A simple principal component analysis example

Brian Russell, August, 2011.

Introduction

In this tutorial, we will look at the basics of principal component analysis using a simple numerical example. In the first section, we will first discuss eigenvalues and eigenvectors using linear algebra. In the second section, we will look at eigenvalues and eigenvectors graphically. Finally, in the last two sections, we will show how an understanding of the eigenvalue/eigenvector problems leads us to principal component analysis.

Eigenvalues and eigenvectors – the linear algebra approach

The example we will be using is taken from seismic analysis, were we consider how to compute the principal components of M seismic attributes, each with N samples. Before discussing principal component analysis, we need to understand the concepts of eigenvectors and eigenvalues.

Let us start by considering the following two three-sample attribute traces (that is, N = 3 and M = 2):

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \text{ and } Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

These two attribute vectors can be combined into the matrix S as

$$S = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

From the matrix of attributes, we then compute the following covariance matrix (note that to make this problem numerically easier we have not normalized C by dividing it by N = 3):

A simple principal component analysis example

Brian Russell, August, 2011.

$$C = S^T S = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

The covariance matrix C can also be written as

$$C = \begin{bmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{YX} & \sigma_{YY} \end{bmatrix},$$

where σ_{XX} is the auto-covariance of attribute X with itself (and also equal to the square of the variance of X), σ_{YY} is the auto-covariance of attribute Y with itself, σ_{XY} is the cross-covariance of attribute X with attribute Y, and σ_{YX} is the cross-covariance of attribute X with attribute Y. Notice that $\sigma_{XY} = \sigma_{YX}$ so that the covariance matrix is symmetric. This is always the case. The fact that $\sigma_{XX} = \sigma_{YY}$ is simply due to the fact that attributes X and Y have the same root-mean-square amplitude, which is not always the case.

We now want to compute the eigenvalues and eigenvectors of C. The eigenvalue equation is given by

$$Cu = \lambda u, \qquad (1)$$

where λ is an eigenvalue (a scalar) and $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is an eigenvector. If all the attribute vectors are independent there are M eigenvalues and eigenvectors, where M is the number of attributes, in this case 2. Furthermore, for the correlation matrix, which is real and symmetric, all of the eigenvalues are real and positive. Note that equation (1) also has a visual interpretation. It says that if we multiply the matrix C by the eigenvector u we get exactly the same result as multiplying the eigenvector by a single number, the eigenvalue λ. We will get back to the visual interpretation in the second part of this tutorial. But first we will discuss the linear algebra solution of equation (1).

For those not familiar with matrix algebra, let us write the eigenvalue equation out in a little more detail:

$$\begin{bmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{XY} & \sigma_{YY} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \lambda \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \lambda u_1 \\ \lambda u_2 \end{bmatrix}. \qquad (2)$$

This can be simplified to give

$$\begin{bmatrix} \sigma_{XX} - \lambda & \sigma_{XY} \\ \sigma_{XY} & \sigma_{YY} - \lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \qquad (3)$$

If we plug in the numbers from our initial example, we therefore get:

$$\begin{bmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Often, it is useful to guess at an answer before actually doing any calculations. For example, let's guess that $\lambda = 1$. This gives us:

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Is there any vector *u* that would make this equation correct? A little thought will tell you that there are an infinite number, each satisfying the relationship that $u_1 = u_2$. We have actually solved for one eigenvalue/eigenvector pair without even doing any mathematics.

To find the second eigenvalue/eigenvector pair, let's turn the problem around a little bit and ask: is there another simple matrix/vector pair similar to the one above? What about

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}?$$

In this case, the set of vectors that solve the problem are given by $v_1 = -v_2$ (note that we have called the second eigenvector *v*). A little more thought will show you that the eigenvalue that produces this result is $\lambda = 3$. This turns out to be the correct form of the second eigenvalue/eigenvector pair. Thus, we have solved this problem just by trial and error.

Of course, in most cases, especially if there are more than two attributes, the solution is much more difficult and we need a more foolproof method. To solve for the eigenvalues, we use the determinant of the matrix in equation (3) to give a quadratic equation which can be solved as follows

$$\begin{vmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - 4\lambda + 3 = 0 \Rightarrow \lambda_1 = 3, \text{ and } \lambda_2 = 1.$$

As expected, the eigenvalues are as we predicted earlier, although we have called the larger of the two the first eigenvalue. To solve for the eigenvectors, we simply substitute the two eigenvalues into the matrix equation (3), as before. It is also general practice to find the simplest eigenvector in each case by normalizing it so that the sum of the squares of its components equals 1. Thus, we get:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

These two eigenvectors can be put into matrix form as

$$U = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

The eigenvector matrix is orthonormal, which means that when it is multiplied by its transpose we get the identity matrix, or

$$UU^T = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Also, the transpose and inverse of U are identical. That is:

$$U^T = U^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Eigenvalues and eigenvectors – the visual approach

Let us now go back and fine a visual interpretation of equation (1), which you recall was written

$$Cu = \lambda u,$$

where $C = \begin{bmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{XY} & \sigma_{YY} \end{bmatrix}$ is the symmetric covariance matrix, $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is an eigenvector and λ is an eigenvalue. Most textbooks teach us to perform matrix multiplication by multiplying each row of the matrix C by the vector u in the following way:

$$\begin{bmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{XY} & \sigma_{YY} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \sigma_{XX} u_1 + \sigma_{XY} u_2 \\ \sigma_{XY} u_1 + \sigma_{YY} u_2 \end{bmatrix}.$$

But Gilbert Strang, in his book Linear Algebra, points out that this is equivalent to multiplying each of the columns of C by the two components of u and summing the result, as follows:

$$\begin{bmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{XY} & \sigma_{YY} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = u_1 c_1 + u_2 c_2 = u_1 \begin{bmatrix} \sigma_{XX} \\ \sigma_{XY} \end{bmatrix} + u_2 \begin{bmatrix} \sigma_{XY} \\ \sigma_{YY} \end{bmatrix} = \begin{bmatrix} \sigma_{XX} u_1 + \sigma_{XY} u_2 \\ \sigma_{XY} u_1 + \sigma_{YY} u_2 \end{bmatrix},$$

where $c_1 = \begin{bmatrix} \sigma_{XX} \\ \sigma_{XY} \end{bmatrix}$ and $c_2 = \begin{bmatrix} \sigma_{XY} \\ \sigma_{YY} \end{bmatrix}$ are the two columns of C.

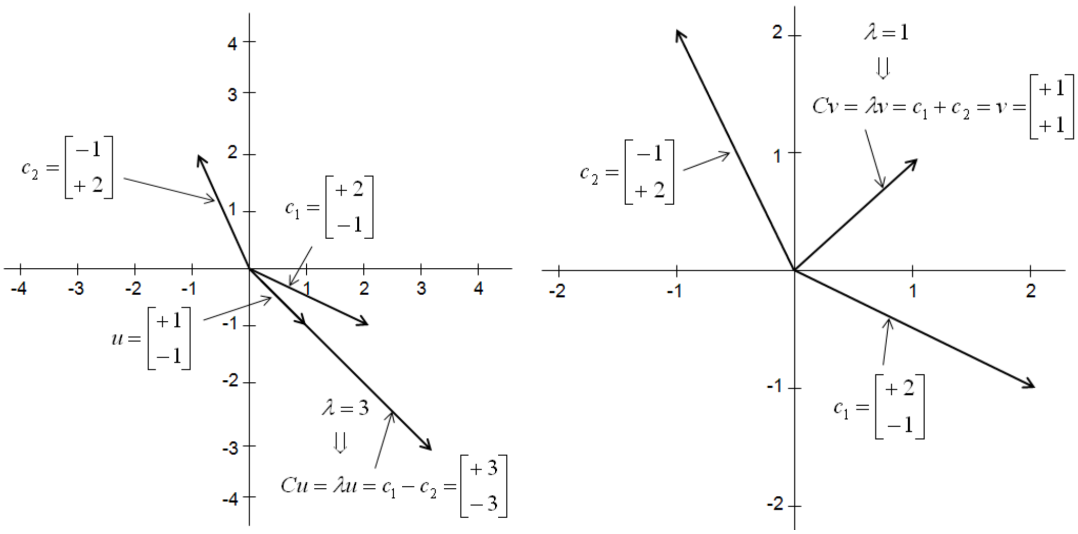
This is actually a much more intuitive way to think of matrix multiplication than the way we are normally taught, and leads us to a visual way of interpreting the eigenvalue and eigenvector problem. By plugging in the values from our covariance matrix and the first eigenvalue/eigenvector pair, we get (note that we have left off the normalization scaling to make this a simpler problem, since this is still a valid eigenvalue/eigenvector pair):

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ -1 \end{bmatrix} \Rightarrow 1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} - 1 \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}.$$

This is shown visually in Figure 1(a). Next, consider the second eigenvalue/eigenvector pair, which can be written as

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow 1\begin{bmatrix} 2 \\ -1 \end{bmatrix} + 1\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

This is shown visually in Figure 1(b).



(a) (b)

Figure 1. This figure shows a visual interpretation of eigenvalues and eigenvectors described in the text, where (a) shows the first eigenvalue/eigenvector pair and (b) shows the second eigenvalue/eigenvector pair.

Principal Components

Now we are in a position to compute the principal components of *S*. The principal components are created by multiplying the components of each eigenvector by the attribute vectors and summing the result. That is, for the two principal components, P_1 and P_2, we can write

$$P_1 = u_1 X + u_2 Y, \text{ and}$$
$$P_2 = v_1 X + v_2 Y.$$

Substituting the values for the eigenvectors, we see that the principal components of the two attribute vectors are simply the scaled difference and sums given by:

$$P_1 = \frac{1}{\sqrt{2}}(X - Y), \text{ and } P_2 = \frac{1}{\sqrt{2}}(X + Y).$$

In matrix form, the principal component matrix is the product of the attribute matrix A and the eigenvector matrix U:

$$P = SU = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix} \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}.$$

Numerically, the individual principal component traces or vectors are

$$P_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} \text{ and } P_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Note also that the principal component matrix has the property that when it is multiplied by its transpose we recover the eigenvalues in diagonal matrix form:

$$P^T P = \frac{1}{2} \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

We have thus transformed the original attributes into new traces which are orthogonal (that is, their cross-correlation is equal to 0) and auto-correlate to the successive eigenvalues. As a final point, note that we can also recover the attributes from the principal components by a linear sum. First, recall that we showed that the inverse and transpose of the eigenvector matrix are identical. Therefore, we can write

$$PU^T = SUU^T = S \Rightarrow S = \frac{1}{2} \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

In other words, we can recover the attributes as follows:

$$X = \frac{1}{2}(P_1 + P_2), \text{ and } Y = \frac{1}{2}(P_2 - P_1).$$

Principal component analysis using geometry

A simple principal component analysis example

Brian Russell, August, 2011.

In this section, we will discuss the same problem as in the previous section. However, this time we will look at the problem from a geometrical point of view. Recall that we initially defined the matrix a in terms of two column vectors, which we called attributes, and which were written as

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix},\ Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \Rightarrow S = \begin{bmatrix} X & Y \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

An alternate interpretation of the matrix S is as three row vectors, which are two-dimensional points in attribute space. This is written:

$$S = \begin{bmatrix} A^T \\ B^T \\ C^T \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}, \text{ where } A = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \text{ and } C = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

These three points are plotted in Figure 2(a). Also shown in this figure is a dashed line drawn at -45^o to the vertical axis, which goes through point 1 and bisects the other two points. Using this line as reference, notice that point 1 is at a distance $\sqrt{2}$ from the origin, and points 2 and 3 are a distance $\sqrt{2}/2$ from the line.

Now, recall that to compute the two principal components, we used the equation

$$P = AU, \text{ where } U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

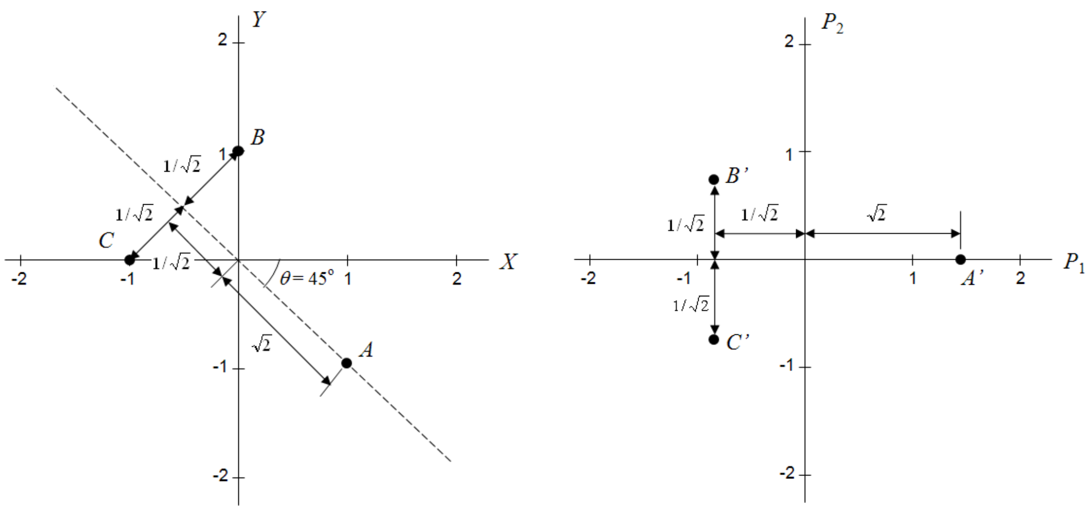
An alternate interpretation of the matrix U is as the rotation matrix R, where the rotation angle is -45^o:

$$P = AR, \text{ where } R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \text{ and } \theta = \frac{-\pi}{4} = -45^0.$$

After application of the rotation, we can interpret the rotated points as

$$P = \begin{bmatrix} A'^T \\ B'^T \\ C'^T \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}, \text{ where } A' = \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}, B' = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \text{ and } C' = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

This geometric interpretation is shown in Figure 2(b). Note that this transform has preserved the same distance relationships that were shown in Figure 2(a).



(a) (b)

Figure 2. This figure shows a geometrical interpretation of principal component analysis (PCA), where (a) shows the points in our example before PCA and (b) shows the points after PCA. Note that the transform involves a rotation of the axes.