MCGILL UNIVERSITY

FALL SEMESTER, 2024
Campus: Montreal - Downtown

COMPUTER SCIENCE

Test of Prerequisites for Applied Machine Learning Course

(Time allowed: TWO hours)

NOTE: This is a practice quiz to check the required background for this course. The equations here are recurring blocks throughout the lectures, and not being comfortable with them will make following the lectures very difficult.

1. For which values of x, the function $x^3 - 6x^2 + 9x + 15$ has a local minimum?

(a) 1

(b) 3

(c) $1, 3$

(d) $1, 0, 3$

(e) $1, 2$

2. What is the derivative of e^x?

(a) $ln(x)$

(b) e^x

(c) e^{-x}

(d) $e^x \ln(x)$

3. What is the derivative of $\log(x^y(1-x)^{(1-y)})$ with respect to x?

(a) $\frac{y}{x} - \frac{1-y}{1-x}$

(b) $\frac{y}{x} + \frac{1-y}{1-x}$

(c) $y\log(x) + (1-y)\log(1-x)$

(d) $y\log(x) - (1-y)\log(1-x)$

4. What is the derivative of the function $y = \frac{1}{1+e^{-x}}$?

(a) $(1-y)\ln(y)$

(b) $y(1+y)$

(c) $y(1-y)$

(d) $y\ln(y)$

5. If $\sigma(x) = \frac{1}{1+e^{-x}}$, what is the derivative of the function $\tanh(x) = 2\sigma(2x) - 1$?

(a) $\tanh(x)(1+\tanh(x))$

(b) $1+\tanh^2(x)$

(c) $1-\tanh^2(x)$

(d) $\tanh(x)(1-\tanh(x))$

(e) $4\tanh(x)(1-\tanh(x))$

(f) $(\tanh(x)+1)^2$

(g) $(\tanh(x)-1)^2$

6. Consider $\mathbf{x_1} = \left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right)$ and $\mathbf{x_2} = \left(\begin{smallmatrix}0\\1\\1\end{smallmatrix}\right)$, what is the acute angle between these two vectors?

(a) $\pi/2$

(b) $\pi/4$

(c) 0

(d) $\pi/3$

7. Consider $\mathbf{x_1} = \left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right)$ and $\mathbf{x_2} = \left(\begin{smallmatrix}0\\1\\1\end{smallmatrix}\right)$, what is the outer product of these two vectors and its trace, i.e.$\mathbf{x_1} \otimes \mathbf{x_2}$, & $tr(\mathbf{x_1} \otimes \mathbf{x_2})$?

(a) $\left(\begin{smallmatrix}0&1&1\\0&1&1\\0&0&0\end{smallmatrix}\right)$, 1

(b) $\left(\begin{smallmatrix}0&1&1\\0&1&1\\0&0&0\end{smallmatrix}\right)$, 0

(c) $\left(\begin{smallmatrix}0&1&0\\0&1&0\\0&0&0\end{smallmatrix}\right)$, 0

(d) $\left(\begin{smallmatrix}0&1&0\\0&1&0\\0&1&0\end{smallmatrix}\right)$, 1

8. What is the value of $(y - \mathbf{w}^T\mathbf{x})^2$ if $\mathbf{w} = \begin{pmatrix}1\\1\\0\end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix}1\\1\\1\end{pmatrix}$ and $y = 0$?

(a) 0

(b) 1

(c) 4

(d) 2

(e) $\sqrt(2)$

9. What is the gradient of $f(\mathbf{x}) = (\mathbf{w}^T\mathbf{x} + w_0)^2$ with respect to $\mathbf{x}$, i.e. $\nabla_{\mathbf{x}} f(\mathbf{x})$? Here, $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^n$ are both n dimensional vectors of real values and $w_0 \in \mathbb{R}$ is a scalar.

(a) $2(\mathbf{w}^T\mathbf{x} + w_0)$

(b) $2\mathbf{w}^T\mathbf{x}$

(c) $2(\mathbf{w}^T\mathbf{x} + w_0)\mathbf{w}$

(d) $2(\mathbf{w}^T\mathbf{x} + w_0)\mathbf{x}$

(e) $2(\mathbf{w}^T\mathbf{w} + w_0)\mathbf{x}$

(f) $2(\mathbf{x}^T\mathbf{x} + w_0)\mathbf{w}$

(g) $2(\mathbf{w}^T\mathbf{w})\mathbf{x}$

(h) $2(\mathbf{w}^T\mathbf{x})\mathbf{w}$

10. What is $\frac{dJ}{d\mathbf{w}}$, the derivative of
$J(\mathbf{w}) = (y_1 - \mathbf{w}^T\mathbf{x}_1)^2 + (y_2 - \mathbf{w}^T\mathbf{x}_2)^2$ with respect to $\mathbf{w}$? Here, $\mathbf{w} \in \mathbb{R}^n$ is a n dimensional vector of real values, same as $\mathbf{x}_1$ and $\mathbf{x}_2$; and $y_1 \in \mathbb{R}$ is a scalar value, same as y_2.

(a) $2(y_1 - \mathbf{w}^T\mathbf{x}_1) + 2(y_2 - \mathbf{w}^T\mathbf{x}_2)$

(b) $2(y_1 - \mathbf{w}^T\mathbf{x}_1)\mathbf{w} + 2(y_2 - \mathbf{w}^T\mathbf{x}_2)\mathbf{w}$

(c) $2(y_1 - \mathbf{w}^T\mathbf{x}_1)\mathbf{x}_1 + 2(y_2 - \mathbf{w}^T\mathbf{x}_2)\mathbf{x}_2$

(d) $2(\mathbf{w}^T\mathbf{x}_1)\mathbf{x}_1 + 2(\mathbf{w}^T\mathbf{x}_2)\mathbf{x}_2$

(e) $2(\mathbf{w}^T\mathbf{x}_1)y_1 + 2(\mathbf{w}^T\mathbf{x}_2)y_2$

(f) $2(\mathbf{w}^T\mathbf{x}_1 - y_1) + 2(\mathbf{w}^T\mathbf{x}_2 - y_2)$

(g) $2(\mathbf{w}^T\mathbf{x}_1 - y_1)\mathbf{w} + 2(\mathbf{w}^T\mathbf{x}_2 - y_2)\mathbf{w}$

(h) $2(\mathbf{w}^T\mathbf{x}_1 - y_1)\mathbf{x}_1 + 2(\mathbf{w}^T\mathbf{x}_2 - y_2)\mathbf{x}_2$

(i) $2(\mathbf{w}^T\mathbf{x}_1)\mathbf{x}_1 + 2(\mathbf{w}^T\mathbf{x}_2)\mathbf{x}_2$

(j) $-2(\mathbf{w}^T\mathbf{x}_1)y_1 - 2(\mathbf{w}^T\mathbf{x}_2)y_2$

11. If an X-ray test that is used to detect lung cancer gives positive results for 90% of people who actually have cancer, and falsely gives positive results for 5% of healthy patients, what is the probability of having cancer after receiving a positive test result? given that we also know in general, 1 in 1000 people have lung cancer.

(a) 90%

(b) 18.9%

(c) 8%

(d) 9%

(e) 9.9%

(f) 8.9%

(g) 1.7%

(h) 4.7%

12. What is the expected value of $x + y$, if x and y are normally distributed random variables defined as: $x = \mathcal{N}(\mu = 3, \sigma = 4)$, and $y = \mathcal{N}(\mu = 2, \sigma = 5)$?

(a) 5

(b) 6

(c) 9

(d) 13

(e) 23

13. If you are given a ball taken at random with equal probability from either bag 1 or bag 2, where before the draw, bag 1 had 3 white and 2 red balls and bag 2 had 4 white and 3 red. Know that a red ball was drawn what is the probability of the ball being from bag 1?

(a) $\frac{7}{29}$

(b) $\frac{1}{2}$

(c) $\frac{2}{5}$

(d) $\frac{1}{5}$

(e) $\frac{7}{58}$

(f) $\frac{29}{35}$

(g) $\frac{14}{29}$

14. $p(x_1, x_2, \ldots x_n|y)$ is equal to :

(a) $\Pi_{i=1}^{n} p(x_i, y)P(x_i)$

(b) $\Pi_{i=1}^{n} p(x_i|y)$

(c) $p(y)\Pi_{i=1}^{n} p(x_i)$

(d) $p(x_1|y)p(x_2|y, x_1)p(x_3|y, x_1, x_2) \ldots$

15. If X is a random variable which takes values in $\{0, 1, 2, 3, 4, 5, 6, 7\}$ where $P(X = x) = \binom{7}{x} 0.25^x 0.75^{7-x}$, what is expected value of X?

(a) 7/5

(b) 28/5

(c) 1/4

(d) 7/4

16. For a bent coin that gives a head 20% of the time, what is the mean and variance of the probability of observing a tail in independent trials?

(a) $4/5, 4/25$

(b) $1/5, 1/5$

(c) $1/4, 1/16$

(d) $4/5, 16/25$

17. With a bent coin that gives a head 20% of the time, what is the probability of getting 3 tails in a row?

(a) 0.512

(b) 0.008

(c) 0.8

(d) 0.08

18. What is the probability of a fair 6-face dice giving 3 rolls with numbers that sum up to 5 or less?

(a) 7/216

(b) 5/72

(c) 10/216

(d) 1/36

19. Which is true for logarithm function?

(a) $\log(ab) = log(a)log(b)$

(b) $\log(ab) = log(a) + log(b)$

(c) $\log(a + b) = log(a)log(b)$

(d) $\log(a + b) = log(a)^{log(b)}$

20. What is the entropy (in bits) of a Bernoulli variable with mean of 0.25?

(a) 0.5

(b) 0.56

(c) 0.24

(d) 0.75

(e) 0.81

21. For which value of x in the interval [0, 0.99] does the function $f(x) = -x \ln x$ reach a minimum?

(a) 0

(b) e^{-1}

(c) e

(d) 0.5

(e) 0.99

(f) 1

22. Let M be a $D \times K$ matrix, N be a $K \times F$ matrix, and W be a $D \times L$ matrix. What is are the values of$(M \cdot N)_{nm}$ and $(W^\top \cdot M \cdot N)_{ij}$, i.e. the (i, j) element of matrix $W^\top \cdot M \cdot N$ where $M \cdot N$ is the matrix multiplication of M and N, and $A^\top$ denotes the transpose of matrix A? Write your answer as function of the indices of the individual matrices, e.g. $\sum_d M_{di} N_{dj}$.

$(M \cdot N)_{nm} =$

$(W^\top \cdot M \cdot N)_{ij} =$