Course › Week 3... › Proble... › Proble...

Problem Set 3

Problems 1-7 correspond to "Linear algebra I: basic notation and dot products"

Problem 1

1/1 point (graded)

A data set consists of 200 points in $\mathbb{R}^{80}$. If we store these in a matrix, with one point per row, what is the dimension of the matrix?

- ● 200×80
- ○ 80×200
- ○ 200×1
- ○ 1×80

✓

Submit

ℹ Answers are displayed within the problem

Problem 2

3/3 points (graded)

For $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ and $B = \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$, compute

a) A^T =

○ $\begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}$

○ $\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$

● $\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

○ $\begin{pmatrix} 6 & 3 \\ 5 & 2 \\ 4 & 1 \end{pmatrix}$

✓

b) $A + B =$

○ $\begin{pmatrix} 2 & 2 & 4 \\ 4 & 6 & 6 \end{pmatrix}$

● $\begin{pmatrix} 0 & 2 & 4 \\ 5 & 4 & 6 \end{pmatrix}$

○ $\begin{pmatrix} 0 & 0 & 2 \\ 5 & 5 & 6 \end{pmatrix}$

○ $\begin{pmatrix} 6 & 3 & 1 \\ 2 & 4 & 7 \end{pmatrix}$

✓

c) $A - B =$

○ $\begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$

● $\begin{pmatrix} 2 & 2 & 2 \\ 3 & 6 & 6 \end{pmatrix}$

○ $\begin{pmatrix} 2 & 2 & 2 \\ 1 & 3 & 6 \end{pmatrix}$

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

✓

Submit

Answers are displayed within the problem

Problem 3

2/2 points (graded)

Let $x = (1, 0, -1)$ and $y = (0, 1, -1)$.

a) What is $x \cdot y$?

1 ✓ **Answer:** 1

1

b) What is the angle between these two vectors, in degrees (give a number in the range 0 to 180)?

60 ✓ **Answer:** 60

60

? Hint (1 of 1): Let θ denote the angle between x and y. Recall that $\cos\theta = \frac{x \cdot y}{\|x\| \ \|y\|}$.

Next Hint

Submit

ℹ Answers are displayed within the problem

Problem 4

2/2 points (graded)

For each pair of vectors below, say whether or not they are orthogonal.

a) $(1, 3, 0, 1)$ and $(-1, -3, 0, -1)$

not orthogonal ✓ **Answer:** not orthogonal

b) $(1, 3, 0, 1)$ and $(1, 3, 0, -10)$

orthogonal ✓ **Answer:** orthogonal

Explanation

They are orthogonal if and only if their dot product is zero.

Submit

ℹ Answers are displayed within the problem

Problem 5

1/1 point (graded)

Find the unit vector in the same direction as $x = (1, 2, 3)$.

○ $(1, 2, 3)/6$

○ $(1, 2, 3)/14$

○ $(1, 2, 3)/\sqrt{7}$

● $(1, 2, 3)/\sqrt{14}$

✓

Explanation

In general, to get a unit vector in the same direction as x, we simply divide by the L_2 norm of x.

Submit

ℹ Answers are displayed within the problem

Problem 6

1/1 point (graded)

Find all unit vectors in $\mathbb{R}^2$ that are orthogonal to $(1, 1)$.

☐ $(1, 1)/\sqrt{2}$ and $(-1, -1)/\sqrt{2}$

☐ $(1, -1)$ and $(-1, 1)$

☐ $(1, 1)/2$ and $(-1, -1)/2$

☑ $(1, -1)/\sqrt{2}$ and $(-1, 1)/\sqrt{2}$

✓

Explanation

The set of points orthogonal to $(1, 1)$ is a line passing through the origin: the line $x_1 + x_2 = 0$. Only two points on this line have length 1.

Submit

Answers are displayed within the problem

Problem 7

1/1 point (graded)

How would you describe the set of all points $x \in \mathbb{R}^d$ with $x \cdot x = 25$? Select all that apply.

- ☑ All points of ℓ_2 length 5.
- ☐ The surface of a sphere that is centered at the origin, of radius 25.
- ☐ All points of ℓ_2 length 25.
- ☑ The surface of a sphere that is centered at the origin, of radius 5.

✓

? **Hint (1 of 1):** Recall from lecture that $x \cdot x = \|x\|^2$. Next Hint

Submit

Answers are displayed within the problem

Problems 8-17 correspond to "Linear algebra II: matrix products and linear functions"

Problem 8

1/1 point (graded)

Which of the following is a linear function of $x \in \mathbb{R}^3$? Select all that apply.

☐ $x_1^2 + 3x_2 + x_3$

☑ $2x_1 - 3x_2$

☐ $x_1x_2 - x_3$

✓

Submit

Answers are displayed within the problem

Problem 9

1/1 point (graded)

True or false: the function $f(x) = 2x_1 - x_2 + 6x_3$ can be written as $w \cdot x$ for $x \in \mathbb{R}^3$, where $w = (2, -1, 6)$.

● True

○ False

✓

Submit

Answers are displayed within the problem

Problem 10

3/3 points (graded)

Consider the linear function that is expressed by the matrix $\begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & -1 \end{pmatrix}$.

This function maps vectors in $\mathbb{R}^p$ to $\mathbb{R}^q$.

a) What is p?

3

3

✓ **Answer:** 3

b) What is q?

2

2

✓ **Answer:** 2

c) Which of the following vectors are mapped to zero?

☑ $(2, -1, 6)$

☑ $(-4, 2, -12)$

☐ $(1, 4, -1)$

☐ $(4, -2, 1)$

✓

Explanation

In general, a matrix of size $r \times c$ sends c-dimensional vectors to r-dimensional vectors.
For part (c), the idea is to just try each of the four options: multiply each vector by the matrix and see what it maps to.

Submit

ℹ Answers are displayed within the problem

Problem 11

3/3 points (graded)

Compute the product: $\begin{pmatrix} 1 & 0 & -1 \\ 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 3 & 0 & 1 \\ 0 & 0 & 1 \\ 2 & 6 & 0 \end{pmatrix}$:

$= \begin{pmatrix} 1 & a & 1 \\ 14 & b & c \end{pmatrix}$

$a =$

-6 ✓ **Answer:** -6

-6

$b =$

24 ✓ **Answer:** 24

24

$c =$

5 ✓ **Answer:** 5

5

Submit

ℹ Answers are displayed within the problem

Problem 12

4/4 points (graded)

For a certain pair of matrices A, B, the product AB has dimension 10×20. Suppose A has 30 columns.

a) $A \in \mathbb{R}^{m \times n}$

$m =$

10 ✓ **Answer:** 10

10

$n =$

30 ✓ **Answer:** 30

30

b) $B \in \mathbb{R}^{r \times s}$

$r =$

30 ✓ **Answer:** 30

30

$s =$

20 ✓ **Answer:** 20

20

Explanation

If AB has size 10×20, it means that A must have size $10 \times c$ and B must have size $c \times 20$, for some number c. The second piece of information tells us that $c = 30$.

Submit

ℹ Answers are displayed within the problem

Problem 13

3/3 points (graded)

We have n data points $x^{(1)}, \ldots, x^{(n)} \in \mathbb{R}^d$ and we store them in a matrix X, one point per row.

a) True or false: X has dimension $d \times n$.

○ True

● False

✓

b) True or false: $X^T X$ has dimension $d \times d$.

● True

○ False

✓

c) Which of the following is a matrix with (i, j) entry $x^{(i)} \cdot x^{(j)}$?

○ XX

○ $X^T X$

● XX^T

○ $X^T X^T$

✓

Explanation

For part (c), notice that the first and last options aren't even valid products, unless $d = n$. Of the middle two, $X^T X$ has dimension $d \times d$ while XX^T has dimension $n \times n$.

Submit

ℹ Answers are displayed within the problem

Problem 14

1/1 point (graded)

Vector x has length 10. What is $x^T xx^T xx^T x$?

1000000 ✓ **Answer:** 10^6

1000000

Explanation

Using the associative property of matrix multiplication, we have

$$x^T xx^T xx^T x = (x^T x)\,(x^T x)\,(x^T x) = (\|x\|^2)\,(\|x\|^2)\,(\|x\|^2) = \|x\|^6 = 10^6.$$

? **Hint (1 of 1):** Use the associative property of matrix/vector multiplication: a product of matrices can be parenthesized in any way, as long as the original order of matrices is maintained.

Next Hint

Submit

ℹ Answers are displayed within the problem

Problem 15

5/5 points (graded)

Suppose $x = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$.

a) What is $x^T x$?

35 ✓ **Answer:** 35

35

b) What is xx^T?

$$xx^T = \begin{pmatrix} 1 & a & b \\ 3 & 9 & c \\ 5 & 15 & d \end{pmatrix}$$

$a =$

3 ✓ **Answer:** 3

3

$b =$

5 ✓ **Answer:** 5

5

$c =$

15 ✓ **Answer:** 15

15

$d =$

25 ✓ **Answer:** 25

25

Submit

ℹ Answers are displayed within the problem

Problem 16

1/1 point (graded)

Vectors $x, y \in \mathbb{R}^d$ both have length 2. If $x^T y = 2$, what is the angle between x and y, in degrees (the answer is an integer in the range 0 to 180)?

60 ✓ **Answer:** 60

60

Explanation

Letting θ denote the angle, we have $\cos\theta = (x \cdot y) / (\|x\| \|y\|) = 0.5$.

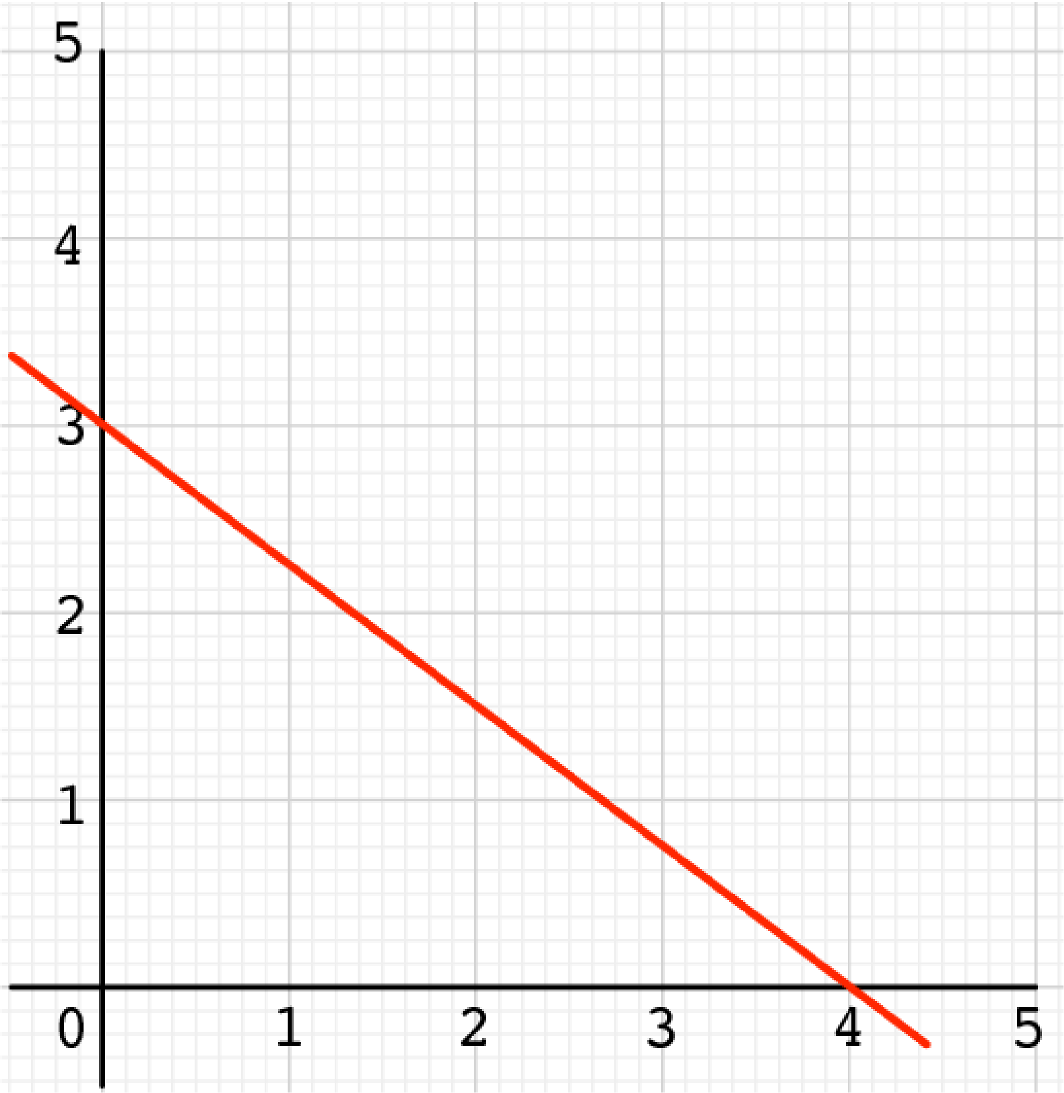
Submit

ℹ Answers are displayed within the problem

Problem 17

2/2 points (graded)

The line shown below can be expressed in the form $w \cdot x = 12$ for $x \in \mathbb{R}^2$. What is w?



$w = (w_1, w_2)$

w_1 =

3

✓ **Answer:** 3

3

w_2 =

4

✓ **Answer:** 4

4

Explanation

We are asked to determine $w = (w_1, w_2)$: there are two unknowns. We can write down two equations corresponding to the fact that the line passes through the points $(4, 0)$ and $(0, 3)$, and then solve for the two unknowns.

In more detail, we have the two equations: $4w_1 = 12$ and $3w_2 = 12$, whereupon $w = (3, 4)$.

Submit

ℹ Answers are displayed within the problem

Problems 18-24 correspond to "Linear algebra III: square matrices as quadratic functions"

Problem 18

4/4 points (graded)

The quadratic function $f : \mathbb{R}^3 \to \mathbb{R}$ given by

$$f(x) = 3x_1^2 + 2x_1x_2 - 4x_1x_3 + 6x_3^2$$

can be written in the form $x^T Mx$ for some **symmetric** matrix M. What are the missing entries in M?

$$M = \begin{pmatrix} a & 1 & b \\ 1 & c & 0 \\ -2 & d & 6 \end{pmatrix}$$

$a =$

3

✓ **Answer:** 3

3

$b =$

-2

✓ **Answer:** -2

-2

$c =$

0 ✓ **Answer:** 0

0

$d =$

0 ✓ **Answer:** 0

0

Explanation

a and c correspond to the coefficients of x_1^2 and x_2^2 and thus are 3,0 respectively. The other entries can be inferred directly from the fact that the matrix is symmetric. Finally, it is worth quickly checking that the matrix does indeed generate the given quadratic function.

Submit

Answers are displayed within the problem

Problem 19

7/7 points (graded)

Answer the following questions about the quadratic function $f : \mathbb{R}^3 \to \mathbb{R}$ associated with the matrices A.

a) True or false: the quadratic function associated with $A = \mathrm{diag}\,(6, 2, -1)$ is $f\,(x_1, x_2, x_3) = 6x_1^2 + 2x_2^2 - x_3^2$.

● True

○ False

✓

b) $A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & -1 & 4 \\ 2 & -2 & 1 \end{pmatrix}$

Find the coefficients of the function $f\left(x_1, x_2, x_3\right) = ax_1^2 + bx_1x_2 + cx_1x_3 + dx_2^2 + ex_2x_3 + fx_3^2$ generated by this matrix.

$a =$

1 ✓ **Answer:** 1

1

$b =$

4 ✓ **Answer:** 4

4

$c =$

6 ✓ **Answer:** 6

6

$d =$

-1 ✓ **Answer:** -1

-1

$e =$

2 ✓ **Answer:** 2

2

$f =$

1 ✓ **Answer:** 1

1

Explanation

Both parts are direct applications of a formula from lecture: any $d \times d$ matrix M can be associated with the quadratic function $x^T Mx$, which maps d-dimensional vectors x to $\sum_{i,j} M_{ij}x_i x_j$.

Submit

Answers are displayed within the problem

Problem 20

1/1 point (graded)

Which of the following matrices is necessarily symmetric? Select all that apply.

- ☑ AA^T for arbitrary matrix A.
- ☑ $A^T A$ for arbitrary matrix A.
- ☑ $A + A^T$ for arbitrary square matrix A.
- ☐ $A - A^T$ for arbitrary square matrix A.

✓

Explanation

The (i, j) entry of AA^T is $\sum_k A_{ik} A_{jk}$.
For $A^T A$, it is $\sum_k A_{ki} A_{kj}$.
For $A + A^T$, it is $A_{ij} + A_{ji}$.

And for $A - A^T$, it is $A_{ij} - A_{ji}$.
From this we see that all except the last are necessarily symmetric.

? **Hint (1 of 1):** In each case, it might be helpful to write down a general formula for the (i, j) entry of the resulting matrix, and to then gauge whether this is necessarily the same as the (j, i) entry.

Next Hint

Submit

Answers are displayed within the problem

Problem 21

2/2 points (graded)
Let $A = \text{diag}\,(1, 2, 3, 4, 5, 6, 7, 8)$.

a) What is $|A|$?

40320 ✓ **Answer:** 40320

40320

b) True or false: $A^{-1} = \text{diag}\,(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8})$

● True

○ False

✓

? **Hint (1 of 2):** The determinant of a diagonal matrix is just the product of the numbers along the diagonal.

Next Hint

Hint (2 of 2): The inverse of a diagonal matrix is obtained by just inverting the numbers along the diagonal. If any of these numbers is 0, the inverse does not exist.

Submit

ℹ Answers are displayed within the problem

Problem 22

2/2 points (graded)

Vectors $u_1, \ldots, u_d \in \mathbb{R}^d$ all have unit length and are orthogonal to each other. Let U be the $d \times d$ matrix whose rows are the u_i.

a) What is UU^T?

○ U

○ U^T

○ U^{-1}

● I_d

✓

b) What is U^{-1}?

○ U

● U^T

○ U^{-1}

○ I_d

✓

? **Hint (1 of 2):** The (i, j) entry of UU^T is just the dot product $u_i \cdot u_j$.

Hint (2 of 2): If $AB = I$, then B is the inverse of A.

Next Hint

Submit

Answers are displayed within the problem

Problem 23

1/1 point (graded)

Matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & z \end{pmatrix}$ is singular. What is z?

6 ✓ **Answer:** 6

6

? **Hint (1 of 2):** Recall that a square matrix is singular if and only if its determinant is zero.

Hint (2 of 2): The determinant of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $ad - bc$.

Next Hint

Submit

Answers are displayed within the problem

Problem 24

1/1 point (graded)

The *trace* of a $d \times d$ matrix A is defined to be $\operatorname{tr}(A) = \sum_{i=1}^{d} A_{ii}$. Which of the following statements is true, for arbitrary $d \times d$ matrices A, B? Select all that apply.

☑ $\operatorname{tr}(A) = \operatorname{tr}(A^T)$.

☑ $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$.

☐ $\operatorname{tr}(AB) = \operatorname{tr}(A)\operatorname{tr}(B)$.

☑ $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

✓

Explanation

Since the trace depends only on diagonal elements, the first two choices are correct. For the next two, we need to derive a formula for $\operatorname{tr}(AB)$. First, the (i, j) entry of AB is

$(AB)_{ij} = \sum_k A_{ik} B_{kj}$

and thus

$\operatorname{tr}(AB) = \sum_i (AB)_{ii} = \sum_i \sum_k A_{ik} B_{ki}$.

Since this depends on the off-diagonal entries of A and B, the third option can't possibly be true. But the final option is true, by this formula.

? **Hint (1 of 2):** For the last two options, obtain a formula for $\operatorname{tr}(AB)$ in terms of the entries of A and B.

Hint (2 of 2): Here's the formula: $\operatorname{tr}(AB) = \sum_i \sum_j A_{ij} B_{ji}$

Next Hint

Submit

Answers are displayed within the problem

Problems 25-27 correspond to "The multivariate Gaussian"

Problem 25

1/1 point (graded)
A spherical Gaussian has mean $\mu = (1, 0, 0)$. At which of the following points will the density be the same as at $(1, 1, 0)$? Select all that apply.

☑ $(0, 0, 0)$

☐ $(1, 1, 1)$

☑ $(2, 0, 0)$

☑ $(1, 0, 1)$

✓

Explanation
Point $(1, 1, 0)$ is at distance 1 from the mean. The density will be the same at all points which are distance 1 from the mean: the first, third, and fourth options.

? **Hint (1 of 1):** In a spherical Gaussian, the density depends only on the L_2 distance from the mean. Next Hint

Submit

Answers are displayed within the problem

Problem 26

1/1 point (graded)
How many real-valued parameters are needed to specify a diagonal Gaussian in $\mathbb{R}^d$?

○ d

● $2d$

○ $\frac{1}{2}d^2$

○ d^2

✓

Explanation

We need d parameters for the mean and another d for the diagonal of the covariance matrix.

Submit

ℹ Answers are displayed within the problem

Problems 28-29 correspond to "Gaussian generative models"

Problem 28

3/3 points (graded)

Suppose we solve a classification problem with k classes by using a Gaussian generative model in which the jth class is specified by parameters π_j, μ_j, Σ_j. In each of the following situations, say whether the decision boundary is `linear`, `spherical`, or `other quadratic`.

a) We compute the empirical covariance matrices of each of the k classes, and then set $\Sigma_1 = \Sigma_2 = \cdots = \Sigma_k$ to the $\mathbf{average}$ of these matrices.

linear ✓ **Answer:** linear

b) The covariance matrices Σ_j are all $\mathbf{diagonal}$, but no two of them are the same.

other quadratic ✓ **Answer:** other quadratic

c) There are two classes (that is, $k = 2$) and the covariance matrices Σ_1 and Σ_2 are multiples of the identity matrix.

spherical ▲▼ ✓ **Answer:** spherical

? **Hint (1 of 1):** If these seem mysterious, just review the final lecture on Gaussian generative modeling. Next Hint

Submit

ℹ Answers are displayed within the problem

Problem 29

2/2 points (graded)

Consider a binary classification problem in which we fit a Gaussian to each class and find that they are both centered at the origin but have different covariances: $\mu_1 = \mu_2 = 0$ and $\Sigma_1 \neq \Sigma_2$. Derive the precise form of the $\textbf{decision boundary}$, that is, the points x for which the two classes are equally likely. You will find that it is

$$x^T \left(\Sigma_2^{-1} - \Sigma_1^{-1}\right) x = a \ln \frac{|\Sigma_1|}{|\Sigma_2|} + b \ln \frac{\pi_1}{\pi_2}.$$

What are a and b?

$a =$

1 ✓ **Answer:** 1

1

$b =$

-2 ✓ **Answer:** -2

-2

Explanation

The decision boundary consists of points x at which $\pi_1 P_1(x) = \pi_2 P_2(x)$. Plugging in the formula for the Gaussian density, and setting the mean to zero, we get

$$\pi_1 \frac{1}{(2\pi)^{d/2}|\Sigma_1|^{1/2}} \exp\left(-\frac{1}{2}x^T\Sigma_1^{-1}x\right) = \pi_2 \frac{1}{(2\pi)^{d/2}|\Sigma_2|^{1/2}} \exp\left(-\frac{1}{2}x^T\Sigma_2^{-1}x\right).$$

Rearranging and taking the log of both sides, we have

$$x^T\left(\Sigma_2^{-1} - \Sigma_1^{-1}\right)x = 2\ln\frac{\pi_2}{\pi_1} + \ln\frac{|\Sigma_1|}{|\Sigma_2|}.$$

? **Hint (1 of 1):** Recall that the decision boundary consists of points x at which $\pi_1 P_1(x) = \pi_2 P_2(x)$. Plug in the formula for the Gaussian density, setting the mean to zero. It will simplify things if you take the log of both sides.

Next Hint

Submit

Answers are displayed within the problem

Problem 30 corresponds to "More generative modeling"

Problem 30

5/5 points (graded)

For each of the situations below, say which of the following distributions would be the best model for the data: `Gaussian`, `gamma`, `beta`, `Poisson`, or `categorical`.

a) You collect the number of airplane landings at Los Angeles International Airport during each one hour interval over the course of a week (thus, a total of 168 data points).

Poisson ✓ **Answer:** Poisson

b) For your favorite sports team, you compute the fraction of games they won each year, during the period 1980-2015 (thus, a total of 36 data points).

beta ✓ **Answer:** beta

c) Your local pet store has mammals, reptiles, birds, amphibians, and fish. You measure the fraction of each (thus, a total of five numbers).

categorical ✓ **Answer:** categorical

d) You collect the pollution levels (positive real numbers reflecting concentrations of particulate matter) recorded in your city over the past year (thus, a total of 365 numbers).

gamma ✓ **Answer:** gamma

e) Like (d), but instead you use the log of these values.

Gaussian ✓ **Answer:** Gaussian

Explanation

For (a), the data consist of positive integer counts, so a Poisson is suitable. For (b), the data consists of fractions, so a Beta is suitable. For (c), we have a probability distribution over a discrete set of options: a categorical. For (d), we have positive reals, so a Gamma would be the closest match. For (e), we get both positive and negative reals, so a Gaussian is the only reasonable choice.

Submit

ℹ Answers are displayed within the problem