CSI 5325 Assignment 0

Greg Hamerly

Assigned: 1/9/2018; Due: 1/16/2018

Instructions

For this assignment, you should write your derivations in LaTeX. They should be concise and easy to follow, with appropriate English descriptions rather than just mathematics (i.e. explain what you are doing as necessary).

Write any programs in Matlab (or Octave), and attach their code with your assignment. Include any relevant figures and graphs (plotted using Matlab / Octave) in your writeup.

Submit your assignment in two ways: hardcopy (in class) and by email (to hamerly@cs.baylor.edu). The email should contain a single attachment as a ZIP file. It should be named "lastname-xx.zip", where lastname is your last name, and xx is the number of the assignment.

Finally, please keep your submitted email attachments small. In particular, make sure you are only submitting things that are necessary (omit datasets I gave you, compiled programs, etc.). Also, try to keep your graphics small by using vector (rather than bitmap) formats (e.g. PDF or EPS rather than JPG or BMP). Vector graphics are generally smaller in size and better quality than bitmap.

This first homework will not be graded in detail, but you will submit it for credit.

1 Mathematics refresher

We will use a lot of basic calculus, linear algebra, and statistics in this course. I will presume you have taken courses where you learned these topics, so this part should be considered a refresher. However, if you have not yet studied these, then you need to study a bit and learn the basics so you can follow this course.

1.1 Calculus refresher

We will use derivative calculus a lot this semester to minimize (or maximize) functions that we want to optimize. For instance, we often will have a model we want to fit to data, and a measure of quality (i.e. how well the model fits the data). We want to change the model's parameters to improve the quality. Taking the (perhaps partial) derivative and setting it to zero allows us to move in that direction.

You should be familiar with the basics of derivative calculus. We will not really use much integral calculus, except when it comes to computing expectations of random variables (see below). Any good calculus book will do. Here's a page I found that looks reasonable:

http://www.stat.wisc.edu/~ifischer/calculus.pdf

1.2 Linear algebra refresher

We often use vectors and matrices to organize data and model parameters in machine learning. Please see this page for a refresher on working with matrices and vectors:

http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/intro.html

Not all the sections of that website are required; these first two cover most of what we will need (including concepts like products, trace, inverse, eigenvectors, etc.):

- http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/property.html
- http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/eigen.html

For further explanations of the concepts, please see a good text or other web pages.

1.3 Statistics refresher

Modern machine learning uses a lot of statistics, and often statistics paired with linear algebra. You should understand basics like: random variables and their associated distributions, probability density, cumulative density, simple random variable transformations, expectation of a random variable, variance. Also, you should be familiar with standard probability distributions like Gaussian (aka Normal), multivariate Gaussian, Bernoulli, binomial, multinomial, exponential, etc. and understand how their parameters affect their distributions.

Here is a refresher on probability and statistics by Peter J. Haas from Stanford:

http://cs.baylor.edu/~hamerly/courses/5325_18s/resources/haas_prob_stats_refresher.pdf

2 Matlab / Octave

Familiarize yourself with Matlab (or Octave, a Matlab-like free software alternative). You may find a resource, or try this online tutorial from the makers of Matlab: https://www.mathworks.com/academia/student_center/tutorials/register.html

Make sure you try out using Matlab for loading data, saving data, writing a .m function file with parameters and calling it, plotting data, and printing a plot to PDF or other vector (not bitmap) format.

3 Simple exercises for understanding

Do these exercises and check your answers with your classmates. Write the derivations carefully in LaTeX for practice.

1. Find the value x that maximizes $f(x) = -3x^2 + 24x - 30$.

2. Find the partial derivatives of $g(x)$ with respect to x_0 and x_1:

$$g(x) = 3x_0^3 - 2x_0x_1^2 + 4x_1 - 8$$

3. What is the value of $AB^T + C^{-1}$, if the following define A, B, and C? Use Matlab to check your answer.

$$A = \begin{bmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

4. Write down the mathematical definitions of the simple Gaussian, multivariate Gaussian, Bernoulli, binomial and exponential distributions. Do this in LaTeX for practice.

5. What is the relationship between the Bernoulli and binomial distributions?

6. Suppose that random variable $X \sim N(1, 3)$. What is its expected value?

7. Suppose that random variable Y has distribution

$$p(Y = y) = \begin{cases} \exp(-y), \text{ if } y \geq 0 \\ 0, \text{ otherwise} \end{cases}$$

   - Verify that $\int_{y=-\infty}^{\infty} p(Y = y) = 1$
   - What is $\mu_Y = E[Y] = \int_{y=-\infty}^{\infty} p(Y = y)y \; dy$? (The expected value of Y.)
   - What is $\sigma^2 = Var[Y] = \int_{y=-\infty}^{\infty} p(Y = y)(y - \mu_Y)^2 \; dy$? (The variance of Y.)
   - What is $E[Y|Y \geq 10]$? (That is, the expected value of Y, given that (or conditioned on) $Y \geq 10$.)

4 Programming

If you don't have one, create a Kattis account at `https://baylor.kattis.com/register`. Then, register for the course `https://baylor.kattis.com/courses/CSI5325/18s` by clicking on "I am a student taking this course and I want to register for it on Kattis.". Then do the following problems linked from the problem group:

- Hello World!
- Estimating the Area of a Circle
- Support Vector Machine

These should be fairly trivial, just to become familiar with the platform.