Close

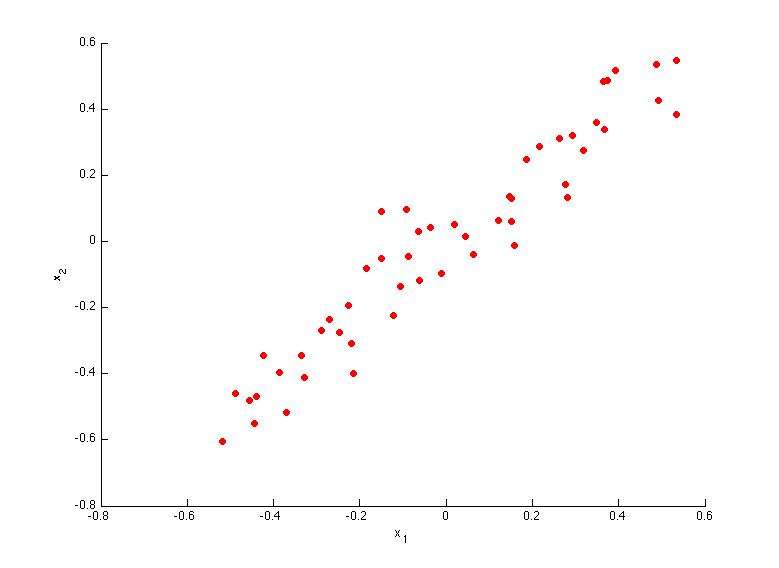
Principal Component Analysis

5 questions

1  
point

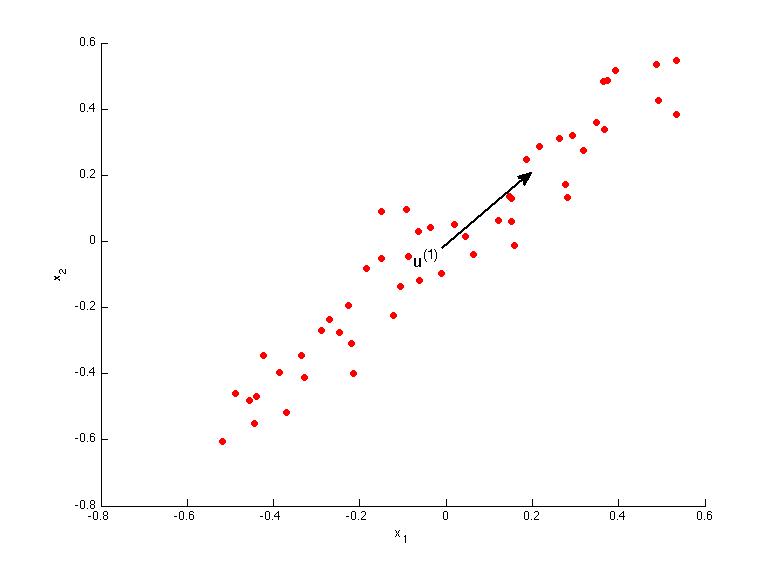
1.

Consider the following 2D dataset:

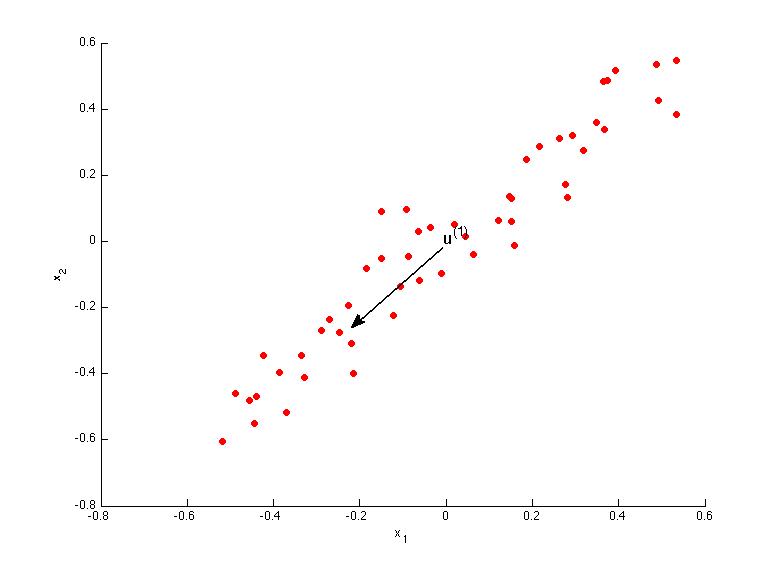


Which of the following figures correspond to possible values that PCA may return for *u*(1) (the first eigenvector / first principal component)? Check all that apply (you may have to check more than one figure).

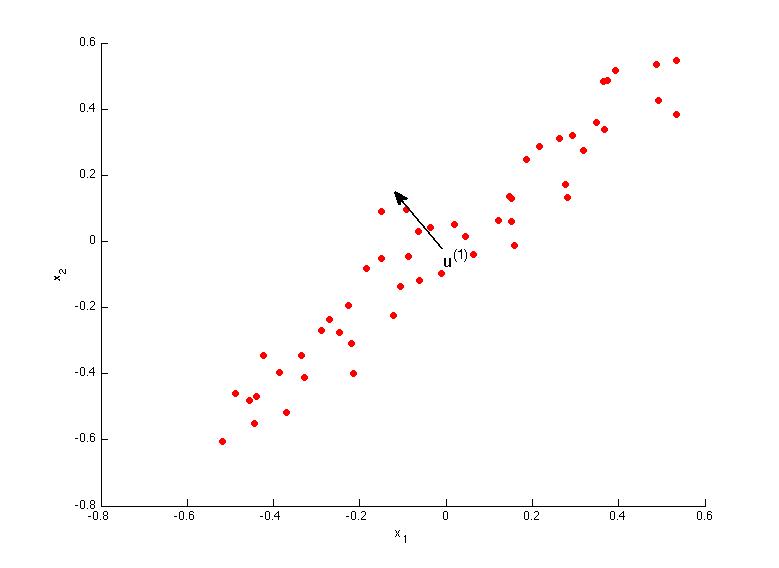




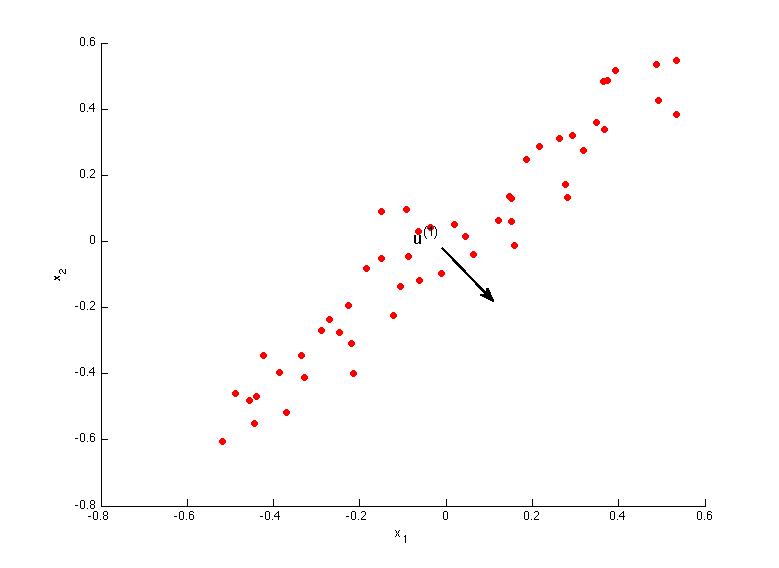












1  
point

2.

Which of the following is a reasonable way to select the number of principal components *k*?

(Recall that *n* is the dimensionality of the input data and *m* is the number of input examples.)



Choose *k* to be 99% of *n* (i.e., *k*=0.99∗*n*, rounded to the nearest integer).



Choose *k* to be the smallest value so that at least 1% of the variance is retained.



Choose *k* to be the smallest value so that at least 99% of the variance is retained.



Choose the value of *k* that minimizes the approximation error 1*m*∑*mi*=1||*x*(*i*)−*x*(*i*)approx||2.

1  
point

3.

Suppose someone tells you that they ran PCA in such a way that "95% of the variance was retained." What is an equivalent statement to this?



1*m*∑*mi*=1||*x*(*i*)||21*m*∑*mi*=1||*x*(*i*)−*x*(*i*)approx||2≥0.05



1*m*∑*mi*=1||*x*(*i*)||21*m*∑*mi*=1||*x*(*i*)−*x*(*i*)approx||2≤0.95



1*m*∑*mi*=1||*x*(*i*)||21*m*∑*mi*=1||*x*(*i*)−*x*(*i*)approx||2≤0.05



1*m*∑*mi*=1||*x*(*i*)−*x*(*i*)approx||21*m*∑*mi*=1||*x*(*i*)||2≤0.05

1  
point

4.

Which of the following statements are true? Check all that apply.



PCA can be used only to reduce the dimensionality of data by 1 (such as 3D to 2D, or 2D to 1D).



Given an input *x*∈R*n*, PCA compresses it to a lower-dimensional vector *z*∈R*k*.



If the input features are on very different scales, it is a good idea to perform feature scaling before applying PCA.



Feature scaling is not useful for PCA, since the eigenvector calculation (such as using Octave's svd(Sigma) routine) takes care of this automatically.

1  
point

5.

Which of the following are recommended applications of PCA? Select all that apply.



To get more features to feed into a learning algorithm.



Data compression: Reduce the dimension of your input data *x*(*i*), which will be used in a supervised learning algorithm (i.e., use PCA so that your supervised learning algorithm runs faster).



Clustering: To automatically group examples into coherent groups.



Data visualization: Reduce data to 2D (or 3D) so that it can be plotted.

2 questions unanswered

Submit Quiz