Problem Set 9

Problems 1-6 correspond to "Linear Projections"

Problem 1

1/1 point (graded)

In $\mathbb{R}^2$, what is the unit vector corresponding to the x_1-direction?

- ( ) $(0, 0)$
- (•) $(1, 0)$
- ( ) $(0, 1)$
- ( ) $(1, 1)$

Submit

Answers are displayed within the problem

Problem 2

1/1 point (graded)

What is the unit vector in the same direction as $(3, 2, 2, 2, 2)$?

- ( ) $(1.5, 1, 1, 1, 1)$
- ( ) $(1, 0.67, 0.67, 0.67, 0.67)$
- (x) $(0.6, 0.4, 0.4, 0.4, 0.4)$
- ( ) $(0.5, 0.33, 0.33, 0.33, 0.33)$

✔

? **Hint (1 of 1):** To get a unit vector in the same direction as x, simply divide by $\|x\|$.

Next Hint

Submit

Answers are displayed within the problem

Problem 3

1/1 point (graded)

What is the projection of the vector $(3, 5, -9)$ onto the direction $(0.6, -0.8, 0)$?

-2.2

✔ **Answer:** -2.2

-2.2

? **Hint (1 of 1):** If u is a unit vector, then the projection of x onto direction u is simply $u \cdot x$.

Next Hint

Submit

Answers are displayed within the problem

Problem 4

1/1 point (graded)

What is the (unit) direction along which the projection of $(4, -3)$ is largest?

- [x] $(0.8, -0.6)$
- [ ] $(-0.6, -0.8)$
- [ ] $(-0.8, 0.6)$
- [ ] $(0.8, 0.6)$

✔

Explanation

The projection of $x = (4, -3)$ is going to be largest in the direction of x itself.

Submit

Answers are displayed within the problem

Problem 5

1/1 point (graded)

What is the (unit) direction along which the projection of $(4, -3)$ is smallest?

- [ ] $(0.8, -0.6)$
- [ ] $(-0.6, -0.8)$
- [x] $(-0.8, 0.6)$

- [ ] $(0.8, 0.6)$

✔

Explanation

The projection of $x = (4, -3)$ will be smallest in the direction opposite to x, that is, the direction of $-x$.

Submit

Answers are displayed within the problem

Problem 6

1/1 point (graded)

The projection of vector x onto direction u is exactly zero. Which of the following statements is necessarily true? Select all that apply.

- [x] u is orthogonal to x.
- [ ] u is in the opposite direction to x.
- [x] u is at right angles to x.
- [ ] It is not possible to have a projection of zero.

✔

Submit

Answers are displayed within the problem

Problems 7-8 correspond to "Principal component analysis I: one-dimensional projection"

Problem 7

4/4 points (graded)

A three-dimensional data set has covariance matrix

$$\Sigma = \begin{pmatrix} 4 & 2 & -3 \\ 2 & 9 & 0 \\ -3 & 0 & 9 \end{pmatrix}.$$

a) What is the variance of the data in the x_1-direction?

4

Answer: 4

4

b) What is the correlation between x_1 and x_3?

-1/2

Answer: -0.5

$-\frac{1}{2}$

c) What is the variance in the direction $(0, -1, 0)$?

9

Answer: 9

9

d) What is the variance in the direction of $(1, 1, 0)$?

17/2

Answer: 8.5

$\frac{17}{2}$

Hint (1 of 3): For part (a): the diagonal entry Σ_{ii} is the variance of X_i.

Hint (2 of 3): For part (b): the entry Σ_{ij} is the *covariance* between X_i and X_j. This is not the same as the *correlation*. Do you remember how to get from one to the other?

Hint (3 of 3): For part (c,d): the variance in direction u, where u is a unit vector, is given by $u^T \Sigma u$.

Next Hint

Submit

Answers are displayed within the problem

Problem 8

1/1 point (graded)

Which of the following covariance matrices has the property that the variance is the same in any direction? Select all that apply.

- [x] The all-zeros matrix.
- [ ] The all-ones matrix.
- [x] The identity matrix.
- [ ] Any diagonal matrix.

✓

Explanation

Let u be any unit vector in d-dimensional space.
If A is the all-zeros matrix, then $u^T A u = 0$, the same for all u.
If B is the all-ones matrix, then $u^T B u = \sum_{ij} u_i u_j = (\sum_i u_i)^2$, which is not the same for all u.
With the identity matrix: $u^T I u = u^T u = 1$, the same for all u.

Let D be the diagonal matrix where $D_{11} = 1$ and all other diagonal entries are zero. Then $u^T Du = u_1^2$, not the same for all u.

Submit

Answers are displayed within the problem

Problems 9-11 correspond to "Principal component analysis II: the top k directions"

Problem 9

8/8 points (graded)

Let $u_1, u_2 \in \mathbb{R}^d$ be two vectors with $\|u_1\| = \|u_2\| = 1$ and $u_1 \cdot u_2 = 0$. Define U to be the matrix whose columns are u_1 and u_2.

What are the dimensions of the following matrices?

a) U

# of Rows =

d

Answer: d

# of Columns =

2

Answer: 2

2

b) U^T

# of Rows =

2

Answer: 2

2

# of Columns

d

✔ **Answer:** d

c) UU^T

# of Rows =

d

✔ **Answer:** d

# of Columns =

d

✔ **Answer:** d

d) $u_1 u_1^T$

# of Rows =

d

✔ **Answer:** d

# of Columns =

d

✔ **Answer:** d

Submit

Answers are displayed within the problem

Problem 10

1/1 point (graded)

Continuing from the previous problem, let $u_1, u_2 \in \mathbb{R}^d$ be two vectors with $\|u_1\| = \|u_2\| = 1$ and $u_1 \cdot u_2 = 0$, and define U to be the matrix whose columns are u_1 and u_2.

Which of the following linear transformations sends points $x \in \mathbb{R}^d$ to their (two-dimensional) projections onto directions u_1 and u_2? Select all that apply.

- [x] $x \mapsto (u_1 \cdot x, u_2 \cdot x)$
- [ ] $x \mapsto (u_1 \cdot x)\, u_1 + (u_2 \cdot x)\, u_2$
- [x] $x \mapsto U^T x$
- [ ] $x \mapsto UU^T x$

Explanation

The first and third maps send 4-d to 2-d. The second and fourth maps send 4-d to 4-d.

Submit

Answers are displayed within the problem

Problem 11

2/2 points (graded)

For a particular four-dimensional data set, the top two eigenvectors of the covariance matrix are:

$$\frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}.$$

a) What is the PCA projection of point $(2, 4, 2, 6)$ into two dimensions? Write it in the form (a, b).

- ○ $(2,2)$
- ○ $(2,3)$
- ◉ $(7,3)$
- ○ $(4,6)$

✔

b) What is the reconstruction, from this projection, to a point in the original four-dimensional space? Write it in the form (a,b,c,d)

- ◉ $(2,5,2,5)$
- ○ $(2,1,2,2)$
- ○ $(4,2,2,2)$
- ○ $(2,6,2,4)$

✔

Submit

Answers are displayed within the problem

Problems 12-14 correspond to "Linear algebra V: eigenvalues and eigenvectors"

Problem 12

2/2 points (graded)

Consider the 2×2 matrix $M = \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}$.

a) One of its eigenvectors is $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. What is the corresponding eigenvalue?

6

✔ **Answer:** 6

6

b) Its other eigenvector is $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. What is the corresponding eigenvalue?

4

✔ **Answer:** 4

4

Submit

Answers are displayed within the problem

Problem 13

6/6 points (graded)

A 2×2 matrix M has eigenvalues 10 and 5.

a) What are the eigenvalues of $2M$ (that is, each entry of M is multiplied by 2)?

Larger eigenvalue =

20

✔ **Answer:** 20

20

Smaller eigenvalue =

10

✔ **Answer:** 10

10

b) What are the eigenvalues of $M + 3I$, where I is the 2×2 identity matrix?

Larger eigenvalue =

13 ✔ **Answer:** 13

13

Smaller eigenvalue =

8 ✔ **Answer:** 8

8

c) What are the eigenvalues of $M^2 = MM$?

Larger eigenvalue =

100 ✔ **Answer:** 100

100

Smaller eigenvalue =

25 ✔ **Answer:** 25

25

Explanation

Suppose (u, λ) is an (eigenvector, eigenvalue) pair for M, that is, $Mu = \lambda u$.
Part (a): for any constant c, we have $(cM)\,u = M\,(cu) = c\lambda u$. Thus $(u, c\lambda)$ is an (eigenvector, eigenvalue) pair for cM.
Part (b): for any constant c, we have $(M + cI)\,u = Mu + cu = (\lambda + c)\,u$. Thus $(u, \lambda + c)$ is an (eigenvector, eigenvalue) pair for $M + cI$.
Part (c): For any positive integer c, we have $M^c u = \lambda^c u$, and thus (u, λ^c) is an (eigenvector, eigenvalue) pair for M^c.

? **Hint (1 of 3):** For part (a): if $Mu = \lambda u$, what do we know about $(2M)\,u = M\,(2u)$?

Hint (2 of 3): For part (b): Note that $(M + 3I)\,u = Mu + 3u$

Hint (3 of 3): For part (c): $M^2 u = M\,(Mu)$

Next Hint

Submit

Answers are displayed within the problem

Problem 14

6/7 points (graded)

A certain three-dimensional data set has covariance matrix

$$\begin{pmatrix} 5 & -3 & 0 \\ -3 & 5 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

.

a) Consider the direction $u = (1, 1, 1)\,/\sqrt{3}$. What is variance of the projection of the data onto direction u?

8/3 ✔

$\frac{8}{3}$

b) Which of the following are eigenvectors of the covariance matrix? Select all that apply.

☐ $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

- [ ] $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$
- [x] $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$
- [x] $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$
- [ ] $\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$
- [x] $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$

✔

c) Find the eigenvalues of the covariance matrix. List them in decreasing order.

8 ✔

8

4 ✔

4

2 ✔

2

d) Suppose we used principal component analysis (PCA) to project points into *two* dimensions. What would be the resulting two-dimensional projection of the point $x = (\sqrt{2}, -3\sqrt{2}, 2)$?

- ( ) $(1, 0)$
- (x) $(4, 2)$
- ( ) $(1, 4)$
- ( ) $(4, 1)$

✔

e) Now suppose we use the projection in (d) to reconstruct a point $\widehat{x}$ in the original three-dimensional space. What is the Euclidean distance between x and $\widehat{x}$, that is, $\|x - \widehat{x}\|$?

12 ✖

12

? **Hint (1 of 5):** Part (a): If Σ is the covariance matrix of a data set, then the projection of the data into the direction given by unit vector u has variance $u^T \Sigma u$.

Next Hint

Hint (2 of 5): Part (b): To check if v is an eigenvector of matrix M, just check whether Mv is a multiple of v, that is, of the form λv.

Hint (3 of 5): Part (c): Given that v is an eigenvector of matrix M, the corresponding eigenvalue is the number λ such that $Mv = \lambda v$.

Hint (4 of 5): Part (d): PCA will project data points x onto the top two eigenvectors (that is, the eigenvectors with the two largest eigenvalues). If these are u_1 and u_2 then the projection of x is $(x \cdot u_1, x \cdot u_2)$.

Hint (5 of 5): Part (e): The reconstruction from the projection of x is $(x \cdot u_1)\, u_1 + (x \cdot u_2)\, u_2$.

Submit

✱ Partially correct (6/7 points)

Problems 15-17 correspond to "Linear algebra VI: spectral decomposition"

Problem 15

1/1 point (graded)

M is a 2×2 real-valued symmetric matrix with eigenvalues $\lambda_1 = 6, \lambda_2 = 1$ and corresponding eigenvectors

$$u_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

What is M?

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

✔

? **Hint (1 of 1):** This is a direct application of the spectral decomposition theorem. We went through an example just like this in lecture.

Next Hint

Submit

Answers are displayed within the problem

Problem 16

1/1 point (graded)

For a certain data set in d-dimensional space, the covariance matrix has the following interesting property: there are k positive eigenvalues and the rest are zero (where $k < d$). What can we conclude from this? Select all that apply.

- [ ] Each of the data points has at most k nonzero coordinates.
- [x] The data can be perfectly reconstructed from their PCA projection onto k dimensions.
- [x] Each data point can be expressed as a linear combination of the top k eigenvectors.
- [ ] It is possible to discard $d - k$ of the coordinates without losing any of the variance in the data.

✔

? **Hint (1 of 1):** Intuitively, the data lies in a k-dimensional subspace, but this subspace need not be aligned with the coordinate axes.

Next Hint

Submit

Answers are displayed within the problem

Problem 17

1/1 point (graded)

A data set in $\mathbb{R}^d$ has a covariance matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. Under which of the following conditions is PCA most likely to be effective as a form of dimensionality reduction? Select all that apply.

- [ ] When the λ_i are approximately equal.
- [x] When most of the λ_i are close to zero.
- [ ] When most of the λ_i are close to 1.
- [x] When the sequence $\lambda_1, \lambda_2, \ldots$ is rapidly decreasing.

Explanation

The overall variance in the data is $\lambda_1 + \lambda_2 + \cdots + \lambda_d$. When PCA is used to reduce the dimension to k, the amount of variance in the projected points is $\lambda_1 + \lambda_2 + \cdots + \lambda_k$. PCA is most effective when this second quantity is not too much smaller than the first, in other words, when the fraction of variance lost, $(\lambda_{k+1} + \cdots + \lambda_d) \,/\, (\lambda_1 + \cdots + \lambda_d)$, is small.

Submit

Answers are displayed within the problem