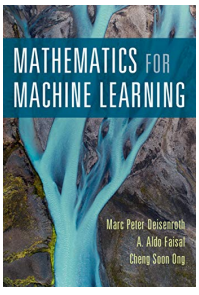
Lecture 1: Introduction

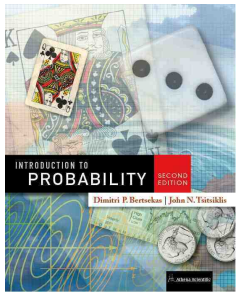
Yi, Yung (이융)

Mathematics for Machine Learning
https://yung-web.github.io/home/courses/mathml.html
KAIST EE

April 8, 2021

Textbook





- Mathematics for Machine Learning[1], Cambridge University Press, Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong
- Other books
  - Convex Optimization, Cambridge University Press, by Stephen Boyd and Lieven Vandenberghe
  - Introduction to Probability, 2nd edition, Athena Scientific, by Dimitri P. Bertsekas and John N. Tsitsiklis

[1] The entire textbook can be downloaded at https://mml-book.github.io/

Organization

- Part I: Math
  1. Linear Algebra
  2. Analytic Geometry
  3. Matrix Decomposition
  4. Vector Calculus
  5. Probability and Distributions
  6. Optimization
- Part II: 4 Basic Machine Learning Problems
  1. When Models Meet Data
  2. Dimensionality Reduction with Principal Component Analysis
  3. Density Estimation with Gaussian Mixture Models
  4. Classification with Support Vector Machines

Suggestions on Course Schedules

Total 16 weeks

- Part I: Math
  1. Linear Algebra (2 weeks)
  2. Analytic Geometry (1 week)
  3. Matrix Decomposition (1 week)
  4. Vector Calculus (1 week)
  5. Probability and Distributions (2 weeks)
  6. Optimization (2 weeks)
- Part II: 4 Basic Machine Learning Problems
  1. When Models Meet Data (1 week)
  2. Dimensionality Reduction with Principal Component Analysis (1 week)
  3. Density Estimation with Gaussian Mixture Models (1 week)
  4. Classification with Support Vector Machines (1 week)
- Total 13 weeks + Midterm (1 week) + Final (1 week) + Extra (1 week)

Target Audience

- Undergraduate
  - They may have partial backgrounds on the math (e.g., only vector calculus + linear algebra). Depending on the students' background, the amount of time for math can be adjusted.
  - Some mathematical parts may need to be provided with some degree of rigorous proofs.
- Graduate
  - Graduate students have already taken the basic math courses on linear algebra, vector calculus, probability, optimization, but they don't have almost no background on machine learning.
  - Math parts can be just reviewed by minimizing the proofs, but additional ML problems can be added to the course, so that they can have more exposure to the ML part.

File Organization

- In each chapter, there is a `main.tex` which you can compile.
- Common files for all chapters
  - `myhead.tex`: common headers, e.g., including necessary packages
  - `mydefault.tex`: default values of many latex environments
  - `mymacro.tex`: macros related to linear algebra, e.g., matrix, transpose, inverse, etc
  - `mymath.tex`: other misc. math macros
  - `compile.sh`: shell script which compiles and generate the pdfs of all chapters
  - `print.sh`: shell script which generates the pdfs of 2/1, 4/1 printed formats
- Just type "./compile.sh" if you want to get all the pdfs[2].

[2]Please make compile.sh and print.sh executable, if not, by typing `chmod u+x compile.sh`

Slide vs. Handout

- Handout

  `\documentclass[handout,fleqn,aspectratio=169]{beamer}`

- Slide

  `\documentclass[fleqn,aspectratio=169]{beamer}`

- Difference between Handout and Slide? If you want to use the functionality of "beamer overlay" to add animations to the slides, you need to compile without handout option. Please visit the following url if you are interested.
  `https://youtu.be/kkM_VPSM8kA`

Letter vs. A4

- A4 In the `myhead.tex` file:

```
\usepackage{pgfpages}
\pgfpagesuselayout{resize to}[a4paper,landscape,border shrink=5mm]
```

- Letter In the `myhead.tex` file:

```
\usepackage{pgfpages}
\pgfpagesuselayout{resize to}[letterpaper,landscape,border shrink=5mm]
```

Basic Notations

- Scalars: $a, b, c, \alpha, \beta, \gamma$
- Vectors: $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$
- Matrices: $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$
- Sets: $\mathcal{A}, \mathcal{B}, \mathcal{C}$
- (Ordered) tuple: $B = (\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3)$
- Matrix of column vectors: $\boldsymbol{B} = [\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3]$ or $\boldsymbol{B} = \begin{pmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \boldsymbol{b}_3 \end{pmatrix}$
- Set of vectors: $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$
- $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}$, $\mathbb{N}$, $\mathbb{R}^n$, etc
- Probability: We use both $p(\cdot)$, $\mathbb{P}[\cdot]$.

Enjoy!

When you modify the latex files for your convenience, if you have any question on macros or pdf generation, feel free to send an email to `yiyung@gmail.com`.