

CS731 Homework 1

Due 1/26/2011 before class

What to hand in: the .pdf file, the .tex source file, and any other files required in the questions below (such as your photo). You do not need to hand in any code. To save you time, you do not need to include intermediate derivation steps, nor a copy of the questions themselves – just the final answer is fine. See course webpage for hand-in instructions.

1. (10 points) Let $f(x) = \log(1 + \exp(x))$. Define $f^*(y) = \sup_{x \in \mathbb{R}} (xy - f(x))$ for $y \in \mathbb{R}$. Write down $f^*(y)$ solely as a function of y .

Solution: $f^*(y) = -H(y) = y \log y + (1 - y) \log(1 - y)$ for $y \in [0, 1]$, and ∞ otherwise. Note the boundary $y = 0, 1$.

2. (10 points) Let A be a symmetric $n \times n$ matrix, and x be a n -vector. Express $\text{trace}(Axx^\top)$ concisely with the elements of A and x .

Solution: $\sum_{i,j} A_{ij} x_i x_j$

3. (10 points) What is $\sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$ for $\lambda > 0$? Hint: Taylor expansion.

Solution: e^λ

4. (10 points) Let

$$p(x) \propto \begin{cases} 1 & x \in (a, b) \\ 0 & \text{otherwise} \end{cases}$$

for $a, b \in \mathbb{R}$, $a < b$. Compute its variance.

Solution: $(a - b)^2/12$

5. (10 points) Let A be an $n \times n$ real symmetric invertible matrix with eigenvalue, eigenvector pairs $\{(\lambda_i, \phi_i)\}$ for $i = 1, \dots, n$. What are the eigenvalues and eigenvectors for the matrix $(cA)^{-1}$ for $c \in \mathbb{R}$, $c \neq 0$?

Solution: $\{(c\lambda_i)^{-1}, \phi_i\}$

6. (10 points) What is $\lim_{p \rightarrow 0^+} p \log p$?

Solution: 0

7. (10 points) Compute $\frac{\partial}{\partial x} \left(\frac{e^{xz}}{xy+w} + v \right)$

Solution: $\frac{\partial}{\partial x} \left(\frac{e^{xz}}{xy+w} + v \right) = \frac{ze^{xz}}{xy+w} - \frac{ye^{xz}}{(xy+w)^2}$

8. (10 points) Your house has an alarm, which is either on (A) or off ($\neg A$). There can either be a burglary (B) or not ($\neg B$) in your house. There can either be an earthquake (E) or not ($\neg E$) in your town. Your friend either tweets “quake!” to you (F) or not ($\neg F$). The corresponding Bayesian Network has the structure

$$B \rightarrow A \leftarrow E \rightarrow F$$

Assume that

- $P(B) = 0.01, P(E) = 0.02$
- $P(A|B, E) = 1, P(A|B, \neg E) = 0.5, P(A|\neg B, E) = 0.5, P(A|\neg B, \neg E) = 0.01$
- $P(F|E) = 1, P(F|\neg E) = 0.1$

What is the probability that your friend will tweet “quake!” to you, given that your alarm is on?

Solution: 0.468

9. (10 points) You tune k for k NN classification with training set accuracy. What k will you get?

Solution: 1

10. (3 points) Read Tom Mitchell’s *The Discipline of Machine Learning* (<http://www.cs.cmu.edu/~tom/pubs/MachineLearning.pdf>). According to the article, what are the five top conferences and journals in machine learning?

Solution: ICML, NIPS, COLT, JMLR, MJL

11. (3 points) A claim was made on the course webpage that “They ... worth more than gold of equal weight.” In one short paragraph, assess this claim in any way you see appropriate.

Solution: Anything you say is great. I’m most impressed by people who actually computed the price of gold: it’s around \$170K – you may well be earning that much a year after graduating! Just don’t be a professor.

12. (2 points) Sign up to the following machine learning mailing lists:

- The Machine Learning News <http://groups.google.com/group/ML-news>
- The Machine Learning List <http://c11.stanford.edu/ml1list/>
- The Connectionists <http://www.cnbc.cmu.edu/connectionists>

State that you have done so.

13. (2 points) Hand in a 150×200 (pixels) JPEG photo of you, name it as Lastname.jpg.