1)What is a projection and how is it used in PCA?

Ans- Projected Principal Component Analysis (Projected-PCA), which employees principal component analysis to the projected (smoothed) data matrix onto a given linear space spanned by covariates. When it applies to high-dimensional factor analysis, the projection removes noise components.

The last step of PCA is we need to multiply Q tranpose of Q with the original data matrix in order to get the projection matrix. We go from the (d x k) Q matrix and Q transpose of Q results in d x d dimension. By multiplying the (d x n) X matrix, the projection matrix is d x n.

2) How does the optimization problem in PCA work, and what is it trying to achieve?

Ans- PCA seeks to solve a sequence of optimization problems. The first in the sequence is the unconstrained problem maximizeuTSuuTu,u∈Rp. Since uTu=‖u‖22=‖u‖‖u‖, the above unconstrained problem is equivalent to the constrained problem maximizeuTSusubject touTu=1.

3) What is the relationship between covariance matrices and PCA?

Ans- Covariance-based PCA is equivalent to MLPCA whenever the variance-covariance matrix of the measurement errors is assumed diagonal with equal elements on its diagonal. The measurement error variance parameter can then be estimated by applying the probabilistic principal component analysis (PPCA) model

PCA is simply described as “diagonalizing the covariance matrix”. What does diagonalizing a matrix mean in this context? It simply means that we need to find a non-trivial linear combination of our original variables such that the covariance matrix is diagonal.

4) How does the choice of number of principal components impact the performance of PCA?

Ans- If our sole intention of doing PCA is for data visualization, the best number of components is 2 or 3. If we really want to reduce the size of the dataset, the best number of principal components is much less than the number of variables in the original dataset

5) How can PCA be used in feature selection, and what are the benefits of using it for this purpose?

Ans- PCA, generally called data reduction technique, is very useful feature selection technique as it uses linear algebra to transform the dataset into a compressed form. We can implement PCA feature selection technique with the help of PCA class of scikit-learn Python library.

6) What are some common applications of PCA in data science and machine learning?

Ans-

PCA is used to visualize multidimensional data.

It is used to reduce the number of dimensions in healthcare data.

PCA can help resize an image.

It can be used in finance to analyze stock data and forecast returns.

PCA helps to find patterns in the high-dimensional datasets.

7) What is the relationship between spread and variance in PCA?

Ans- Variance is the spread of the data in a dataset. In PCA, the variables are transformed in such a way that they explain variance of the dataset in decreasing manner. Co-variance: Covariance provides a measure of the strength of the correlation between two or more sets of random variates.

8) How does PCA use the spread and variance of the data to identify principal components?

Ans- PCA works by finding the directions of maximum variance in the data set and projecting the data onto these directions. The principal components are ordered by the amount of variance they explain and are used for feature selection, data compression, clustering, and classification.

9) How does PCA handle data with high variance in some dimensions but low variance in others?

Ans- PCA generally tries to find the lower-dimensional surface to project the high-dimensional data. PCA works by considering the variance of each attribute because the high attribute shows the good split between the classes, and hence it reduces the dimensionality.