Statistical and Mathematical Methods for Artificial Intelligence a.a. 2024-25

Elena Loli Piccolomini

Outline

Teacher Module I:
Prof. Elena Loli Piccolomini (elena.loli@unibo.it)
Teacher module II : dott. Davide Evangelista (davide.evangelista5@unibo.it)

Learning outcomes At the end of the course, the student masters the basic mathematical and statistical methods needed to acquire skill in artifical intelligence foundations, theories and applications.

Course contents

Linear algebra for AI

- Linear Algebra basics: eigenvalues and eigenvectors.
- Matrix factorizations: Singular Value Decomposition and LU factorization.
- Numerical solution of linear systems and linear least squares.
- Singular Values Decomposition, Principal Component Analysis.

Multivariate analysis and optimization

- Multivariate and multivalued functions. Gradient, Jacobian, Hessian.
- The chain rule for the derivation of multivariate functions.
- The backtracking algorithm.
- Introduction to unconstrained optimization. Extrema of multivariate functions. Optimality conditions.
- Descent methods. Gradient type methods and Newton type methods.
- Basic concepts of stochastic optimization.
- Again on Regularization.

Course contents

Elements of probability and statistics

- Probability and Bayes theorem.
- Random variables. Continuous and discrete distributions of random variables. Normal and Poisson distributions.
- Independent and dependent variables. Covariance and correlation.
- Estimates: Maximum Likelihood and Maximum a Posteriori estimates.

Bibliography

- G. Thomas, Mathematics for Machine Learning
  http://gwthomas.github.io/docs/math4ml.pdf
- R. Johanson, Numerical Python, Apress
- Lessons of Evangelista Davide available at:
  https://devangelista2.github.io

Laboratory exercises and course material

- Some guided Laboratory Python lessons will be given. Laboratory homeworks will be assigned and they must be completed before the exam.
- The course material will be upoloaded at: `https://virtuale.unibo.it`.

Assessment methods

There is a SINGLE exam for the two modules.
It is mandatory to complete the homeworks assigned in the Laboratory lessons which must be uploaded on virtuale before the exam.
The exam consists in a written test and a brief oral discussion about the assigned homeworks, that take place on the same day.
The final score is the sum of:

1. the score of the written test (quiz) (maximum 21/30)
2. the score of the exercises (maximum 11/30)

If the score of the written part is less than 15 the exam must be repeated. If the student do not get the grade, BOTH parts of the exam must be repeated.

Office hours

- **Elena Loli Piccolomini:** Monday 11,00-13,00 or by email appointment: elena.loli@unibo.it
- **Davide Evangelista:** Wednsday 9,00 - 12,00 by email appointment: davide.evangelista5@unibo.it

Motivation

Why a course on mathematics and statistics in master degree on Artifical Intelligence?

- We introduce the main mathematical and statistical concepts to talk about the three main components of machine learning: data, models and learning.
- **Data** are represented as vectors: linear algebra is the setting for manipulating and using vectors.
- **Model**. A good model can used to predict what happen in the real world without performing real-world experiments. Optimization and probability are two different perspective for building models.
- **Learning** from available data is the basis of Machine Learning. Multivariate analysis and optimization are necessary tools for learning.

Motivation

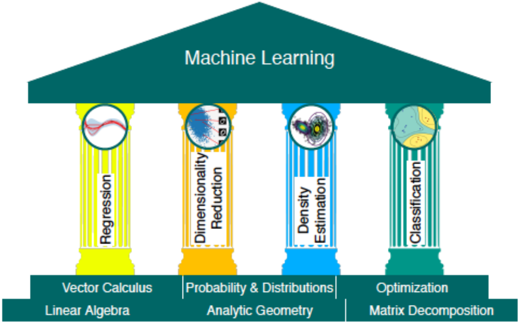
- Machine learning is about designing algorithms that automatically extract valuable information from data. The emphasis here is on "automatic", i.e., machine learning is concerned about general-purpose methodologies that can be applied to many datasets, while producing something that is meaningful.
- There are three concepts that are at the core of machine learning: data, a model, and learning.

Motivation

Why a course on mathematics and statistics in master degree on Artifical Intelligence?

- We introduce the main mathematical and statistical concepts to talk about the three main components of machine learning: data, models and learning.
- **Data** are represented as vectors: linear algebra is the setting for manipulating and using vectors.
- **Model**. A good model can be used to predict what happens in the real world without performing real-world experiments. Optimization and probability are two different perspective for building models.
- **Learning** from available data is the basis of Machine Learning. Multivariate analysis and optimization are necessary tools for learning.

Motivation


Machine Learning
Regression
Dimensionality Reduction
Density Estimation
Classification
Vector Calculus
Probability & Distributions
Optimization
Linear Algebra
Analytic Geometry
Matrix Decomposition

Data and linear algebra

- Since machine learning is inherently data driven, data is at the core data of machine learning.
- The goal of machine learning is to design general purpose methodologies to extract valuable patterns from data, ideally without much domain-specific expertise.
- We represent numerical data as vectors and represent a table of such data as a matrix. The study of vectors and matrices is called *linear algebra*

Models and probability

- A model is typically used to describe a process for generating data, similar to the dataset at hand.
- Therefore, good models can also be thought of as simplified versions of the real (unknown) data-generating process, capturing aspects that are relevant for modeling the data and extracting hidden patterns from it.
- A good model can then be used to predict what would happen in the real world without performing real-world experiments. We often consider data to be noisy observations of some true underlying signal.
- We often would also like to have predictors that allow us to express some sort of uncertainty, e.g., to quantify the confidence we have about the value of the prediction at a particular test data point.
- Quantification of uncertainty is the realm of *probability theory*

Learning and optimization

- The crux of the matter is the learning component of machine learning.
- Assume we are given a dataset and a suitable model. Training the model means to 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.
- Most training methods can be thought of as an approach analogous to climbing a hill to reach its peak. In this analogy, the peak of the hill corresponds to a maximum of some performance measures.

By summarizing...

- We represent data as vectors.
- We choose an appropriate model, either using the probabilistic or optimization view.
- We learn from available data by using numerical optimization methods with the aim that the model performs well on data not used for training.