

Course > Week 5... > Proble... > Proble...

Problem Set 5

Problems 1-4 correspond to "Unconstrained optimization I"

Problem 1

1/1 point (graded)

Let F be a function from \mathbb{R}^d to \mathbb{R} . Which of the following is the most accurate description of the derivative ∇F ?

- lt is a real number.
- It is a d-dimensional vector.
- For any point $u \in \mathbb{R}^d$, the derivative at that point, $\nabla F(u)$, is a real number.
- lacktriangle For any point $u \in \mathbb{R}^d$, the derivative at that point, $\nabla F(u)$, is a d-dimensional vector.



Submit

Problem 2

6/6 points (graded)

Consider the following loss function on vectors $w \in \mathbb{R}^3$:

$$L(w) = w_1^2 - 2w_1w_2 + w_2^2 + 2w_3^2 + 3.$$

a) Compute $\nabla L(w)$. Match each of its coordinates to the following list:

\circ	ption	1:	$4w_2$
\circ	puon	٠.	TW 2

Option 2:
$$2w_1 - 2w_2$$

Option 3:
$$-2w_1 + 2w_2$$

What is dL/dw_1 ? (Just answer 1,2,or 3)



$$dL/dw_2 =$$



$$dL/dw_3 =$$



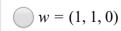
b) What is the minimum value of L(w)?



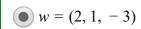
c) Is there is a unique solution w at which this minimum is realized?



d) Suppose we use gradient descent to minimize this function, and that the current estimate is w = (1, 2, 3). If the step size is $\eta = 0.5$, what is the next estimate?







$$w = (0, -1, -1.5)$$



Submit

Problem 3

1/1 point (graded)

We are given a set of data points $x^{(1)}, ..., x^{(n)} \in \mathbb{R}^d$, and we want to find a single point $z \in \mathbb{R}^d$ that minimizes the loss function

$$L(z) = \sum_{i=1}^{n} \|x^{(i)} - z\|^{2}.$$

Use calculus to determine z, in terms of the $x^{(i)}$. (Hint: It might help to just start by looking at one particular coordinate.) Then select which of the following correctly describes the solution.

- \bigcirc The sum of the $x^{(i)}$ vectors
- The average of the $x^{(i)}$ vectors
- The average of the $x^{(i)}$ vectors, times a constant $c \neq 1$
- \bigcirc Zero, regardless of what the $x^{(i)}$ vectors are



Submit

Problem 4

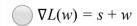
2/2 points (graded)

Given a set of data points $x^{(1)}, ..., x^{(n)} \in \mathbb{R}^d$, we want to find the vector $w \in \mathbb{R}^d$ that minimizes this loss function:

$$L(w) = \sum_{i=1}^{n} (w \cdot x^{(i)}) + \frac{1}{2} c \|w\|^{2}.$$

Here c > 0 is some constant.

a) Let s denote the sum of the data points, that is, $s = \sum_{i=1}^{n} x^{(i)}$. Express $\nabla L(w)$ in terms of *s*, *c*, and *w*.





Correct: The derivative is $\nabla L(w) = \sum_{i} x^{(i)} + cw = s + cw$

b) What value of w minimizes L(w)? Give the answer in terms of s and c.

$$w = -\frac{s}{c}$$

$$w = cs$$

$$\bigcirc_{w} = \frac{s}{4c}$$

$$w = -\frac{s}{2c}$$



Correct: This results from setting $\nabla L(w) = 0$.

Submit

Problems 5-7 correspond to "Convexity I"

Problem 5

7/7 points (graded)

For each of the following functions of one variable, say whether it is convex, concave, both, or neither.

$$a) f(x) = x^2$$

convex

Answer

Correct: f''(x) = 2

$$b) f(x) = -x^2$$

concave •

Answer

Correct: f''(x) = -2

c)
$$f(x) = x^2 - 2x + 1$$

convex

Answer

Correct: f''(x) = 2

$$\mathsf{d}) f(x) = x$$

both

Answer

Correct: f''(x) = 0

e)
$$f(x) = x^3$$

neither •

Answer

Correct: f''(x) = 6x, which is sometimes positive, sometimes negative.

$$f) f(x) = x^4$$

convex

Answer

Correct: $f''(x) = 12x^2$

$$g) f(x) = \ln x$$

concave



Answer

Correct: $f''(x) = -1/x^2$

Submit

Problem 6

1/1 point (graded)

Consider the function $f: \mathbb{R}^3 \to \mathbb{R}$ given by

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2 - 4x_1x_2 + 6x_2x_3.$$

Compute and select the matrix of second derivatives (the Hessian) H(x).



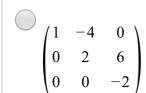
$$\begin{pmatrix} 1 & -2 & 0 \\ -2 & 1 & 3 \\ 0 & 3 & -1 \end{pmatrix}$$



$$\begin{pmatrix} 2 & -4 & 0 \\ -4 & 2 & 6 \\ 0 & 6 & 0 \end{pmatrix}$$



$$\begin{pmatrix} 2 & -4 & 0 \\ -4 & 2 & 6 \\ 0 & 6 & -2 \end{pmatrix}$$





Submit

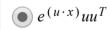
Problem 7

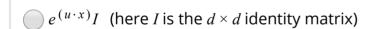
1/1 point (graded)

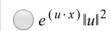
For some fixed vector $u \in \mathbb{R}^d$, define the function $F: \mathbb{R}^d \to \mathbb{R}$ by

$$F(x) = e^{u \cdot x}$$
.

Which of the following is the Hessian H(x)?







$$\bigcirc e^{(u\cdot x)}(u\cdot x)^2$$



Submit

Problems 8-11 correspond to "Positive semidefinite matrices"

Problem 8

1/1 point (graded)

Is the matrix $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ positive semidefinite?

- Yes, because every entry in the matrix is ≥ 0
- No, because not every entry is > 0
- Yes, because $u^T M u \ge 0$ for all vectors u
- \bullet No, because there is a vector u for which $u^T M u < 0$



Submit

Problem 9

1/1 point (graded)

Is the matrix $M = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ positive semidefinite?

- $\widehat{}$ No, because not every entry is $\,\geq 0$
- Yes, because $u^T M u \ge 0$ for all vectors u
- No, because there is a vector u for which $u^T M u < 0$
- No, because there is a vector u for which $u^T M u = 0$



Submit

Problem 10

1/1 point (graded)

For a fixed set of vectors $v^{(1)}, ..., v^{(n)} \in \mathbb{R}^d$, let M be the $n \times n$ matrix of all pairwise dot products: that is, $M_{ii} = v^{(i)} \cdot v^{(j)}$. Do you see why M is positive semidefinite? Think about it a little bit, and then choose one of the following options (you'll get marked as correct whichever you choose).

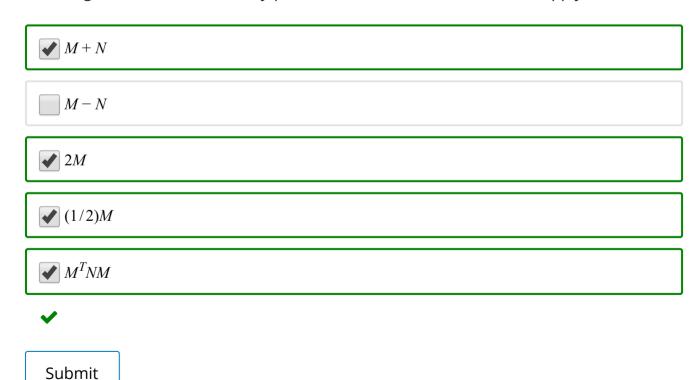
Yes, the entire argument is clear to me. That sounds right, but I can't fully construct the argument. I don't get it.

Problem 11

Submit

1/1 point (graded)

Suppose M and N are positive semidefinite matrices of the same size. Which of the following matrices are necessarily positive semidefinite? Select all that apply.



Problems 12-13 correspond to "Convexity II"

Problem 12

2/2 points (graded) For some fixed vector $u \in \mathbb{R}^d$, define

$$F(x) = \|x - u\|^2.$$

We wish to determine whether F(x) is a convex function of x.

a) The Hessian matrix H(x) is of the form cI, where I is the $d \times d$ identity matrix and c is some constant. What is *c*?

2

b) Is F(x) a convex function?



 $\overline{}$ It depends on the specific vector u

Submit

No

Problem 13

2/3 points (graded)

Let $p = (p_1, p_2, ..., p_m)$ be a probability distribution over m possible outcomes. The *entropy* of *p* is a measure of how much randomness there is in the outcome. It is defined as

$$F(p) = -\sum_{i=1}^{m} p_i \ln p_i,$$

where In denotes natural logarithm. We wish to ascertain whether F(p) is a convex function of p. As usual, we begin by computing the Hessian.

a) Consider the specific point p = (1/m, 1/m, ..., 1/m). What is the (1, 1) entry of the Hessian at this point? Your answer should be a function of m.



b) Continuing, what is the (1, 2) entry of the Hessian at this specific point?



c) Is the function F(p) convex, concave, both, or neither?



© All Rights Reserved