Northeastern University, Data Mining Techniques - CS6220
Fall 2017

Solutions to Self Test

Rashmi Dwaraka

September 13, 2017

Contents

1 Bayes' Rule

The Weatherly app predicts rain tomorrow. In recent years, it has rained only 73 days each year. When it actually rains, the Weatherly app correctly forecasts rain 70% of the time. When it doesnt rain, the app incorrectly forecasts rain 30% of the time. What is the probability that it will rain tomorrow?

$$Hint : P(H|D) = \frac{P(H)P(D|H)}{P(D)}$$

Solution:

$Probability\ of\ raining \rightarrow P(A) = 73/365 = 0.2$
$Probability\ of\ not\ raining \rightarrow P(B) = 0.8$
$Probability\ that\ the\ Weatherly\ app\ predicts\ correctly \rightarrow P(C|A) = 0.7$
$Probability\ that\ the\ Weatherly\ app\ predicts\ incorrectly\ \rightarrow P(C|B) = 0.3$
$We\ need\ to\ find\ the\ probability\ that\ it\ will\ rain\ tomorrow?\ \rightarrow P(A|C) = ?$

$By\ Bayes'\ rule,$

$$P(A|C) = \frac{P(A)P(C|A)}{P(A).P(C|A) + P(B).P(C|B)}$$

$$P(A|C) = \frac{0.2 * 0.7}{(0.2 * 0.7) + (0.8 * 0.3)}$$

$$P(A|C) = \frac{0.14}{0.24 + 0.14}$$

$$P(A|C) = 0.3684$$

There is 36.84% chance of raining tomorrow.

2 Probability Distributions

Given the following probability density function (PDF) of a random variable x . . .

$$p(x) = \begin{cases} 4x & 0 \leq x \leq \frac{1}{2} \\ -4x + 4 & \frac{1}{2} \leq x \leq 1 \end{cases}$$

What is the equation and graph of the corresponding cumulative density function (CDF)?

Solution:

Case 1: $\mathbf{0 \leq x \leq \frac{1}{2}}$

$$F(x) = \int_0^x 4x.dx$$

$$F(x) = \frac{4x^2}{2}\Big|_0^x$$

$$F(x) = 2x^2$$

Case 2: $\mathbf{\frac{1}{2} \leq x \leq 1}$

$$F(x) = Integral\ of\ second\ equation + Case\ 1\ equation\ substituting\ x\ =\ \frac{1}{2}$$

$$F(x) = \int_{\frac{1}{2}}^x (-4x + 4).dx + \frac{1}{2}$$

$$F(x) = (-\frac{4x^2}{2} + 4x)\Big|_{\frac{1}{2}}^x$$

$$F(x) = 4x - 2x^2 - 1$$

Substituting x in Case 2 equation with maximum value 1, we get

$$F(x) = 1$$

The CDF equation is

$$F(x) = \begin{cases} 0 & x < 0 \\ 2x^2 & 0 \leq x \leq \frac{1}{2} \\ 4x - 2x^2 - 1 & \frac{1}{2} \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

3 Discrete Expectation

Calculate the expected value of X, E[X], where X is a random variable representing the outcome of a roll of a trick die. Use the sample space x $\in \{1,2,3,4,5,6\}$ (i.e. six-sided die) and let

$$P(X=x)=\begin{cases}\frac{1}{2} & x=1\\ \frac{1}{10} & x\neq 1\end{cases}$$

Solution:

Based on the given P(X=x), we can calculate expected value E[X] as below:

$$E(X)=P_1.x_1+P_2.x_2+P_3.x_3+P_4.x_4+P_5.x_5+P_6.x_6$$

$$E(X)=1*\frac{1}{2}+2*\frac{1}{10}+3*\frac{1}{10}+4*\frac{1}{10}+5*\frac{1}{10}+6*\frac{1}{10}$$

$$\mathbf{E(X)\ =\ 2.5}$$

4 Expectation Properties

Use the properties of expectation to show that we can rewrite the variance of a random variable X . . .
$Var[X] = E[(X - \mu)^2]$
as
$Var[X] = E[X^2] - (E[X])^2$

Solution:

By definition, $Var[X] = E[(X - \mu)^2]$

$$Var[X] = E[(X - \mu)^2]$$
$$Var[X] = E[X^2 + \mu^2 - 2\mu X]$$

We know that $\mu = E[X]$,

$$Var[X] = E[X^2 + \mu^2 - 2E[X^2]]$$
$$Var[X] = E[X^2] - \mu^2$$
$$Var[X] = E[X^2] - (E[X])^2$$

5 Matrices/Linear Equations

Consider the following system of equations...

$$2x_1 + x_2 + x_3 = 3$$

$$4x_1 + 2x_3 = 10$$

$$2x_1 + 2x_2 = -2$$

Solution:

a. **Write the system as a matrix equation of the form Ax = b.**

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix} . \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 10 \\ -2 \end{bmatrix}$$

b. **Write the solution of the system as a column s and verify by matrix multiplication that As = b.**
Solving the given equations,

$$s = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$$

Verifying s by checking if As = b,

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix} . \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 10 \\ -2 \end{bmatrix}$$

$$\begin{bmatrix} 2 - 2 + 3 \\ 4 + 6 \\ 2 - 4 \end{bmatrix} . = \begin{bmatrix} 3 \\ 10 \\ -2 \end{bmatrix}$$

L.H.S. = R.H.S

c. **Write b as a linear combination of the columns in A.**

$$x_1 . \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix} + x_2 . \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + x_3 . \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 10 \\ -2 \end{bmatrix}$$

6 Matrices

Consider the following matrix . . .

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{pmatrix}$$

Solution:

a. **What is the determinant, det(A) or $|A|$, of the matrix?**

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$det(A) = a_{11}.det(\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}) - a_{12}.det(\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}) + a_{13}.det(\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix})$$

$$det(A) = a_{11}.det(\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}) - a_{12}.det(\begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix}) + a_{13}.det(\begin{pmatrix} 1 & 4 \\ 1 & 3 \end{pmatrix})$$

$$det(A) = (1.(16 - 9)) - (2 * (4 - 3)) + (3 * (3 - 4))$$

$$det(A) = 7 - 2 - 3$$

$$\mathbf{det(A) = 2}$$

b. **Is the matrix invertible?**
The given matrix A is invertible, as it has full rank in its echelon form and the det(A) is not zero.

c. **What is the rank of the matrix?**
By finding the echelon matrix of A, we can find the maximum number of linearly independent row/column vectors which is the rank of A.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

By above result, we can say 3 is the rank of the matrix.

7 Eigenvalues/Eigenvectors

The eigenvalues of the following matrix . . .

$$A = \begin{pmatrix} 3 & 6 \\ 1 & 4 \end{pmatrix}$$

are λ=6 and λ=1. Which of the following is an eigen vector for λ=1?

a. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

b. $\begin{pmatrix} -3 \\ 1 \end{pmatrix}$

c. $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$

d. $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Solution:

An eigen vector should satisfy the equatio Ax = λ x

$$\begin{pmatrix} 3 & 6 \\ 1 & 4 \end{pmatrix} . \begin{pmatrix} x_1 \\ x2 \end{pmatrix} = \lambda . \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

When $\lambda = 1$,

$$\begin{pmatrix} 3.x_1 + 6.x_2 \\ 1.x_1 + 4.x_2 \end{pmatrix} = 1. \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Solving the above, we get $x_1 = -3.x_2$. Only option (b) satisfies the above equation and hence is an eigen vector for A when $\lambda = 1$.

8 Norms

Find the 0, 1, 2, and ∞ norms of x ...

$$x = \begin{pmatrix} 2 \\ 1 \\ -4 \\ -2 \end{pmatrix}$$

Solution:

0 norm: The zero norm of x is the number of non-zero coordinates of x.

$$||(2, 1, -4, -2)||_0 = 4$$

1 norm: 1 norm is the sum of absolute values of its components

$$||(2, 1, -4, -2)||_1 = |2| + |1| + |-4| + |-2|$$

$$||(2, 1, -4, -2)||_1 = 9$$

2 norm: 2 norm is the square root of the sum of the squares of absolute values of the components.

$$||(2, 1, -4, -2)||_2 = \sqrt{|2|^2 + |1|^2 + |-4|^2 + |-2|^2}$$

$$||(2, 1, -4, -2)||_2 = \sqrt{4 + 1 + 16 + 4}$$

$$||(2, 1, -4, -2)||_2 = 5$$

∞ norm: ∞ norm is the maximum of absolute values of its components

$$||(2, 1, -4, -2)||_\infty = max(|2|, |1|, |-4|, |-2|)$$

$$||(2, 1, -4, -2)||_\infty = 4$$

9 Naive Bayes

Your task is to implement a verbose Naive Bayes classifier. The goals for this task:

a. **Self-assess/self-teach Python comfort - this is a very useful language in the data mining/science community**

b. **Gain experience implementing a simple probabilistic model, one which is quite handy for many tasks, including spam classification and sentiment analysis**

Solution

Naive Bayes Classifier is a simple classifier which is based on the Bayes' rule. Naive Bayes assumes that the effect of the value of the parameter is independent of the other parameters.

$$P(c|x) = \frac{P(c)P(x|c)}{P(x)}$$

where,
$P(c|x) \leftarrow$ Posterior Probability
$P(x|c) \leftarrow$ likelyhood
$P(c) \leftarrow$ prior probability
$P(x) \leftarrow$ predictor prior probability