



“Industrial mathematics” seminar at Chebyshev Laboratory

Tuesday, 8 June 2021, 15:30 (Moscow time, GMT+3)

Zoom ID: 870-6449-5461, password: ind

Geostatistical Learning: Challenges and Opportunities



Júlio Hoffmann

IMPA (Instituto de Matemática Pura e Aplicada), Brazil

Statistical learning theory provides the foundation to applied machine learning, and its various successful applications in computer vision, natural language processing and other scientific domains. The theory, however, does not take into account the unique challenges of performing statistical learning in geospatial settings. For instance, it is well known that model errors cannot be assumed to be independent and identically distributed in geospatial (a.k.a. regionalized) variables due to spatial correlation; and trends caused by geophysical processes lead to covariate shifts between the domain where the model was trained and the domain where it will be applied, which in turn harm the use of classical learning methodologies that rely on random samples of the data. In this work, we introduce the geostatistical (transfer) learning problem, and illustrate the challenges of learning from geospatial data by assessing widely-used methods for estimating generalization error of learning models, under covariate shift and spatial correlation. Experiments with synthetic Gaussian process data as well as with real data from geophysical surveys in New Zealand indicate that none of the methods are adequate for model selection in a geospatial context. We provide general guidelines regarding the choice of these methods in practice while new methods are being actively researched.

Everyone is welcome!

All information about past and future seminars can be found on our web-site:
<https://sites.google.com/view/industrial-math-seminar/main-en>