Course: Machine Learning - Foundations
Test Questions
Lecture Details: Week 10

---

1. (1 point) For a given point $x = 2$, is the function $f(x) = xe^{-3x}$ increasing or decreasing?
   A. increasing
   B. decreasing

   Answer: B

   If the first derivative of $f(x)$ at any point x is positive or negative, then we call $f(x)$ as increasing or decreasing respectively.
   $f'(x) = e^{-3x}(1 - 3x)$, at $x = 2$ we get $f'(2) = -0.012$. Hence decreasing.

2. (1 point) Find the value of a function$f(x) = x + 3x^2$ at its global minimum point.

   Answer: -0.0833

   range: -0.09,-0.08

   To get global minimum, $f'(x) = 0$. Therefore $1 + 6x = 0$, $x = -\frac{1}{6}$.

   $$f(\frac{1}{6}) = -0.0833$$

3. (1 point) Choose the largest interval of x in which a function $f(x) = xe^{x^2}$ is convex.
   A. (-0.5,0.5)
   B. $(0, \infty)$
   C. $(\infty,0)$
   D. (1,0)

   Answer: B

   In order to find the interval over which $f(x)$ is convex, let us find where $f''(x) > 0$.

   $$f'(x) = 2x^2 e^{x^2}$$
   $$f''(x) = e^{x^2}(4x^3 + 6x)$$

   Therefore $f''(x) > 0$ implies $e^{x^2}(4x^3 + 6x) > 0$.
   To satisfy this inequality, we want an interval of x for which $e^{x^2} > 0$ and $(4x^3 + 6x) > 0$.
   Because e raised to any power will be positive, first condition can be satisfied for any value of x.
   Second condition can be written as $2x(x^2 + 3) > 0$.
   $(x^2 + 3)$ will be greater than 0 for any values of x.
   $2x$ will be positive only when $x > 0$.
   Thus in interval notation, the largest interval of x for which $f(x)$ is convex is $(0, \inf)$.

4. (1 point) (Multiple select) Let the composition of two functions $f(x) = sin(x) - 2x^2 + 1$ and $g(x) = e^x$ be $h = fog$. At a point $x = 5$, Select the true statement(s).

    A. $h(x)$ is a convex function.

    B. $h(x)$ is a concave function.

    C. $h(x)$ is a non-decreasing function.

    D. $h(x)$ is a decreasing function.

Answer: B,D

$$h(x) = f(g(x)) = sin(e^x) - 2e^{2x} + 1$$

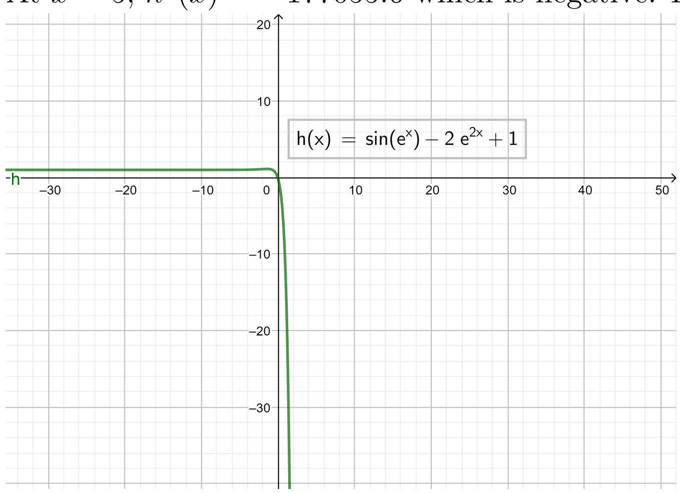
First derivative:

$$h'(x) = e^x cos(e^x) - 4e^{2x}$$

At $x = 5$; $h'(x) = -88232.28$ which is less than 0. Hence h(x) is decreasing.
Second derivative:

$$h''(x) = e^{2x}(cos(e^x) - sin(e^x)) - 8e^{2x}$$

At $x = 5$; $h''(x) = -177055.6$ which is negative. Hence h(x) is concave.



5. (2 points) Find the minimum value of $f(x, y) = x + y$ subject to $x^2 + y^2 = 1$, where x, y are the coordinates of points on the circumference of the unit circle.

Answer: -1.414

range: -1,-2 The Lagrangian function for the problem is

$$L(x, y, \lambda) = x + y + \lambda x^2 + \lambda y^2 - \lambda$$

Take partial derivatives with respect to x,y, λ and equate to zero.

$$\frac{\partial L}{\partial x} = 2\lambda x + 1 = 0$$

$$x = -\frac{1}{2\lambda} \tag{1}$$

$$\frac{\partial L}{\partial y} = 2\lambda y + 1 = 0$$

$$y = -\frac{1}{2\lambda} \tag{2}$$

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 = 1 \tag{3}$$

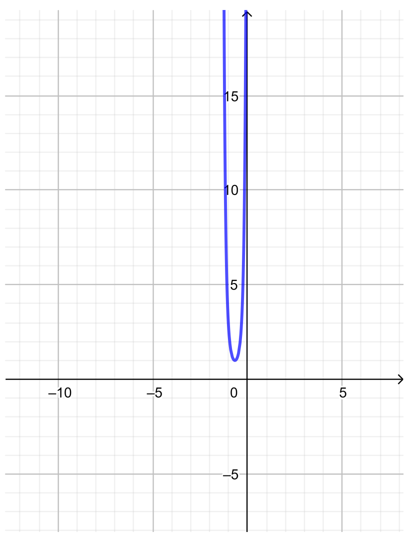
Substituting (1) and (2) in (3), we get

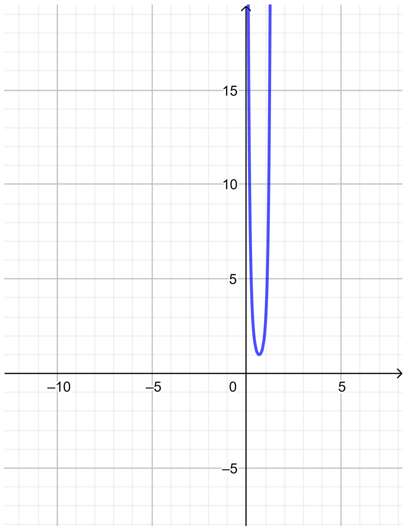
$$\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} = 1$$

$$\lambda = \pm\frac{1}{\sqrt{2}} \tag{4}$$

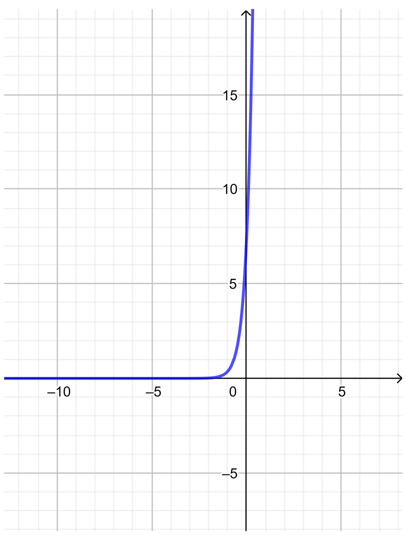
Using (3) in (1) and (2) we get, $x = \mp\frac{1}{\sqrt{2}}$ and $y = \mp\frac{1}{\sqrt{2}}$. Since we want to minimize $f(x, y)$, we shall consider $x = -\frac{1}{\sqrt{2}}$ and $y = -\frac{1}{\sqrt{2}}$.
Minimum value of $f(x, y) = -1.414$.

6. (1 point) For the functions $g(x) = (3x + 2)^2$ and $f(x) = e^x$, select the plot that corresponds to the correct composition $h = fog$.

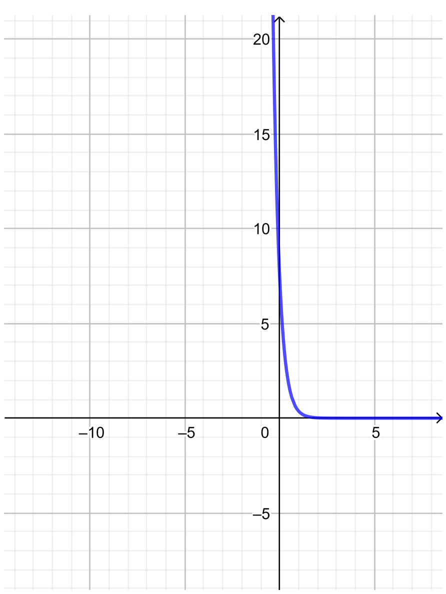
A. 



B.



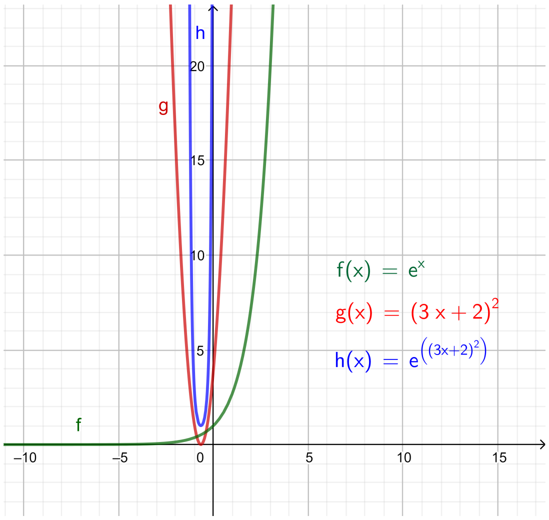
C.



D.

Answer: A

The functions f(x), g(x) and their composition h(x) are plotted as follows:



(Common data for Q7, Q8) Linear programming deals with the problem of finding a vector x that minimizes a given linear function $c^T x$, where x ranges over all vectors $(x \geq 0)$ satisfying a given system of linear equations $Ax = b$. Here A is a $m \times n$ matrix, $c, x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$.

7. (1 point) Choose the correct dual program with y as the dual variable for the above linear program from the following.

   A. $$\min_y by \quad \text{subject to} \quad A^T y \geq c$$

   B. $$\max_y b^T y \quad \text{subject to} \quad A^T y \leq c$$

   C. $$\max_y b^T y \quad \text{subject to} \quad A^T y \geq c$$

   D. $$\max_y by \quad \text{subject to} \quad A^T y \leq c$$

   Answer: B

8. (1 point) From the below given statements regarding constraints and decision variables related to the primal and dual problems of the linear program, choose the correct statement.

   A. Primal problem has m constraints and m decision variables whereas dual problem has n constraints and n decision variables.

   B. Primal problem has m constraints and n decision variables whereas dual problem has n constraints and m decision variables.

C. Primal problem has n constraints and n decision variables whereas dual problem has m constraints and m decision variables.

D. Primal problem has n constraints and m decision variables whereas dual problem has m constraints and n decision variables.

Answer: B

9. (2 points) Let a set of data points with five samples and two features per sample be $X = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 2.5 \\ 6 & 4 \\ 7.5 & 5 \end{bmatrix}$ and the corresponding labels be $y = \begin{bmatrix} 1.5 \\ 2 \\ 2.5 \\ 3 \\ 4 \end{bmatrix}$. Perform linear regression on this data set and choose the optimal solution for w^* to minimize the sum of squares error.

A. $\begin{bmatrix} 0.2763 \\ 1.2039 \end{bmatrix}$

B. $\begin{bmatrix} 0.0691 \\ 0.3010 \end{bmatrix}$

C. $\begin{bmatrix} 0.1382 \\ 0.6019 \end{bmatrix}$

D. $\begin{bmatrix} 0.0276 \\ 0.1204 \end{bmatrix}$

Answer: C

optimal $w^* = (X^T X)^{-1}(X^T y)$

$X^T X = \begin{bmatrix} 113.25 & 79.5 \\ 79.5 & 60.25 \end{bmatrix}$

$X^T y = \begin{bmatrix} 63.5 \\ 47.25 \end{bmatrix}$

$w^* = \begin{bmatrix} 0.1382 \\ 0.6019 \end{bmatrix}$

A, B and D are scalar multiples of C.

(Common data for Q10, Q11) Krishna runs a steel fabrication industry and produces steel products. He regularly purchases raw steel for Rs.500 per ton. His revenue is modeled by a function $R(s) = 100\sqrt{s}$, where s is the tons of steel purchased. His budget for steel purchase is Rs.150000.

10. (1 point) Using Lagrangian function, find the amount of raw steel to be purchased to get maximum revenue?

Answer: 300

range: 298, 302

Lagrangian function: $L(s, \lambda) = 100s^{0.5} + \lambda(s - 150000)$

$$L(s, \lambda) = 100s^{0.5} + s\lambda - \lambda 150000$$

Take partial derivatives with respect to s and λ and equate to zero.

$$\frac{\partial L}{\partial s} = 0 = 50s^{-0.5} + \lambda 500 = 0$$

$$\frac{\partial L}{\partial \lambda} = 0 = s500 - 150000 = 0$$

$$s = 300$$

11. (1 point) What is the maximum revenue value in Rs?

Answer: 1732.05

range: 1730,1734

Maximum revenue: $100 * \sqrt{300} = 1732.05$

12. (1 point) Consider a vector $\hat{w} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$ in $\mathbb{R}^3$. In $\mathbb{R}^3$, there are many unit vectors. Use Lagrange method to find the unit vector which gives the minimum dot product.

A. $\hat{u} = \frac{1}{2\lambda} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$, with $\lambda \geq 0$

B. $\hat{u} = \frac{-1}{3\lambda} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$, with $\lambda \geq 0$

C. $\hat{u} = \frac{-1}{4\lambda} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$, with $\lambda \geq 0$

D. $\hat{u} = \frac{-1}{2\lambda} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$, with $\lambda \geq 0$

Answer: D

Let unit vector be $\vec{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

Objective is to minimize $\vec{u}.\vec{w}$ subject to $x^2 + y^2 + z^2 = 1$.

$$f(x, y, z) = \vec{u}.\vec{w} = 2x + 4y + 3z$$

Lagrangian function can be written as follows:

$$L(x, y, z, \lambda) = (2x + 4y + 3z) + \lambda(x^2 + y^2 + z^2 - 1)$$

Take partial derivatives with respect to x, y, z and λ and equate to zero.

$$\frac{\partial L}{\partial x} = 2 + 2\lambda x = 0$$

$$x = \frac{-1}{2\lambda}2$$

$$\frac{\partial L}{\partial y} = 4 + 2\lambda y = 0$$

$$y = \frac{-1}{2\lambda}4$$

$$\frac{\partial L}{\partial z} = 3 + 2\lambda z = 0$$

$$z = \frac{-1}{2\lambda}3$$

(Common data for Q13, Q14) Solve the following linear program using KKT conditions.

$$\text{minimize} \quad v = 24y_1 + 60y_2$$

subject to

$$0.5y_1 + y_2 \geq 6$$

$$2y_1 + 2y_2 \geq 14$$

$$y_1 + 4y_2 \geq 13$$

$$y_1 \geq 0, y_2 \geq 0$$

13. (2 points) Choose the optimal solution for $[y_1^*, y_2^*]$

    A. [8,2.5]

    B. [10,6]

    C. [11,0.5]

    D. [10.5,1]

    Answer: C

    ]$[y_1^*, y_2^*] = [11, 0.5]$, satisfies all the constraints and gives minimum value of v.

14. (1 point) Enter the minimum value of v.

    Answer: 294

$$v = 24(11) + 60(0.5) = 294$$