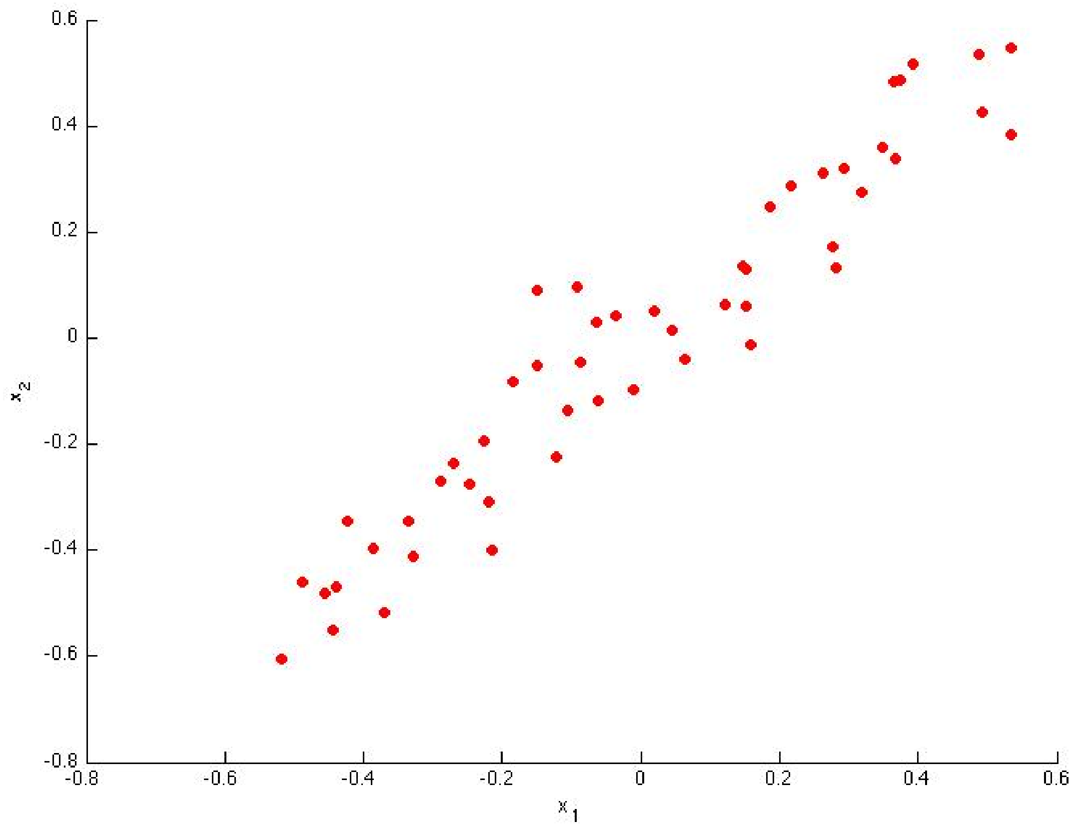
Feedback — XIV. Principal Component Analysis

Help

You submitted this quiz on **Tue 12 Aug 2014 8:48 AM PDT**. You got a score of **5.00** out of **5.00**.

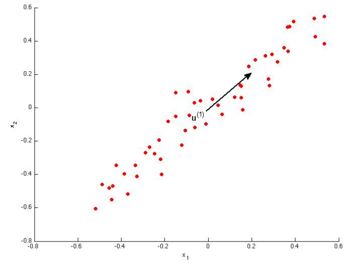
Question 1

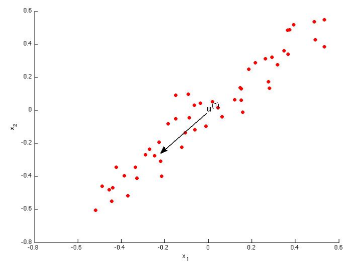
Consider the following 2D dataset:



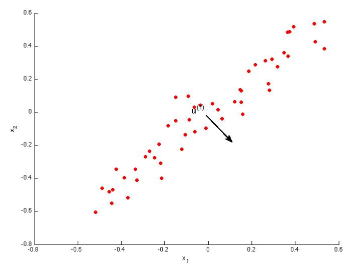
Which of the following figures correspond to possible values that PCA may return for $u^{(1)}$ (the first eigenvector / first principal component)? Check all that apply (you may have to check more than one figure).

Your Answer		Score	Explanation
✔	✔	0.25	The maximal variance is along the y = x line, so this option is correct.

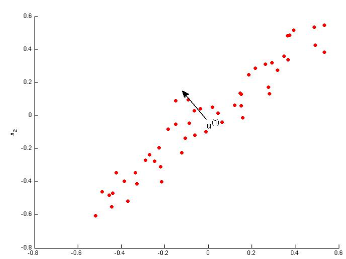




0.25 The maximal variance is along the y = x line, so the negative vector along that line is correct for the first principal component.



0.25 The first principal component is aligned with the direction of maximal variance, but this is aligned with the direction of minimal variance.



0.25 The first principal component is aligned with the direction of maximal variance, but this is aligned with the direction of minimal variance.

Total 1.00 / 1.00

Question 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.)

Your Answer		Score	Explanation
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.	✔	1.00	This is correct, as it maintains the structure of the data while maximally reducing its dimension.
Choose k to be the smallest value so that at least 1% of the variance is retained.			
Choose k to be 99% of n (i.e., $k = 0.99 * n$, rounded to the nearest integer).			
Total		1.00 / 1.00	

Question 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?

Your Answer		Score	Explanation
$\frac{\frac{1}{m}\sum_{i=1}^{m} \lvert\lvert x^{(i)} - x_{\text{approx}}^{(i)} \rvert\rvert^2}{\frac{1}{m}\sum_{i=1}^{m} \lvert\lvert x^{(i)} \rvert\rvert^2} \geq 0.95$			
$\frac{\frac{1}{m}\sum_{i=1}^{m} \lvert\lvert x^{(i)} \rvert\rvert^2}{\frac{1}{m}\sum_{i=1}^{m} \lvert\lvert x^{(i)} - x_{\text{approx}}^{(i)} \rvert\rvert^2} \leq 0.95$			
$\frac{\frac{1}{m}\sum_{i=1}^{m} \lvert\lvert x^{(i)} \rvert\rvert^2}{\frac{1}{m}\sum_{i=1}^{m} \lvert\lvert x^{(i)} - x_{\text{approx}}^{(i)} \rvert\rvert^2} \leq 0.05$			
$\frac{\frac{1}{m}\sum_{i=1}^{m} \lvert\lvert x^{(i)} - x_{\text{approx}}^{(i)} \rvert\rvert^2}{\frac{1}{m}\sum_{i=1}^{m} \lvert\lvert x^{(i)} \rvert\rvert^2} \leq 0.05$	✔	1.00	This is the correct formula.
Total		1.00 / 1.00	

Question 4

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

Your Answer		Score	Explanation
☑ 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.)	✔	0.25	The reasoning given is correct: with $k = n$, there is no compression, so PCA has no use.
☐ Feature scaling is not useful for PCA, since the eigenvector calculation (such as using Octave's `svd(Sigma)` routine) takes care of this automatically.	✔	0.25	Octave's routine does not perform feature scaling, so you should do so yourself.
☐ Given only $z^{(i)}$ and U_{reduce}, there is no way to reconstruct any reasonable approximation to $x^{(i)}$.	✔	0.25	You can easily reconstruct an approximation of $x^{(i)}$ by computing $U_{\text{reduce}}\ z^{(i)}$ where $z^{(i)}$ is padded with $n - k$ zeros in the computation.
☑ 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.	✔	0.25	If you do not perform mean normalization, PCA will rotate the data in a possibly undesired way.
Total		1.00 / 1.00	

Question 5

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

Your Answer		Score	Explanation
☑ Data compression: Reduce the dimension of your data, so that it takes up less memory / disk space.	✔	0.25	If memory or disk space is limited, PCA allows you to save space in exchange for losing a little of the data's information. This can be a reasonable tradeoff.
☐ As a replacement for (or alternative to) linear regression: For most learning applications, PCA and linear regression give substantially similar results.	✔	0.25	PCA is not linear regression. They have different goals (and cost functions), so they give different results.
☑ Data visualization: Reduce data to 2D (or 3D) so that it can be plotted.	✔	0.25	This is a good use of PCA, as it can give you intuition about your data that would otherwise be impossible to see.
☐ Data visualization: To take 2D data, and find a different way of plotting it in 2D (using k=2).	✔	0.25	You should use PCA to visualize data with dimension higher than 3, not data that you can already visualize.
Total		1.00 / 1.00	