

Stochastic Gradient Methods

for students of Master Data Science

Stochastic gradient methods (SG) are used to solve non-linear optimisation problems with a special structure. Examples are non-linear least squares problems with a very large number of summands or optimisation problems in which the expected value of a target variable over a large data-set is minimised. Prominent applications can be found in the area of machine learning, inverse problems or optimal design. In the seminar, the basics of stochastic gradient methods will first be worked out on the basis of the overview article [1]. In addition to convergence analysis, algorithmic aspects and complexity considerations, the connection to and comparison with deterministic gradient methods will also play a role. In line with the latter, the continuous stochastic gradient method (CSG) is examined [2], which is formally an SG method, but shares many properties of deterministic descent methods, such as those discussed in courses and text books on "Nonlinear Optimisation". In the sequel, contributions and papers from different application areas are possible.

If you have any further questions about the seminar or need more information, please feel free to contact me and/or Dr. Pflug by e-mail (michael.stingl@fau.de, lukas.pflug@fau.de).

Literature (examples)

[1] Léon Bottou, Frank E. Curtis, and Jorge Nocedal: Optimization Methods for Large-Scale Machine Learning, *SIAM Review* 60(2), 2018

[2] Lukas Pflug, Niclas Bernhardt, Max Grieshammer, and Michael Stingl: CSG: A new stochastic gradient method for the efficient solution of structural optimization problems with infinitely many states, *Structural and Multidisciplinary Optimization*, 61, 2020