PRINCIPAL COMPONENTS ANALYSIS (PCA)

What is PCA ?

Principal Component Analysis (PCA) is the general name for a technique which uses mathematical tools from Linear Algebra to transform a complex data set with large number of variables into a smaller number of variables called principal components.

It is therefore often the case that an examination of the reduced dimension data set will allow the user to spot trends, patterns and outliers in the data, far more easily than would have been possible without performing the principal component analysis.

What do we need under our BELT

- Basics of statistical measures , e.g. variance and covariance.

- Basics of linear algebra:
  - Matrices
  - Vector space
  - Basis
  - Eigenvectors and Eigenvalues

- **Variance** measure of the deviation from the mean for points in one dimension.

- **Covariance** a measure of how much each of the dimensions varies from the mean with respect to each other.

$$\operatorname{var}(X) = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)\left(X_i - \overline{X}\right)}{(n-1)}$$

$$\operatorname{cov}(X, Y) = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{(n-1)}$$

- Covariance is measured between 2 dimensions to see if there is a relationship between the 2 dimensions.

❑ What is the interpretation of covariance calculations?

Say you have a 2- dimensional data set
X: number of hours studied for a subject
Y : marks obtained in that subject

- And assume the covariance value (between X and Y) is: 104.5

What does this mean?

Exact value is not as important as sign

❑ A positive value of covariance indicates that both the dimensions increase or decrease together

❑ A negative value indicates that while one increases, the other decreases and vice-versa.

❑ A zero value indicates that both the variables are independent of each other.

Covariance Matrix

- Representing covariance among dimensions as a matrix , e.g. , for 3 dimensions :

$$C = \begin{bmatrix} \operatorname{cov}(X,X) & \operatorname{cov}(X,Y) & \operatorname{cov}(X,Z) \\ \operatorname{cov}(Y,X) & \operatorname{cov}(Y,Y) & \operatorname{cov}(Y,Z) \\ \operatorname{cov}(Z,X) & \operatorname{cov}(Z,Y) & \operatorname{cov}(Z,Z) \end{bmatrix}$$

❖ Properties:

- Diagonal : Variances of the variables
- cov(X,Y) = cov(Y,X), hence matrix is symmetrical about the diagonal
- m-dimensional data will result in m x m covariance matrix

Inner Product

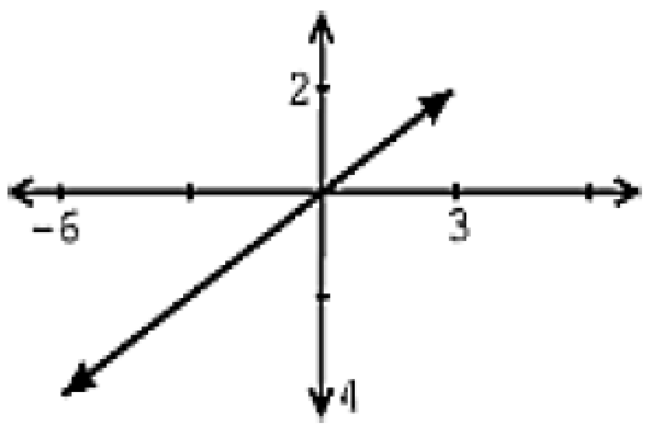
- ❑ Inner (dot) product: $a^T \cdot b = \sum_{i=1}^{n} a_i b_i$
- ❑ Length (Euclidian norm) of a vector : $\|a\| = \sqrt{a^T \cdot a} = \sqrt{\sum_{i=1}^{n} a_i^{\ 2}}$
- ❑ The angle between two n-dimensional vectors : $\cos\theta = \dfrac{a^T \cdot b}{\| a \| \| b \|}$
- ❑ An inner product is a measure of collinearity :

  - ○ a and b are orthogonal iff : $a^T \cdot b = 0$
  - ○ a and b are collinear iff : $a^T \cdot b = \| a \| \| b \|$

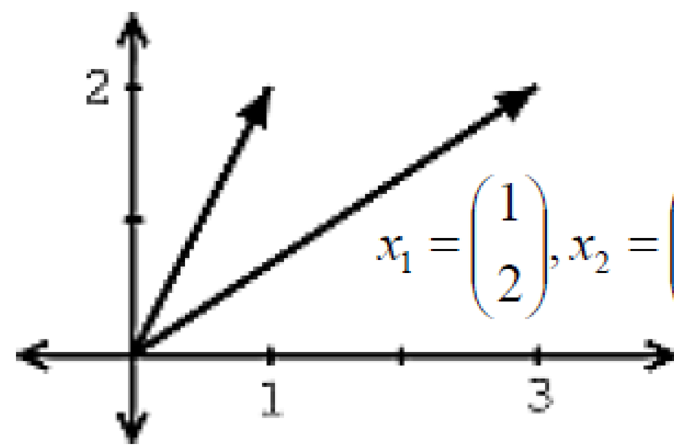
Linear Independence

- A set of n – dimensional vectors xi Є Rn , are said to be linearly independent if none of them can be written as a linear combination of the others.
- In other words , $c_1 x_1 + c_2 x_2 + \ldots + c_k x_k = 0$
  iff $c_1 = c_2 = \ldots = c_k = 0$
- Another approach to reveal a vectors independence is by graphing the vectors .

$$x_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, x_2 = \begin{pmatrix} -6 \\ 4 \end{pmatrix}$$



Not linearly independent vectors



$$x_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, x_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Linearly independent vectors

Span and Basis

❖ A span of a set of vectors $x_1, x_2, \ldots, x_k$ is the set of vectors that can be written as a linear combination of $x_1, x_2, \ldots, x_k$.

$$span(x_1, x_2, \ldots, x_k) = \{c_1 x_1 + c_2 x_2 + \ldots + c_k x_k \mid c_1, c_2, \ldots, c_k \in \Re\}$$

❖ A basis for $\mathbf{R}^n$ is a set of vectors which :

➢ Spans $\mathbf{R}^n$, i.e. any vector in this n –dimensional space can be written as linear combination of these basis vectors.

➢ Are linearly independent.

❑ For example vectors $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ form the basis of $\mathbf{R}^3$

They are also linearly independent, because if $a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0$, then

a = b = c = 0. hence the basis $\{e_1 , e_2 , e_3\}$ are called the standard basis of $\mathbf{R}^3$

Orthogonal/Orthonormal Basis

- An orthonormal basis of a vectors pace V with an inner product , is a set of basis vectors whose elements are mutually orthogonal and of magnitude 1 (unit vectors).
- Two vectors are orthogonal if they are perpendicular, i.e., they form a right angle, i.e. if their inner product is zero.

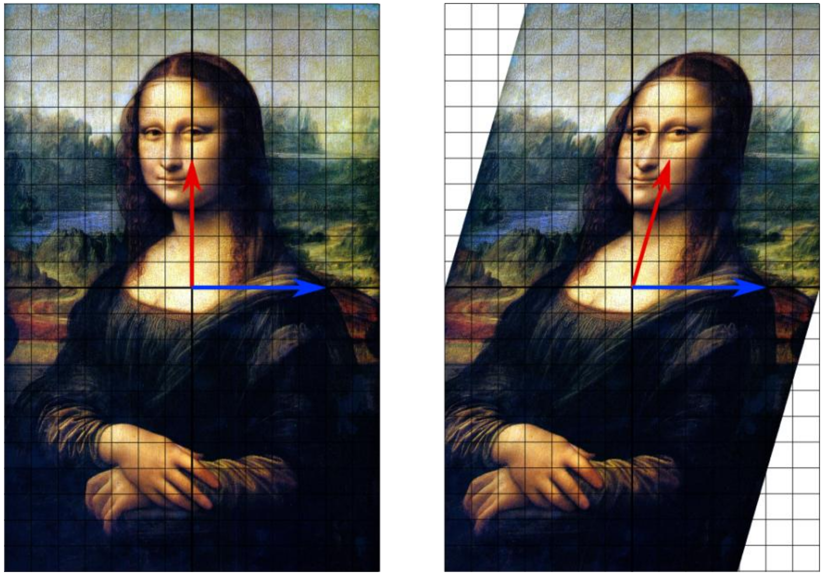
$$a^T \cdot b = \sum_{i=1}^{n} a_i b_i = 0 \quad \Rightarrow \quad a \perp b$$

- The standard basis of the 3–dimensional Euclidean space R^3 is an example of orthonormal (and ordered) basis.

Transformation Matrices

- Eigenvectors make understanding linear transformations easy. They are the "axes" (directions) along which a linear transformation acts simply by "stretching/compressing" and/or "flipping"; eigenvalues give you the factors by which this compression occurs.

- In linear algebra, an **eigenvector** or **characteristic vector** of a square matrix is a vector that points in a direction which is invariant under the associated linear transformation.



Eigenvalues and Eigenvectors

The eigenvalue problem is any problem having the following form:

$$\mathbf{A} \,.\, \mathbf{v} = \boldsymbol{\lambda} \,.\, \mathbf{v}$$

A: m x m matrix

v: m x 1 non-zero vector

λ:scalar

Any value of λ for which this equation has a solution is called the eigenvalue of A and the vector v which corresponds to this value is called the eigenvector of A.

Calculating Eigenvectors & Eigenvalues

Simple matrix algebra shows that:

$$A \, . \, v = \lambda \, . \, v$$

$$\Leftrightarrow \quad A \, . \, v - \lambda . \, \mathbf{I} . \, v = 0$$

$$\Leftrightarrow \quad (A - \lambda \, . \, \mathbf{I}) \, . \, v = 0$$

- Finding the roots of $| \mathbf{A} - \lambda . \, \mathbf{I} |$ will give the eigenvalues and for each of these eigenvalues there will be an eigenvector

EXAMPLE

$$A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$$

- To do this, we find the values of λ which satisfy the characteristic equation of the matrix A , namely those values of λ for which

$$\det(A-\lambda I) = 0$$

- Form the matrix $(A - \lambda I)$:

$$A - \lambda I = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} 1-\lambda & -3 & 3 \\ 3 & -5-\lambda & 3 \\ 6 & -6 & 4-\lambda \end{pmatrix}.$$

- Calculate $\det(A - \lambda I)$:

$$\begin{aligned} \det(A-\lambda I) &= (1-\lambda)\begin{vmatrix} -5-\lambda & 3 \\ -6 & 4-\lambda \end{vmatrix} - (-3)\begin{vmatrix} 3 & 3 \\ 6 & 4-\lambda \end{vmatrix} + 3\begin{vmatrix} 3 & -5-\lambda \\ 6 & -6 \end{vmatrix} \\ &= (1-\lambda)\left((-5-\lambda)(4-\lambda) - (3)(-6)\right) + 3(3(4-\lambda) - 3\times 6) + 3(3\times(-6) - (-5-\lambda)6) \\ &= (1-\lambda)(-20+5\lambda-4\lambda+\lambda^2+18) + 3(12-3\lambda-18) + 3(-18+30+6\lambda) \\ &= (1-\lambda)(-2+\lambda+\lambda^2) + 3(-6-3\lambda) + 3(12+6\lambda) \\ &= -2+\lambda+\lambda^2+2\lambda-\lambda^2-\lambda^3-18-9\lambda+36+18\lambda \\ &= 16+12\lambda-\lambda^3. \end{aligned}$$

- **REQUIRED**: To find solutions to $\det(A-\lambda I) = 0$ i.e., to solve:

$$\boldsymbol{\lambda^3 - 12\lambda - 16 = 0}$$

- By solving we find that the eigevalues of A are : 4, -2 ($\lambda = -2$ is a repeated root of the characteristic equation).

Finding Eigenvectors

- Once the eigenvalues of a matrix (A) have been found, we can find the eigenvectors by Gaussian Elimination

- **STEP 1** : For each eigenvalue λ, we have

$$(A - \lambda I)x = 0$$

Where x is the eigenvector associated with eigenvalue λ.

- **STEP 2** : We find x by Gaussian elimination, i.e, convert the augmented matrix $(A - \lambda I : 0)$ to row echelon form and solve the resulting linear system by back – substitution.

Case 1 : $\lambda = 4$

We must find vectors x which satisfy $(A - \lambda I)x = 0$

First form the matrix $(A - 4I)$

$$A - 4I = \begin{pmatrix} -3 & -3 & 3 \\ 3 & -9 & 3 \\ 6 & -6 & 0 \end{pmatrix}.$$

- Construct the augmented matrix (A – λI : 0) and convert it to row echelon form.

$$\left(\begin{array}{cccc} -3 & -3 & 3 & 0 \\ 3 & -9 & 3 & 0 \\ 6 & -6 & 0 & 0 \end{array}\right) \begin{array}{c} \mathrm{R1} \\ \mathrm{R2} \\ \mathrm{R3} \end{array} \xrightarrow{R1 \to -1/3 \times R3} \left(\begin{array}{cccc} 1 & 1 & -1 & 0 \\ 3 & -9 & 3 & 0 \\ 6 & -6 & 0 & 0 \end{array}\right) \begin{array}{c} \mathrm{R1} \\ \mathrm{R2} \\ \mathrm{R3} \end{array}$$

$$\xrightarrow{\substack{R2 \to R2 - 3 \times R1 \\ R3 \to R3 - 6 \times R1}} \left(\begin{array}{cccc} 1 & 1 & -1 & 0 \\ 0 & -12 & 6 & 0 \\ 0 & -12 & 6 & 0 \end{array}\right) \begin{array}{c} \mathrm{R1} \\ \mathrm{R2} \\ \mathrm{R3} \end{array}$$

$$\xrightarrow{R2 \to -1/12 \times R2} \left(\begin{array}{cccc} 1 & 1 & -1 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & -12 & 6 & 0 \end{array}\right) \begin{array}{c} \mathrm{R1} \\ \mathrm{R2} \\ \mathrm{R3} \end{array}$$

$$\xrightarrow{R3 \to R3 + 12 \times R2} \left(\begin{array}{cccc} 1 & 1 & -1 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{c} \mathrm{R1} \\ \mathrm{R2} \\ \mathrm{R3} \end{array}$$

$$\xrightarrow{R1 \to R1 - R2} \left(\begin{array}{cccc} 1 & 0 & -1/2 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{c} \mathrm{R1} \\ \mathrm{R2} \\ \mathrm{R3} \end{array}$$

- Rewriting the augmented matrix as a linear system gives :

$$\begin{aligned} x_1 - 1/2x_3 &= 0 \\ x_2 - 1/2x_3 &= 0 \end{aligned}$$

So the eigenvector x is given by :

$$\mathbf{x} = \begin{pmatrix} x_1 = \frac{x_3}{2} \\ x_2 = \frac{x_3}{2} \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

For any real number $x_3 \neq 0$, those are the eigenvectors associated with the eigenvalue $\lambda = 4$.

Case 2 : $\lambda = -2$

We must find vectors x which satisfy $(A - \lambda I)x = 0$

First form the matrix $A - (-2)I = A + 2I$

$$A + 2I = \begin{pmatrix} 3 & -3 & 3 \\ 3 & -3 & 3 \\ 6 & -6 & 6 \end{pmatrix}$$

Construct the augmented matrix $(A - \lambda I : 0)$ and convert it to row echelon form.

$$\begin{array}{cc} \begin{pmatrix} 3 & -3 & 3 & 0 \\ 3 & -3 & 3 & 0 \\ 6 & -6 & 6 & 0 \end{pmatrix} & \begin{array}{l} \mathrm{R1} \\ \mathrm{R2} \\ \mathrm{R3} \end{array} \end{array} \xrightarrow{R1 \to 1/3 \times R3} \begin{array}{cc} \begin{pmatrix} 1 & -1 & 1 & 0 \\ 3 & -3 & 3 & 0 \\ 6 & -6 & 6 & 0 \end{pmatrix} & \begin{array}{l} \mathrm{R1} \\ \mathrm{R2} \\ \mathrm{R3} \end{array} \end{array}$$

$$\xrightarrow[R3 \to R3 - 6 \times R1]{R2 \to R2 - 3 \times R1} \begin{array}{cc} \begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} & \begin{array}{l} \mathrm{R1} \\ \mathrm{R2} \\ \mathrm{R3} \end{array} \end{array}$$

- Rewriting the augmented matrix as a linear system gives :

$$x_1 + x_2 - x_3 = 0,$$

- So the eigenvector x associated with λ = -2 is given by :

$$\mathbf{x} = \begin{pmatrix} x_1 = x_3 - x_2 \\ x_2 \\ x_3 \end{pmatrix}$$

Thus, $\mathbf{x} = \begin{pmatrix} x_3 - x_2 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ for any x3, x2 Є **R**\{0}

Are the eigenvectors associated with λ = -2

Coming back to Principal Component Analysis

Change of Basis

- Let **X** and **Y** be *m* x *n* matrices related by a linear transformation *P.*
- **X** is the original recorded data set and **Y** is a re-representation of that data set

$$\mathbf{PX = Y}$$

p_i are the rows of P.

x_i are the columns of X.

y_i are the columns of Y.

➢ What does this mean?

- P is a matrix that transforms X into Y.
- The rows of P, $\{p_1, p_2, \ldots, p_m\}$ are a set of new basis vectors for expressing the columns of X.

Lets write out the explicit dot products of **PX.**

$$\mathbf{PX} = \begin{bmatrix} \mathbf{p_1} \\ \vdots \\ \mathbf{p_m} \end{bmatrix} \begin{bmatrix} \mathbf{x_1} & \cdots & \mathbf{x_n} \end{bmatrix}$$

We can note the form of each column of **Y**.

$$\mathbf{Y} = \begin{bmatrix} \mathbf{p_1 \cdot x_1} & \cdots & \mathbf{p_1 \cdot x_n} \\ \vdots & \ddots & \vdots \\ \mathbf{p_m \cdot x_1} & \cdots & \mathbf{p_m \cdot x_n} \end{bmatrix}$$

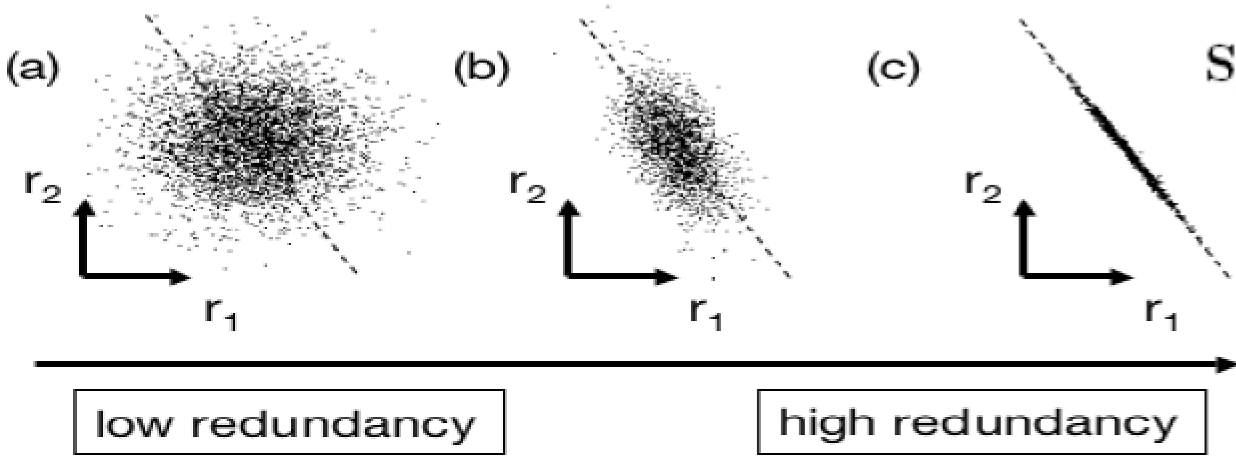
$$y_i = \begin{bmatrix} \mathbf{p_1 \cdot x_i} \\ \vdots \\ \mathbf{p_m \cdot x_i} \end{bmatrix}$$

We can see that each coefficient of $\mathbf{y_i}$ is a dot – product of $\mathbf{x_i}$ with the corresponding row in **P**.

1. In other words , the jth coefficient of $\mathbf{y_i}$ is a projection on to the jth row of **P**. This is in fact the very form of an equation where $\mathbf{y_i}$ is a projection on to the basis of $\{\mathbf{p_1}, \mathbf{p_2}, \ldots.\ \mathbf{p_m}\}$.
2. Therefore, the rows of **P** are a new set of basis vectors for representing the columns of **X**.

- Later we will see that the row vectors $\{\mathbf{p_1}, \mathbf{p_2}, \ldots.\ \mathbf{p_m}\}$ in this transformation will become the principal components of X

Redundancy



$$S_X \equiv \frac{1}{n-1} X X^T$$

- Panel(a) depicts two recordings with no redundancy, i.e. they are un - correlated, e.g. person's height and his GPA.
- However, in panel(c) both recordings appear to be strongly related, i.e. one can be expressed in terms of the other.

Computing the covariance matrix quantifies the correlations between all possible pairs of measurements. Between one pair of measurements, a large covariance corresponds to a situation like panel(c), while zero covariance corresponds to entirely uncorrelated data as in panel(a).

Why Diagonalize Covariance Matrix?

- Our goals are to find an optimized covariance matrix $\mathbf{S_y}$ so that it:

1. Minimizes redundancy, measured by covariance. (off-diagonal), i.e. we would like each variable to co-vary as little as possible with other variables.
2. Maximizes the variance. (the diagonal)

- Evidently, in an optimized matrix all of diagonal elements in $\mathbf{S_y}$ are zero. Therefore, removing redundancy diagonalizes $\mathbf{S_y}$

PCA Assumptions

1. PCA assumes that all basis vectors $\{\mathbf{p_1}, \mathbf{p_2}, \ldots. \mathbf{p_m}\}$ are orthonormal. In the language of Linear Algebra, PCA assumes P is an orthonormal matrix.

2. Secondly, PCA assumes the directions with the largest variances are the most "important" or in other words, most principal.

Solving PCA

- **<u>Objective</u>** : Find some orthonormal matrix **P** where **Y** = **PX** such that $\mathbf{S_Y} = 1/(n-1)\mathbf{Y}\,\mathbf{Y}^{\mathbf{T}}$ is diagonalized. The rows of **P** are the principal components of **X**.

- where **X** is a *m* x *n* matrix
- n is the number of variables
- m is the number of observations.

We begin by writing $\mathbf{S_Y}$ in the terms of our variable of choice **P**

$$
\begin{aligned}
\mathbf{S_Y} &= \frac{1}{n-1}\mathbf{Y}\mathbf{Y}^T \\
&= \frac{1}{n-1}(\mathbf{PX})(\mathbf{PX})^T \\
&= \frac{1}{n-1}\mathbf{P}\mathbf{X}\mathbf{X}^T\mathbf{P}^T \\
&= \frac{1}{n-1}\mathbf{P}(\mathbf{X}\mathbf{X}^T)\mathbf{P}^T \\
\mathbf{S_Y} &= \frac{1}{n-1}\mathbf{P}\mathbf{A}\mathbf{P}^T
\end{aligned}
$$

- **THEOREM** : A symmetric matrix is diagonalized by a matrix of its orthonormal eigenvectors.

Hence $\mathbf{A = EDE^T}$

- **D** is a diagonal matrix
- **E** is a matrix of eigenvectors of **A** arranged as column

- Now comes the trick. We select the matrix **P** to be a matrix where each row $\mathbf{p_i}$ is an eigenvector of $\mathbf{XX^T}$.
- By this selection, $\mathbf{P = E^T}$. Hence $\mathbf{A = P^TDP}$.

- **THEOREM** : The inverse of an orthogonal matrix is its transpose.

Hence , $\mathbf{P^T = P^{-1}}$

$$\begin{aligned} \mathbf{S_Y} &= \frac{1}{n-1}\mathbf{PAP}^T \\ &= \frac{1}{n-1}\mathbf{P}(\mathbf{P}^T\mathbf{DP})\mathbf{P}^T \\ &= \frac{1}{n-1}(\mathbf{PP}^T)\mathbf{D}(\mathbf{PP}^T) \\ &= \frac{1}{n-1}(\mathbf{PP}^{-1})\mathbf{D}(\mathbf{PP}^{-1}) \\ \mathbf{S_Y} &= \frac{1}{n-1}\mathbf{D} \end{aligned}$$

- It is evident that the choice of **P** diagonalizes $\mathbf{S_Y}$. This was the goal for PCA.

Computing sample covariance matrix

Original data: 100 (x, y) points in a 2 x100 matrix M_0:

$$M_0 = \begin{bmatrix} x_1 & \cdots & x_{100} \\ y_1 & \cdots & y_{100} \end{bmatrix} = \begin{bmatrix} 3.0858 & 0.8806 & 9.8850 & \cdots & 4.4106 \\ 12.8562 & 10.7804 & 8.7504 & \cdots & 13.5627 \end{bmatrix}$$

Centered data: subtract $\bar{x}$ from x's and $\bar{y}$ from y's to get M; here $\bar{x} = 5$, $\bar{y} = 10$:

$$M = \begin{bmatrix} -1.9142 & -4.1194 & 4.8850 & \cdots & -0.5894 \\ 2.8562 & 0.7804 & -1.2496 & \cdots & 3.5627 \end{bmatrix}$$

Sample covariance:

$$C = \frac{MM'}{100-1} = \begin{bmatrix} 31.9702 & -16.5683 \\ -16.5683 & 13.0018 \end{bmatrix}$$

$$= \begin{bmatrix} s_{XX} & s_{XY} \\ s_{YX} & s_{YY} \end{bmatrix} = \begin{bmatrix} s_X^{\,2} & s_{XY} \\ s_{XY} & s_Y^{\,2} \end{bmatrix}$$

Diagonalizing the sample covariance matrix C

- C is a real-valued symmetric matrix, so it can be diagonalized C = VDV' where V'= V transpose. D includes the eigenvalues as the diagonal elements.

$$\begin{array}{ccccccc} C & = & V & & D & & V' \\ \begin{bmatrix} 31.9702 & -16.5683 \\ -16.5683 & 13.0018 \end{bmatrix} & = & \begin{bmatrix} -0.8651 & -0.5016 \\ 0.5016 & -0.8651 \end{bmatrix} & & \begin{bmatrix} 41.5768 & 0 \\ 0 & 3.3952 \end{bmatrix} & & \begin{bmatrix} -0.8651 & 0.5016 \\ -0.5016 & -0.8651 \end{bmatrix} \end{array}$$

- It is conventional to put the eigenvalues into D in decreasing order

$\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$.

Principal axes

- The columns of V are the right eigenvectors of C.

Eigenvalue	Eigenvector
41.5768	$\begin{bmatrix} -0.8651 \\ 0.5016 \end{bmatrix}$
3.3952	$\begin{bmatrix} -0.5016 \\ -0.8651 \end{bmatrix}$

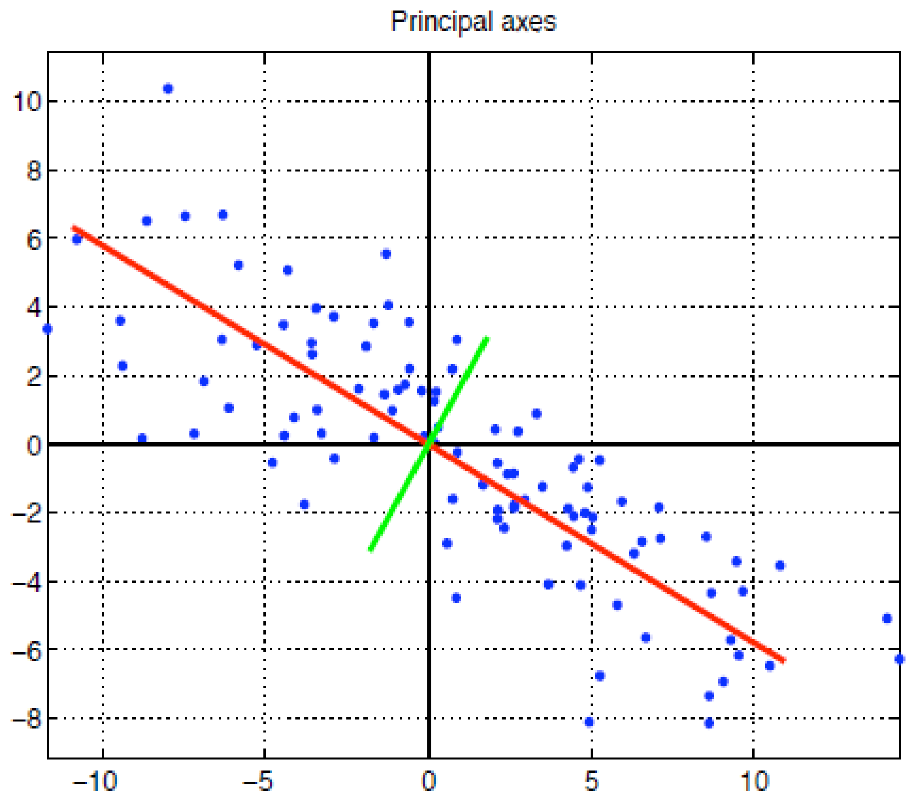
- Put them into the columns of a matrix:

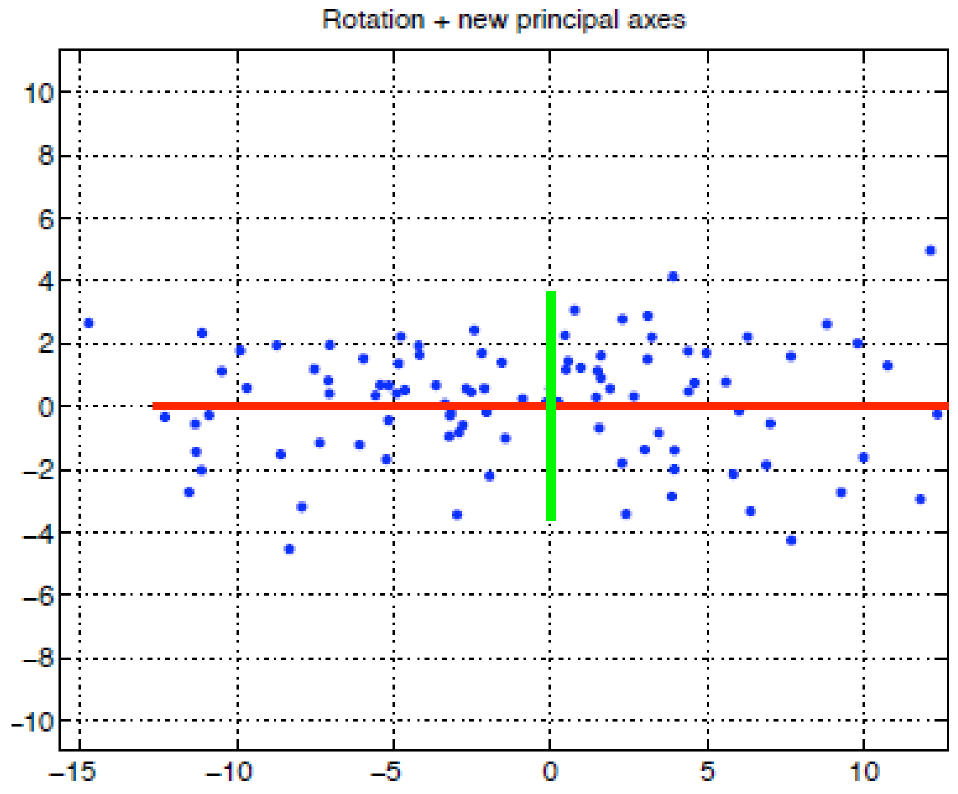
PC1	PC2
-0.8651	-0.5016
0.5016	-0.8651

- The axis of the principal components will be the span of eigenvectors of the covariance matrix .

Principal axes

- Plot the centered data with lines along the principal axes:
- Sum of squared perpendicular distances of data points to first PC line (red) is minimum among all lines through origin.





New coordinates

- The rotated data has new coordinates $(t_1, u_1), \ldots, (t_{100}, u_{100})$ and covariance matrix D:

$$\begin{bmatrix} \mathrm{Var}(T) & \mathrm{Cov}(T,U) \\ \mathrm{Cov}(T,U) & \mathrm{Var}(U) \end{bmatrix} = \begin{bmatrix} 41.5768 & 0 \\ 0 & 3.3952 \end{bmatrix}$$

- The total variance is $\lambda 1 + \lambda 2 + \ldots = \mathrm{Tr}(D) = \mathrm{Tr}(C)$.
- Here, the total variance is Var(T) + Var(U) = 44.9720.
- The part of the variance explained by each axis is λi/total variance:

Eigenvector	Eigenvalue	Explained
$\begin{bmatrix} -0.8651 \\ 0.5016 \end{bmatrix}$	41.5768	$41.5768/44.9720 = 92.45\%$
$\begin{bmatrix} -0.5016 \\ -0.8651 \end{bmatrix}$	3.3952	$3.3952/44.9720 = 7.55\%$
Total	44.9720	100%

New Data

- **Final Data = RowFeatureVector x RowZeroMeanData**

- RowFeatureVector is the matrix with the eigenvectors in the columns transposed so that the eigenvectors are now in the rows, with the most significant eigenvector at the top.

- RowZeroMeanData is the mean –adjusted data transposed, i.e., the data items are in each column, with each row holding a separate dimension.

This gives us the original data solely in terms of the vectors we chose.