Operator Theory and Machine Learning

Maths for Machine Learning: A syllabus under discussion

Ana Moura Santos[1],

[1] Instituto Superior Técnico

Linear algebra and probability/statistics and optimization are the mathematical pillars of machine learning. We represent numerical data as vectors, and matrices can represent both a table of data and a model to describe a process for generating data, similar to the dataset at hand. So, good models can be thought as simplified versions of the real (unknown) data-generating process, capturing aspects that are relevant and extracting hidden patterns in data. Many times, matrix decomposition is useful in machine learning, since they allow to interpret the data and find the "right angle" to approach a more efficient learning model.

The learning component of machine learning is "training" the model, which means we use the data available to optimize some parameters of the model with respect to a utility function that evaluates how well the model predicts the training data. Some training models can be seen as an approach analogous of climbing a hill to reach its peak, a maximum of a desired performance measure. We are using an optimization method. Often, we want to distinguish the signal in the data from the noise and we need to quantify what "noise" means. Or we want to have predictors to express a sort of uncertainty. In both cases, we use probability/statistics theory.

In this presentation, I will discuss the syllabus choices made in collaboration with colleagues for the first implementation of the course *Maths for Machine Learning* offered by the Department of Mathematics at Técnico Lisboa.

References

[1] Strang, G., *Linear Algebra and Learning from Data, Wellesley-Cambridge Press, 2019.*

[2] M. Deisenroth, A. Faisal, and C. Ong, *Mathematics for Machine Learning, Cambridge University Press, 2020.*