CS 335 Homework 1

Last Updated: March 25, 2019

Instructions

Due Friday 2/8 at 11:59pm. Complete all exercises and problems below.

How to submit

Your submission consists of three steps:

1. **Upload written solutions to Gradescope**. Create a single high-quality pdf with your solutions to the exercises and Problems 1–3. The solutions can be typed or written and scanned but the resulting pdf *must* be high quality and easily readable. Upload the pdf to Gradescope.

2. **Upload code listing to Gradescope**. You'll need to edit these files in the directory `hw1-files`:

   - `gd.py`
   - `gradient-descent.ipynb`
   - `linear-regression.ipynb`

   Complete the requested code in these files. When you are *completely done* and ready to submit your code listing, run the `pprint_hw1.py` script from the `hw1-files` directory:

   ```
   $ python pprint_hw1.py
   ```

   This will create a new file called `hw1-code.pdf` with a listing of all of your code and results. Open the pdf to make sure it is correct and includes all of your code as well as the notebook outputs and plots. You can run this multiple times if you update your code. Upload this to Gradescope.

3. **Submit a single zip file containing source code to Moodle.** Make sure your code is complete and files are included in the directory. Also include any auxiliary data or code files you created that are needed to run your code.

   Rename your code directory from `hw1-files` to `hw1-<your last name>` and zip it:

   ```
   $ mv hw1-files hw1-sheldon
   $ zip -r hw1-sheldon.zip hw1-sheldon
   ```

   Submit the single zip file to Moodle.

Exercises (2 points each)

These are designed to test your understanding of basic concepts. I strongly recommend attempting these first on your own.

Partial derivatives

Consider the following functions of the variables u, v, and w. Assume the variables x, y, $x^{(i)}$ and $y^{(i)}$ are **constants**: they represent numbers that will not change during the execution of a machine learning algorithm (e.g., the training data).

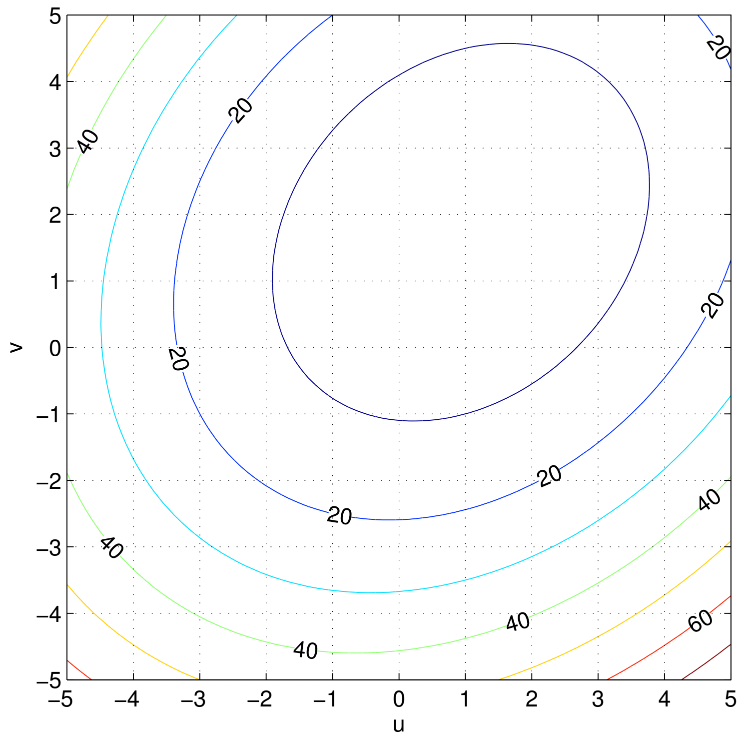
- $f(u, v) = 10u^2v^3 + 3v^2 + 4u$
- $g(u, v, w) = x\log(u) + yuvw^2 + 10x^2$
- $h(u, v) = \sum_{i=1}^{m} \frac{1}{2}\left(x^{(i)}u + y^{(i)}v\right)^2$

Write the following partial derivatives:

1. $\frac{\partial}{\partial u} f(u, v) =$
2. $\frac{\partial}{\partial v} f(u, v) =$
3. $\frac{\partial}{\partial u} g(u, v, w) =$
4. $\frac{\partial}{\partial v} g(u, v, w) =$
5. $\frac{\partial}{\partial w} g(u, v, w) =$
6. $\frac{\partial}{\partial u} h(u, v) =$
7. $\frac{\partial}{\partial v} h(u, v) =$

Partial derivative intuition

Consider the following contour plot of a function $f(u, v)$:



For each of the following partial derivatives, state whether it is positive, negative, or equal to zero. Briefly explain. These questions can be answered from the contour plot without knowing the formula for the function.

(**Note**: for two numbers a and b we will use the notation $\frac{\partial}{\partial u} f(a, b)$ to mean "the partial derivative of $f(u, v)$ with respect to u at the point where $u = a$ and $v = b$". This notation is succinct but obfuscates the original variable names. A more explicit way to write the same thing is $\frac{\partial}{\partial u} f(u, v)|_{u=a,v=b}$.)

1. $\dfrac{\partial}{\partial u} f(-2, -2)$
2. $\dfrac{\partial}{\partial v} f(-2, -2)$
3. $\dfrac{\partial}{\partial u} f(3, -3)$
4. $\dfrac{\partial}{\partial v} f(3, -3)$
5. To the nearest integer, estimate the values of u and v that minimize $f(u, v)$.

Problems (10 points each)

The first few problems will consider the "slope-only" (i.e., $\theta_0 = 0$) linear regression model from class, which has the following hypothesis and cost function:

$$\text{Hypothesis}: h_\theta(x) = \theta_1 x$$

$$\text{Cost function}: J(\theta_1) = \frac{1}{2} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

Problem 1. Consider the following small data set:

x	y
1	3
-1	-2
2	4

Solve for the value of θ_1 that minimizes the cost function by substituting the values from the training set into the cost function, setting the derivative equal to zero, and solving for θ_1. Show your work.

Problem 2. Now do the same thing, but don't substitute the values of the training set into the cost function. Instead, leave the $x^{(i)}$ and $y^{(i)}$ variables, take the derivative with respect to θ_1, set it equal to zero, and solve for θ_1. This will give you a general expression for θ_1 in terms of the training data.

Check your answer by plugging in the training data from the previous problem into your expression for θ_1. You should get the same value for θ_1 that you got in that problem.

Problem 3 (5 points extra credit). Do the same thing as in the previous problem, but use the more general hypothesis and cost function:

$$h_\theta(x) = \theta_0 + \theta_1 x, \quad J(\theta_0, \theta_1) = \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

Take partial derivatives with respect to both θ_0 and θ_1 and set them to zero, then solve the system of two equations to come up with expressions for θ_0 and θ_1 in terms of the training data.

Problem 4. In this problem you will implement gradient descent for linear regression. Open the notebook `gradient-descent.ipynb` in Jupyter and follow the instructions to complete the problem.

Problem 5. In this problem you will find a data set of your own (with at least 5 training examples) that is suitable for linear regression with one input variable. Open the file `linear-regression.ipynb` in Jupyter and follow the instructions to complete the problem.

Problem 6. (0 points) How much time did you spend on this homework?