CSC 665: Artificial Intelligence Homework

Uzair Hamed Mohammed

Spring 2024

A compilation of completed homework assignments

San Francisco State University

Introduction

This document contains the work I, Uzair Hamed Mohammed, have done as required by this course. This is my first time using LaTeX to type homework, so strange formatting may occur.

I once read in a math and Linux oriented blog that handwriting your math homework and notes, and then typing it up in LaTeX greatly boosts mastery and memorization of concepts. My plan is to do exactly that, and append homework assignments to this document as they are completed. At the end of the semester, the result should be a nice and comprehensive compilation of assignments.

My current workflow is as follows:

1. Read the homework instructions and try to understand the assignment.
2. Work by hand, on my tablet, in the Samsung Notes app.
3. Beautify my work by transcribing it here using LaTeX.

When it's time to submit an assignment, I will export this document as a PDF file and turn in just the relevant pages. Please let me know what you think of this format!

0 Homework 0

By turning in this assignment, I agree to abide by SFSU's academic integrity code and declare that all of my solutions are my own work:

0.1 About You

a. My pronouns are he/him.

b. I've taken a lot of math and computer science courses. I'm not sure if I can list them, as I took most of them at the College of San Mateo and their course numbers are different. However, off the top of my head, I've taken:

   - Calculus 1
   - Calculus 2
   - Calculus 3
   - Discrete Mathematics
   - Linear Algebra
   - Analysis of Algorithms
   - Data Structures

   This list consists of the courses I think are relevant to this class; I've taken other CS courses, of course.

c. Yes, I am okay with being called on. I won't always know the answer, and I might embarrass myself sometimes, but that's alright because I believe it will force me to pay more attention and learn better, and as a result be more prepared for assignments and exams.

0.2 Optimization

a. Let a, b, and c be positive real numbers. Consider the quadratic function

$$f(\theta) = a\theta^2 + b\theta + c$$

   Note that θ here is a real number. What value of θ minimizes $f(\theta)$?

   Solution: Since this is a positive parabola, we can use the formula for the vertex of a parabola, $\frac{-b}{2a}$, to find the minimum:

$$\boxed{\theta = \frac{-b}{2a}}$$

For the following problems, I'm going to use the general approach of:

1. Take derivative of equation
2. Set derivative equal to 0
3. Solve for θ

b. Let $x_1, \ldots, x_n$ be real numbers. Consider the quadratic function

$$g(\theta) = \sum_{i=1}^{n} (\theta - x_i)^2.$$

What value of θ minimizes $g(\theta)$?
Solution:

$$\begin{aligned} g'(\theta) &= \sum_{i=1}^{n} 2(\theta - x_i) \\ 0 &= \sum_{i=1}^{n} 2(\theta - x_i) \\ \theta \sum_{i=1}^{n} 2 &= \sum_{i=1}^{n} 2x_i \\ &\boxed{\theta = \frac{\sum_{i=1}^{n} x_i}{n}} \end{aligned}$$

c. Let $x_1, \ldots, x_n$ again be real numbers, and let $w_1, \ldots, w_n$ be positive real numbers that we can interpret as representing the importance of each of the x_i's. Consider the weighted quadratic function

$$h(\theta) = \sum_{i=1}^{n} w_i (\theta - x_i)^2.$$

What value of θ minimizes $h(\theta)$?
Solution:

$$h'(\theta) = 2 * \sum_{i=1}^{n} w_i(\theta - x_i)$$

$$0 = 2 * \sum_{i=1}^{n} w_i(\theta - x_i)$$

$$\frac{0}{2} = \frac{\cancel{2} * \sum_{i=1}^{n} w_i(\theta - x_i)}{\cancel{2}}$$

$$0 = \sum_{i=1}^{n} w_i(\theta - x_i)$$

$$0 = \sum_{i=1}^{n} (w_i\theta - w_i x_i)$$

$$0 = \sum_{i=1}^{n} w_i\theta - \sum_{i=1}^{n} w_i x_i$$

$$\sum_{i=1}^{n} w_i x_i = \sum_{i=1}^{n} w_i\theta$$

$$\boxed{\frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} = \theta}$$

d. What issue could arise in the minimization of h if some of the w_i's are negative?
Solution: If some of the w_i's are negative, it can lead to convoluted results and a possible divergence towards ∞

0.3 Probability

a. Consider a standard 52-card deck of cards with 13 card values (Ace, King, Queen, Jack, and 2-10) in each of the four suits (clubs, diamonds, hearts, spades). If a card is drawn at random, what is the probability that it is a spade or a two?

Solution:
→ Let $P(S)$ be the probability of drawing a spade $= \frac{13}{52} = \frac{1}{4}$
→ Let $P(T)$ be the probability of drawing a card that is a two $= \frac{4}{52} = \frac{1}{13}$

$$P(S \cap T) = \frac{1}{52}$$

The overall probability of drawing a spade or a two is:

$$P(S \cup T) = (P(S) + P(T)) - P(S \cap T)$$
$$= (\frac{1}{4} + \frac{1}{13}) - \frac{1}{52}$$
$$\boxed{= \frac{4}{13} \approx 30.7\%}$$

b. Two factories — Factory A and Factory B — design batteries to be used in mobile phones. Factory A produces 60% of all batteries, and Factory B produces the other 40%. 2% of Factory A's batteries have defects, and 4% of Factory B's batteries have defects. What is the probability that a battery is both made by Factory A and defective?

Solution:
→ Let $P(A) = 0.6$ (probability battery is from Factory A)
→ Let $P(D|A) = 0.02$ (probability battery is defective if it's from Factory A)
→ Let $P(A)' = 0.4$ (probability battery is not from Factory A)
→ Let $P(D|A)' = 0.04$ (probability battery is defective if it's not from Factory A)

→ Find $P(A \cap D)$.

...

...

...

$$P(A \cap D) = 0.012 \approx 1.2\%$$

c. Consider the following (made up) facts about COVID incidence and testing:

- In the absence of any special information, the probability that a person has COVID is 1%.
- If a person has COVID, the probability that a test will correctly read positive is 80%.
- If a person does not have COVID, the probability that a test will incorrectly produce a false positive is 10%.

Suppose you take a COVID test and it reads positive. Given the facts above, what is the probability that you have COVID?

Solution:
→ Let C = COVID
→ Let T = Positive Test

→ Let $P(C) = 0.01$ (probability of having COVID)

→ Let $P(T|C) = 0.8$ (probability of a positive test if a person has COVID)
→ Let $P(T|C)' = 0.1$ (probability of a false positive)
→ Find $P(C|T)$ (probability of having COVID with a positive test result)

$$P(T) = P(T|C) * (P(C) + (P(T|C)' * P(C)'))$$
$$P(T) = 0.8 * (0.01 + (0.1 * 0.99))$$
$$P(T) = 0.0872$$

$$P(C|T) = \frac{(0.8 * 0.01)}{0.0872} = \frac{0.008}{0.0872} = \boxed{\frac{10}{109} \approx 9.17\%}$$

d. Suppose you repeatedly roll a fair six-sided die until you roll a 1 (and then you stop). Every time you roll a 3, you win a points, and every time you roll a 6, you lose b points. You do not win or lose any points if you roll a 2, 4, or 5. What is the expected number of points (as a function of a and b) you will have when you stop?

0.4 Counting

a. $n \times n$ grid, rectangle, Big O?
Solution: $O(n^2)$

b. Three rectangles, possible ways?
Solution: $n^2 * n^2 * n^2 = n^6$

0.5 Programming in Python

See attached .py file

1 Homework 1

By turning in this assignment, I agree to abide by SFSU's academic integrity code and declare that all of my solutions are my own work:

1.1 Grid City

Modeled as a search problem:

- $s_0 = (0, 0)$
- $\text{Actions}(s) = \{(+1, 0), (-1, 0), (0, +1), (0, -1)\}$
- $\text{Succ}(s, a) = s + a$
- $\text{Cost}((x, y), a) = 1+ \max(x, 0)$ (it is more expensive as you go further to the right)
- $\text{IsEnd}(s) = \begin{cases} \text{True} & \text{if } s = (m, n) \\ \text{False} & \text{otherwise} \end{cases}$

a. Minimum cost of reaching location (m, n) starting from $(0, 0)$? Describe possible path achieving the minimum cost. Is it unique?

   Solution: Since $(m, n) \geq 0$, we can safely say that (m, n) can't be at $(0, 0)$ in the minimum cost path. The next closest location to $(0, 0)$ that (m, n) could be is $(1, 1)$, as it satisfies the condition listed earlier. With this in mind, the minimum cost would be:

$$\begin{aligned} \text{Cost}((x, y), a) &= 1 + \max(x, 0) \\ &= (1 + \max(m, 0)) + (1 + \max(n, 0)) \\ &= 1 + m + 1 + n \\ &\boxed{= 2 + m + n} \end{aligned}$$

b. True or false (and explain): UCS will never terminate on this problem because the number of states is infinite.

   Solution: False, because UCS explores nodes with the lowest cost first. If (m, n) is available at the minimum cost from $(0, 0)$, then UCS will terminate, and that too fairly quickly.

c. True or false (and explain): UCS will return the minimum cost path and explore only locations between $(0, 0)$ and (m, n); that is, locations (x, y) such that $0 \leq x \leq m$ and $0 \leq y \leq n$.

   Solution: True, if (m, n) is available at the minimum cost from $(0, 0)$, because UCS will terminate very quickly and therefore won't explore other paths. This is because UCS prioritizes the lowest cost first.

d. True or false (and explain): UCS will return the minimum cost path and explore only locations whose path costs are strictly less than the minimum cost from $(0, 0)$ to (m, n).

   Solution: True, because UCS focuses on cumulative costs that are less than the optimal cost to reach (m, n).

Now consider UCS running on an arbitrary graph.

e. True or false (and explain): If you add an edge between two nodes, the cost of the min-cost path cannot go up.

   Solution: True, because adding an edge can only maintain or reduce the cost of the min-cost path.

f. True or false (and explain): If you make the cost of an action from some state small enough (possibly negative), you can guarantee that that action will show up in the minimum cost path.

   Solution: False, because that action isn't guaranteed to lead us to the goal, no matter how cheap we make it.

g. True or false (and explain): If you increase the cost of every action by 1, the minimum cost path does not change (even though its cost does).

   Solution: True. Since we are raising the cost of all costs equally, the minimum cost path will remain the the same even despite the costs going up.

1.2 Six Degrees

See degrees.py

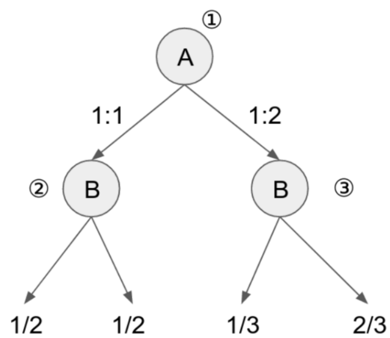
1.3 Tic Tac Toe

See tictactoe.py

2 Homework 2

2.1 Cake

Alice and Bob are sharing a cake....



Write down the utility Alice should expect to receive in each of the game states given the following knowlege of player strategies:

a. Alice and Bob playing adversarially, each trying to maximize the cake they receive.

   1. 1/2
   2. 1/2
   3. 1/2

b. Alice still trying to maximize, Bob now playing collaboratively and helping Alice.

   1. 2/3
   2. 1/2
   3. 2/3

c. Random play.

   1. Node 2 or 3
   2. Doesn't matter because of constant utility
   3. Node 3

d. Alpha Beta pruning

   1. Leaf notes with utility 1/3, 1/3, 1/3, 1/3 on left, and right leaf node with utility 1/6

e. Alpha Beta pruning with tree elements swapped.

   1. No nodes are pruned

2. Logical Formulas

   a. $(P \land \neg Q) \lor (\neg P \land Q)$
   b. $(\exists x \text{Course}(x) \land \text{Enrolled}(A, x) \land \text{Enrolled}(B, x))$
   c. $(\forall x \text{Course}(x) \land \text{Enrolled}(A, x) \rightarrow \text{Enrolled}(B, x))$
   d. $\forall x \forall y (\text{Enrolled}(x, y) \land \text{Course}(y) \rightarrow \text{Student}(x))$
   e. $\forall x (\text{Course}(x) \rightarrow \exists y (\text{Student}(y) \land \text{Enrolled}(y, x)))$

2.2 Knights and Knaves

See attached puzzle.py file

2.3 Odds and Evens

See attached submission.py file

3 Homework 3

3.1 Probability

a. An urn contains w white balls and b black balls...

1. What is the probability that the first ball is white?

$$P(\text{first ball is white}) = \boxed{\frac{w}{w+b}}$$

2. What is the probability that the second ball is white?

$$\text{If first ball was white :}$$
$$P(\text{second ball white — first ball white}) = \boxed{\frac{w-1}{w+b-1}}$$
$$\text{If first ball was black :}$$
$$P(\text{second ball white — first ball black}) = \boxed{\frac{w}{w+b+1}}$$

b. Alice has two brothers, Bob and Charlie...

1. What is the probability that Alice is older than Charlie?

   There are 3! = 6 possible scenarios, and Alice is older in 3 of them. So, $\boxed{\frac{3}{6} \text{ or } 0.5}$.

2. Alice tells you that she is older than Bob

   In this case, the probability of Alice being older than Charlie is $\boxed{\frac{2}{3} \text{ or } 0.66}$.

c. The inhabitants of a small island...

Let:

- T = first person tells truth
- L = first person lies
- Y = second person says yes

We know that:

- P(T) = 1/3
- P(L) = 2/3
- P(Y — T) = 1/3

– P(Y — L) = 2/3

We find: $P(T|Y) = \frac{P(Y|T) \cdot P(T)}{P(Y)}$

$$
\begin{aligned}
P(Y) &= P(Y|T) \cdot P(T) + P(Y|L) \cdot P(L) \\
&= \frac{1}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{2}{3} \\
&= \frac{1}{9} + \frac{4}{5} \\
&= \frac{5}{9}
\end{aligned}
$$

$$
\begin{aligned}
P(T|Y) &= \frac{\frac{1}{3} \cdot \frac{1}{3}}{\frac{5}{9}} \\
&= \boxed{\frac{1}{5} \text{ or } 0.2}
\end{aligned}
$$

d. Bag contains one ball...

Let:

– W = original ball was white
– B = original ball was black
– D = white ball is drawn

We know:

– P(W) = 1/2
– P(B) = 1/2
– P(D — W) = 1
– P(D — B) = 1/2

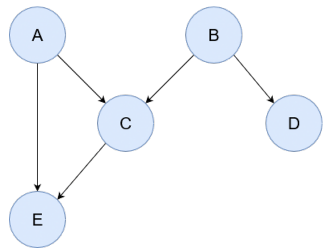
We find: $P(W|D) = \frac{P(D|W) \cdot P(W)}{P(D)}$

$$
\begin{aligned}
P(D) &= P(D|W) \cdot P(W) + P(D|B) \cdot P(B) \\
&= 1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \\
&= \frac{1}{2} + \frac{1}{4} \\
&= \frac{3}{4}
\end{aligned}
$$

$$
\begin{aligned}
P(W|D) &= \frac{1 \cdot \frac{1}{2}}{\frac{3}{4}} \\
&\boxed{= \frac{2}{3}}
\end{aligned}
$$

3.2 Bayesian Networks

a. Given the following Bayesian network, write down a factorization of the joint probability distribution P(A,B, C,D, E) as a product of local conditional distributions, one for each random variable.



$$P(A, B, C, D, E) = \boxed{P(A) \cdot P(B) \cdot P(C|A, B) \cdot P(D|B) \cdot P(E|C)}$$

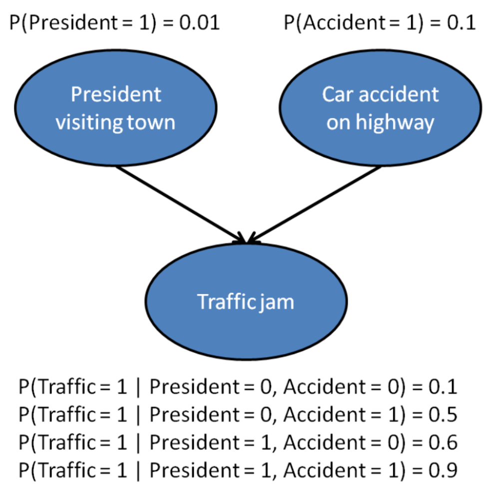
b. For the same Bayesian network above, list all the random variables that are independent of A, assuming that none of the variables have been observed.

  – B

c. For the same Bayesian network above, assume that E is now observed. List all pairs of random variables (not including E) that are conditionally independent given E.

  – (A, B)
  – (A, D)
  – (B, D)

d. Consider the following model for traffic jams in a small town, which can be caused by either a car accident or by a visit from the president (and the accompanying security motorcade).



Compute P(Accident = 1 — Traffic = 1) and P(Accident = 1 — Traffic = 1, President = 1).

Let:

- A = event of accident
- P = event of president's visit
- T = event of traffic jam

We're given:

- P(A) = prior probability of an accident
- P(P) = prior probability of the president's visit
- P(T — A) = probability of a traffic jam given an accident
- P(T — P) = probability of a traffic jam given the president's visit

$$P(\text{Accident} = 1|\text{Traffic} = 1) \rightarrow P(A|T) = \frac{P(T|A) \cdot P(A)}{P(T)}$$

4 Homework 4

4.1 Inference in Bayesian Networks

a - c. Refer to bn.py

d. Likelihood weighting gets more accurate as we use more samples. With more samples, the estimated probabilities get closer to the real chances we calculated before. This shows that using a bigger sample size is helpful for getting better results.

4.2 Gradient Descent

a. Hinge loss is a reasonable cost function for binary classification because it focuses on creating a good margin between classes. It only penalizes models when their predictions fall close to the wrong class, and ignores classified points even with a small margin. This makes the model prioritize a clear separation between positive and negative classes.

b.

$$\frac{\partial c(h(x), y)}{\partial w_0} = 0$$

$$\frac{\partial c(h(x), y)}{\partial w_1} = \begin{cases} -y, & \text{if } w_0 + w_1 x < 1 \\ 0, & \text{if } w_0 + w_1 x \geq 1 \end{cases}$$

c.

Iteration	w_0	w_1
1	0	-0.1
2	0	-0.09
3	0	-0.008
4	0.082	0.074
5	0.164	0.066
6	0.172	0.058

5 Homework 5

5.1 Regularization

a.

$$\frac{d}{dw_0}C(w) = \sum_{i=1}^{n} 2(y_i - (w_0 + w_1 x_i))(-1) = \boxed{-2\sum_{i=1}^{n}(y_i - (w_0 + w_1 x_i))}$$

$$\frac{d}{dw_1}C(w) = \sum_{i=1}^{n} 2(y_i - (w_0 + w_1 x_i))(-x_i) = \boxed{-2\sum_{i=1}^{n} x_i(y_i - (w_0 + w_1 x_i))}$$

b.

$$\begin{aligned}\frac{d}{dw_0}\tilde{C}(w) &= \sum_{i=1}^{n} 2(y_i - (w_0 + w_1 x_i))(-1) + 2\lambda w_0 \\ &= 2\sum_{i=1}^{n}(w_0 + w_1 x_i - y_i) + 2\lambda w_0] \\ &= \boxed{2\sum_{i=1}^{n}(w_0 + w_1 x_i - y_i + \lambda w_0)}\end{aligned}$$

and

$$\begin{aligned}\frac{d}{dw_1}\tilde{C}(w) &= \sum_{i=1}^{n} 2(y_i - (w_0 + w_1 x_i))(-x_i) + 2\lambda w_1 \\ &= \boxed{2\sum_{i=1}^{n} x_i(y_i - (w_0 + w_1 x_i)) + 2\lambda w_1}\end{aligned}$$

c-d. Refer to regularization.py

e. Avg loss, regularization: 1.715104320688656
Avg loss, no regularization: 0.4384015750703455