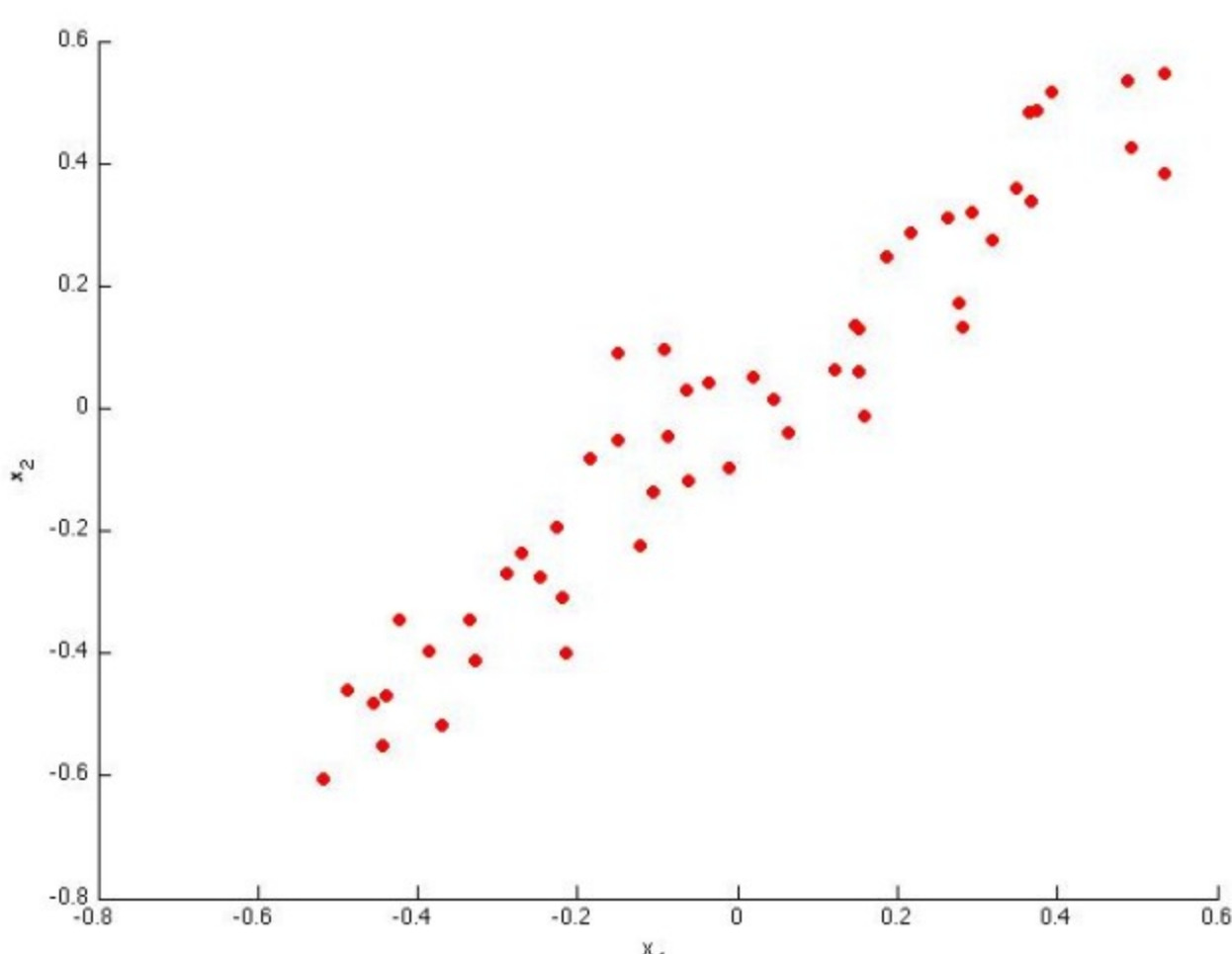


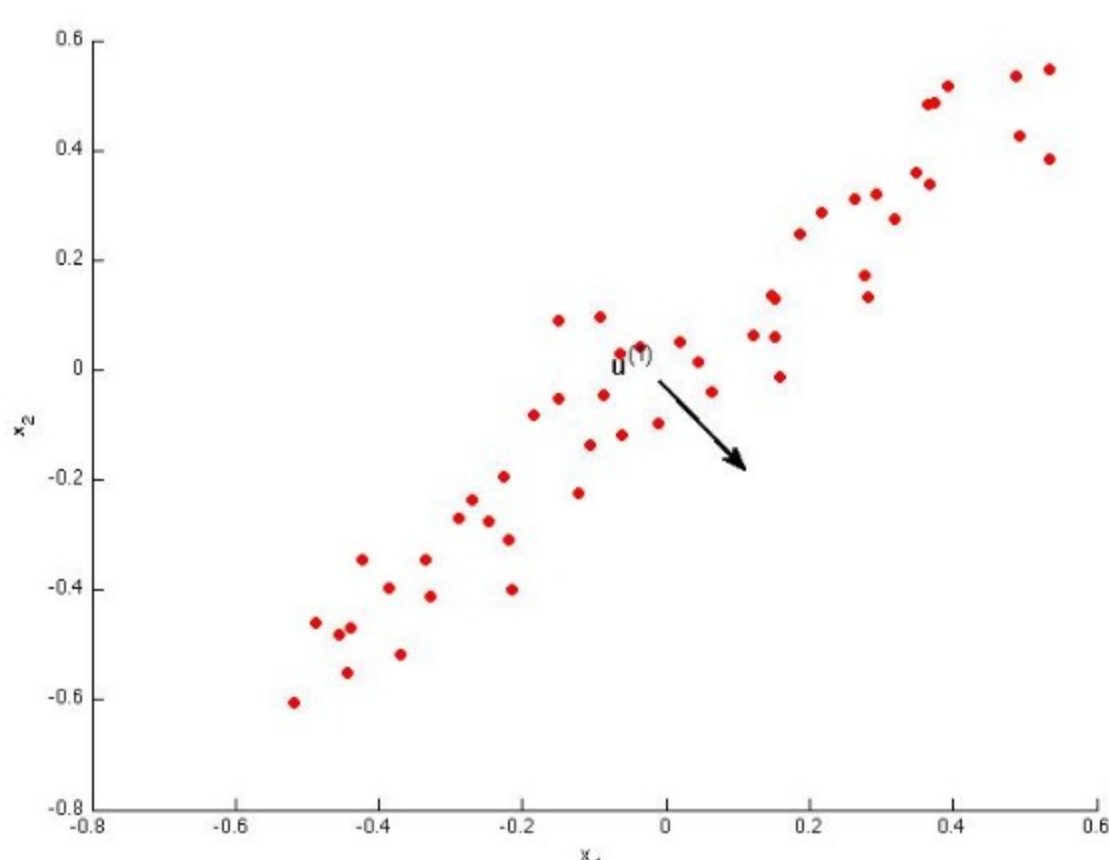
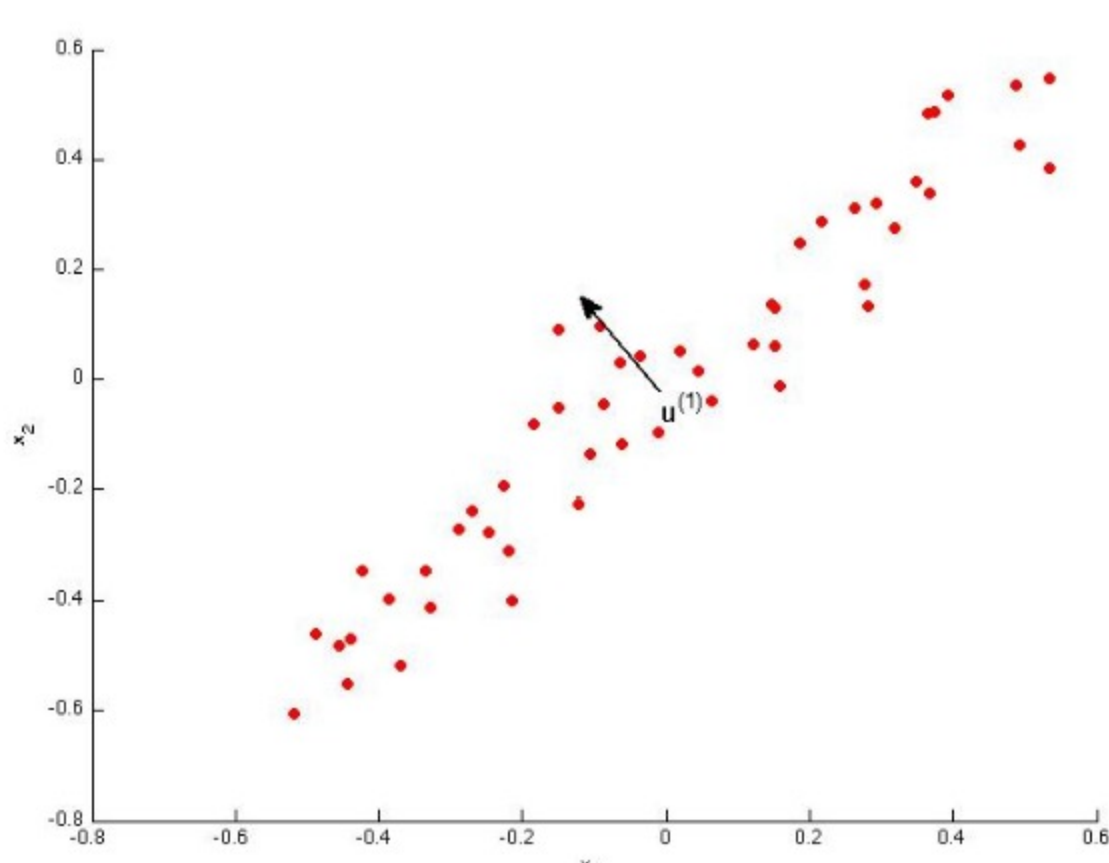
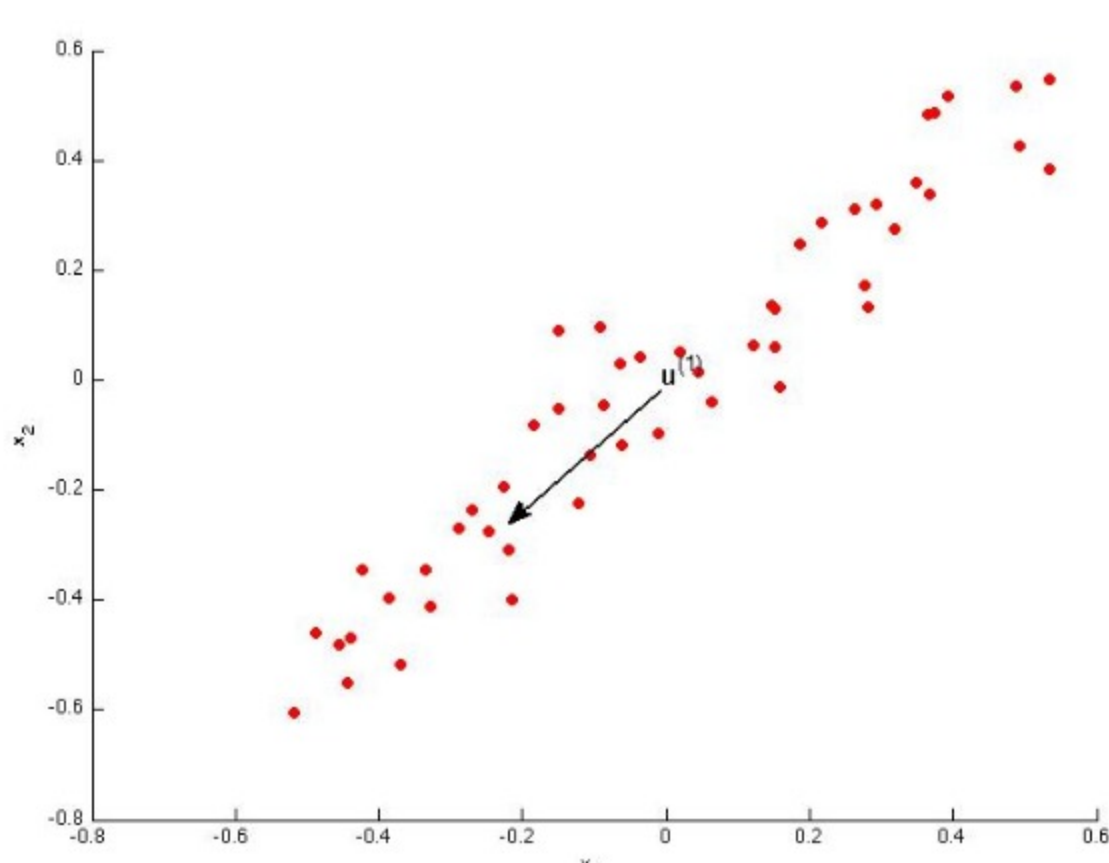
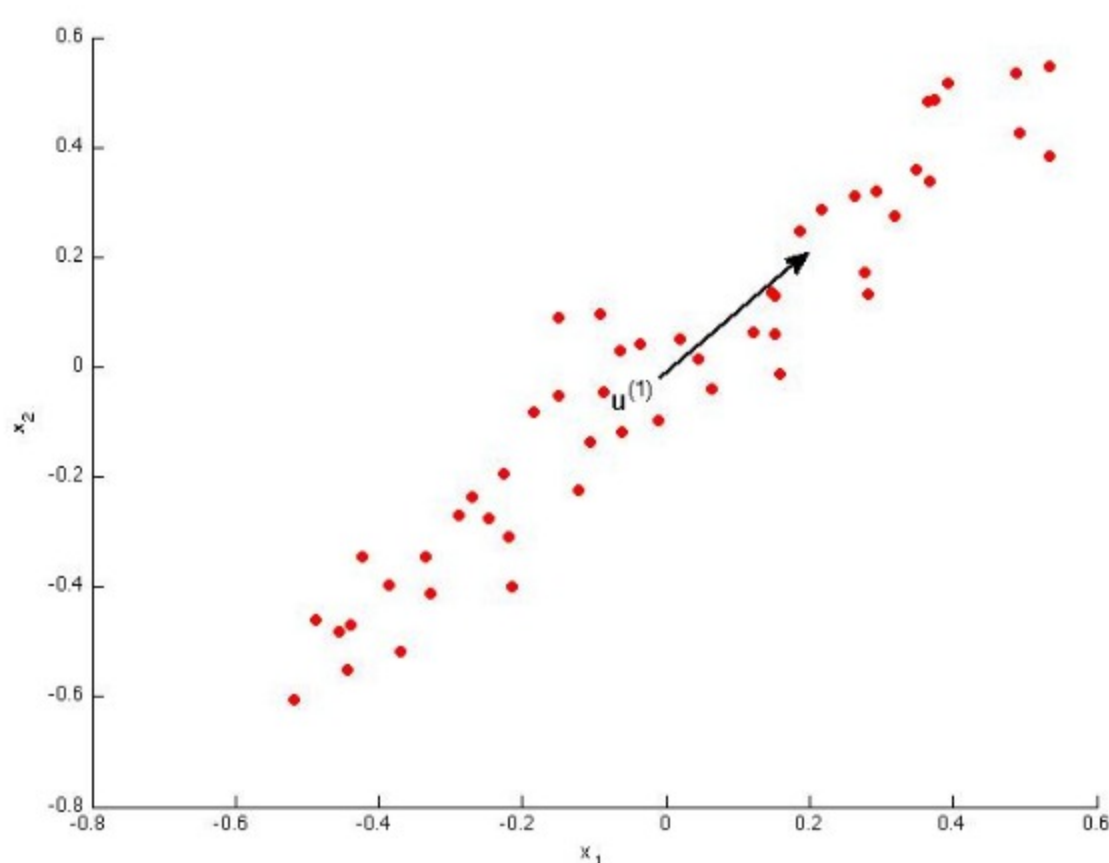


1 point

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 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?

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

1 point

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

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

- ☒ 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.
- ☐ To get more features to feed into a learning algorithm.
- ☒ Data visualization: Reduce data to 2D (or 3D) so that it can be plotted.

☒ I 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](#)

Attila Földvárszky

Use the name on your government issued ID

Submit Quiz