

# Resources

## Resources to Cover Background Material

### 1. Probability

#### I. Courses

- ✓ [Khan Academy lectures on probability](#)
- ✓ [Stat110](#) of Harvard University

#### II. Essential Topics

- ✓ Random variables
- ✓ Expectation
- ✓ Joint distributions
- ✓ Independence
- ✓ Conditional distributions
- ✓ Bayes rule
- ✓ Multivariate normal distribution

### 2. Linear Algebra

#### I. Courses

- ✓ <https://www.3blue1brown.com/essence-of-linear-algebra-page>
- ✓ [Khan Academy lectures on linear algebra](#)
- ✓ [MIT's course](#) as taught by Gilbert Strang

✓

### III. Essential Topics

- ✓ Norms and Inner Products [1](#)
- Orthogonal Vectors and Matrices [1](#) [2](#)
- ✓ Linear Independence [1](#) [2](#)
- ✓ Eigen Values and Eigen Decomposition [1](#) [2](#)
- ✓ Singular Value Decomposition [1](#)

## 3. Multi-variable Calculus I.

### Courses [Khan Academy](#) [Course](#)

- ✓ MIT [OCW Course](#)

### IV. Essential Topics

- ✓ Partial derivatives
- ✓ Gradients ✓ The chain rule

## 4. Programming

- ✓ Experience in any language but homeworks will be in python

## Python Resources

- ✓ [Jupyter Notebook](#) (all python coding assignments should be in this format)
- ✓ [Code Documentation in Matlab](#) (all matlab coding assignments should be in this format)

- ✓
- ✓ [Anaconda Python distribution](#)) (I recommend installing Python 2.7)
- ✓ [Official Python documentation](#)
- ✓ [Official Pytorch Documentation](#)
- ✓ [Google's Python class](#)
- ✓ [Another Python tutorial](#)
- ✓ [SciPy tutorial](#)
- ✓ [NumPy for Matlab users](#)
- ✓ [Python quick reference sheet](#)
- ✓ [scikit-learn](#)
- ✓ Python for Data Analysis, McKinney, 2012. ([Amazon](#))

## Other Useful Machine Learning Courses

- ✓ [Edx.org Course](#) by University of California, San Diego
- ✓ [Courera.org Course](#) as taught by Andrew Ng (You can cover this course in 3,4 weeks)
- ✓ [Machine learning](#) by Standford as taught by Andrew Ng
- ✓ [Machine learning](#) as taught by Hal Daumé
- ✓ [Machine learning](#) as taught by Yaser Abu-Mostafa (the Caltech course based on Learning from Data)
- ✓ [Talking machines](#): A surprisingly interesting podcast about machine learning
- ✓ [Winter 2017 441 / 841 Classification](#) by Ali Ghodsi

✓

## Papers

- ✓ [Introduction to Statistical Learning Theory](#), Bousquet, Boucheron, and Lugosi, 2004.
- ✓ [A global geometric framework for nonlinear dimensionality reduction](#), Tenenbaum, de Silva, and Langford, 2000.
- ✓ [Nonlinear dimensionality reduction by locally linear embedding](#), Roewis and Saul, 2000.
- [Learning the parts of objects by non-negative matrix factorization](#), Lee and Seung, 1999.
- ✓ [Computational methods for sparse solution of linear inverse problems](#), Tropp and Wright, 2010.
- ✓ [K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation](#), Aharon, Elad, and Bruckstein, 2006.
- ✓ [An overview of low-rank matrix recovery from incomplete observations](#), Davenport and Romberg, 2016.
- ✓ [Robust Principal Component Analysis?](#) Candes, Li, Ma, and Wright, 2011.

## Books on statistical learning

- ✓ Learning from Data, Abu-Mostafa, Magdon-Ismail, and Lin, 2012. This book is short and sweet. It doesn't cover everything we will talk about, but has a fantastic and very accessible overview of VC theory. (Amazon)
- ✓ The Elements of Statistical Learning, Hastie, Tibshirani, and Friedman, 2009. This book covers most of the material we will be covering in the class and is probably the best overall resource that is freely available on the internet. ([Free online version](#), Amazon)
- ✓ Machine Learning: A Probabilistic Perspective, Murhpy. A close second to Hastie et al., this book is also great and is an excellent resource with some material covering more modern topics such as deep learning. (Amazon)

✓

- ✓ Pattern Classification, Duda, Hart, and Stork, 2000. Another good overview of much of what we will cover in this course. This book was another close second to Hastie et al., but it isn't available online. (Amazon)
- ✓ Pattern Recognition and Machine Learning, Bishop, 2006. This is a classic. A good introduction to machine learning from a Bayesian perspective. I view Murphy to be a slightly more modern update, but for a lot of classic material this is still my go-to resource. (Amazon)
- ✓ A Probabilistic Theory of Pattern Recognition, Devroye, Györfi, and Lugosi, 1996. A (somewhat intense) book that considers the more theoretical aspects of statistical learning. A great resource if you want to know the details. (Amazon)
- ✓ Learning with Kernels, Scholkopf and Smola, 2001. The book on support vector machines and related kernel methods. (Amazon)

Reinforcement Learning: An Introduction, Sutton and Barto, 1998. We will focus mostly on supervised and unsupervised learning, but this book is a good introduction to the key ideas in reinforcement learning. (Amazon)

- ✓ The Signal and the Noise, Silver, 2012. A higher level book that examines several applications of the kinds of statistical learning techniques, including weather, earthquakes, athletic performance, and more. (Amazon)
- ✓ The Master Algorithm, Domingos, 2015. A popular introduction to machine learning that also hits on some portions of the machine learning community that we will not emphasize much in this course. (Amazon)

## Books covering background material

- ✓ How to Solve It, Polya, 1945. Classic introduction to mathematical problem solving. (You can find in Lincon's Inn Corner, 2nd Floor)(Amazon)
- ✓ Introduction to Probability, Bertsekas and Tsitsiklis, 2008. Good introduction to elementary probability theory. (Amazon)

✓

- ✓ An Introduction to Probability Theory and its Applications, Feller, 1950. An absolute classic. A more advanced (but still accessible) introduction to probability theory. (Amazon)
- ✓ The Art of Probability, Hamming, 1994. Another introduction to probability theory written by a fellow engineer. (Amazon)
- ✓ All of Statistics, Wasserman, 2004. Exactly what the title says. (Amazon)
- ✓ Linear Algebra and Its Applications, Strang, 2005. A great introduction to linear algebra. (Amazon)
- ✓ Convex Optimization, Boyd, 2004. We will cover the tools from convex optimization that we need as part of the course, but if you want to know more, this is a great resource targeted towards electrical engineers. ([Free online version](#), Amazon)