Mathematical Background for ML - Solutions

Nelson Kshetrimayum

August 2021

1 Minimum Background Test

1.1 Vectors and Matrices

1. $y.z = \begin{bmatrix}1\\3\end{bmatrix} \cdot \begin{bmatrix}2\\3\end{bmatrix} = 1 \times 2 + 3 \times 3 = 11$

2. $Xy = \begin{bmatrix}2 & 4\\1 & 3\end{bmatrix} \cdot \begin{bmatrix}1\\3\end{bmatrix} = \begin{bmatrix}2 \times 1 + 4 \times 3\\1 \times 1 + 3 \times 3\end{bmatrix} = \begin{bmatrix}14\\10\end{bmatrix}$

3. $\left|X\right| = \begin{vmatrix}2 & 4\\1 & 3\end{vmatrix} = 6 - 4 = 2 \Rightarrow$ Inverse exist.
   $X^{-1} = \frac{adjX}{|X|} = \frac{1}{2}\begin{bmatrix}3 & -1\\-4 & 2\end{bmatrix}$

4. X is of rank 2 as it has two linearly independent columns (or rows).

1.2 Calculus

1. $\frac{dy}{dx} = 3x^2 + 1$

2. $d_{x_1}f = sin(x_2)e^{-x_1}(1 - x_1)$ & $d_{x_2}f = x_1e^{-x_1}cos(x_2)$
   Thus $\Delta f(x) = e^{-x_1}\begin{pmatrix}sin(x_2)(1 - x_1)\\x_1cos(x_2)\end{pmatrix}$

1.3 Probability and Statistics

1. $\bar{x} = \frac{\sum x_i}{n} = \frac{3}{5}$

2. $\sigma^2 = \frac{\sum(x - \bar{x})^2}{N - 1} = \frac{2.\frac{9}{25} + 3.\frac{4}{25}}{4} = \frac{3}{10}$

3. $\frac{1}{2}.\frac{1}{2}.\frac{1}{2}.\frac{1}{2}.\frac{1}{2} = \frac{1}{32}$

4. If x is the Pr(head) then probability of the given sample to be obtained is $X_0 = x^3(1-x)^2$. Maxima of the expression is the value of x at which $\dfrac{dX_0}{dx}$ is 0. By calculating, x is 3/5 (or 0.6).

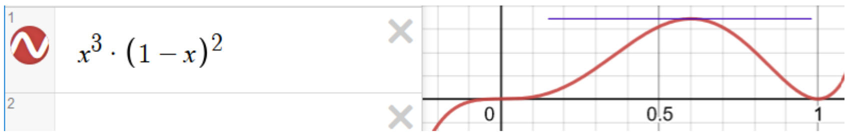


Figure 1: Plot showing maxima of X_0

5.
   - p(z = T AND y = b) = 0.1
   - p(z = T|y = b) = 0.2 + 0.1 + 0.2 + 0.15 = 0.65

1.4 Big-O Notation

1. Both are true. ln can be expressed in log along with a constant factor.
2. $g(n) = O(f(n))$ is true. g is an exponential and f is a polynomial.
3. $g(n) = O(f(n))$ is true. $3^n = 2^{nlog_2 3}$.
4. $f(n) = O(g(n))$ is true. g is a degree 3 polynomial while f is degree 2.

2 Medium Background Test

2.1 Algorithms

1. We can run an algorithm that count the index from the start and continue with the next one. If the next element in the array is 1 then return the count else continue. This has a complexity of $O(n)$

2.2 Probability and Random Variables

2.2.1 Probability

(a) False

(b) True

(c) False

(d) False

(e) True

2.2.2 Discrete and Continuous Distributions

RHS matches LHS in the order

1. Bernoulli
2. Uniform
3. Binomial
4. Multivariate Gaussian

2.2.3 Mean, Variance and Entropy

(a)

$$
\begin{aligned}
Var(X) &= E((X - E(X))^2) \\
&= E(X^2 - 2XE(X) + E(X)^2) \\
&= E(X^2) - 2E(X)E(X) + E(X)^2 \\
&= E(X^2) - E(X)^2
\end{aligned}
$$

(b)
- Mean = p
- Variance = p(1-p)
- Entropy = $-p.lnp - (1-p)ln(1-p)$

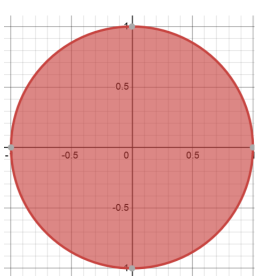
2.2.4 Law of Large Numbers and Central Limit Theorem

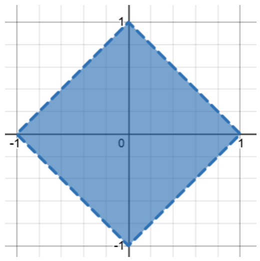
(a) The probability of getting a particular number (say 3) in rolling a dice is $\frac{1}{6}$. If we interpret what it means, it is that if we roll the dice 6 times we are very likely to get at least one 3. Law of large number states that if we extend is procedure numerous times (say 6000), number of times 3 shows up will be close to $\frac{1}{6} \times 6000 = 1000$.

2.3 Linear Algebra

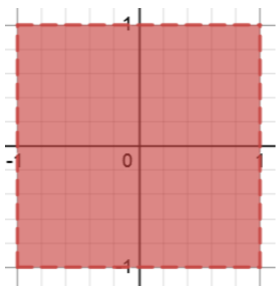
2.3.1 Vector norms

(a)





(b)



(c)