensitivity in detecting gaps of lower density between clusters. We extend the theoretical study to the case where the data is lying on a smooth manifold with a bounded reach.

Darina Dvinskikh (WIAS),
e-mail: dviny.d@yandex.ru

Complexity bounds for optimal distributed primal and dual methods for finite sum minimization problems

We study the optimal complexity bounds of first-order methods for the problem of minimizing the finite sum of convex (strongly convex) functions. To address this problem, we appeal to distributed optimization and consider the network of multiple agents (processors) which cooperatively minimize the target function. We compare two different approaches toward the finding of the minimum depending on the functions properties in terms of oracle calls and communication rounds: primal approach and dual approach, based on the assumption of existence of the Fenchel-Legendre transform of the target function. In particular we propose new method (more simple than Lan et al.,2017, and Scaman et al.,2018) based on penalty functions for non-smoothed convex optimisation problems with stochastic oracle.

Nikita Gladkov (HSE),
e-mail: gladkovna@gmail.com, alekszm@gmail.com

Variants of the Monge-Kantorovich problem for the cost function xyz

We consider variants of the Monge-Kantorovich problem on the unit cube $[0,1]^3$ with the cost function $c(x,y,z) = xyz$. Even the classical (multimarginal) problem with uniform one-dimensional marginals turns out to be a non-trivial example with unexpected properties. We describe geometry of solutions in the context of existing structural results for the multimarginal problem and discuss related open questions. Further we discuss the primal and dual Monge-Kantorovich problems with 2-dimensional uniform marginals and explain their relations to discrete operations and Sierpinski-type fractals.

Leonid Iosipoi (HSE-Skoltech),
e-mail: iosipoileonid@gmail.com

MCMC estimation for distributions with known characteristic function

In this paper, we propose a novel approach on how to use Langevin type MCMC algorithms to sample from distributions when only a characteristic function is known analytically. The main idea of the proposed approach is to replace the gradient of the log-density by the gradient of the logarithm of the characteristic function. A subsequent application of the Parseval formula leads to an efficient algorithm for the computation of mean values with respect to analytically intractable densities. We illustrate our approach by several numerical examples including multivariate elliptically contoured stable distributions and pricing of multi-asset options.

Maxim Kaledin (HSE-Skoltech),
e-mail: maxkaledin@gmail.com

Approximate Dynamic Programming with Approximation of Transition Density

The theory of optimal stopping is concerned with the problem of choosing optimal time to take action in order to maximize the reward. Such problems naturally arise in statistics, economics and mathematical finance. The natural approach to solve optimal stopping is backward dynamic programming with Monte Carlo simulations. Despite regression methods are state-of-the-art and adopted by most of practitioners, they suffer from the curse of time discretization (CoTD): if one desires to approximate continious optimal stoppping with a discrete one, the overall complexity of dynamic programming becomes exponential in ϵ. In my talk I will present the algorithm influenced by the ideas of Rust, Broadie and Glasserman which does not have CoTD and possesses similar convergence rates to stochastic mesh methods.

Alexey Kroshnin (HSE, IITP),
e-mail:

TBA

TBA

Nikita Puchkin (HSE),
e-mail: npuchkin@gmail.com

Manifold estimation from noisy observations

We consider a problem of manifold learning when one observes a sample $Y_1, \dots, Y_n$, generated from an additive noise model $Y_i = X_i^* + \varepsilon_i$, where the points $X_1^*, \dots, X_n^*$ lie on a low-dimensional manifold and $\varepsilon_1, \dots, \varepsilon_n$ are independent zero-mean random errors. Our goal is to recover the points $X_1^*, \dots, X_n^*$. Such a problem finds plenty of applications in image analysis, biology, medicine, etc. However, classical methods of manifold learning, such as Isomap (Tenenbaum, Langford (2000)), LLE (Roweis, Saul (2000)), Laplacian eigenmaps (Belkin, Niyogi (2003)), along with methods of interpolation on manifolds (e.g. Aamari, Levrard (2017), Li, Shi, Sun (2015)) assume that the noise magnitude tends to zero as the number of observations tends to infinity. Another approach (Genovese et. al. (2012)) is based on a singular deconvolution and, though does not have the previously mentioned drawback, assumes a known noise distribution. We address both these issues and propose a new scalable procedure, which provably recovers the underlying manifold under mild assumptions.

Sergey Samsonov (HSE-Skoltech),
e-mail: samsonovsergey97@mail.ru

Concentration inequalities for functionals of Markov Chains with applications to variance reduction

TBA

Valentina Shumovskaia (HSE-Skoltech),
e-mail: valentina.shumovskaia@skoltech.ru

Hypothesis Testing for Random Graphs with a Low-Rank Structure

A random graph has a low-rank structure if its probability matrix has a low rank. It is a more general structure for probability matrices that covers not only the stochastic block model (SBM) but also a mixed membership stochastic block model (MMSBM), on which community detection algorithms fail. Observing two graphs G_1 and G_2, which are generated from unknown probability matrices with a low rank, we would like to test between hypotheses $\mathcal{H}_0$: probability matrices coincide and $\mathcal{H}_1$: probability matrices differ. We propose an approach of such a problem solving which is both computationally inexpensive and applicable to graphs with probability matrices with general low-rank structure (not only SBM). Now there exist approaches that cover only one of those wishes (D. Ghoshdastidar et al, 2018; M. Tang et al, 2015).