Machine Learning and Data Mining - Homework 0

Philip Warton

September 23, 2021

1 Probability

Problem 1

It typically rains 73/365 days of the year. The forecast correctly perdicts rain 70% of the time, and falsely perdicts rain 30% of the time. How likely is it to rain given that rain has been forecasted.

Let $F, \neg F$ denote a forecast of rain or no rain respectively, and let $R, \neg R$ denote the presence of rain, or lack of rain, respectively. We know the following to be true from the problem statement:

$$P(R) = \frac{73}{365} = 0.2$$
$$P(F \mid R) = 0.7$$
$$P(F \mid \neg R) = 0.3$$

Then using this information we can compute the desired probability, $P(R \mid F)$. We invoke Bayes' Theorem, giving us

$$\begin{aligned}
P(R \mid F) &= \frac{P(F \mid R)P(R)}{P(F)} \\
&= \frac{P(F \mid R)P(R)}{P(F \mid R) \cdot P(R) + P(F \mid \neg R) \cdot P(\neg R)} \\
&= \frac{0.7 \cdot 0.2}{0.7 \cdot 0.2 + 0.3 \cdot 0.8} \\
&= \frac{7}{19} \approx 0.3684
\end{aligned}$$

Problem 2

You have a fair 6 sided die with a payout as follows:

$$\text{payout} = \begin{cases} 1 & x = 1 \\ -0.25 & x \neq 1 \end{cases}$$

We can simply compute the expectation to be

$$E[X] = \frac{1 - 0.25 - 0.25 - 0.25 - 0.25 - 0.25}{6} = -\frac{0.25}{6} = -\frac{1}{24}$$

Problem 3

Let $p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ be a probability mass function for some random variable X which has a mean of 0 and a variance of 1. Compute $\int_{-\infty}^{\infty} p(x)(ax^2 + bx + c)dx$.

First recall that $V[X] = \int_{\mathbb{R}} p(x)x^2 dx$ and that $E[X] = \int_{\mathbb{R}} p(x)x dx$ and finally that $1 = \int_{\mathbb{R}} p(x)dx$. Then, after distributing $p(x)$, the integral follows trivially. That is,

$$\begin{aligned}\int_{-\infty}^{\infty} p(x)(ax^2+bx+c)dx &= \int_{-\infty}^{\infty} p(x)ax^2 + p(x)bx + p(x)cdx \\ &= a\int_{-\infty}^{\infty} p(x)x^2dx + b\int_{-\infty}^{\infty} p(x)xdx + c\int_{-\infty}^{\infty} p(x)dx \\ &= aV[X] + bE[X] + c(1) \\ &= a(1) + b(0) + c(1) = a + c\end{aligned}$$

Problem 4

Let X be a continuous random variable over $[0, 1]$ with the following probability mass function:

$$p(x) = \begin{cases} 4x & x \in [0, 0.5] \\ -4x+4 & x \in (0.5, 1] \end{cases}$$

Compute the cumulative mass function.

We take the integral $\int_0^t p(x)dx$ to be $C(t)$. So for $x \in [0, 0.5]$ we get

$$C(t) = \int_0^t 4xdx = 2x^2\Big|_0^t = 2t^2$$

Then we know that $C(0.5) = 0.5$ so we can simply append this to the following integral:

$$\int_{0.5}^{t} -4x+4dx = -2x^2+4x\Big|_0 .5^t = [-2t^2+4t] - [-0.5+2] = -2t^2+4t-1.5$$

Which in doing so gives us our desired function,

$$C(x) = \begin{cases} 2x^2 & x < \frac{1}{2} \\ -2x^2+4x-1 & x \geqslant \frac{1}{2} \end{cases}$$

2 Linear Algebra

Problem 1

Let $B = bb^T$ where $b \in \mathbb{R}^{d\times 1} \neq \mathbf{0}$. Show that $\forall x \in \mathbb{R}^{d\times 1}, x^TBx \geqslant 0$.

Proof. We will simply demonstrate the proof by rewriting the term x^TBx. So we have,

$$\begin{aligned}x^TBx &= x^T(bb^T)x \\ &= (x^Tb)(b^Tx) \\ &= \left(\begin{pmatrix} x_1 & x_2 & \cdots & x_d \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_d \end{pmatrix}\right)\left(\begin{pmatrix} b_1 & b_2 & \cdots & b_d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}\right) \\ &= (x_1b_1 + x_2b_2 + \cdots + x_db_d)(x_1b_1 + x_2b_2 + \cdots + x_db_d) \\ &= (x_1b_1 + x_2b_2 + \cdots + x_db_d)^2 \geqslant 0\end{aligned}$$

The last line follows since the square of any real number is non-negative. □

Problem 2

Solve the following system of equations:

$$\begin{aligned}2x_1 + x_2 + x_3 &= 3\\ 4x_1 + 2x_3 &= 10\\ 2x_1 + 2x_2 &= -2\end{aligned}$$

Let $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Then we have the following system of equations:

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix} x = \begin{pmatrix} 3 \\ 10 \\ -2 \end{pmatrix}$$

So let $A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}$ and $b = \begin{pmatrix} 3 \\ 10 \\ -2 \end{pmatrix}$ so we have our system in the form $Ax = b$. Then we compute the inverse of A to be

$$A^{-1} = \frac{1}{2}\begin{pmatrix} -2 & 1 & 1 \\ 2 & -1 & 0 \\ 4 & -1 & -2 \end{pmatrix}$$

So then we have

$$Ax = b$$

$$\begin{aligned}A^{-1}Ax &= A^{-1}b\\ x = A^{-1}b &= \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}\end{aligned}$$

3 Proving Things

Problem 1

Show that $\ln x \leqslant x - 1 \quad \forall x > 0$ and that $\ln x = x - 1$ if and only if $x = 1$.

Proof. For the first part, consider $f(x) = \ln(x) - (x+1)$. Then $f'(x) = \frac{1}{x} - 1$. Suppose that $0 < x < 1$. Then we have

$$\begin{aligned}x &< 1\\ \frac{1}{x} &> 1\\ \frac{1}{x} - 1 &> 0\end{aligned}$$

Which means that $f(x)$ is increasing on $(0,1)$. Let $x > 1$, then by the same chain of logic $f'(x) < 0$ meaning that $f(x)$ is decreasing on $(1,\infty)$. Then since $f'(1) = 1 - 1 = 0$, it follows that $x = 1$ is a maximum for $f(x)$ on $(0,\infty)$. This is the property we wish to show. Since the derivative $f'(x)$ is only zero at exactly $x = 1$, the inequality must be strict. $\square$

Problem 2

Let $\sum_{i=1}^{k} p_i = \sum_{i=1}^{k} q_i = 1$ and $KL(p||q) = \sum_{i=1}^{k} p_i \ln\left(\frac{p_i}{q_i}\right)$. Show that $KL(p||q) \geqslant 0$.

Proof. Recall that $\ln(a/b) = -\ln(b/a)$. Then we begin with the following:

$$1 - 1 = 0$$
$$\sum_{i=1}^{k} q_i - \sum_{i=1}^{k} p_1 = 0$$
$$\sum_{i=1}^{k} q_i - p_i = 0$$
$$\sum_{i=1}^{k} p_i \left(\frac{q_i}{p_i} - 1 \right) = 0$$

Then since $\ln(x) \leqslant x - 1$ it follows that $0 = \sum_{i=1}^{k} p_i \left(\frac{q_i}{p_i} - 1 \right) \geqslant \sum_{i=1}^{k} p_i \ln \left(\frac{q_i}{p_i} \right)$. This term can be rewritten as

$$-\sum_{i=1}^{k} p_i \ln \left(\frac{q_i}{p_i} \right) \leqslant 0$$
$$\sum_{i=1}^{k} p_i \ln \left(\frac{p_i}{q_i} \right) \geqslant 0$$
$$KL(p||q) \geqslant 0$$

□

4 Debriefing

```
Time Spent | 3:07:17 h:m:s (yes I timed it)
Difficulty | Moderate (the probability portion took some thinking)
Worked alone, looked up the linearity of expectation property.
Understanding | 40% (my understanding of probability is quite shallow)
Comments | Good assignment, I think as a mathematician I will have much more
           difficulty with the coding/programming assignments.
```