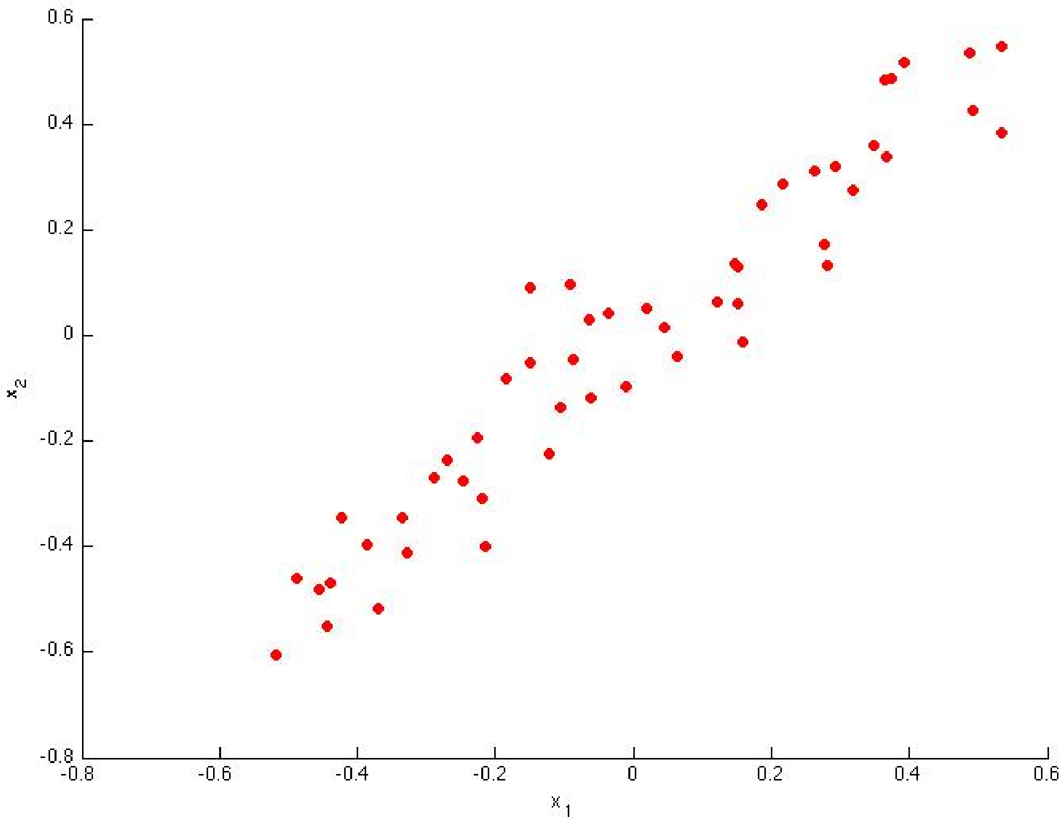
Feedback — XIV. Principal Component Analysis

You submitted this quiz on **Sat 15 Jun 2013 12:21 PM PDT (UTC -0700)**. You got a score of **5.00** out of **5.00**.

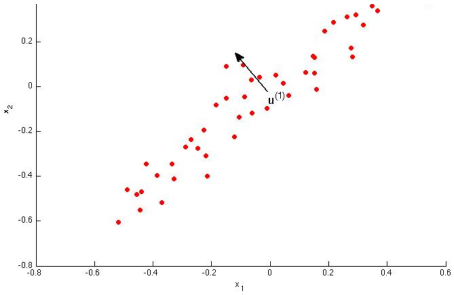
Question 1

Consider the following 2D dataset:



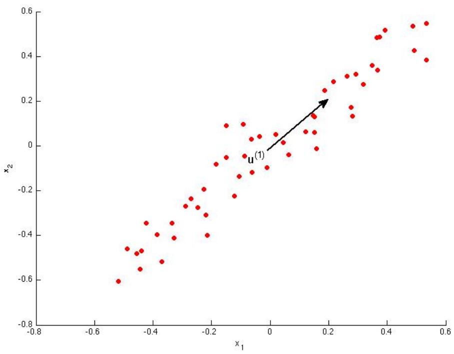
Which of the following figures correspond to possible values that PCA may return for $u^{(1)}$ (the first eigenvector / first principal component)? Check all that apply (you may have to check more than one figure).

Your Answer		Score	Explanation
	✔	0.25	The first principal component is aligned with the direction of maximal variance, but this is aligned with the



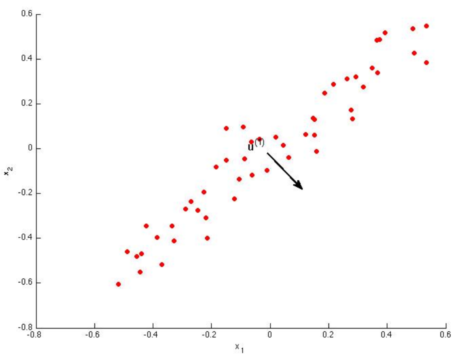
direction of minimal variance.

☑



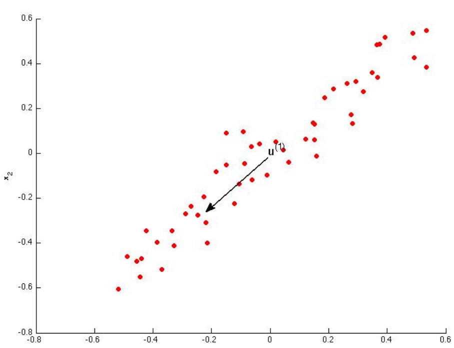
✔ 0.25 The maximal variance is along the y = x line, so this option is correct.

☐



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

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 99% of m (i.e., $k = 0.99 * m$, rounded to the nearest integer).			
Choose k to be the smallest value so that at least 1% 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 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
$\frac{\frac{1}{m}\sum_{i=1}^{m}\|x^{(i)}\|^2}{\frac{1}{m}\sum_{i=1}^{m}\|x^{(i)}-x_{\text{approx}}^{(i)}\|^2} \leq 0.05$			
$\frac{\frac{1}{m}\sum_{i=1}^{m}\|x^{(i)}-x_{\text{approx}}^{(i)}\|^2}{\frac{1}{m}\sum_{i=1}^{m}\|x^{(i)}\|^2} \leq 0.05$	✔	1.00	This is the correct formula.

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

Total 1.00 / 1.00

Question 4

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

Your Answer		Score	Explanation
☑ 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.
☑ If the input features are on very different scales, it is a good idea to perform feature scaling before applying PCA.	✔	0.25	Feature scaling prevents one feature dimension from becoming a strong principal component only because of the large magnitude of the feature values (as opposed to large variance on that dimension).
☐ PCA can be used only to reduce the dimensionality of data by 1 (such as 3D to 2D, or 2D to 1D).	✔	0.25	PCA can reduce data of dimension n to any dimension $k < n$.
☐ Feature scaling is not useful for PCA, since the eigenvector calculation (such as using Octave's $\texttt{svd(Sigma)}$ routine) takes care of this automatically.	✔	0.25	Octave's routine does not perform feature scaling, so you should do so yourself.
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.
☑ 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.
☐ To get more features to feed into a learning algorithm.	✔	0.25	PCA will reduce the number of features, not expand it.
☐ 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	