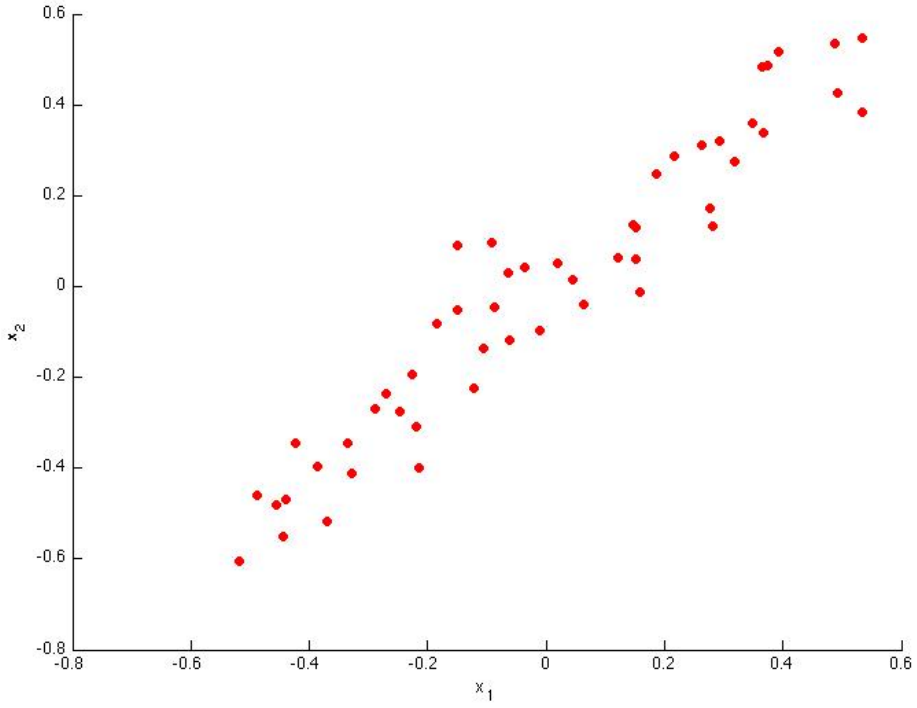
1 point

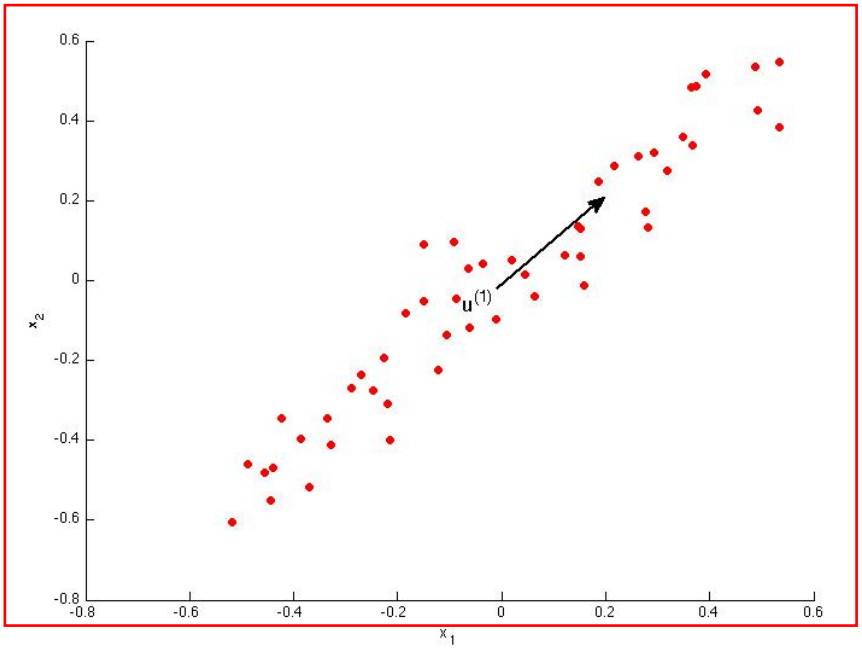
1.

Consider the following 2D dataset:

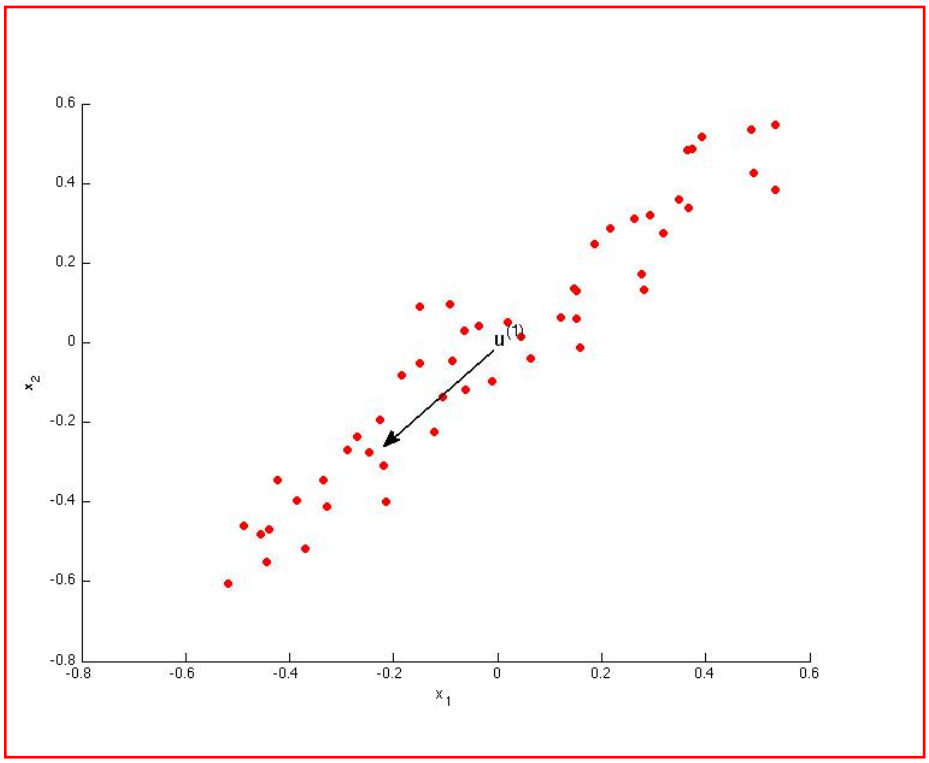


Which of the following figures correspond to possible values that PCA may return for $u^{(1)}$ (the first eigenvector / first principal component)? Check all that apply (you may have to check more than one figure).

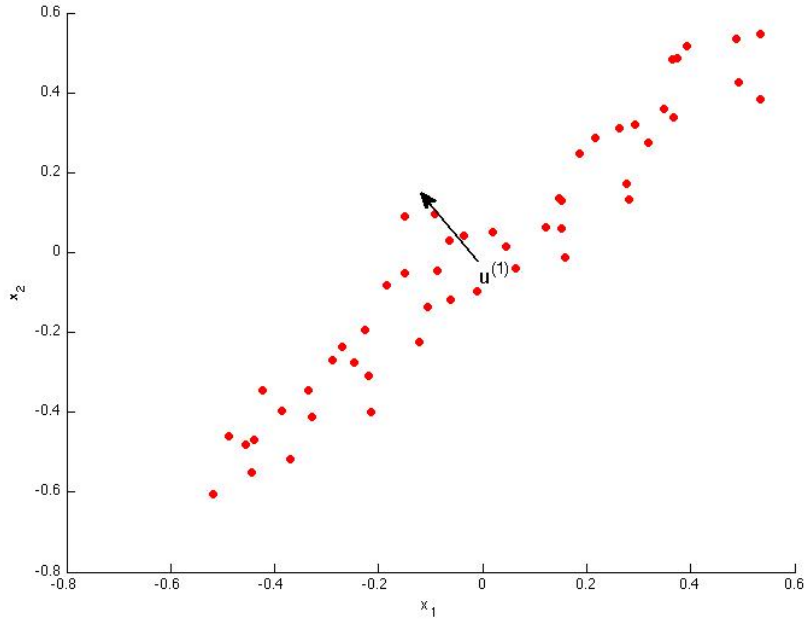
☐


0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
x_2
x_1
$u^{(1)}$

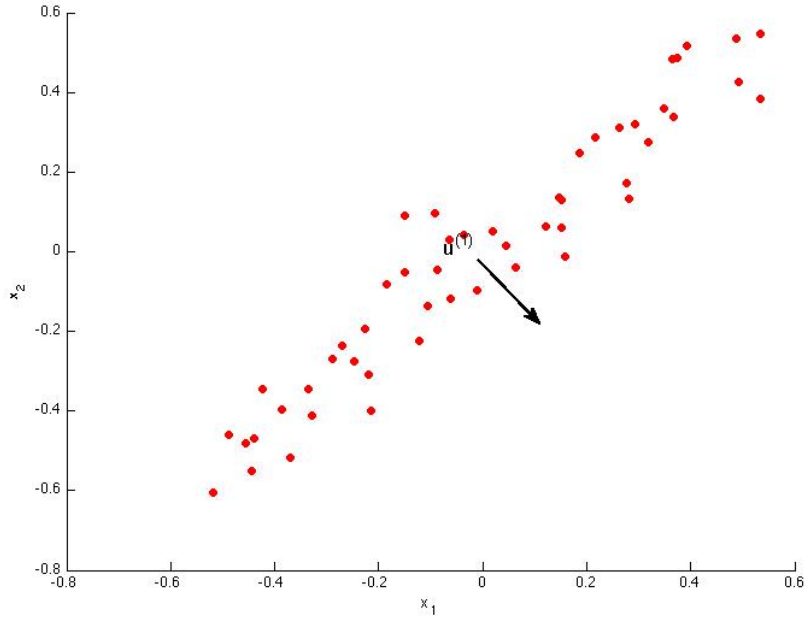
☐


0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
x_2
x_1
$u^{(1)}$

☐


0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
x2
x1
u(1)

☐


0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
x2
x1
u(1)

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.)

- Use the elbow method.
- Choose k to be the largest value so that at least 99% of the variance is retained
- Choose k to be the smallest value so that at least 99% of the variance is retained.
- Choose k to be 99% of m (i.e., $k = 0.99 * m$, rounded to the nearest integer).

Principal Component Analysis

Quiz, 5 questions

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?

- $\frac{\frac{1}{m}\sum_{i=1}^{m}||x^{(i)}-x_{\text{approx}}^{(i)}||^2}{\frac{1}{m}\sum_{i=1}^{m}||x^{(i)}||^2} \geq 0.05$
- $\frac{\frac{1}{m}\sum_{i=1}^{m}||x^{(i)}||^2}{\frac{1}{m}\sum_{i=1}^{m}||x^{(i)}-x_{\text{approx}}^{(i)}||^2} \geq 0.95$
- $\frac{\frac{1}{m}\sum_{i=1}^{m}||x^{(i)}-x_{\text{approx}}^{(i)}||^2}{\frac{1}{m}\sum_{i=1}^{m}||x^{(i)}||^2} \geq 0.95$
- $\frac{\frac{1}{m}\sum_{i=1}^{m}||x^{(i)}-x_{\text{approx}}^{(i)}||^2}{\frac{1}{m}\sum_{i=1}^{m}||x^{(i)}||^2} \leq 0.05$

1 point

4.

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

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

- [ ] 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 $x \in \mathbb{R}^n$, PCA compresses it to a lower-dimensional vector $z \in \mathbb{R}^k$.

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

Principal Component Analysis

Quiz, 5 questions

1 point

5.

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

- [ ] Data compression: Reduce the dimension of your data, so that it takes up less memory / disk space.

- [ ] To get more features to feed into a learning algorithm.

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

- [ ] Preventing overfitting: Reduce the number of features (in a supervised learning problem), so that there are fewer parameters to learn.

- [x] I, **Jun-Chieh Wang**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account. Learn more about Coursera's Honor Code

Submit Quiz