1
point

1.

Consider the problem of predicting how well a student does in her second year of college/university, given how well she did in her first year.

Specifically, let x be equal to the number of "A" grades (including A-. A and A+ grades) that a student receives in their first year of college (freshmen year). We would like to predict the value of y, which we define as the number of "A" grades they get in their second year (sophomore year).

Here each row is one training example. Recall that in linear regression, our hypothesis is $h_\theta(x) = \theta_0 + \theta_1 x$, and we use m to denote the number of training examples.

x	y
3	2
1	2
0	1
4	3

For the training set given above (note that this training set may also be referenced in other questions in this quiz), what is the value of m? In the box below, please enter your answer (which should be a number between 0 and 10).

Enter answer here

1
point

2.

Many substances that can burn (such as gasoline and alcohol) have a chemical structure based on carbon atoms; for this reason they are called hydrocarbons. A chemist wants to understand how the number of carbon atoms in a molecule affects how much energy is released when that molecule combusts (meaning that it is burned). The chemist obtains the dataset below. In the column on the right, “kJ/mol” is the unit measuring the amount of energy released.

Name of molecule	Number of hydrocarbons in molecule (x)	Heat release when burned (kJ/mol) (y)
methane	1	-890
ethene	2	-1411
ethane	2	-1560
propane	3	-2220
cyclopropane	3	-2091
butane	4	-2878
pentane	5	-3537
benzene	6	-3268
cycloexane	6	-3920
hexane	6	-4163
octane	8	-5471
napthalene	10	-5157

You would like to use linear regression ($h_\theta(x) = \theta_0 + \theta_1 x$) to estimate the amount of energy released (y) as a function of the number of carbon atoms (x). Which of the following do you think will be the values you obtain for θ_0 and θ_1? You should be able to select the right answer without actually implementing linear regression.

- $\theta_0 = -1780.0, \theta_1 = 530.9$
- $\theta_0 = -1780.0, \theta_1 = -530.9$
- $\theta_0 = -569.6, \theta_1 = -530.9$
- $\theta_0 = -569.6, \theta_1 = 530.9$

1
point

3.

Suppose we set $\theta_0 = -1, \theta_1 = 2$ in the linear regression hypothesis from Q1. What is $h_\theta(6)$?

Enter answer here

1
point

4.
Let f be some function so that

$f(\theta_0, \theta_1)$ outputs a number. For this problem,

f is some arbitrary/unknown smooth function (not necessarily the

cost function of linear regression, so f may have local optima).

Suppose we use gradient descent to try to minimize $f(\theta_0, \theta_1)$

as a function of θ_0 and θ_1. Which of the

following statements are true? (Check all that apply.)

- [ ] If θ_0 and θ_1 are initialized at

  a local minimum, then one iteration will not change their values.

- [ ] If θ_0 and θ_1 are initialized so that $\theta_0 = \theta_1$, then by symmetry (because we do simultaneous updates to the two parameters), after one iteration of gradient descent, we will still have $\theta_0 = \theta_1$.

- [ ] Even if the learning rate α is very large, every iteration of

  gradient descent will decrease the value of $f(\theta_0, \theta_1)$.

- [ ] If the learning rate is too small, then gradient descent may take a very long

  time to converge.

1
point

5.

Suppose that for some linear regression problem (say, predicting housing prices as in the lecture), we have some training set, and for our training set we managed to find some θ_0, θ_1 such that $J(\theta_0, \theta_1) = 0$.

Which of the statements below must then be true? (Check all that apply.)

- [ ] We can perfectly predict the value of y even for new examples that we have not yet seen.

  (e.g., we can perfectly predict prices of even new houses that we have not yet seen.)

- [ ] This is not possible: By the definition of $J(\theta_0, \theta_1)$, it is not possible for there to exist

  θ_0 and θ_1 so that $J(\theta_0, \theta_1) = 0$

- [ ] For these values of θ_0 and θ_1 that satisfy $J(\theta_0, \theta_1) = 0$,

  we have that $h_\theta(x^{(i)}) = y^{(i)}$ for every training example $(x^{(i)}, y^{(i)})$

- [ ] For this to be true, we must have $\theta_0 = 0$ and $\theta_1 = 0$

  so that $h_\theta(x) = 0$

- [ ] I understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

Learn more about Coursera's Honor Code

Enter your legal name

Submit Quiz