George Dallas

UK based Information Engineer/Internet Social Scientist

Principal Component Analysis 4 Dummies: Eigenvectors, Eigenvalues and Dimension Reduction

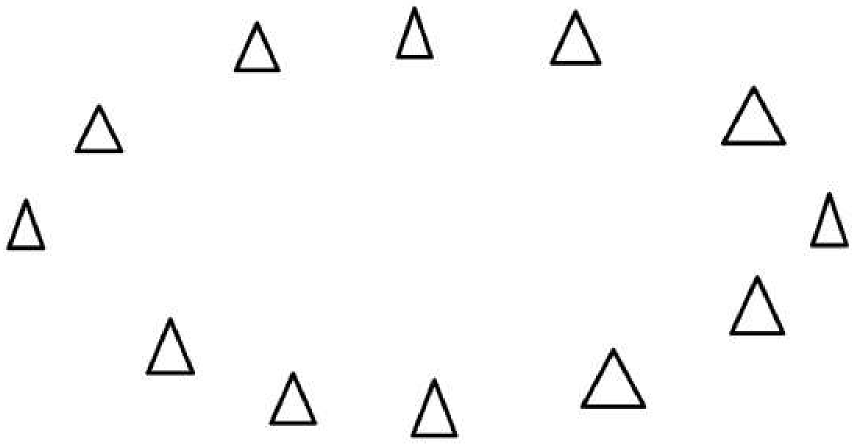
Having been in the social sciences for a couple of weeks it seems like a large amount of quantitative analysis relies on Principal Component Analysis (PCA). This is usually referred to in tandem with eigenvalues, eigenvectors and lots of numbers. So what's going on? Is this just mathematical jargon to get the non-maths scholars to stop asking questions? Maybe, but it's also a useful tool to use when you have to look at data. This post will give a very broad overview of PCA, describing eigenvectors and eigenvalues (which you need to know about to understand it) and showing how you can reduce the dimensions of data using PCA. As I said it's a neat tool to use in information theory, and even though the maths is a bit complicated, you only need to get a broad idea of what's going on to be able to use it effectively.

There's quite a bit of stuff to process in this post, but i've got rid of as much maths as possible and put in lots of pictures.

What is Principal Component Analysis?

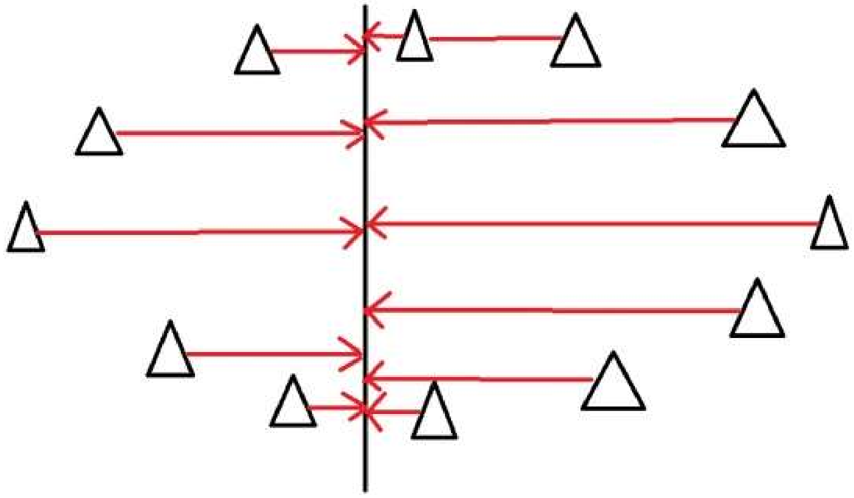
First of all Principal Component Analysis is a good name. It does what it says on the tin. PCA finds the principal components of data.

It is often useful to measure data in terms of its principal components rather than on a normal x-y axis. So what are principal components then? They're the underlying structure in the data. They are the directions where there is the most variance, the directions where the data is most spread out. This is easiest to explain by way of example. Here's some triangles in the shape of an oval:



(https://georgemdallas.files.wordpress.com/2013/10/pca3.jpg)

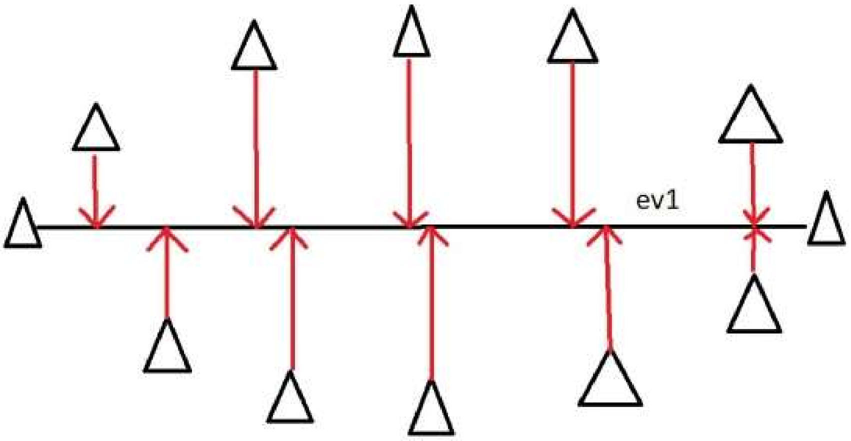
Imagine that the triangles are points of data. To find the direction where there is most variance, find the straight line where the data is most spread out when projected onto it. A vertical straight line with the points projected on to it will look like this:



(https://georgemdallas.files.wordpress.com/2013/10/pca9.jpg)

The data isn't very spread out here, therefore it doesn't have a large variance. It is probably not the principal component.

A horizontal line are with lines projected on will look like this:



(https://georgemdallas.files.wordpress.com/2013/10/pca8.jpg)

On this line the data is way more spread out, it has a large variance. In fact there isn't a straight line you can draw that has a larger variance than a horizontal one. A horizontal line is therefore the principal component in this example.

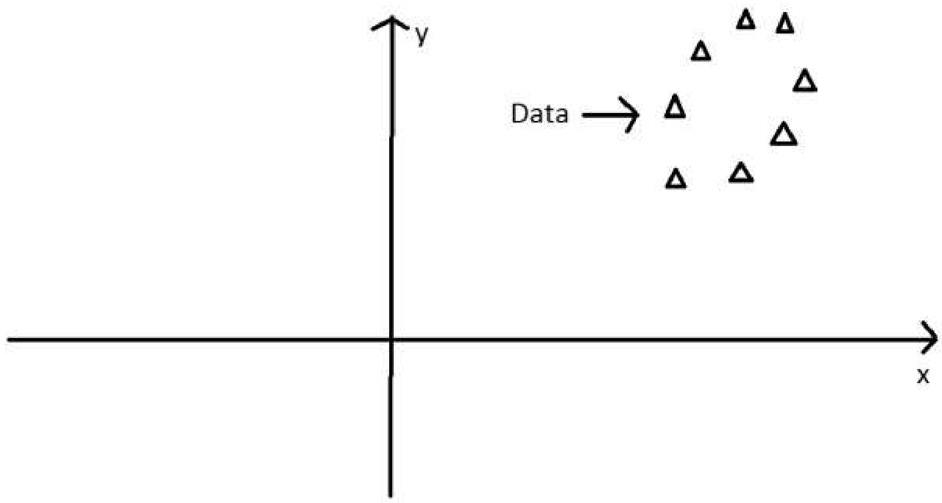
Luckily we can use maths to find the principal component rather than drawing lines and unevenly shaped triangles. This is where eigenvectors and eigenvalues come in.

Eigenvectors and Eigenvalues

When we get a set of data points, like the triangles above, we can deconstruct the set into eigenvectors and eigenvalues. Eigenvectors and values exist in pairs: every eigenvector has a corresponding eigenvalue. An eigenvector is a direction, in the example above the eigenvector was the direction of the line (vertical, horizontal, 45 degrees etc.) . An eigenvalue is a number, telling you how much variance there is in the data in that direction, in the example above the eigenvalue is a number telling us how spread out the data is on the line. The eigenvector with the highest eigenvalue is therefore the principal component.

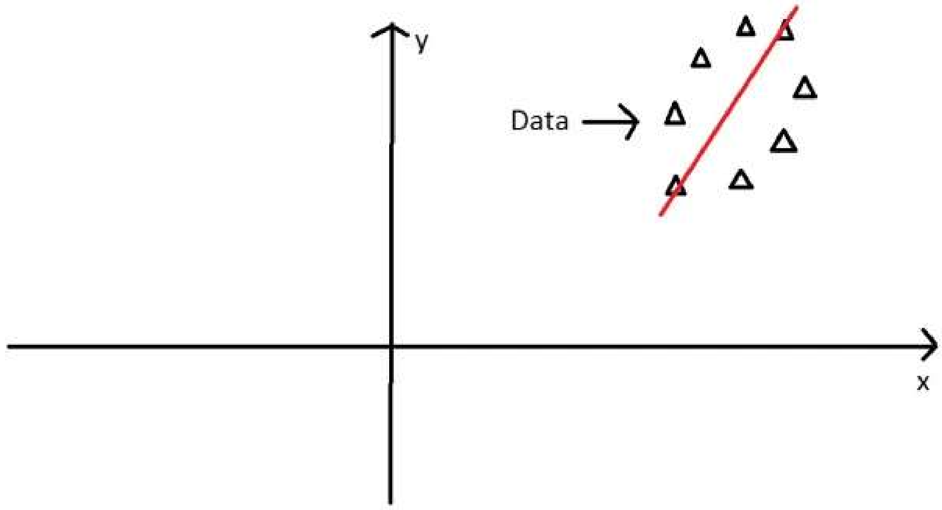
Okay, so even though in the last example I could point my line in any direction, it turns out there are not many eigenvectors/values in a data set. In fact the amount of eigenvectors/values that exist equals the number of dimensions the data set has. Say i'm measuring age and hours on the internet. there are 2 variables, it's a 2 dimensional data set, therefore there are 2 eigenvectors/values. If i'm measuring age, hours on internet and hours on mobile phone there's 3 variables, 3-D data set, so 3 eigenvectors/values. The reason for this is that eigenvectors put the data into a new set of dimensions, and these new dimensions have to be equal to the original amount of dimensions. This sounds complicated, but again an example should make it clear.

Here's a graph with the oval:



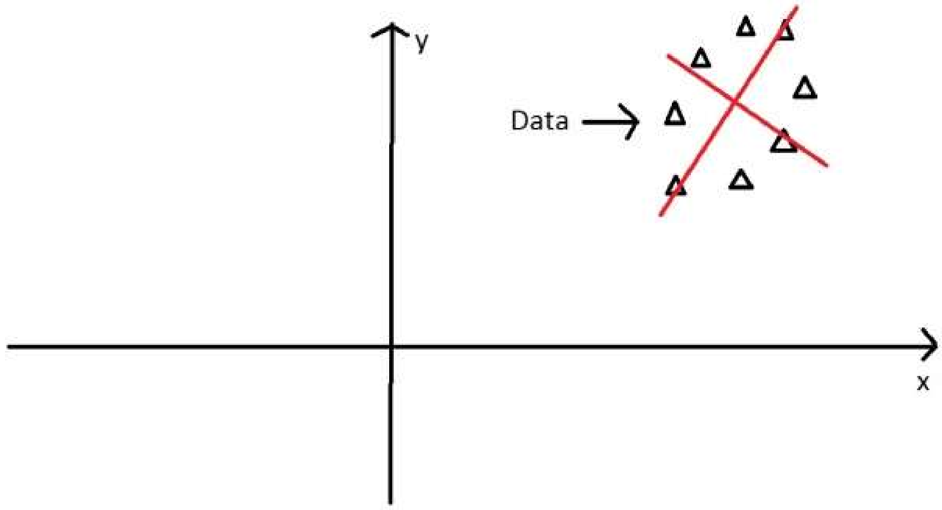
[(https://georgemdallas.files.wordpress.com/2013/10/pca2.jpg)](https://georgemdallas.files.wordpress.com/2013/10/pca2.jpg)

At the moment the oval is on an x-y axis. x could be age and y hours on the internet. These are the two dimensions that my data set is currently being measured in. Now remember that the principal component of the oval was a line splitting it longways:



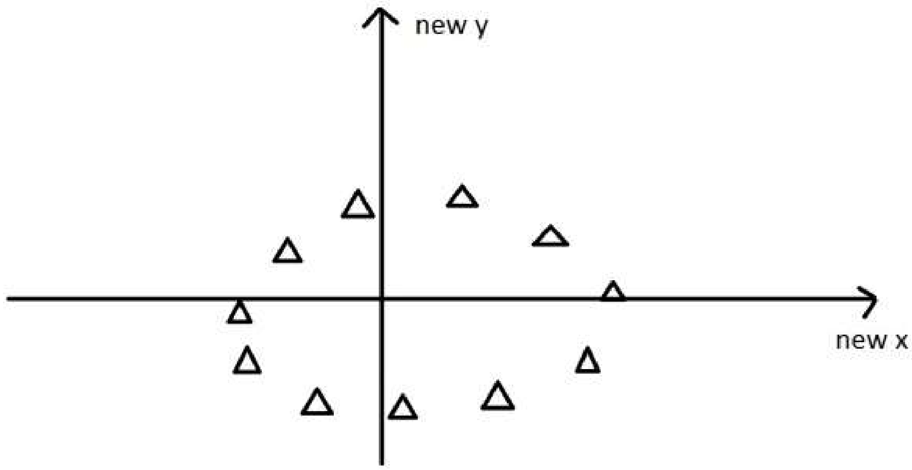
[(https://georgemdallas.files.wordpress.com/2013/10/pca10.jpg)](https://georgemdallas.files.wordpress.com/2013/10/pca10.jpg)

It turns out the other eigenvector (remember there are only two of them as it's a 2-D problem) is perpendicular to the principal component. As we said, the eigenvectors have to be able to span the whole x-y area, in order to do this (most effectively), the two directions need to be orthogonal (i.e. 90 degrees) to one another. This why the x and y axis are orthogonal to each other in the first place. It would be really awkward if the y axis was at 45 degrees to the x axis. So the second eigenvector would look like this:



(https://georgemdallas.files.wordpress.com/2013/10/pca11.jpg)

The eigenvectors have given us a much more useful axis to frame the data in. We can now re-frame the data in these new dimensions. It would look like this::



(https://georgemdallas.files.wordpress.com/2013/10/pca1.jpg)

Note that nothing has been done to the data itself. We're just looking at it from a different angle. So getting the eigenvectors gets you from one set of axes to another. These axes are much more intuitive to the shape of the data now. These directions are where there is most variation, and that is where there is more information (think about this the reverse way round. If there was no variation in the data [e.g. everything was equal to 1] there would be no information, it's a very boring statistic – in this scenario the eigenvalue for that dimension would equal zero, because there is no variation).

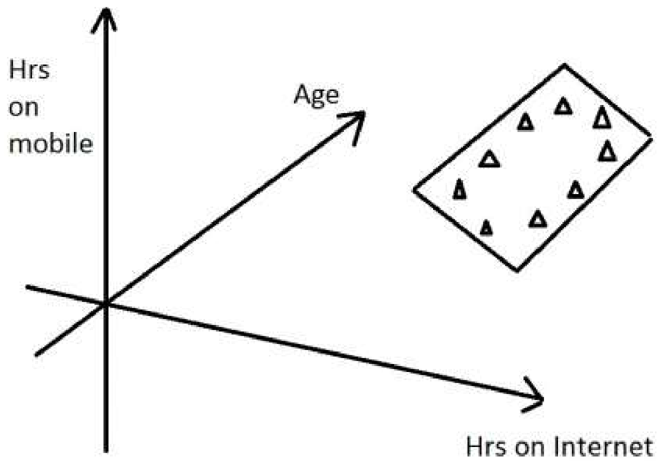
But what do these eigenvectors represent in real life? The old axes were well defined (age and hours on internet, or any 2 things that you've explicitly measured), whereas the new ones are not. This is where you need to think. There is often a good reason why these axes represent the data better, but maths won't tell you why, that's for you to work out.

How does PCA and eigenvectors help in the actual analysis of data? Well there's quite a few uses, but a main one is dimension reduction.

Dimension Reduction

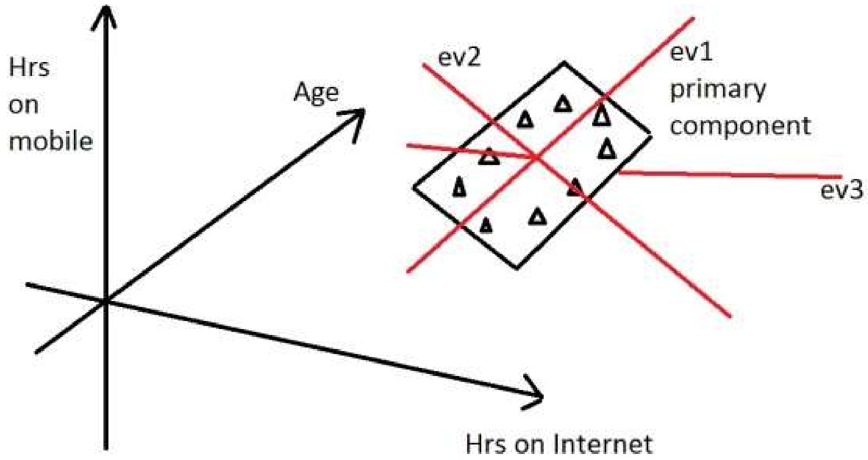
PCA can be used to reduce the dimensions of a data set. Dimension reduction is analogous to being philosophically reductionist: It reduces the data down into it's basic components, stripping away any unnecessary parts.

Let's say you are measuring three things: age, hours on internet and hours on mobile. There are 3 variables so it is a 3D data set. 3 dimensions is an x,y and z graph, It measure width, depth and height (like the dimensions in the real world). Now imagine that the data forms into an oval like the ones above, but that this oval is on a plane. i.e. all the data points lie on a piece of paper within this 3D graph (having width and depth, but no height). Like this:



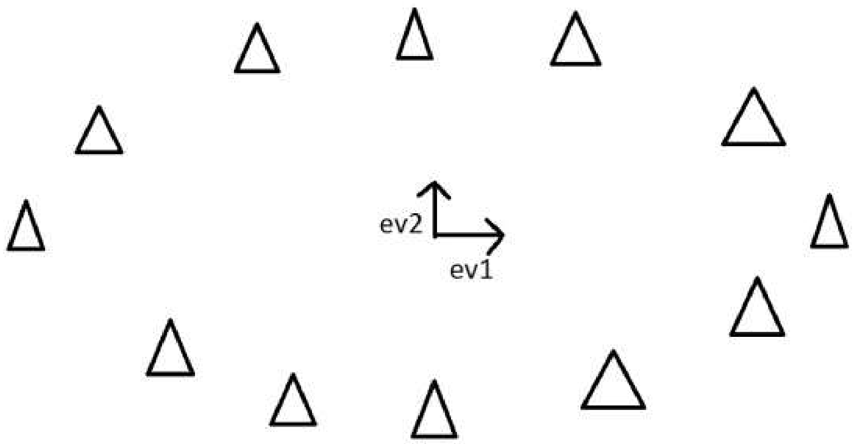
(https://georgemdallas.files.wordpress.com/2013/10/pca12.jpg)

When we find the 3 eigenvectors/values of the data set (remember 3D probem = 3 eigenvectors), 2 of the eigenvectors will have large eigenvalues, and one of the eigenvectors will have an eigenvalue of zero. The first two eigenvectors will show the width and depth of the data, but because there is no height on the data (it is on a piece of paper) the third eigenvalue will be zero. On the picture below ev1 is the first eignevector (the one with the biggest eigenvalue, the principal component), ev2 is the second eigenvector (which has a non-zero eigenvalue) and ev3 is the third eigenvector, which has an eigenvalue of zero.



(https://georgemdallas.files.wordpress.com/2013/10/pca13.jpg)

We can now rearrange our axes to be along the eigenvectors, rather than age, hours on internet and hours on mobile. However we know that the ev3, the third eigenvector, is pretty useless. Therefore instead of representing the data in 3 dimensions, we can get rid of the useless direction and only represent it in 2 dimensions, like before:



(https://georgemdallas.files.wordpress.com/2013/10/pca7.jpg)

This is dimension reduction. We have reduced the problem from a 3D to a 2D problem, getting rid of a dimension. Reducing dimensions helps to simplify the data and makes it easier to visualise.

Note that we can reduce dimensions even if there isn't a zero eigenvalue. Imagine we did the example again, except instead of the oval being on a 2D plane, it had a tiny amount of height to it. There would still be 3 eigenvectors, however this time all the eigenvalues would not be zero. The values would be something like 10, 8 and 0.1. The eigenvectors corresponding to 10 and 8 are the dimensions where there is alot of information, the eigenvector corresponding to 0.1 will not have much information at all, so we can therefore discard the third eigenvector again in order to make the data set more simple.

Example: the OxIS 2013 report

The OxIS 2013 report asked around 2000 people a set of questions about their internet use. It then identified 4 principal components in the data. This is an example of dimension reduction. Let's say they asked each person 50 questions. There are therefore 50 variables, making it a 50-dimension data set. There will then be 50 eigenvectors/values that will come out of that data set. Let's say the eigenvalues of that data set were (in descending order): 50, 29, 17, 10, 2, 1, 1, 0.4, 0.2….. There are lots of eigenvalues, but there are only 4 which have big values – indicating along those four directions there is alot of information. These are then identified as the four principal components of the data set (which in the report were labelled as enjoyable escape, instrumental efficiency, social facilitator and problem generator), the data set can then be reduced from 50 dimensions to only 4 by ignoring all the eigenvectors that have insignificant eigenvalues. 4 dimensions is much easier to work with than 50! So dimension reduction using PCA helped simplify this data set by finding the dominant dimensions within it.

OCTOBER 30, 2013 CHECKDETECTOR

267 thoughts on "Principal Component Analysis 4 Dummies: Eigenvectors, Eigenvalues and Dimension Reduction"

1. fareena says:
   Wow.Amazing Explanation. Thanks Dude

   REPLY APRIL 11, 2016 AT 11:31 AM
2. Vinoj John Hosan says:
   Thank u for this nice explanation. Really awesome material. Without a single equation, you have explained PCA perfectly, and conveyed the reasons why we are using.

   REPLY APRIL 14, 2016 AT 8:50 AM

3. Vincent says:

I've seen many posts on PCA and eigenvectors but none of them even come close to this. Really awesome stuff!!!

REPLY APRIL 14, 2016 AT 6:02 PM

4. muatik says:

you explain it intuitively, thank you.

REPLY APRIL 26, 2016 AT 11:00 AM

5. Lisette Pool says:

This is really useful! Thank you 🙂

REPLY APRIL 26, 2016 AT 12:00 PM

6. Tariq says:

Have read many textbook explanations and if only they would explain it as clearly as you have. Well done!

REPLY APRIL 27, 2016 AT 1:28 PM

7. Dhiraj says:

Good Explanation.. Look for its implementation in R

REPLY APRIL 30, 2016 AT 5:25 AM

8. operationunplugtexas says:

Great explanation, thank you!

REPLY APRIL 30, 2016 AT 11:17 PM

9. MIruna says:

I admit to not havin looked into any other sources on PCA yet,and it's my first brush with it, but i felt compelled to say that was a great explanation: straightforward, visual and concise.
Thank you

REPLY MAY 6, 2016 AT 9:51 PM

10. Anton says:

Great article! Really helped me understand the intuition behind this stuff.

Just a small correction: You wrote:

…the eigenvectors have to be able to span the whole x-y area, in order to do this (most effectively), the two directions need to be orthogonal (i.e. 90 degrees) to one another. This why the x and y axis are orthogonal to each other in the first place. It would be really awkward if the y axis was at 45 degrees to the x axis.

This is not really true, since there is a result that says:

Any two vectors in R^2 that are not scalar multiples of each other will span all of R^2

Hence, orthogonality is not required. In other words, if the y axis is 45 degrees to the x axis, both vectors would still span R^2 since they are not scalar multiples of each other.

I guess orthogonality is probably required to make sure the vectors can never be scalar multiples of each other. I think this is what you meant in brackets as most effectively.

REPLY MAY 12, 2016 AT 4:38 PM

- Дархан Медеуов says:

  Well, i well may be wrong, but I think that PC being a sort of linear transformation will not end up with collinear vectors anyway. So, may be the idea behind orthogonality is probably more related to general social theory concerns, rather than technical necessity in the way that it is easier to interpret orthogonal, that is uncorrelated, components than non-orthogonal and therefore correlated ones.

  REPLY AUGUST 30, 2016 AT 7:10 AM

11. Vinod K Subramani (@noddyvinod) says:

    Very well explained. Thank you for making it easy.

    REPLY MAY 13, 2016 AT 2:07 AM

12. gzm says:

    This is very helpful, thank you 🙂

    REPLY MAY 24, 2016 AT 6:11 PM

13. O.rka says:

    you explained it like I'm 5 and it made perfect intuitive sense. great work with the analogies.

    REPLY JUNE 3, 2016 AT 12:12 AM

14. Ashutosh Bhardwaj says:

    Excellent explanation in the most lucid way.

    REPLY JUNE 8, 2016 AT 11:45 AM

15. Prithvi Bhardwaj says:

    Too Good. It is not simple but you did a phenomenon job of making it look so simple. Great stuff.

    REPLY JUNE 15, 2016 AT 5:02 PM

16. Pingback: Principal Component Analysis (PCA) – People's Revolutionary Committee of West New York
17. waiguru says:

    big time ……this what I have been looking for …than you

    REPLY JUNE 29, 2016 AT 11:55 AM

18. J.C. says:

    This article saved my life!!!!! THANK YOU SOOO MUCH!

    REPLY JULY 1, 2016 AT 3:56 AM

19. Mohamed Nour l. Ismail says:

    Concise & precise. Great.

    REPLY JULY 4, 2016 AT 1:04 AM

20. Peter Kamerman says:

    Thanks for the great explanation.

REPLY JULY 4, 2016 AT 6:20 PM

21. Sanjit says:

    Very nice. Neat and clean with specific example.

    REPLY JULY 11, 2016 AT 10:42 AM

22. Sohini says:

    Excellently explained. So intuitive that it would stay in one's mind for a long time. Kudos!

    REPLY JULY 18, 2016 AT 9:37 AM

23. Pingback: Making Sense of: Principle Component Analysis (PCA) | Silver Rice Bowl
24. Pingback: Making Sense of : Principal Component Analysis (PCA) | Silver Rice Bowl
25. shimg says:

    Wow. very diffult concepts are explained in very easy way.

    REPLY JULY 19, 2016 AT 5:33 AM

26. GoodJob says:

    Amazing explanation! Good Job!

    REPLY JULY 23, 2016 AT 2:02 PM

27. Pingback: Principle Component Understanding | Humble. Learner.
28. A. Setijanto says:

    Excellent. People just need to read this before studying PCA.

    REPLY AUGUST 11, 2016 AT 4:34 AM

29. ILAN LIVNE says:

    Bravo!

    REPLY AUGUST 13, 2016 AT 10:36 PM

30. UPUL Liyanage says:

    Excellent. I have read lot of document to understand PCA but most of them are useless and very complicated. This is a very simple and easy to understand document. Thanks again

    REPLY AUGUST 14, 2016 AT 4:48 PM

31. Sriharsha Vinnakota says:

    Perfect..Thanks a lot

    REPLY AUGUST 24, 2016 AT 3:36 PM

32. white says:

    awesome article i have been studying pca since last six months yet could not find this like explanation and easy for understanding thing about pca, thankyou so much for helping me,,

    REPLY AUGUST 25, 2016 AT 6:29 AM

33. Vivek Patil says:

    Great article for the beginners. Thank you.

    REPLY SEPTEMBER 7, 2016 AT 8:33 PM

34. Farai Leboho says:

Well, excellent rendition of this hard to understand topic, thank you a lot, a very lot.

REPLY SEPTEMBER 8, 2016 AT 9:03 AM

35. Pingback: Using Principle Component Analysis to Understand the College Football AP Poll | mantena notes
36. Aing says:

    Wow Excellent explanation

    REPLY OCTOBER 21, 2016 AT 10:16 AM
37. claire says:

    Good explanation but saying it is for dummies is rude.

    REPLY NOVEMBER 3, 2016 AT 2:14 PM
38. yonas says:

    nice explanation!!

    REPLY NOVEMBER 4, 2016 AT 2:17 PM
39. Trevor Dunn says:

    Really good, concise and easy to understand explanation.

    REPLY NOVEMBER 24, 2016 AT 12:15 PM
40. Samuel says:

    Just perfect! Thanks a lot!

    REPLY NOVEMBER 25, 2016 AT 11:55 AM
41. ali erdogan says:

    Perfect explanation. You've helped me a lot, thanks!

    REPLY NOVEMBER 29, 2016 AT 7:26 PM
42. Andrew Kennett says:

    thanks that is a great explanation

    REPLY JANUARY 6, 2017 AT 2:46 AM
43. puneet says:

    basic and crisp; just like it should be. Though i need to read more for detailed understanding; but that wouldn't be possible without the basics.
    thank you.

    REPLY JANUARY 6, 2017 AT 6:46 AM
44. Pavan says:

    Thanks a lot. good explanation!!!

    REPLY JANUARY 6, 2017 AT 3:00 PM
45. Jobiba says:

    This nailed it! Finally I understand what PCA is all about. Thank you!!!

    REPLY JANUARY 6, 2017 AT 3:49 PM
46. Stephen says:

Hi, thank you very much for this guide! Do you have some suggested reading or another guide which covers when we consider an eigenvalue "insignificant" and some examples of interpreting a PCA? thanks again!

REPLY JANUARY 9, 2017 AT 12:54 AM

47. Matt says:

Thank you very much for this explanation. I am currently writing a journal club where I work focused on using PCA to compare the pathology of people with Alzheimers disease. I've made reference to this page in it and will be sending people here for a better explanation! Thank you for taking the time to explain it all so well and clearly!

REPLY JANUARY 17, 2017 AT 7:15 PM

48. Harikrishnan says:

Nice explanation. Thank you.

REPLY JANUARY 23, 2017 AT 2:04 AM

49. Tally Hatzakis says:

Thanks. Finally, exactly what I was looking for…

REPLY FEBRUARY 3, 2017 AT 6:23 PM