Machine Learning - Homework 0

Parita Pooj (psp2133)

September 12, 2016

Problem 1

I am Computer Vision student currently working on research in the area of Computational Imaging. I believe that Machine Learning provides a mathematical tool to channel human intuition to computation. While being widely applicable in Vision and Imaging techniques, I hope to have fun learning the math behind Machine Learning which helps me understand how and why the techniques work with more certainty.

Problem 2

(a) False

(b) False

Problem 3

(a) Rank of A is 1

(b)

$$Au + Bv = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$$

(c)

$$u^T Av = \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 20$$

(d) $\|u\|_2 = \sqrt{2^2 + 1^2} = \sqrt{5}$

(e)

$$\bigtriangledown f(x) = 2(A+B)x = 2\begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix} x$$

$$\bigtriangledown f(v) = 2\begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 18 \\ 36 \end{bmatrix}$$

(f) Minimize $x^T(A+B)x$ such that $\|x\|_2 = 1$
Minimize $L(x,\lambda) = x^T(A+B)x - \lambda(x^T x - 1)$
$\frac{\partial L}{\partial x} = (A+B)x - \lambda x = 0$
$(A+B)x = \lambda x$
Eigenvector for $(A+B)$ gives the value of x which minimizes the function $f(x)$
Eigenvalues for $(A+B)$ are 4 and 9. The lowest eigenvalue minimizes the function. The Eigenvector of unit length corresponding the eigenvalue 4 is $\begin{bmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$
The value of f evaluated at this vector is 4.

Problem 4

(a) A and B are independent. The rest of the pairs are dependent.

(b)

$$\frac{P(CU\ \cap\ TU)}{P(TU)} = \frac{0.2}{0.5} = 0.4$$

(c) P(X = 1) = 11/36
P(X = 2) = 9/36
P(X = 3) = 7/36
P(X = 4) = 5/36
P(X = 5) = 3/36
P(X = 6) = 1/36

(d) E(X) = 91/36 = 2.528

(e) E(number of tosses until head come up) $= \sum_{n=1}^{\infty} n\frac{1}{5}(\frac{4}{5})^{n-1} = 5$

(f) E(number of times the phrase appears in the sentence) = P(phrase matches starting at position 1) + P(phrase matches starting at position 2) + ... + P(phrase matches starting at position n - 4)
= (n - 4) P(phrase matching at position 1)
= (n - 4) $(\frac{1}{4})^5$

Problem 5

(a) $\lambda = 6.93 \times 10^{-7}$

(b) E(X) = Mean of the Gaussian Probability Distribution = 0
E(Y) = Variance of the Gaussian Probability Distribution = 1

(c)

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2 = 1$$

$$\int_{0}^{0.5} \int_{0}^{1} c\, \mathrm{d}x_1 \mathrm{d}x_2 = 1$$

$$c = 2$$

(d)

$$P(X_2 \geq X_1) = \int_{0}^{0.5} \int_{x_1}^{1} c\, \mathrm{d}x_1 \mathrm{d}x_2 = \int_{0}^{0.5} 2(1-x_1)\mathrm{d}x_1 = (2x_1 - x_1^2)\Big|_{0}^{0.5} = \frac{3}{4}$$

(e) X_1 and Y are not independent.
The probability space for Y (same as $p(x_1, x_2)$ is divided by the line $X_1 - 2X_2 = 0$.

$$E(Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y.p(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2$$

$$E(Y) = \int_{0}^{0.5} \int_{0}^{\frac{x_1}{2}} 2\, \mathrm{d}x_1 \mathrm{d}x_2 - \int_{0}^{0.5} \int_{\frac{x_1}{2}}^{1} 2\, \mathrm{d}x_1 \mathrm{d}x_2$$

$$E(Y) = \frac{1}{8} - \frac{7}{8} = -\frac{3}{4}$$

(f) X_1 and Z are independent. $E(X_1 Z) = E(X_1)E(Z)$
E(Z) = 1/2 - 1/2 = 0
Thus, $E(X_1 Z) = 0$

References

[1] http://cs229.stanford.edu/section/cs229-linalg.pdf

[2] http://cs229.stanford.edu/section/cs229-prob.pdf

[3] Nayar, Shree. "Camera Calibration" Computer Vision(COMSW4731). Columbia University, New York. 10-30-2015. Lecture.