For our eigenvector example, we had

$$\vec{A} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$. It's eigenvalues are $\lambda_1 = 4$ and $\lambda_2 = 2$

Its eigenvectors are $\vec{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\vec{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Note that these eigenvectors are orthogonal

$$\vec{x}_1^T x_2 = (1 \;\; 1)\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0$$

$$x_2^T x_1 = (1 \, -1)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0$$

Let's rewrite the eigenvectors slightly

$$\vec{x}_1 = \sqrt{2}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad \text{and} \quad \vec{x}_2 = \sqrt{2}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

For the first eigenvector we know

$$\vec{A}\vec{x}_1 = \lambda_1 \vec{x}_1 \Rightarrow \vec{A}\cancel{\sqrt{2}}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = 4\cancel{\sqrt{2}}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

Similar for second eigenvector, so an equivalent set of eigenvector for $\vec{A}$ are $\vec{x}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ and $\vec{x}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ with the same eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 2$.

In this form, the eigenvectors are orthonormal, meaning

$$\vec{x}_1^T \vec{x}_2 = \vec{x}_2^T \vec{x}_1 = 0 \quad \text{as before}$$

and

$$\vec{x}_1^T \vec{x}_1 = \left(\frac{1}{\sqrt{2}} \;\; \frac{1}{\sqrt{2}}\right)\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = 1$$

$$\vec{x}_2^T \vec{x}_2 = \left(\frac{1}{\sqrt{2}} \;\; -\frac{1}{\sqrt{2}}\right)\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = 1$$

For our eigenvector example, note also that

$\vec{A} = \vec{A}^t$ meaning the matrix is symmetric about it's diagonal. Symmetric matrices have nice properties.

① They have real eigenvalues

② They can be written in the form $\vec{Q}\vec{\Lambda}\vec{Q}^T = \vec{A}$, where $\vec{Q}$ is an orthogonal matrix (i.e. colums are orthogonal to one another) and $\vec{\Lambda}$ is a diagonal matrix.

For an example $\vec{A} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$

$$\underbrace{\vec{Q} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}}_{\text{colums of eigenvectors}} = \vec{Q}^T \qquad \vec{\Lambda} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \text{ eigenvalues along diagonal}$$

$$\vec{Q}\vec{\Lambda}\vec{Q}^t = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 2\sqrt{2} & 2\sqrt{2} \\ \sqrt{2} & -\sqrt{2} \end{pmatrix}$$

$$= \frac{\sqrt{2}}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 1 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

Suppose we want to represent an arbitrary vector $\begin{pmatrix} a \\ b \end{pmatrix}$ in terms of the orthonormal eigenvectors

$$\begin{pmatrix} a \\ b \end{pmatrix} = c_1 \vec{x}_1 + c_2 \vec{x}_2$$

where c_1 and c_2 are constants to be found. Multiply both sides by $\vec{x}_1^{\,T}$

$$\vec{x}_1^{\,T} \begin{pmatrix} a \\ b \end{pmatrix} = c_1 \underbrace{\vec{x}_1^{\,T} \vec{x}_1}_{1} + c_2 \underbrace{\vec{x}_1^{\,T} \vec{x}_2}_{0}$$

$$c_1 = \vec{x}_1^{\,T} \begin{pmatrix} a \\ b \end{pmatrix} \qquad \text{Same for other constants}$$

Thus we have a nice way of decomposing an arbitrary vector into a set of orthonormal eigen vectors. In turn, $\vec{A}$ acting on these eigen vectors gives the corresponding eigenvalue.

<u>PRINCIPAL COMPONENTS ANALYSIS</u> - This is a technique for analyzing data that utilizes the preceding concepts. It tries to find patterns in the data that correspond to the variance in the data variables and tries to find the directions that account for the variance in rank order. Let's start with some data to illustrate the process

x	y	$x-\bar{x}$	$y-\bar{y}$
1	0.496	-7	-2.16275
3	2.101	-5	-0.55775
5	2.397	-3	-0.26175
7	2.651	-1	-0.00775
9	2.088	1	-0.57075
11	3.111	3	0.45225
13	4.044	5	1.38525
15	4.382	7	1.72325

Note $\bar{x}, \bar{y}$ are means

$\bar{x} = 8$

$\bar{y} = 2.65875$

After subtracting the mean from the data variables, a covariance matrix is formed.

Covariance =

$$\mathrm{cov}(x,y) = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$$ where N is # of data points

Note this is similar to the definition of variance

$$\text{variance} = \mathrm{var}(x) = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

Since we subtracted the mean value off of the original data, these definitions simplify to

$$\mathrm{cov}(x_2, y_2) = \frac{1}{N-1}\sum_{i=1}^{N} x_{2i}\, y_{2i}$$

$$\mathrm{var}(x_2) = \frac{1}{N-1}\sum_{i=1}^{N} x_{2i}^2$$

where $x_2 = x - \bar{x}$ and $y_2 = y - \bar{y}$ and $\bar{x}_2 = \bar{y}_2 = 0$

The covariance matrix looks like

$$\begin{pmatrix} \mathrm{var}(x_2) & \mathrm{cov}(x_2, y_2) \\ \mathrm{cov}(y_2, x_2) & \mathrm{var}(y_2) \end{pmatrix}$$

Note further that $\mathrm{cov}(y_2, x_2) = \mathrm{cov}(x_2, y_2)$ so this matrix is symmetric

For our example

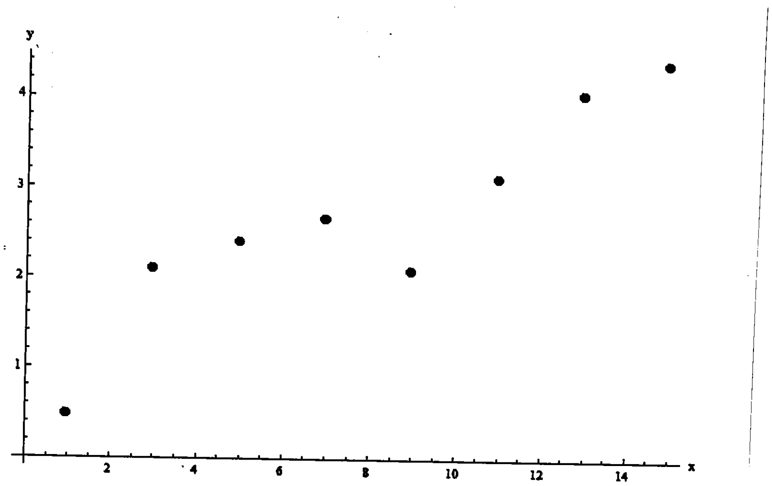
$$\begin{pmatrix} 24 & 5.49943 \\ 5.49943 & 1.49656 \end{pmatrix}$$

The eigenvalues and eigenvectors of the covariance matrix are

$$\lambda_1 = 25.2721 \qquad \vec{x}_1 = \begin{pmatrix} -0.97428 \\ -0.22536 \end{pmatrix}$$

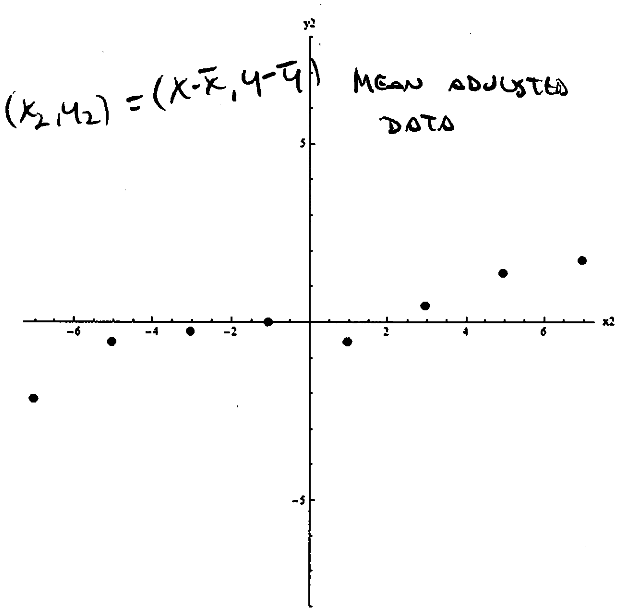
$$\lambda_2 = 0.224509 \qquad \vec{x}_2 = \begin{pmatrix} 0.22536 \\ -0.97428 \end{pmatrix}$$

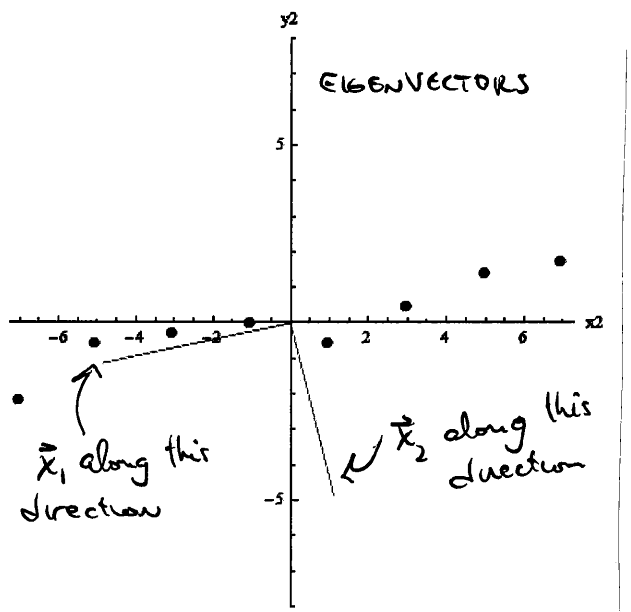
So most of the variance is occurring in the $\vec{x}_1$ direction and only a little in the $\vec{x}_2$ direction.



ORIGINAL DATA

(x, y)





We can also use this technique to compress the data. Let's take each point $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ and multiply it by $\vec{x}_1^T$ and $\vec{x}_2^T$

$$c_1 = \vec{x}_1^T \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \qquad c_2 = \vec{x}_2^T \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = c_1 \vec{x}_1 + c_2 \vec{x}_2$$

but we know that most of the variance occurs in the $\vec{x}_1$ direction, so let's approximate our data as

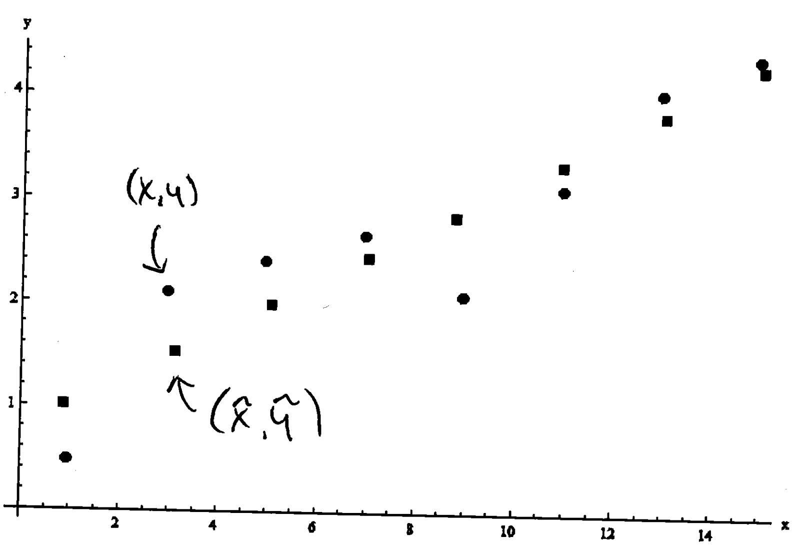
$$\begin{pmatrix} \hat{x}_2 \\ \hat{y}_2 \end{pmatrix} = c_1 \vec{x}_1$$

We can also get back to original data by adding $\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}$ back in

$$\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = c_1 \vec{x}_1 + \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}$$

These estimates lie along the line

$$y = 0.808297 + 0.231307\,x$$



If we did a least squares fit of the original data to a line, we would get

$$y = 0.825667 + 0.229143\,x$$

Former: minimizes $\perp$ distance to the line.

Latter: minimizes Δy to the line.