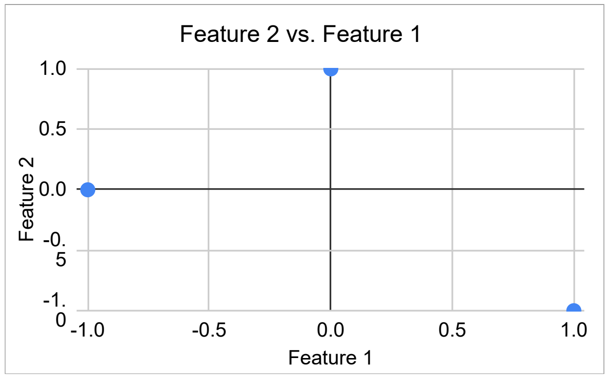
Principal Component Analysis PCA Solved Example

Feature 1	Feature 2
1	-1
0	1
-1	0



Principle components with greater amount of variance are grouped under one category.
Principle components with smaller variance are grouped under the second category.

S =

1	-1
0	1
-1	0

Compute the (un-normalized) covariance matrix, C

C = S'S =

2	-1
-1	2

Compute the eigenvalues and eigenvectors of C by solving

$Cu = \lambda u$ or $(C - \lambda I)u = 0$

If all the attribute vectors are independent, then we have M=2 eigenvalues and M = 2 eigenvectors

Because the Correlation matrix is real and symmetric all the eigenvalues are real and positive.

$Cu = \lambda u$ or $(C - \lambda I)u = 0$

To solve for the eigenvalues, we use the determinant of the matrix to get a quadratic equation

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

$$(2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = 0 \Rightarrow \lambda_1 = 3, \lambda_2 = 1$$

Note that the lambda's are listed in descending order of magnitude
• To solve for the eigenvectors, we simply substitute the two eigenvalues into the matrix equation.
• 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.

Eigenvectors: For $\lambda = 1$ • For $\lambda = 1$ $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ which gives $u_1 = u_2$

Eigenvectors: For $\lambda = 3$ • For $\lambda = 3$ $\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ which gives $u_1 = -u_2$

Normalized Eigenvectors

It is customary to normalize the eigenvectors

• The normalized eigenvectors are

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}; \qquad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Note: u is the eigenvector associated with larger eigenvalue

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 a:ribute vectors and summing the result.

• That is, for the two principal components, P1 and P2, we can write P 1 = u1 X + u2 Y ; and P 2 = v1X + v2Y.

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

$$P_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}; \quad P_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

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:

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

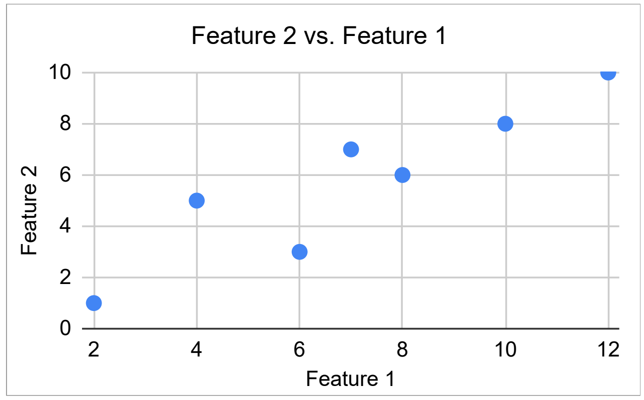
Because the inverse and transpose of the eigenvector matrix are identical, we can write

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

These columns are the values of X & Y recovered!

Solved Example 2: Principal Component Analysis PCA

Feature 1	Feature 2
2	1
4	5
6	3
7	7
8	6
10	8
12	10



Principle components with greater amount of variance are grouped under one category. Principle components with smaller variance are grouped under the second category.