

# Machine Learning 1 - Notes

Selene Baez Santamaria

September 28, 2016

## THINGS WE NEED TO KNOW

### 1. Probability:

- Variable representation
- Conditional probability (Bayes Theorem)
- Likelihood
- Prior
- Normalization constant
- Marginalization

### 2. Conditional models

- Objective function
- Maximum log-likelihood (MLE)
- Maximum posterior (MAP)
- Training and test errors as a function of  $M$
- Overfitting
- Underfitting

### 3. Classification

- Optimal Bayesian decision rule
- Loss function

### 4. Regularization (Theory)

- Bias

- Variance
  - Approximation of bias and variance
5. Conditional classification
- General expression for  $p(\mathcal{C}_2|x)$
  - Conditions for classifying  $x$  as  $\mathcal{C}_1$  (inequality)
  - Decision boundary
6. Regression
- Linear vs logistic functions
  - Stochastic gradient descent algorithm
7. Bayesian linear regression
- Posterior distributions
  - Likelihood
  - Posterior predictive distribution
8. General
- Gaussian distribution and its parameters
  - Conversion from matrix to scalars

# 1 Intro

## 1.1 Notation

1.  $N$ : number of data samples
2.  $D$ : dimensionality of input data
3.  $K$ : number of output predictions (labels or tags)
4.  $x_n$ : row vector for  $n^{th}$  sample

## 1.2 Models, Algorithms, etc

A model:

- Belongs to a `model class`, with corresponding `complexity`
- Depends on `parameters  $\theta$`
- is optimized by minimizing/maximizing an `objective function`