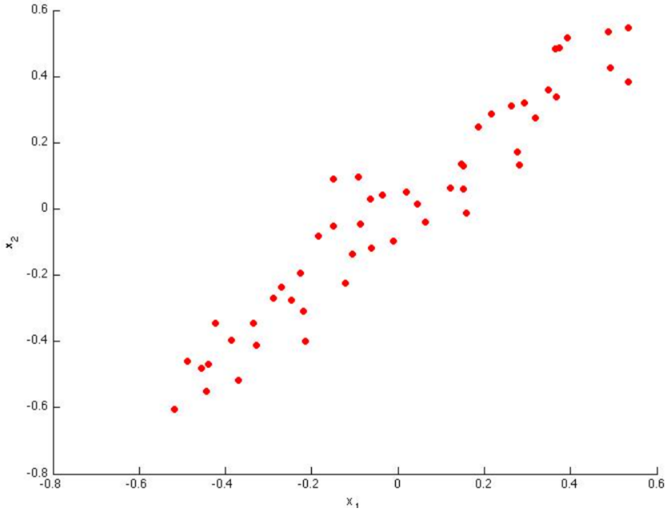
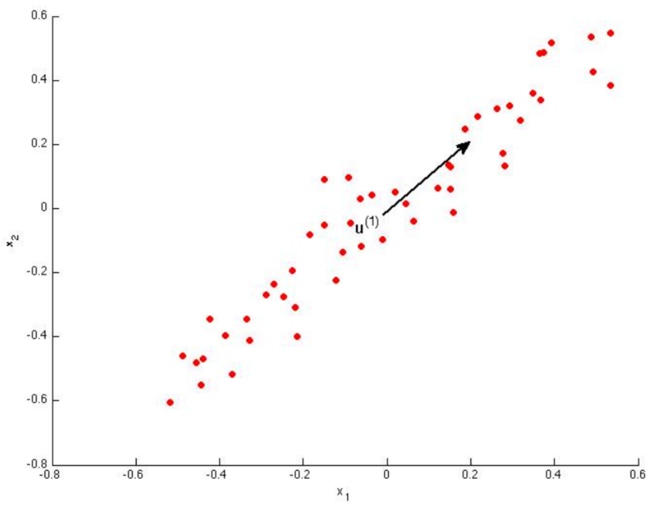
Week 8-2: PCA: Principal Component Analysis

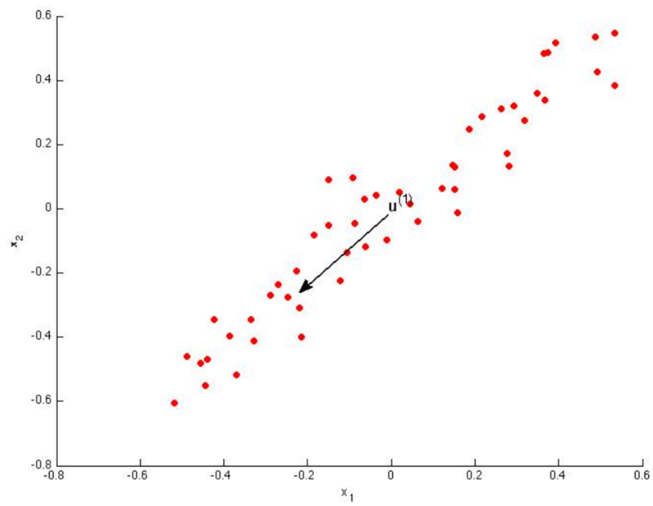
1. Consider the following 2D dataset:



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.

Eigenvectors can face either direction for the PCA; orientation doesn't matter. It looks a bit like linear regression but the algorithm is different.





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 smallest value so that at least 99% of the variance is retained.

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?

> We want the ratio of squared projection error to variation be less than $1 - 0.95 = 0.05$, where the error is $\frac{1}{m}\sum_{i=1}^{m}||x^{(i)} - x^{(i)}_{approx}||^2$ and variation is $\frac{1}{m}\sum_{i=1}^{m}||x^{(i)}||^2$.
>
> $$\boxed{\frac{\frac{1}{m}\sum_{i=1}^{m}||x^{(i)} - x^{(i)}_{approx}||^2}{\frac{1}{m}\sum_{i=1}^{m}||x^{(i)}||^2} \leq 0.05}$$

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

> (a) If the input features are on very different scales, it is a good idea to perform feature scaling before applying PCA.
>
> (b) Given an input $x \in \mathbb{R}^n$, PCA compresses it to a lower-dimensional vector $z \in \mathbb{R}^k$.

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

> (a) 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).
>
> (b) Data compression: Reduce the dimension of your data so that it takes up less memory/disk space.