BerkeleyX: CS190.1x Scalable Machine Learning

PRINCIPAL COMPONENT VECTORS (1/1 point)

Principal component vectors have length

- n -- the length of the dataset
- sqrt(n)
- d -- the number of features ✔
- k -- the number of principal components selected

EXPLANATION

Principal component vectors have length d. The dot product between a principal component and an observation's features creates a score or transformed feature, so the number of features and the length of the principal component vector must match.

CHECK HIDE ANSWER

EIGENVALUES (1/1 point)

The top k principal components correspond to the top k eigenvalues.

- True ✔
- False

EXPLANATION

Eigenvalues are proportional to the variance explained by the corresponding principal component. The top principal component has the largest eigenvalue, the second principal component has the next largest eigenvalue, etc.

CHECK HIDE ANSWER

COMPUTATION AND STORAGE COMPLEXITY (1/1 point)

The distributed PCA algorithm we implemented in the lab used:

- [x] O(d^2) local storage ✔
- [x] O(d^3) local computation ✔
- [ ] O(d) local storage
- [ ] O(d^2) local computation

✔

Note: Make sure you select all of the correct options—there may be more than one!

EXPLANATION

The outer product results in a d by d matrix which requires O(d^2) local storage. Eigendecomposition on the aggegated results of the outer products requires O(d^3) local computation.

CHECK HIDE ANSWER

PLOTTING CORRELATED VARIABLES (1/1 point)

In Visualization 1, what would the data look like if covariance equaled -1.0?

- ( ) A vertical line
- ( ) Similar to covariance of 0
- ( ) Similar to covariance of .9
- (•) A diagonal line ✔

EXPLANATION

A covariance of -1.0 would produce an exact relationship between x1 and x2, which would show up as a diagonal line of points from the top-left to the bottom-right of the graph.

CHECK HIDE ANSWER

HIGH POSITIVE COVARIANCE (1/1 point)

In Visualization 1 when the covariance is .9, where are most of the points on the plot?

- lower-left and lower-right
- upper-right and lower-left ✔
- upper-left and lower-right
- upper-left and upper-right

EXPLANATION

A positive covariance creates the tendency for the points to fall around a diagonal line with a positive slope, leading to most points falling in the lower-left and upper-right of the graph.

CHECK HIDE ANSWER

PCA FUNCTION (1/1 point)

When running the pca function what is the largest k we should use?

- 1
- n -- length of dataset
- d -- number of features ✔

EXPLANATION

We can generate up to d orthonormal eigenvectors from the covariance matrix, so k should be less than or equal to d.

CHECK HIDE ANSWER

TIME BASED AGGREGATION (1/1 point)

In Visualization 9, does the resulting spatial map appear symmetric or asymmetric about the midline (horizontal line across the middle of the brain)?

- (•) Symmetric ✔
- ( ) Asymmetric

EXPLANATION

The time aggregated information shows a symmetric response about the midline.

CHECK HIDE ANSWER

DIRECTION BASED AGGREGATION (1/1 point)

In Visualization 10, does the resulting spatial map appear symmetric or asymmetric about the midline (horizontal line across the middle of the brain)?

- ( ) Symmetric
- (•) Asymmetric ✔

EXPLANATION

The direction aggregated information shows an asymmetric response about the midline.

CHECK HIDE ANSWER

SURVEY: LAB5 COMPLETION TIME (1/1 point)

How long did Lab FIVE take you to complete (in hours - decimals are OK)?

3 ✔

3

Please click "Check" to save your answers.

CHECK SHOW ANSWER