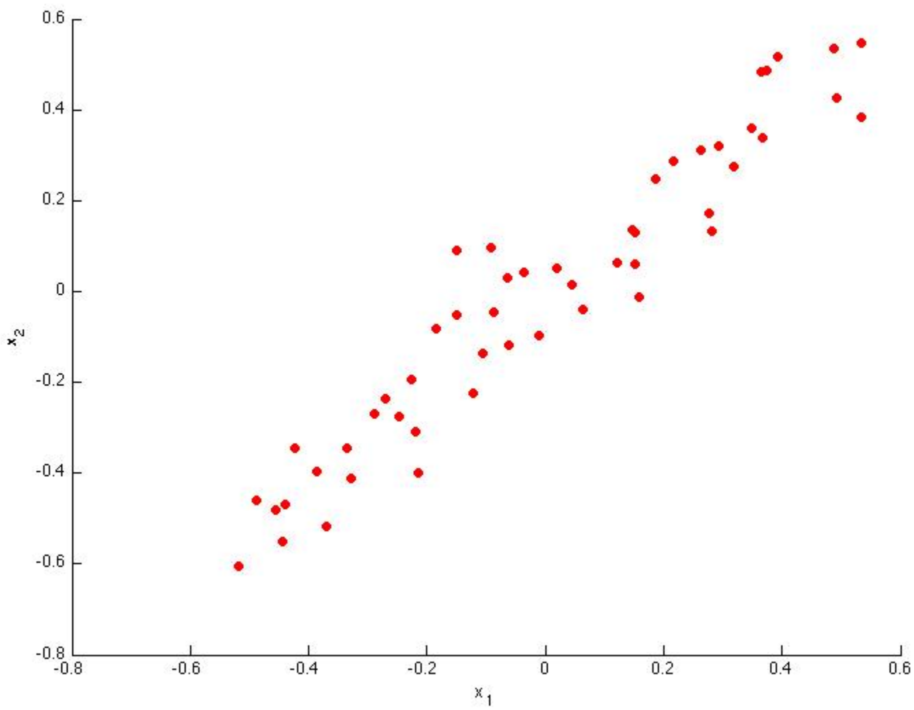
Principal Component Analysis

Quiz, 5 questions

1 point

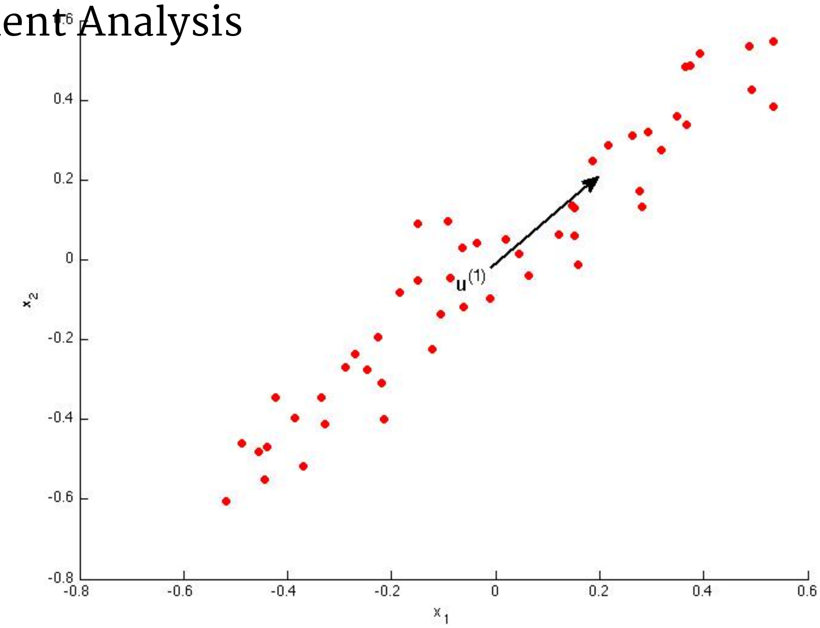
1.

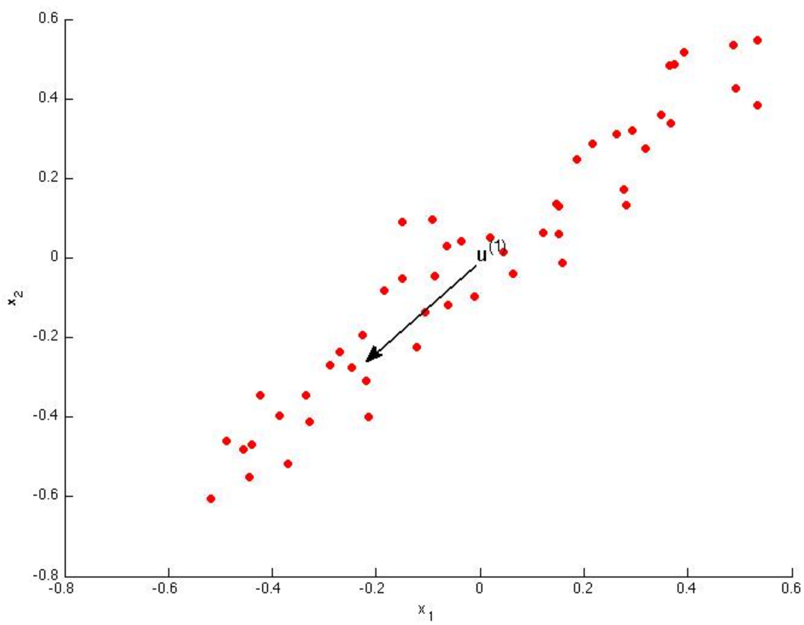
Consider the following 2D dataset:

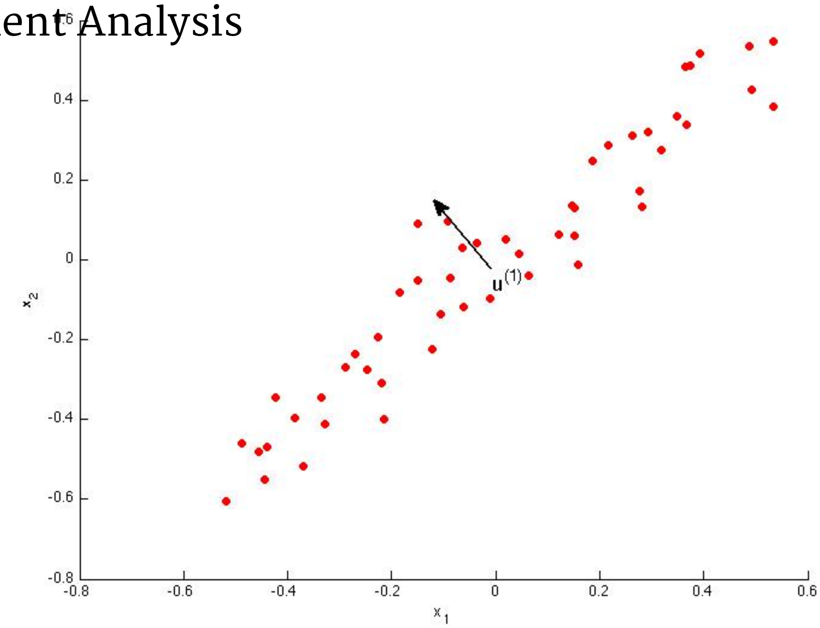


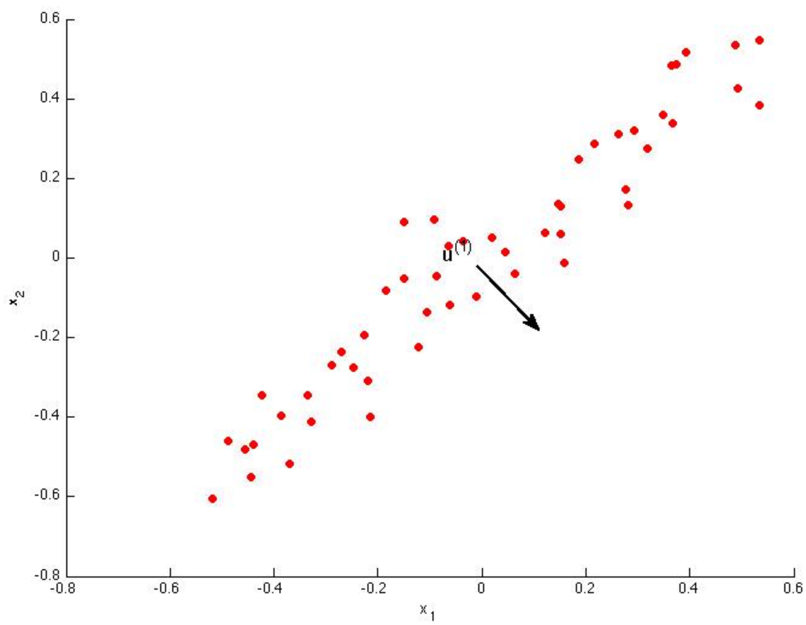
Which of the following figures correspond to possible values that PCA may return for $u^{(1)}$ (the first eigenvector / first principal component)? Check all that apply (you may have to check more than one figure).

☐


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
$u^{(1)}$





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 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).
- Use the elbow method.

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?

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

1
point

4.
Which of the following statements are true? Check all that apply.

Given input data $x \in \mathbb{R}^n$, it makes sense to run PCA only with values of k that satisfy $k \leq n$. (In particular, running it with $k = n$ is possible but not helpful, and $k > n$ does not make sense.)

- [ ] Even if all the input features are on very similar scales, we should still perform mean normalization (so that each feature has zero mean) before running PCA.

- [ ] Given only $z^{(i)}$ and U_{reduce}, there is no way to reconstruct any reasonable approximation to $x^{(i)}$.

- [ ] PCA is susceptible to local optima; trying multiple random initializations may help.

1
point

5.
Which of the following are recommended applications of PCA? Select all that apply.

- [ ] Data visualization: To take 2D data, and find a different way of plotting it in 2D (using k=2).

- [ ] 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).

- [ ] As a replacement for (or alternative to) linear regression: For most learning applications, PCA and linear regression give substantially similar results.

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

- [ ] I, **George Wolberg**, 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