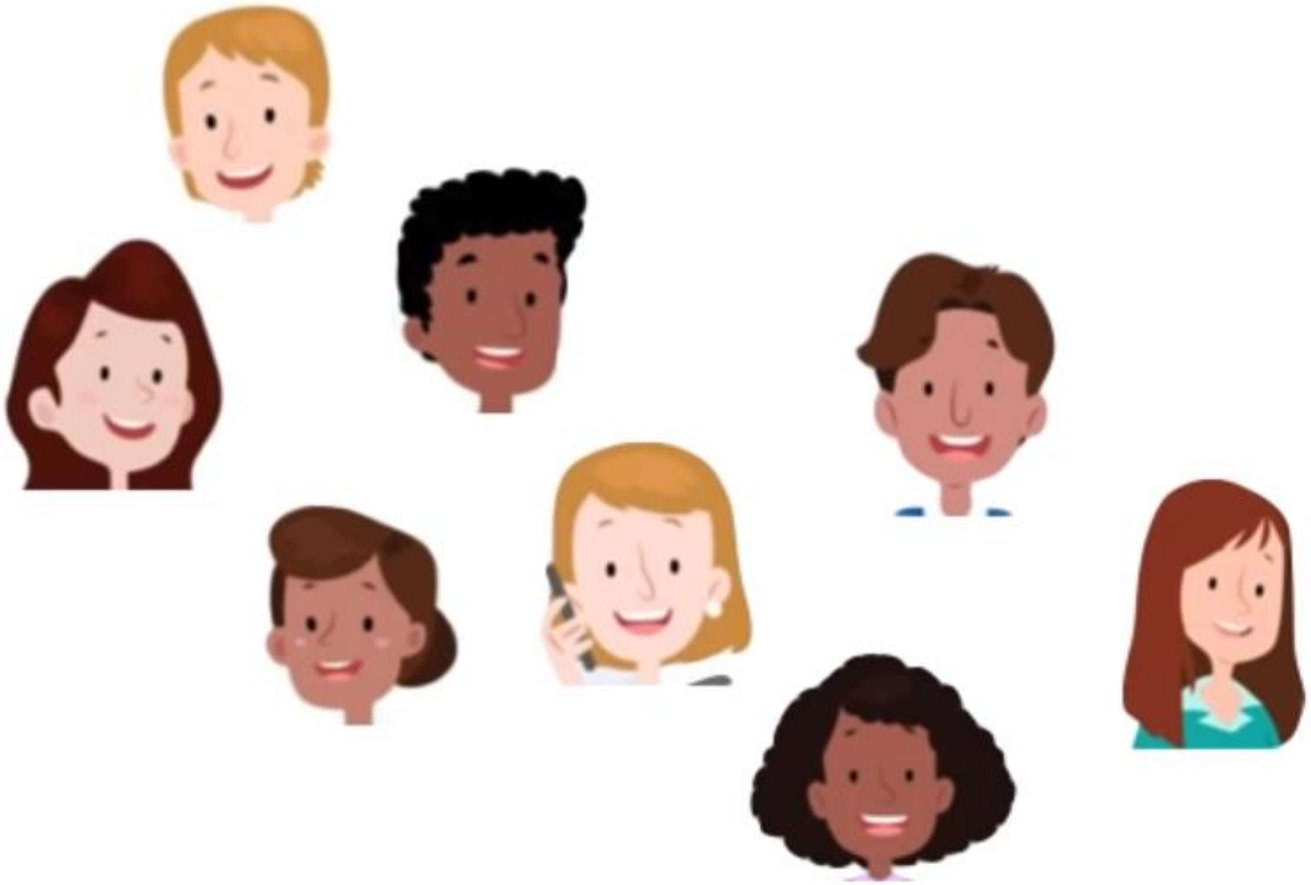
Principal Component Analysis

Taking a picture

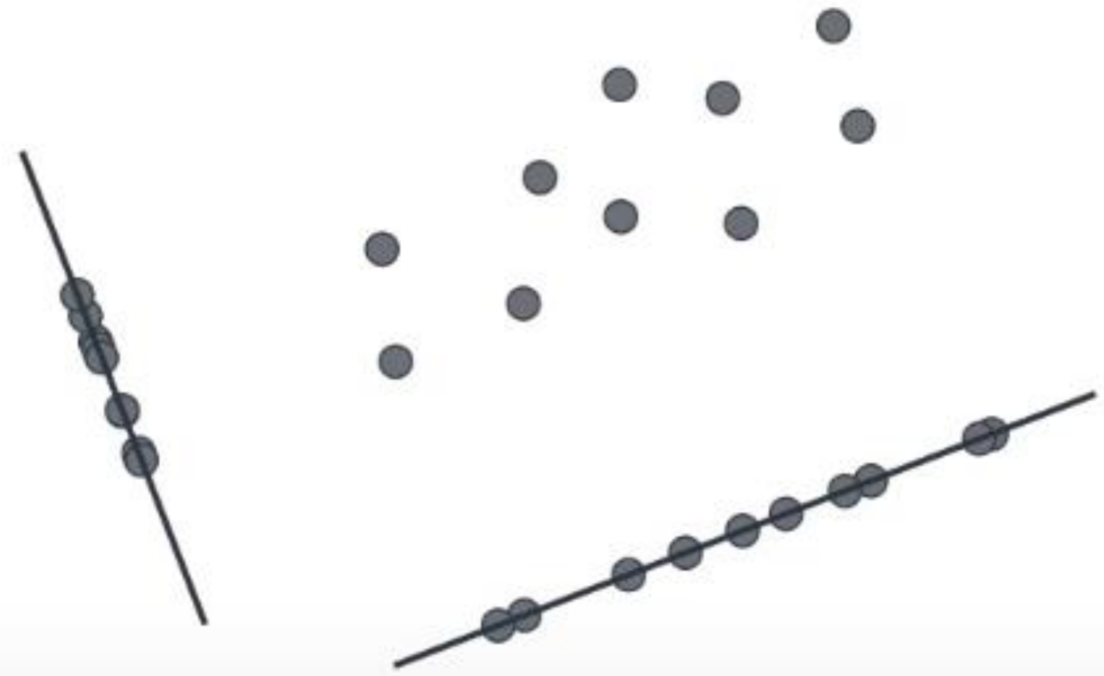


Taking a picture

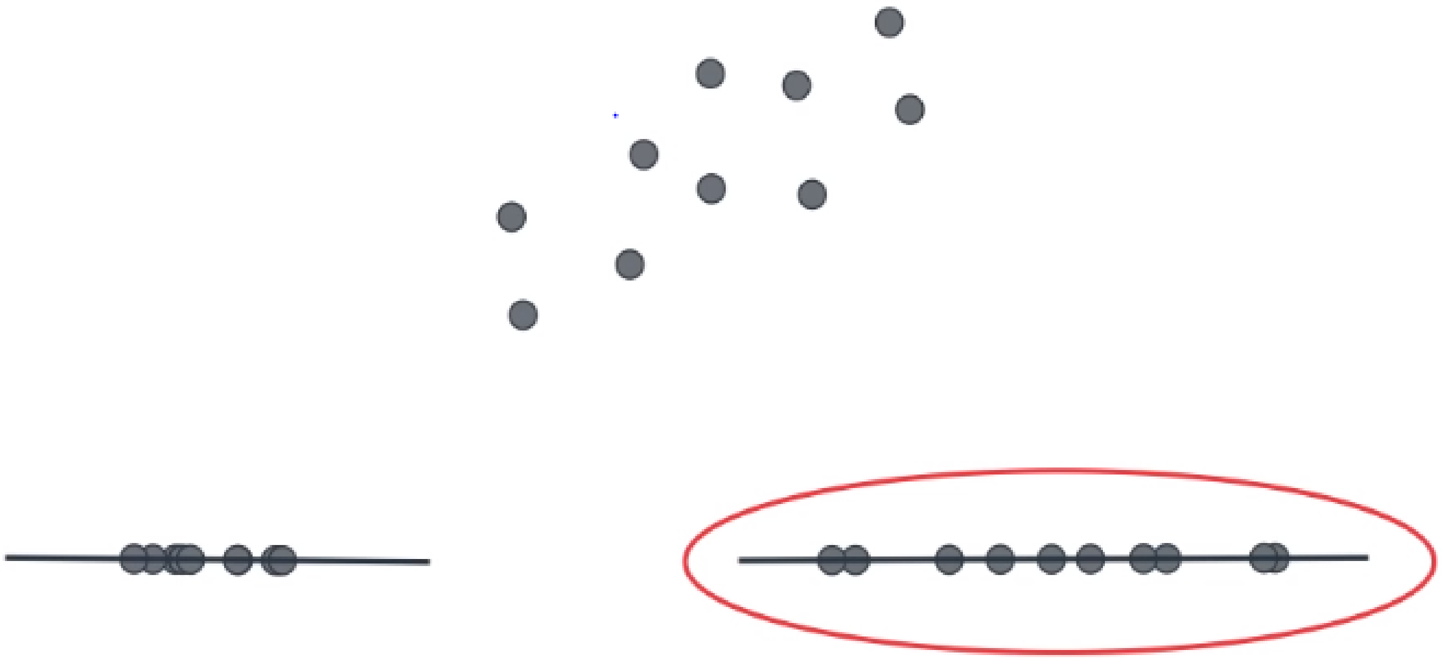
Dimensionality reduction:

Taking a picture of data and try to keep as much information as possible

Dimensionality Reduction



Dimensionality Reduction



Let’s learn how to project the data into a line that can keep the points as far as possible.

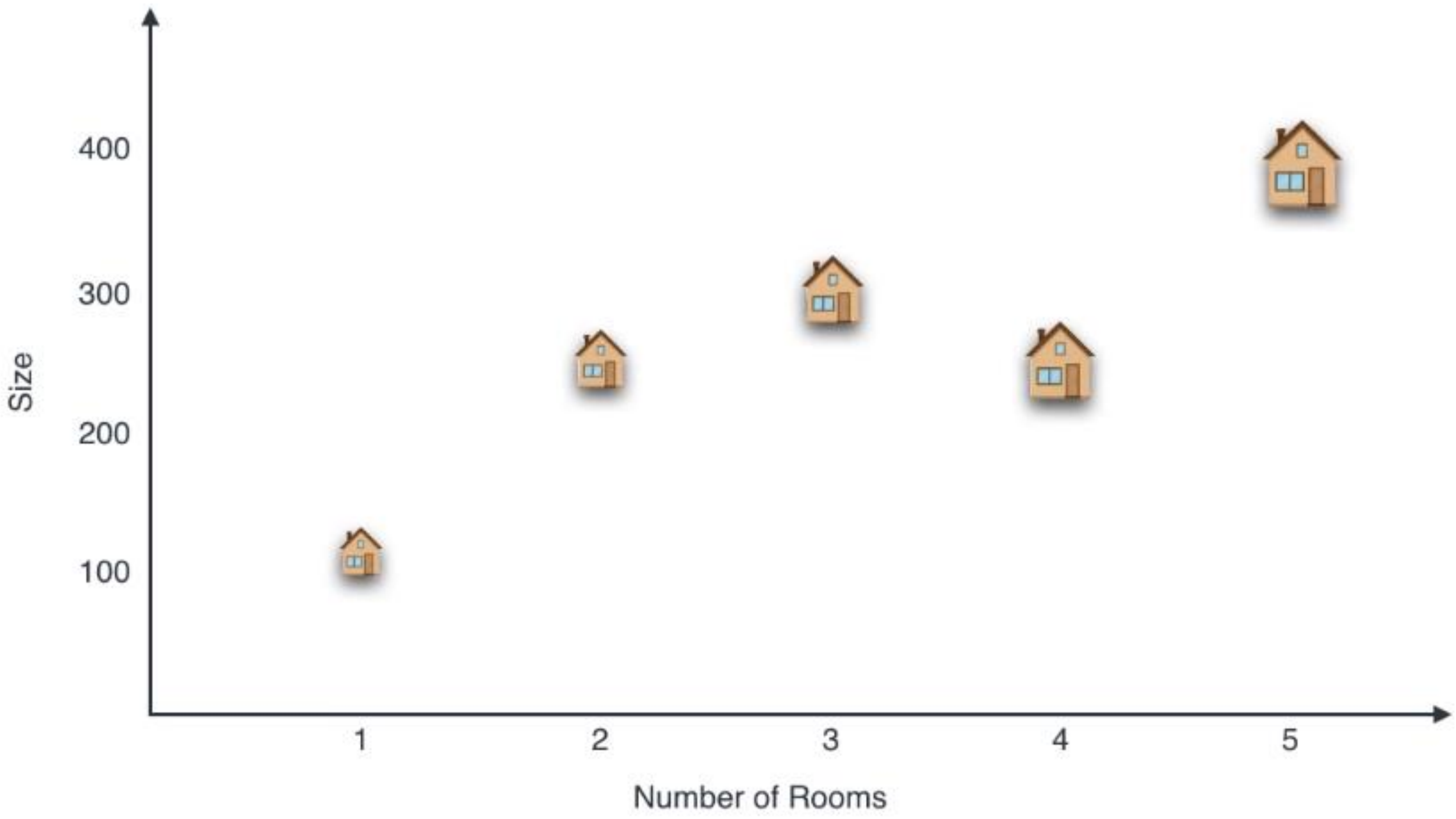
Housing Data

Size
Number of rooms
Number of bathrooms
Schools around
Crime rate

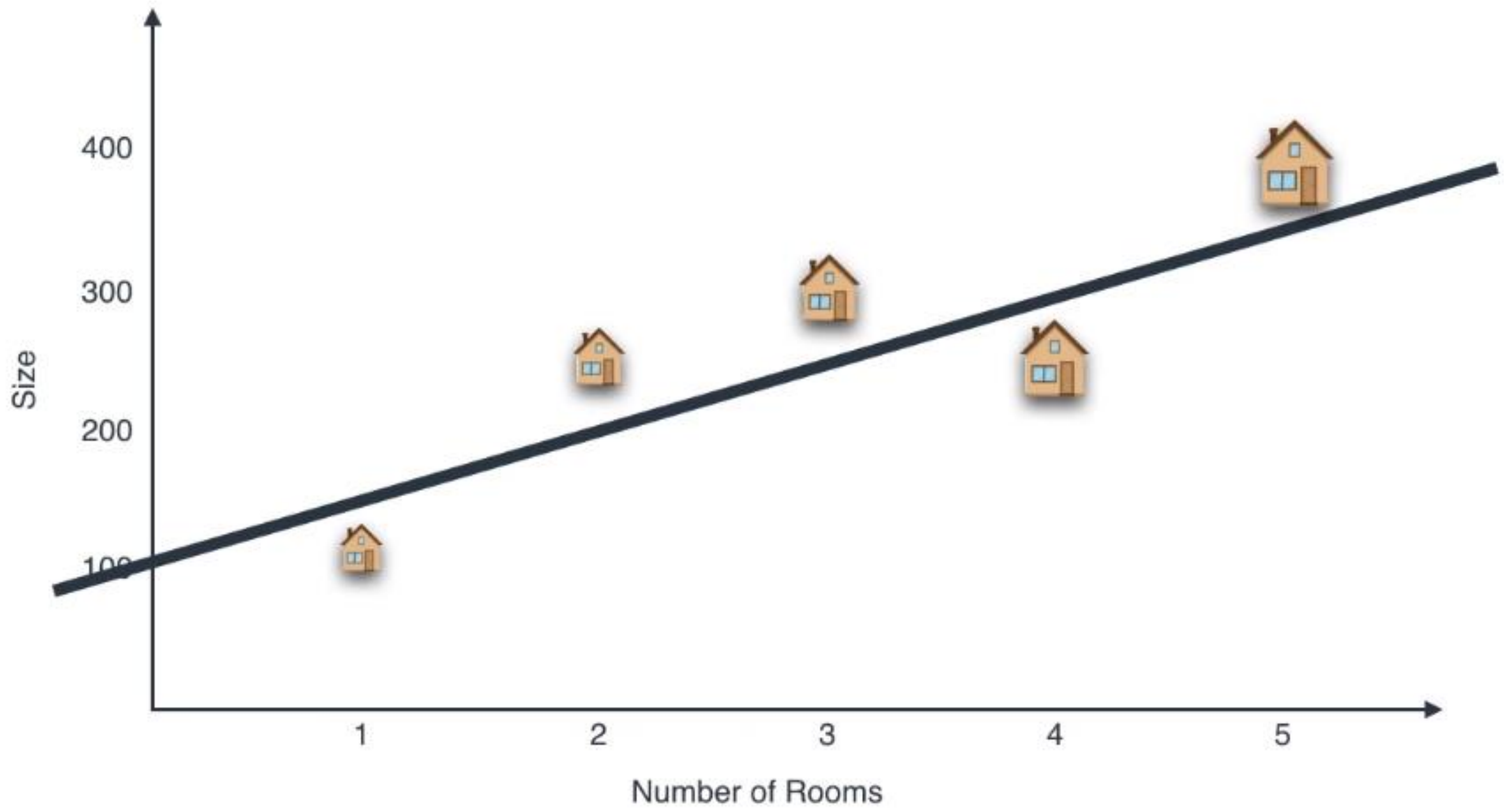
Housing Data

Size
Number of rooms → Size feature
Number of bathrooms

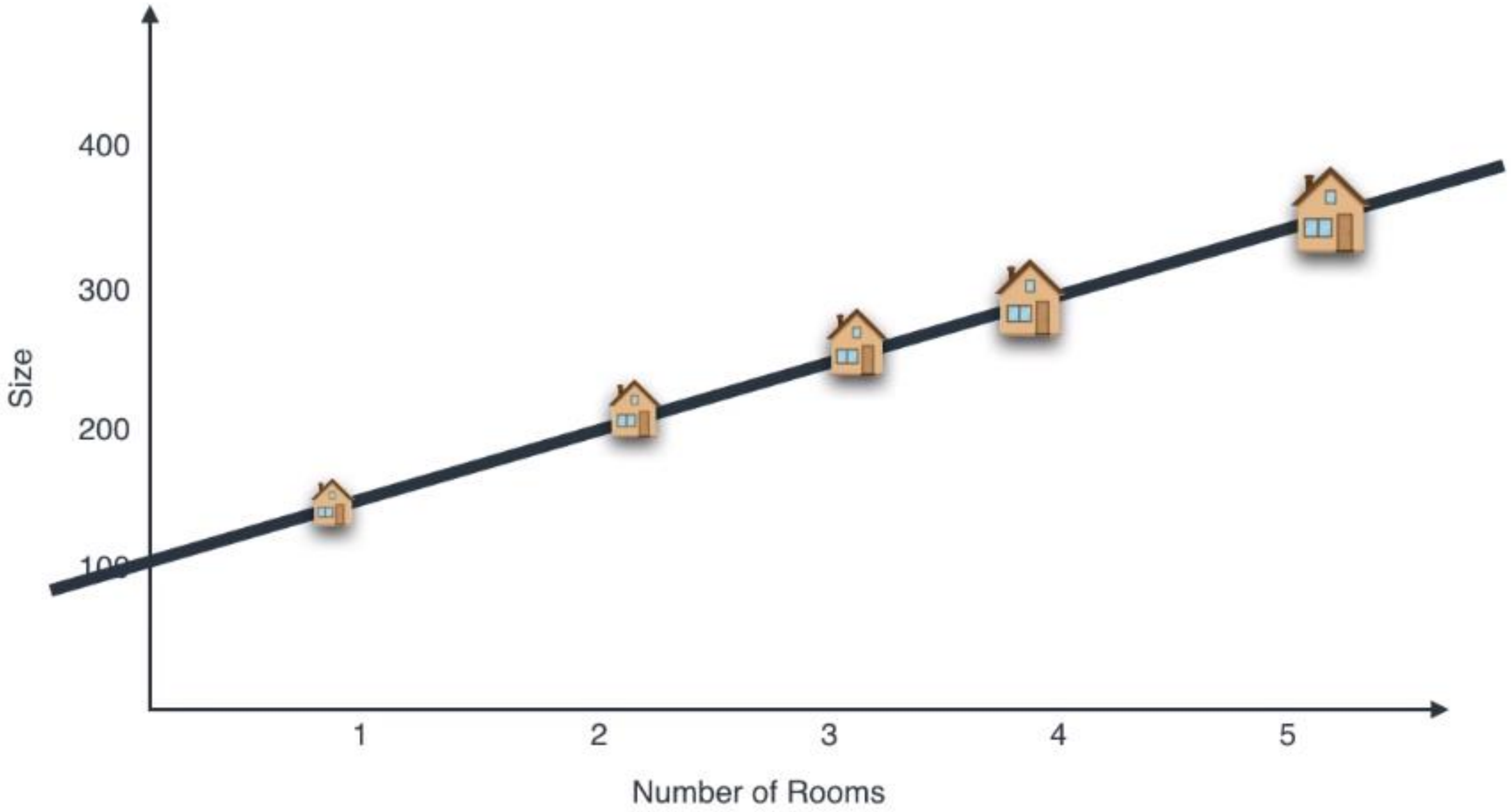
Schools around → Location feature
Crime rate



400
300
200
100
Size
1
2
3
4
5
Number of Rooms



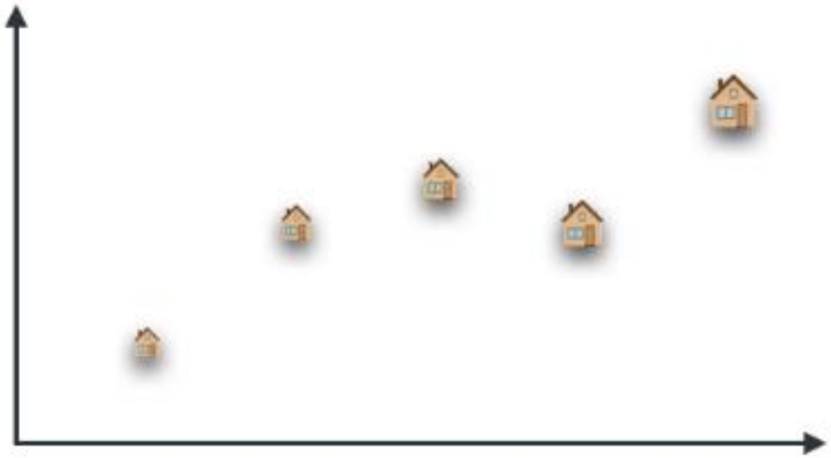
400
300
200
100
Size
1
2
3
4
5
Number of Rooms


400
300
200
Size
1
2
3
4
5
Number of Rooms

Size feature

2 dimensions

size
number of rooms



1 dimension

size feature

Housing Data

5 dimensions

Size
Number of rooms
Number of bathrooms
Schools around
Crime rate

2 dimensions

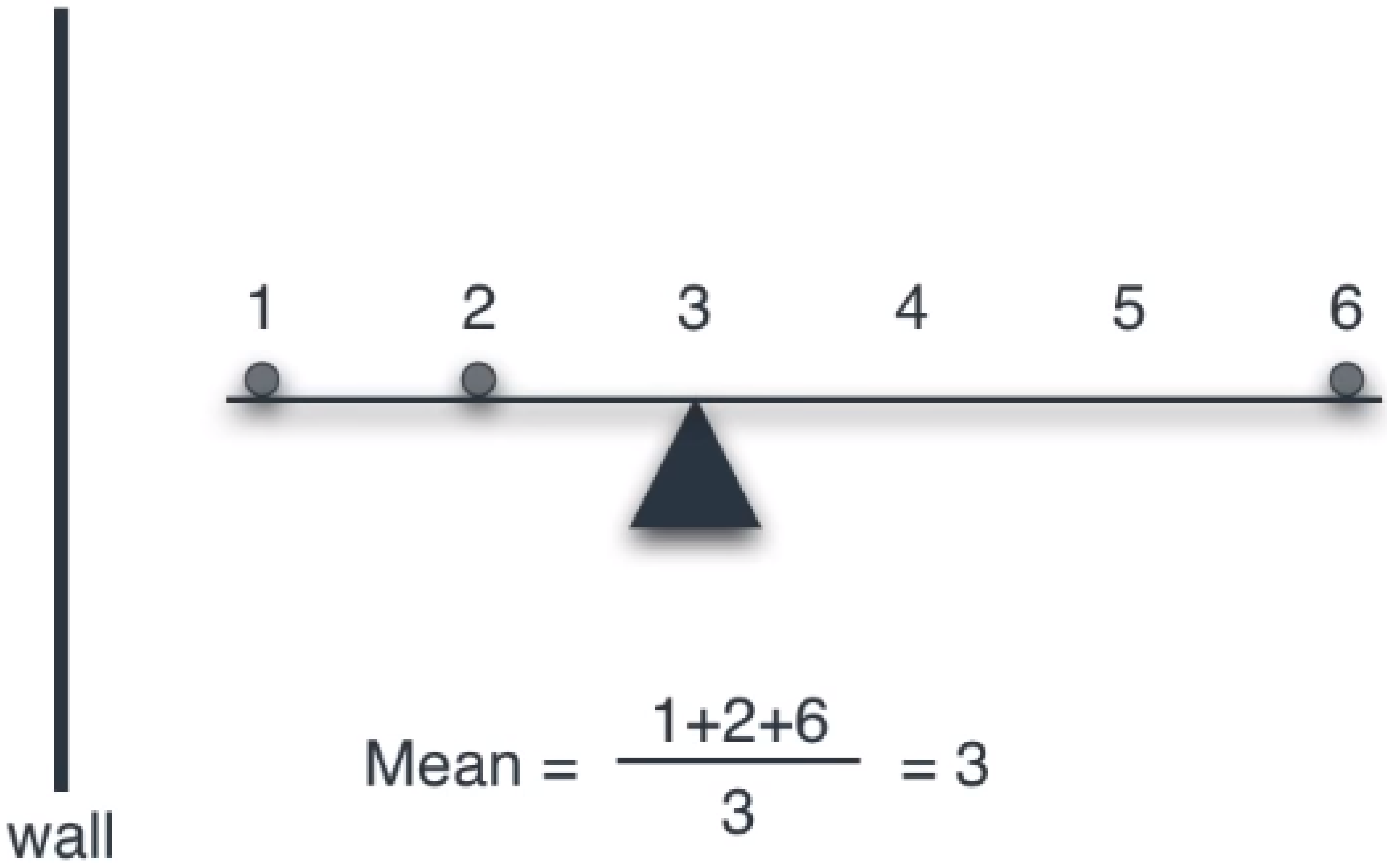
Size feature

Location feature

How do we balance the 3 weights (of exactly same weight), assuming the bar has no weight

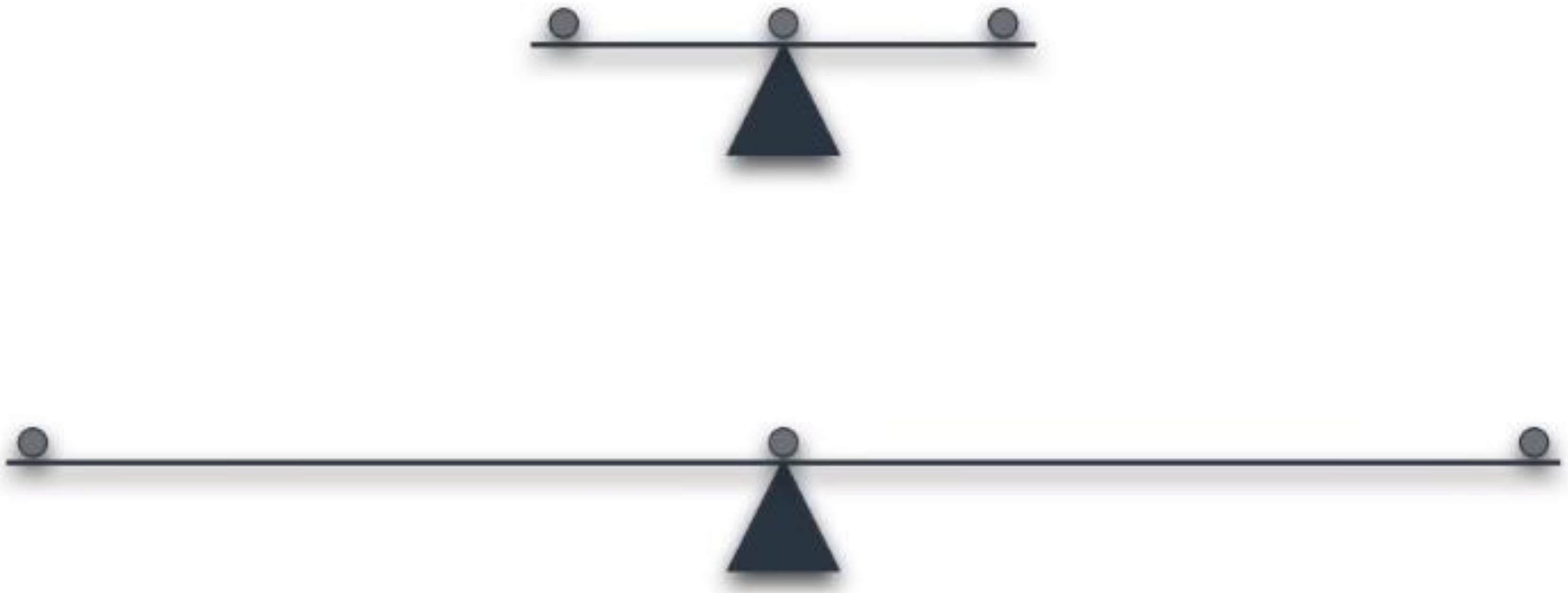
Mean

Mean



$$\text{Mean} = \frac{1+2+6}{3} = 3$$

Variance



How do we measure distance when we have same mean?

Variance

Variance =

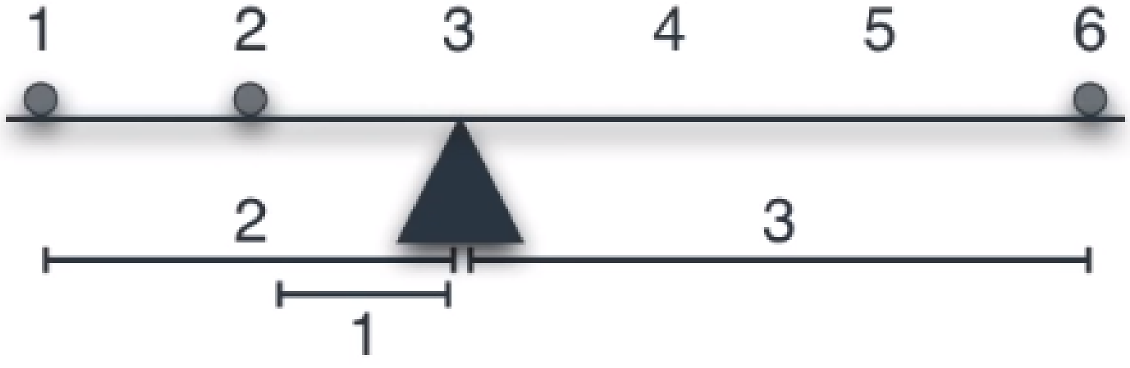
Variance is basically a measure that tells how spread out is a set

Variance

$$\text{Variance} = \frac{1^2+0^2+1^2}{3} = 2/3$$

$$\text{Variance} = \frac{5^2+0^2+5^2}{3} = 50/3$$

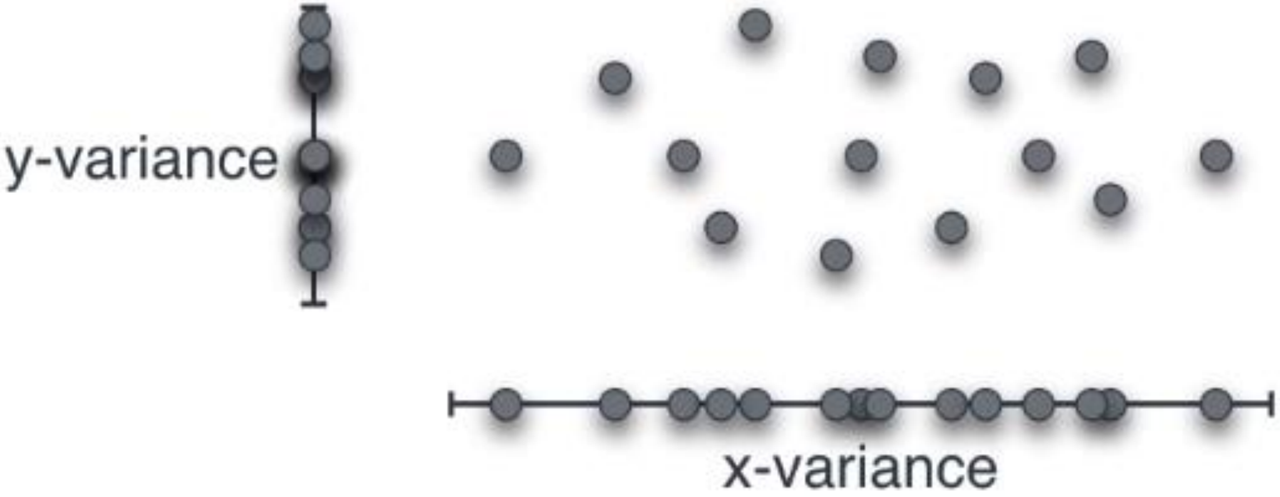
Mean



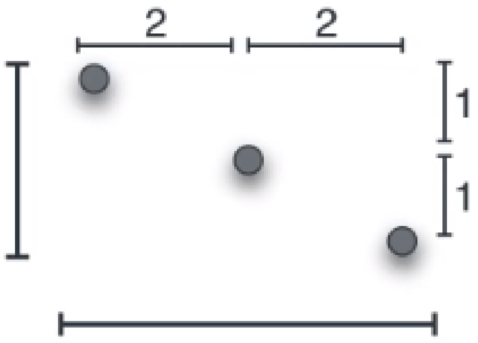
$$\text{Variance} = \frac{2^2+1^2+3^2}{3} = 14/3$$

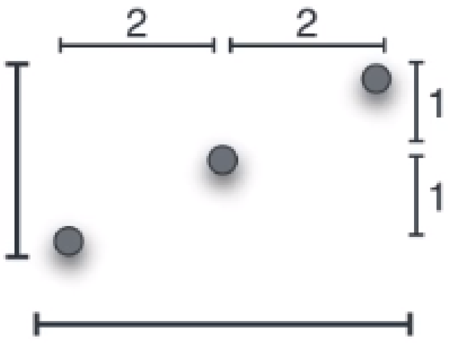
What is the variance for a 2D dataset? i.e. a set is in the plane

Variance?



Variance?



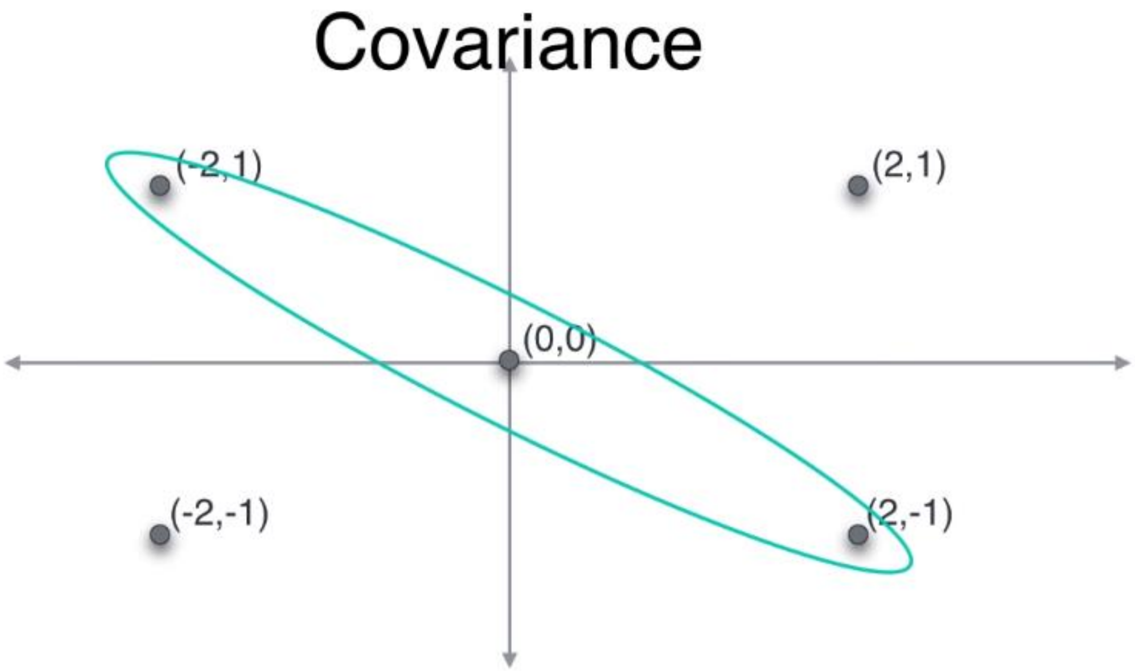


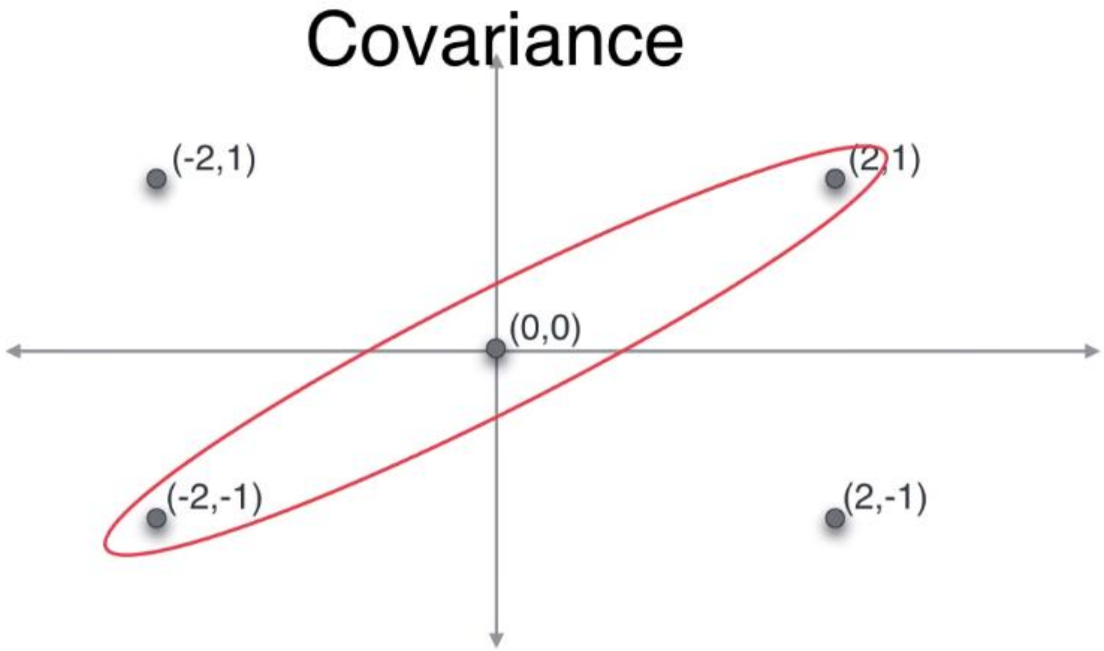
$$\text{x-variance} = \frac{2^2+0^2+2^2}{3} = 8/3$$

$$\text{y-variance} = \frac{1^2+0^2+1^2}{3} = 2/3$$

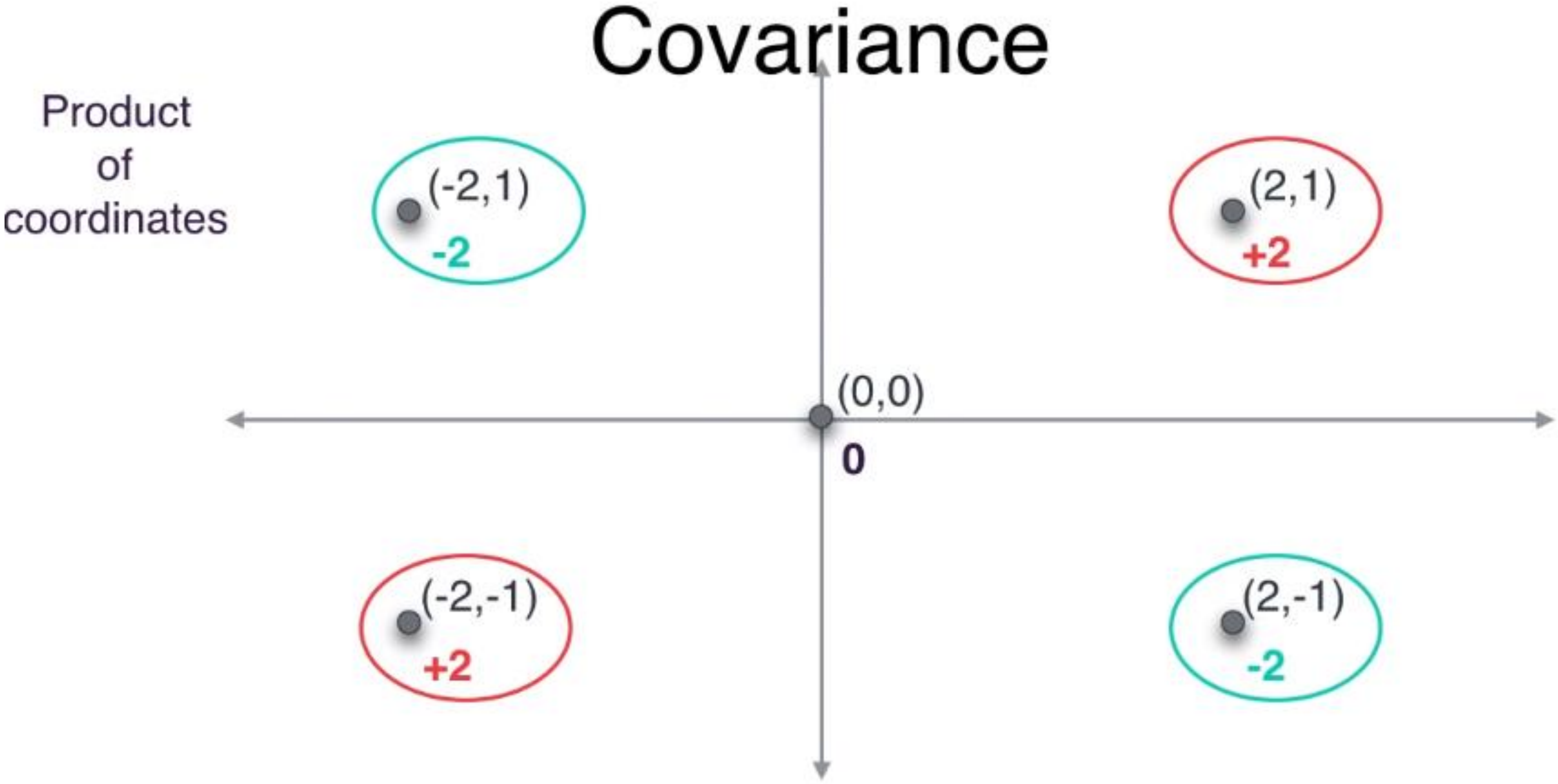
How to differentiate 2 fundamentally different datasets that have same x variance and y variance?

Covariance tells how 3 points (in the right image) are different from 3 points (in the left image)

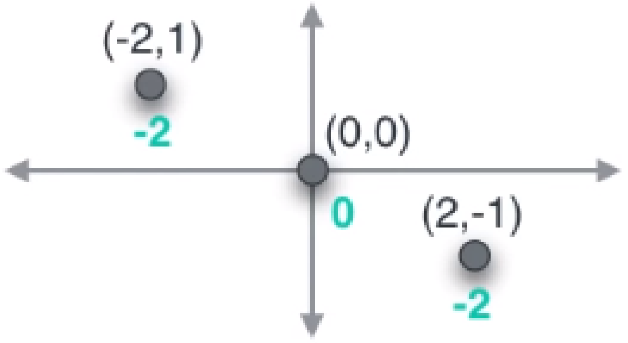


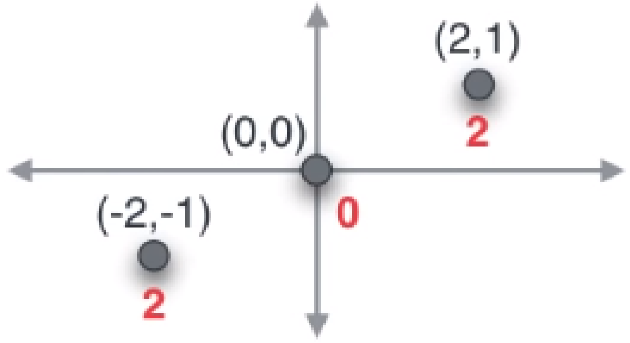


Covariance calculates sum of products of coordinates (Earlier it was a sum of square of one of the coordinates)



Covariance

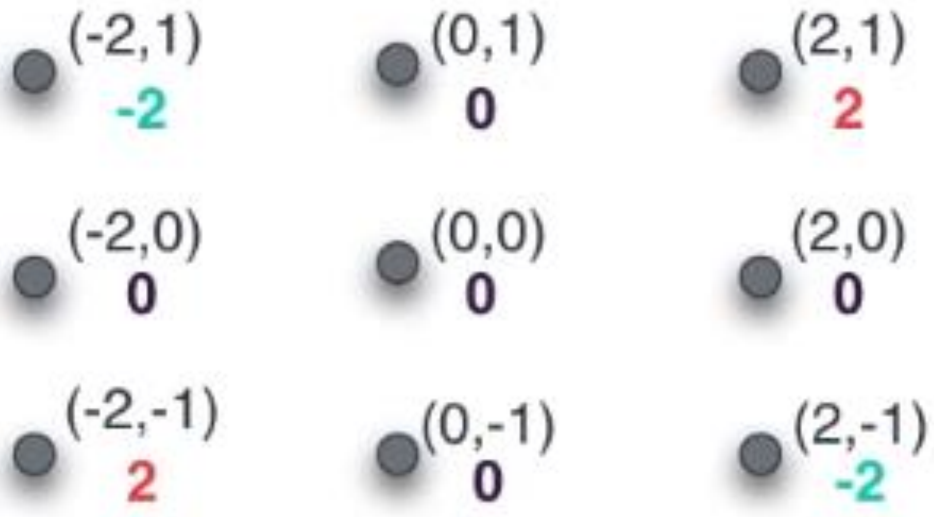




$$\text{covariance} = \frac{(-2) + 0 + (-2)}{3} = -4/3$$

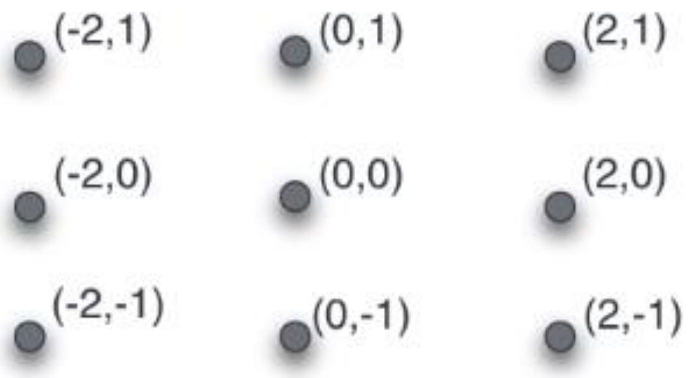
$$\text{covariance} = \frac{2 + 0 + 2}{3} = 4/3$$

Covariance



This set looks neither like positively correlated nor negatively correlated

Covariance

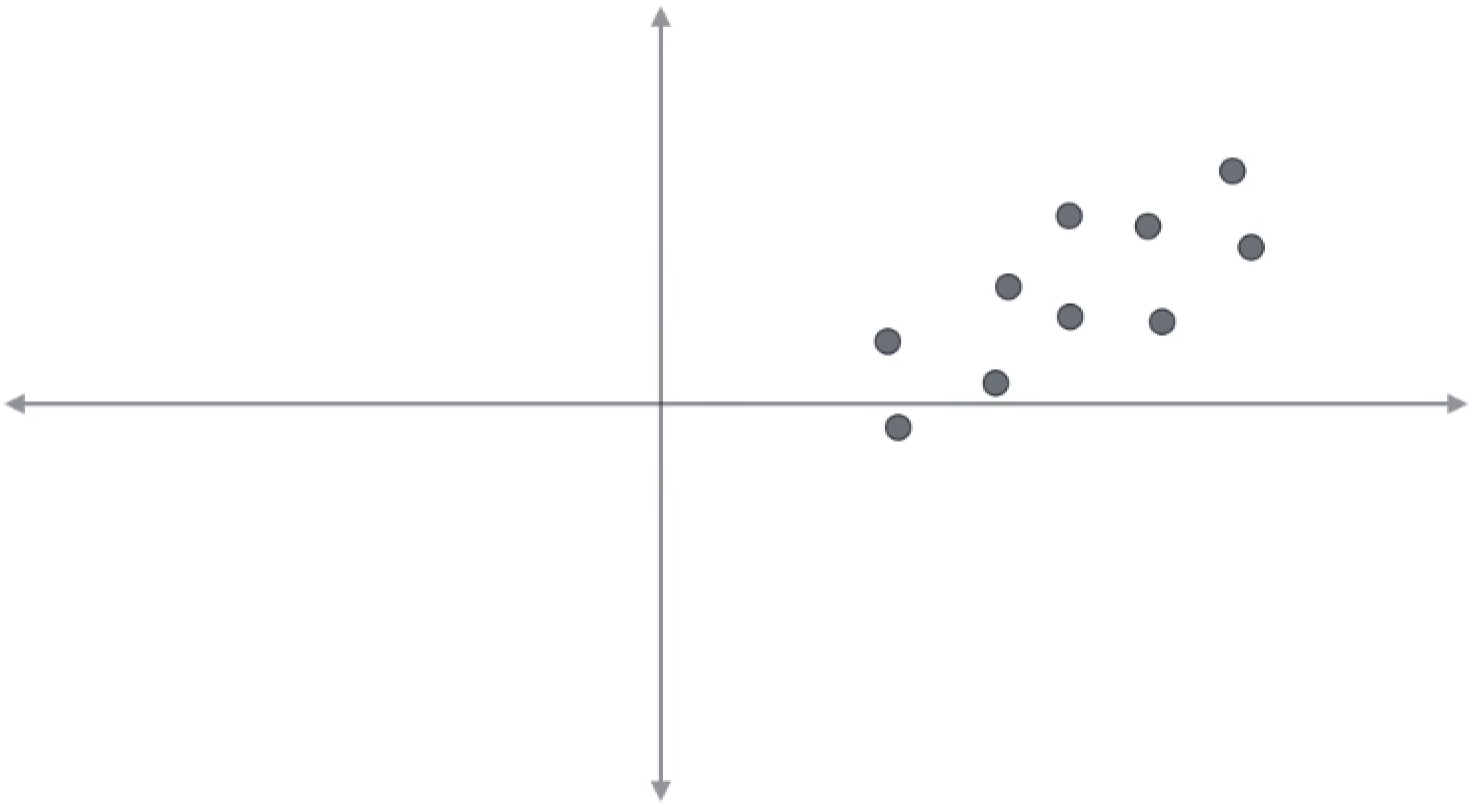


$$\text{covariance} = \frac{-2 + 0 + 2 + 0 + 0 + 0 + 2 + 0 + -2}{9} = 0$$

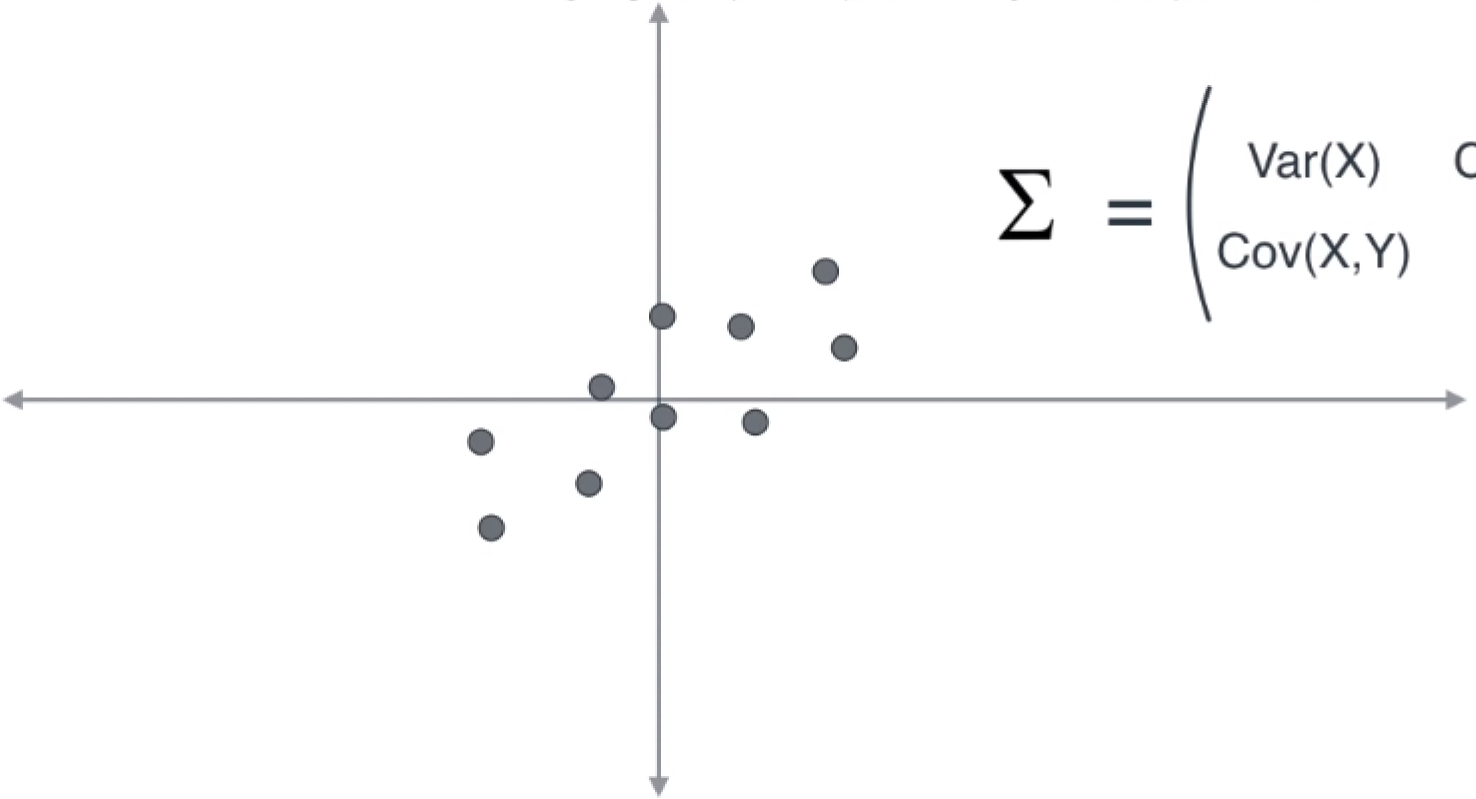
So by looking at the set (or data points) we get an intuition about the covariance

Covariance helps a lot in PCA

Imagine we have a dataset like below. How do we find the perfect projection

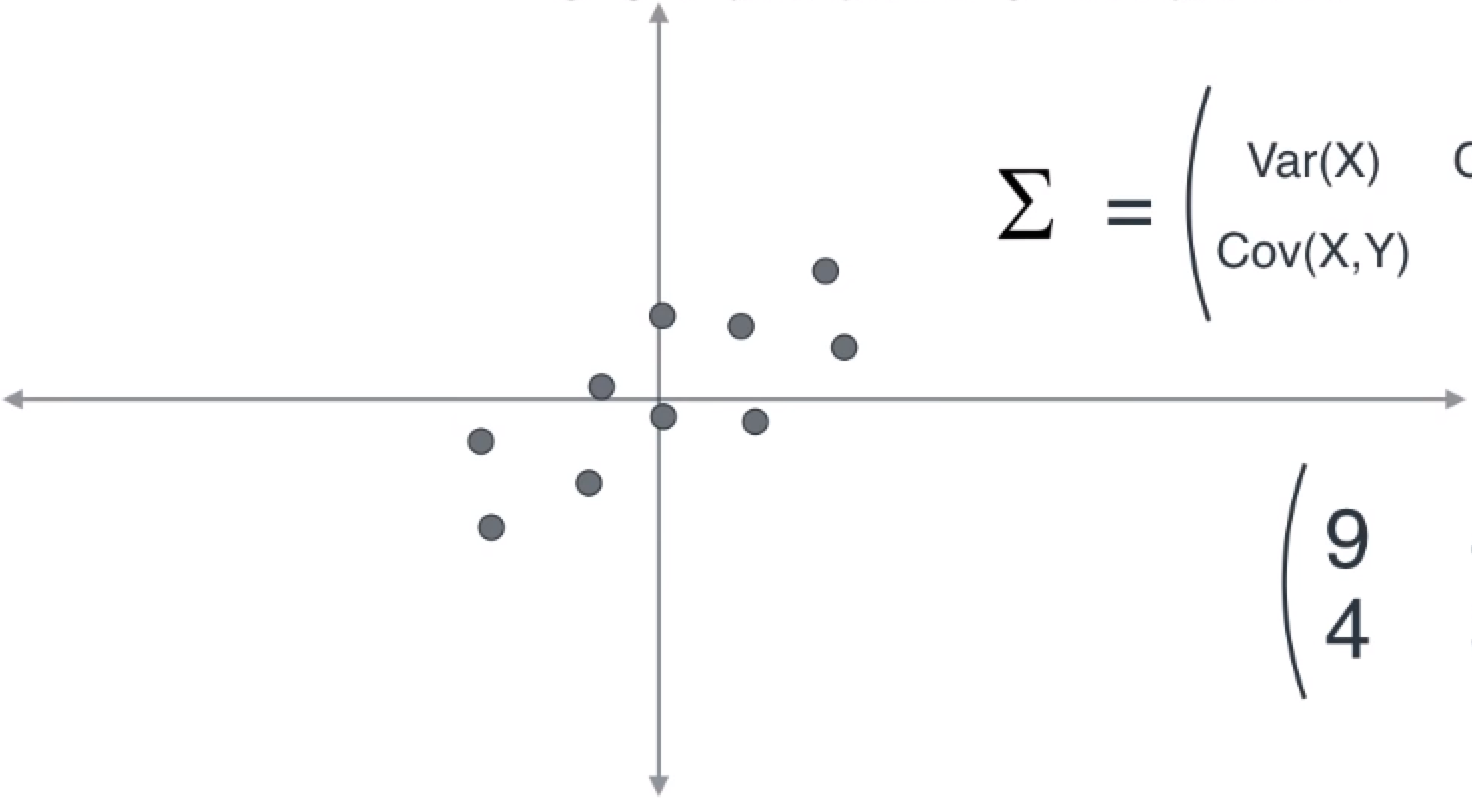


Covariance matrix



$$\Sigma = \begin{pmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \mathrm{Var}(Y) \end{pmatrix}$$

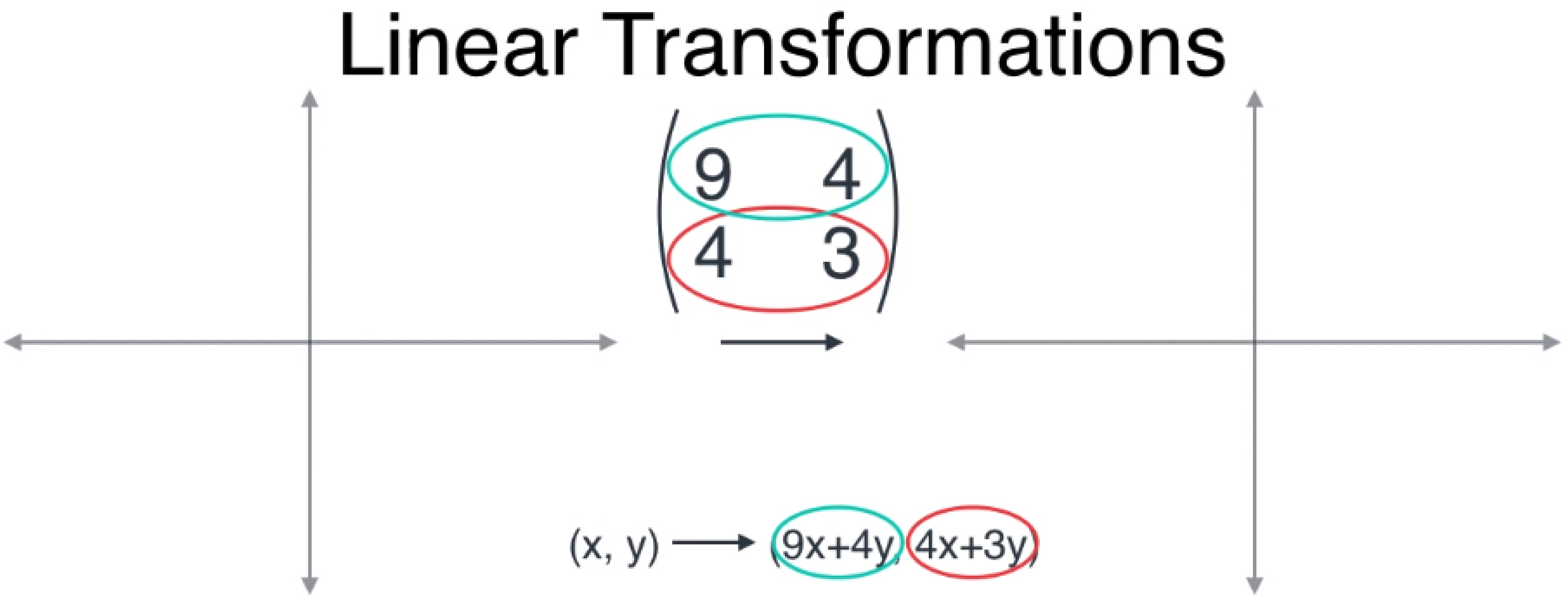
Covariance matrix



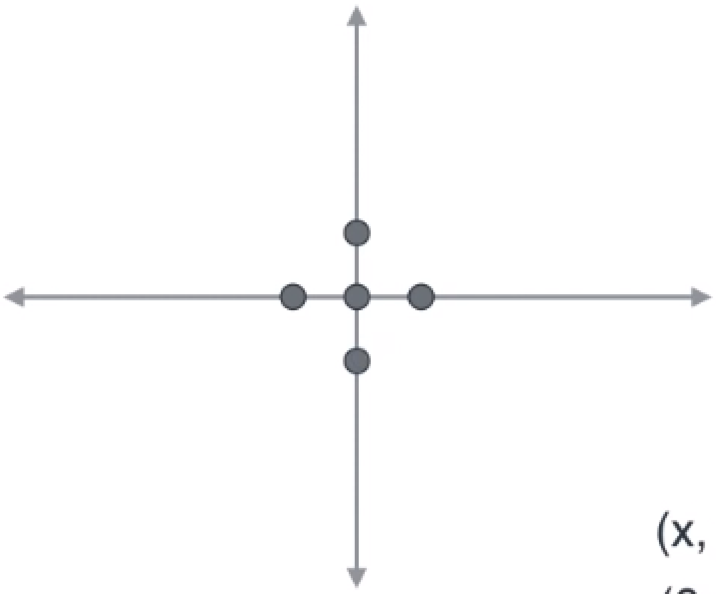
$$\Sigma = \begin{pmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \mathrm{Var}(Y) \end{pmatrix}$$

$$\begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

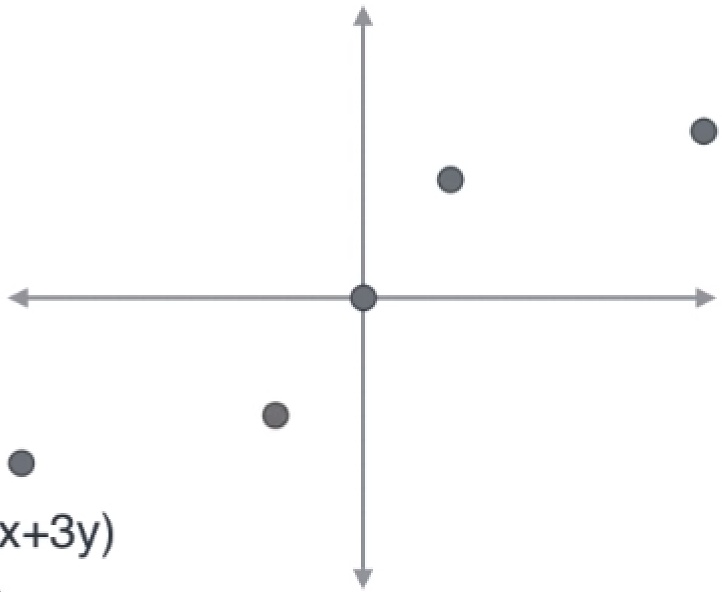
Linear transformation just turns the points of one plane to point on another plane using the number of matrix



Linear Transformations

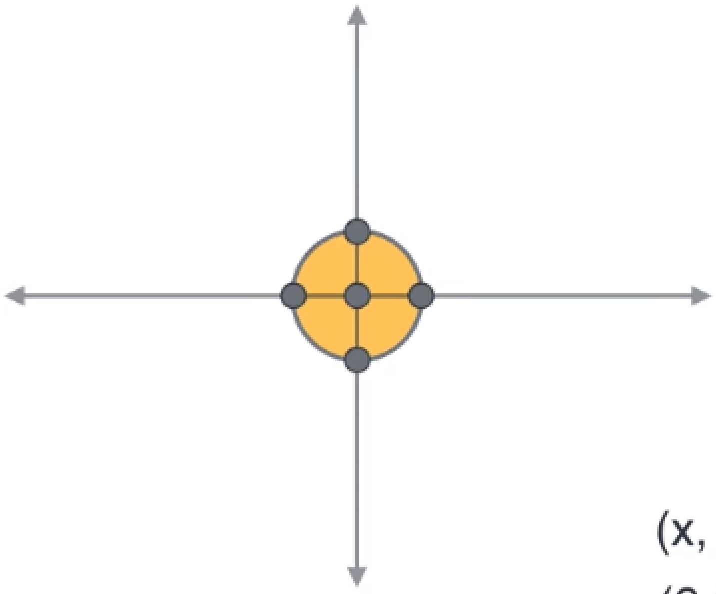


$$\begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix} \longrightarrow$$

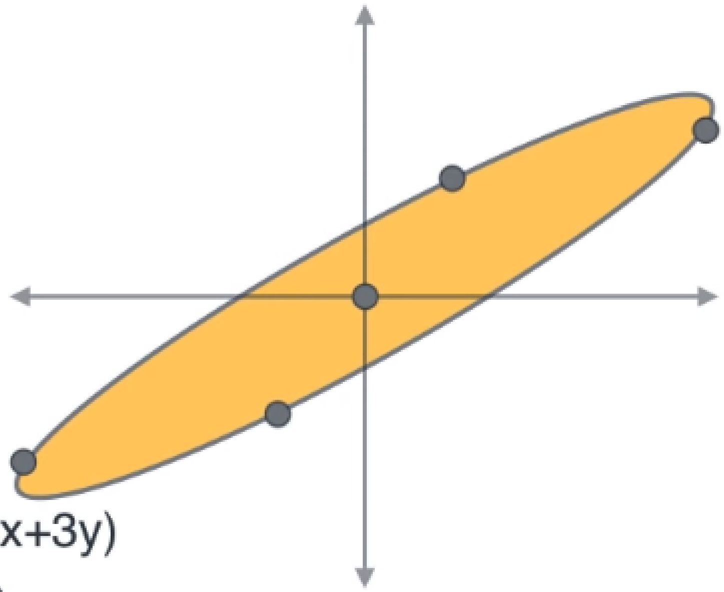


$(x, y) \longrightarrow$	$(9x+4y, 4x+3y)$
(0,0)	(0,0)
(1,0)	(9,4)
(0,1)	(4,3)
(-1,0)	(-9,-4)
(0,-1)	(-4,-3)

Linear Transformations



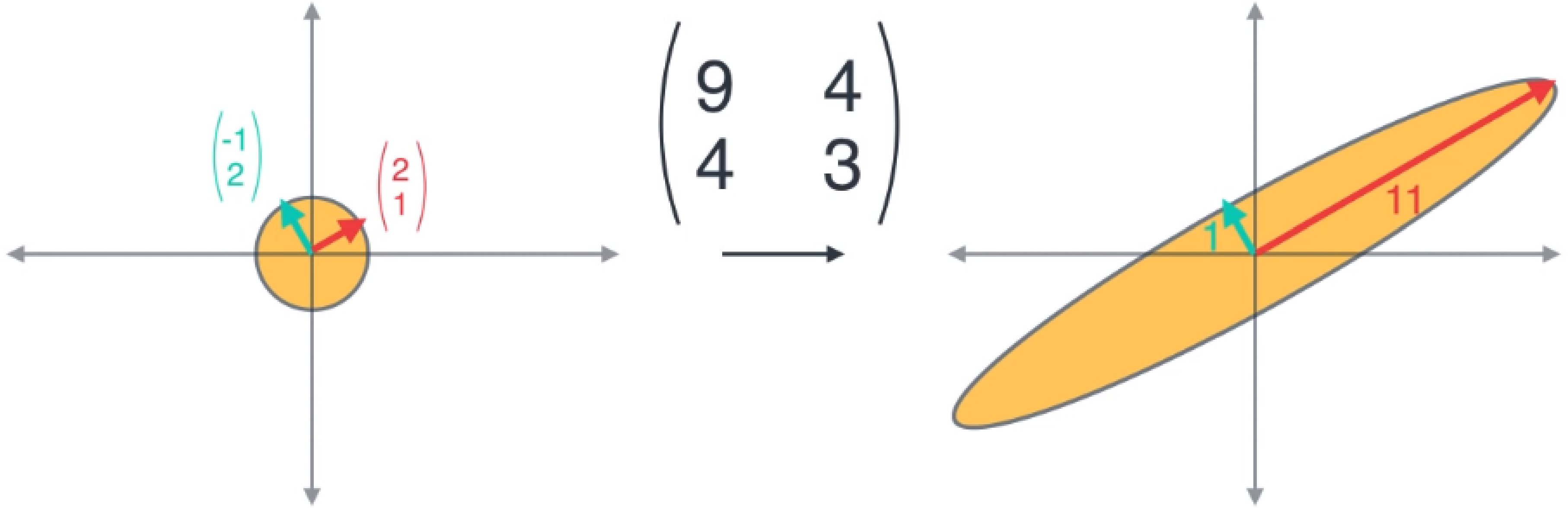
$$\begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix} \longrightarrow$$



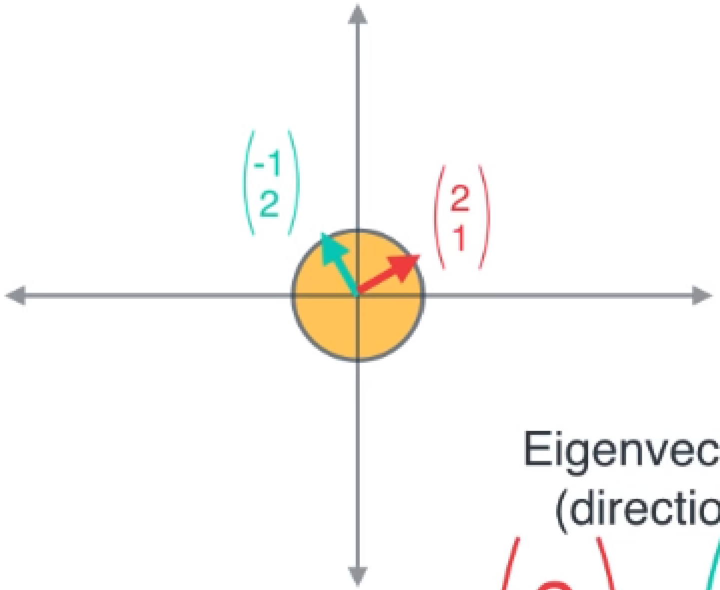
$(x, y) \longrightarrow (9x+4y, 4x+3y)$

(0,0)	(0,0)
(1,0)	(9,4)
(0,1)	(4,3)
(-1,0)	(-9,-4)
(0,-1)	(-4,-3)

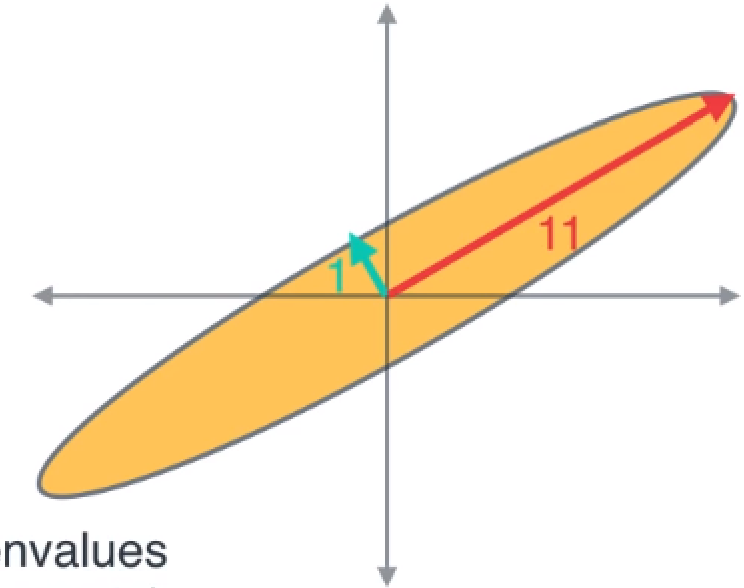
Linear Transformations



Linear Transformations



$$\begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix} \longrightarrow$$



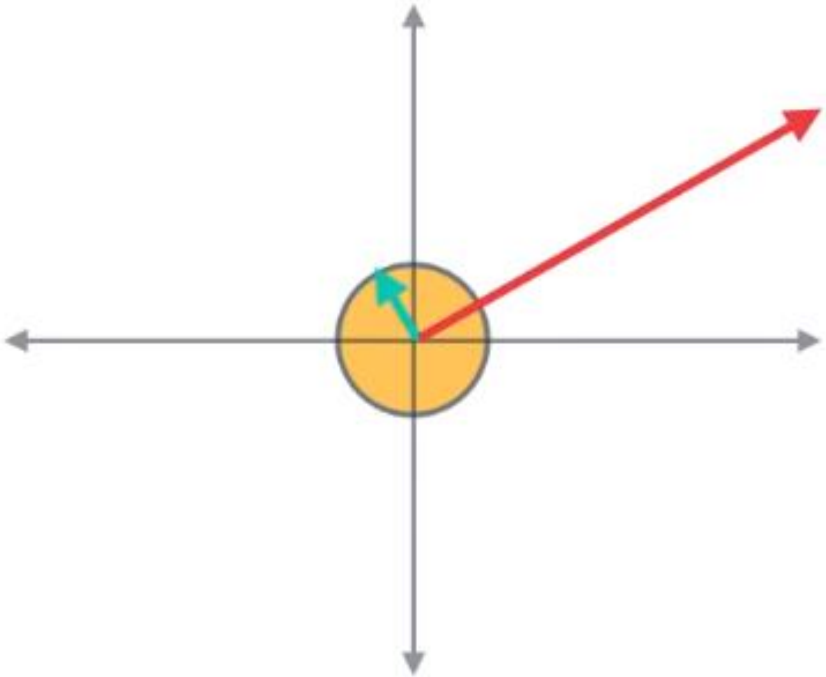
Eigenvectors
(direction)

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Eigenvalues
(magnitude)

11 1

Linear Transformations



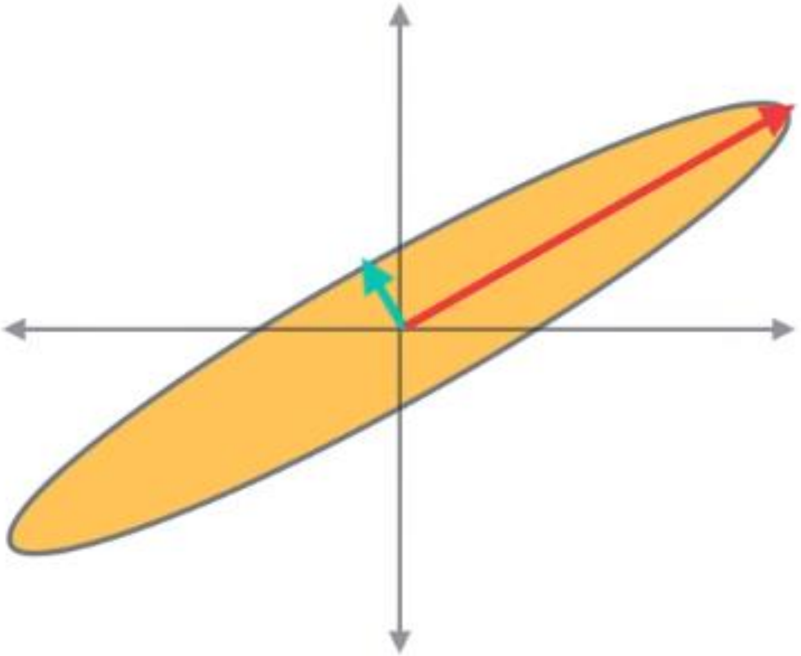
Eigenvectors

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Eigenvalues

11   1

Linear Transformations



Eigenvectors
(direction)

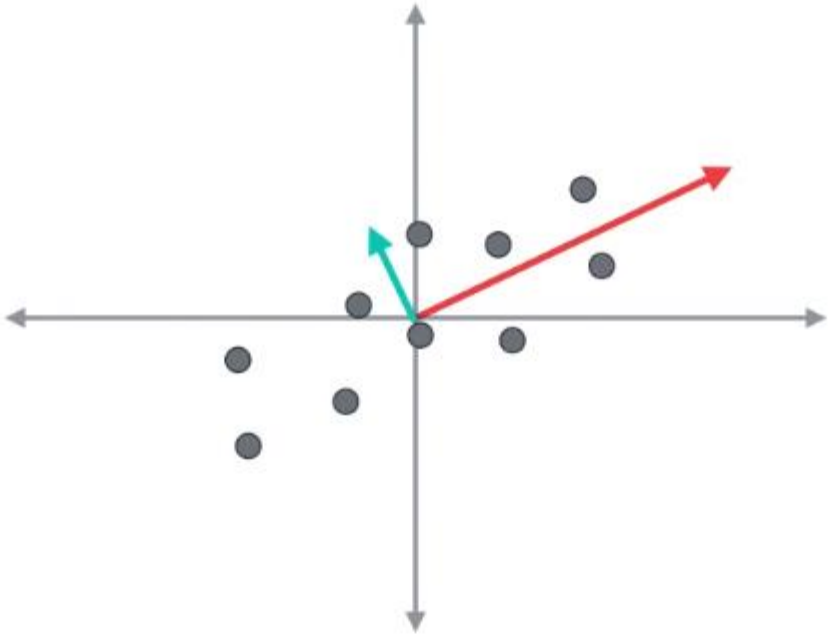
$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Eigenvalues
(magnitude)

11   1

Hence the linear transformation is stretching the plane in two direction, and those directions are given by Eigen vectors and the amount is given by Eigen values

Principal Component Analysis (PCA)



$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

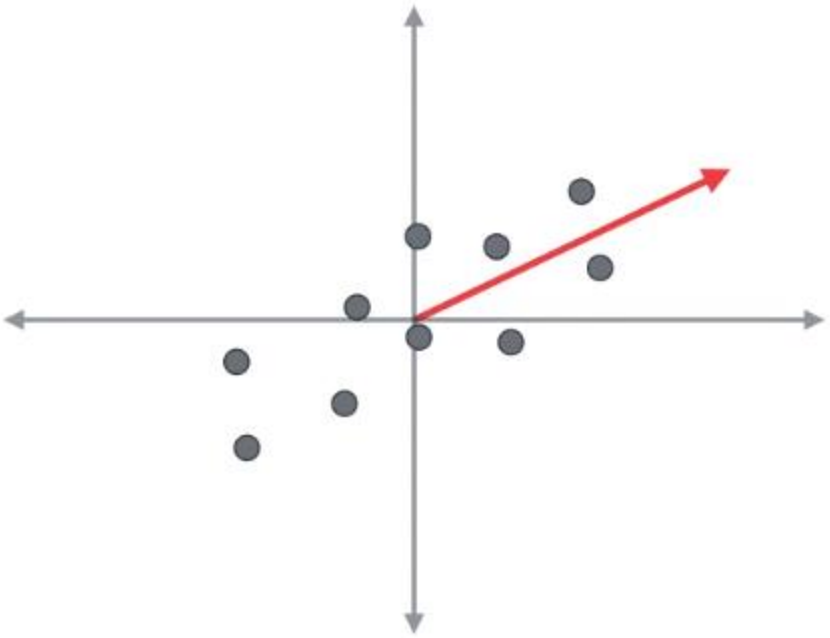
$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Eigenvectors (direction)

$$11 \quad 1$$

Eigenvalues (magnitude)

Principal Component Analysis (PCA)



$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

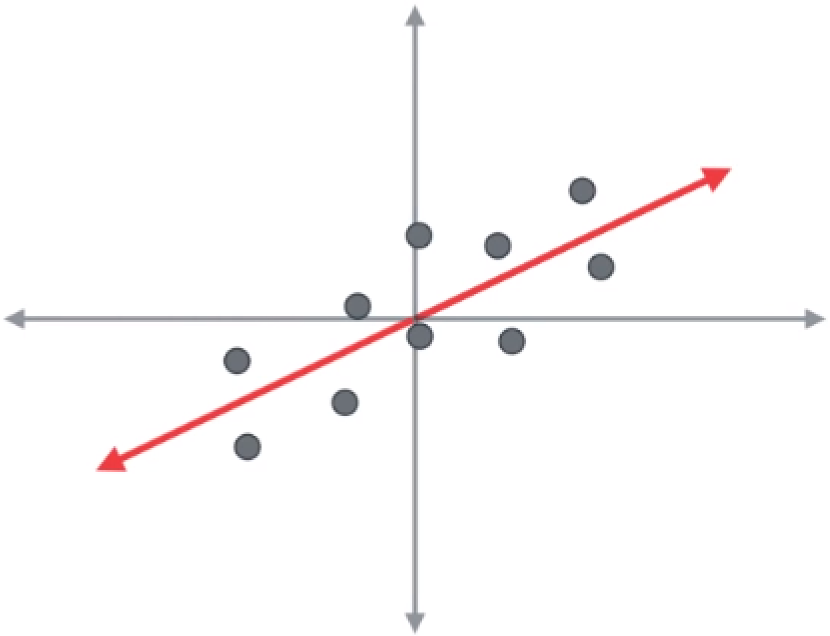
$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Eigenvectors
(direction)

$$11$$

Eigenvalues
(magnitude)


Principal Component Analysis (PCA)

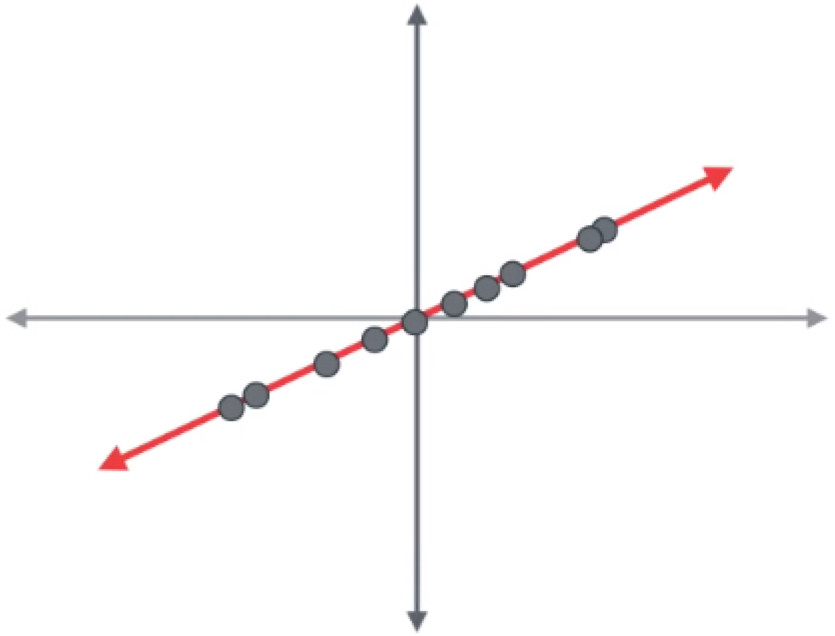


$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ Eigenvectors (direction)

11 Eigenvalues (magnitude)

Principal Component Analysis (PCA)

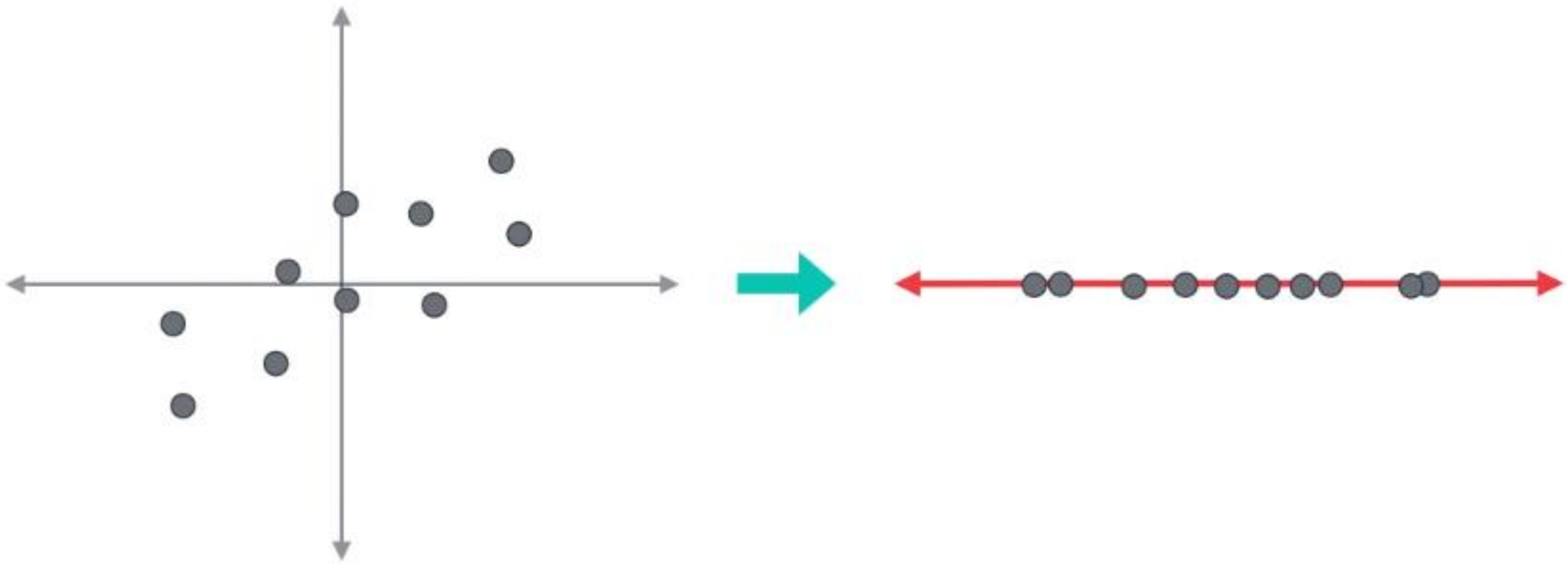


$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

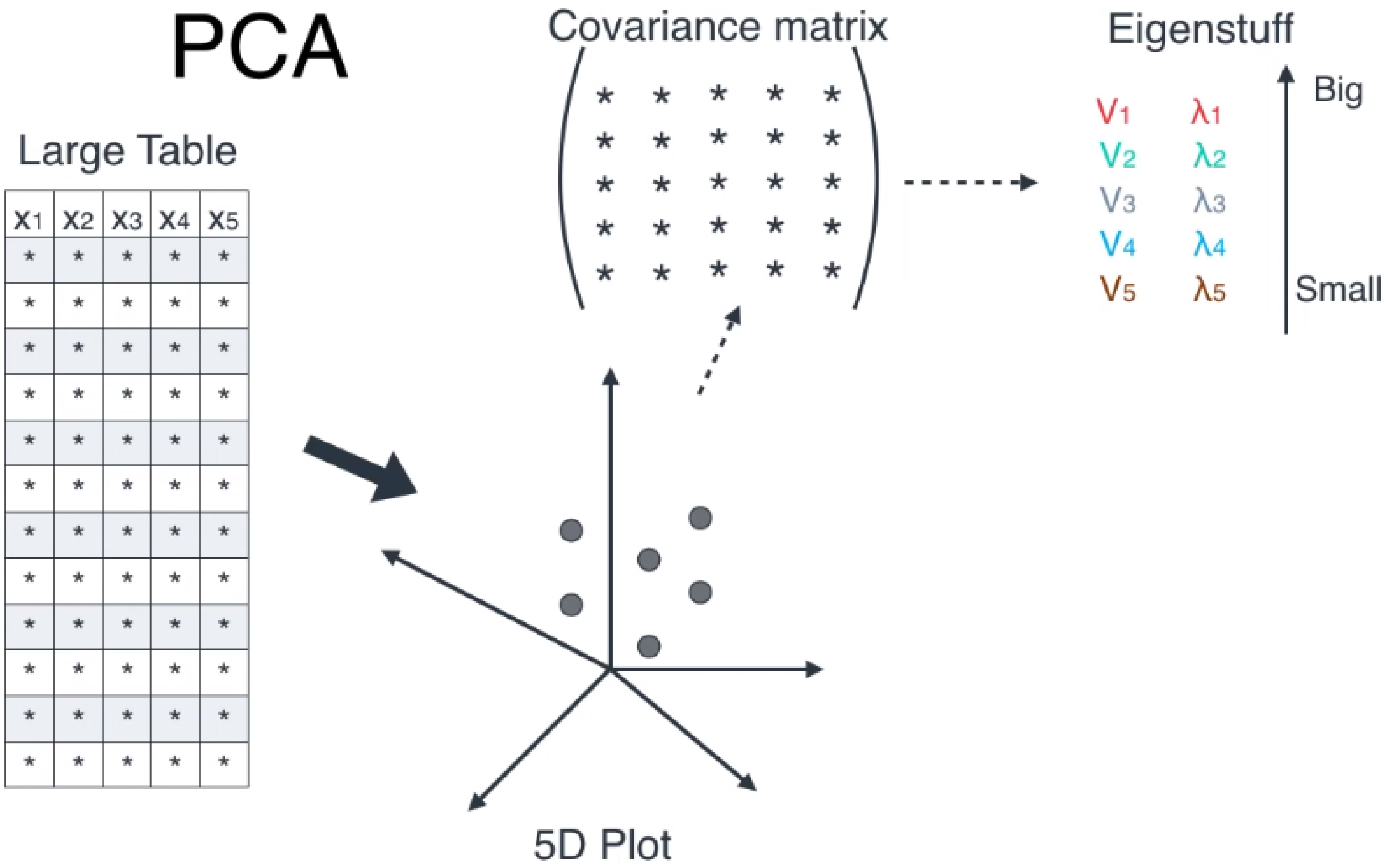
$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ Eigenvectors (direction)

11 Eigenvalues (magnitude)

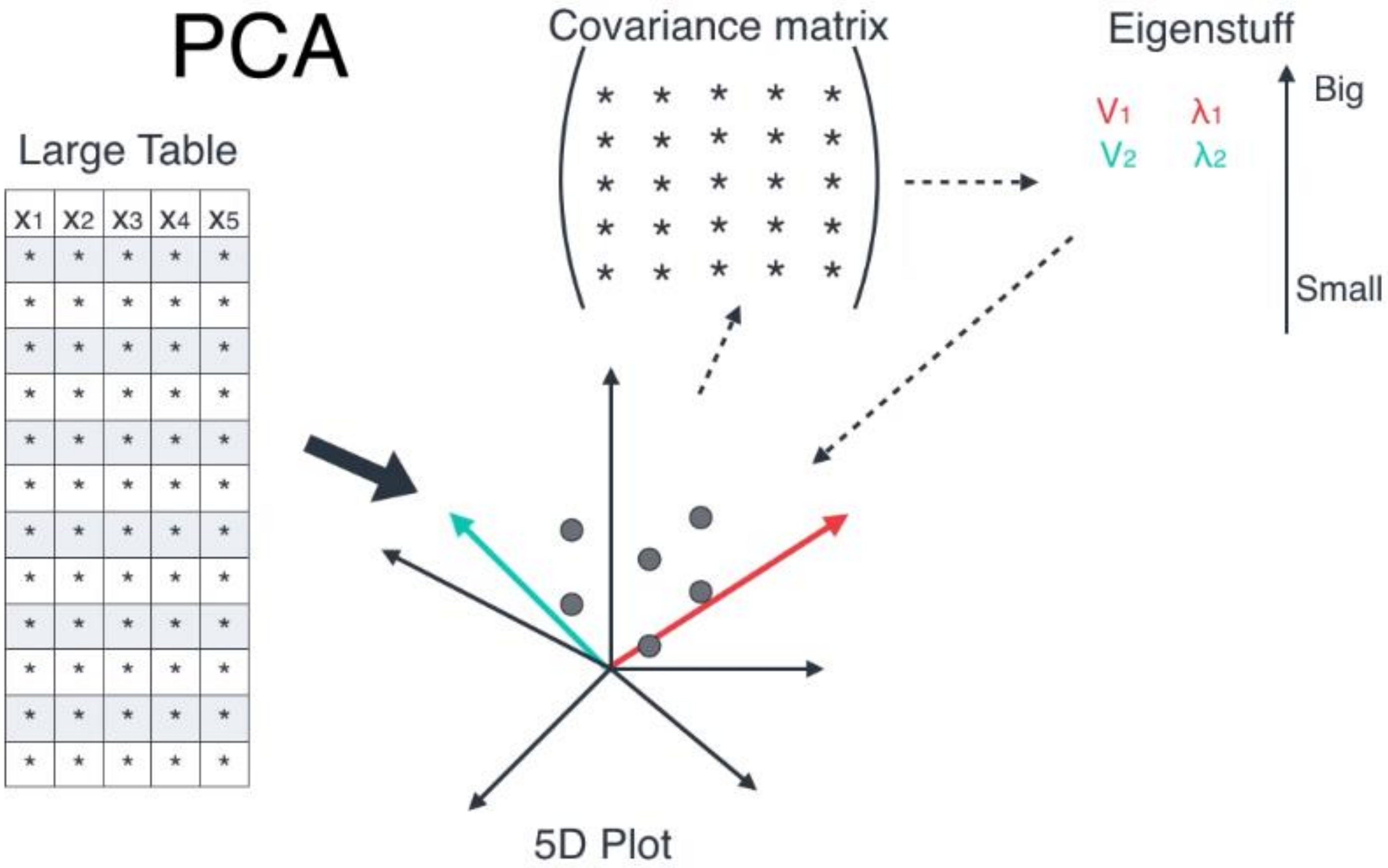
Principal Component Analysis (PCA)



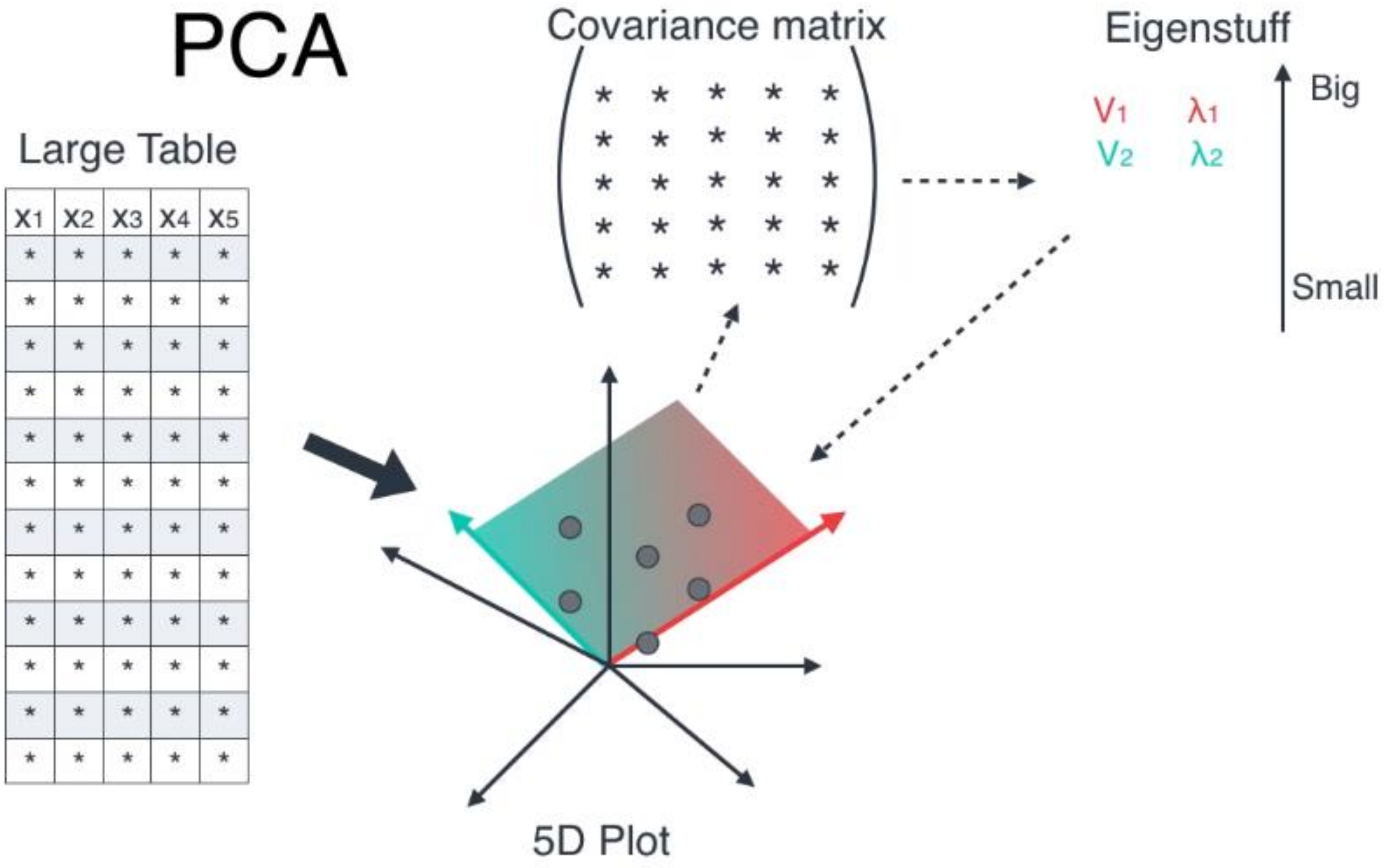
2D dataset is transformed into 1D dataset by picking Eigen vector with high Eigen value i.e. projecting the axis the carries most amount of information

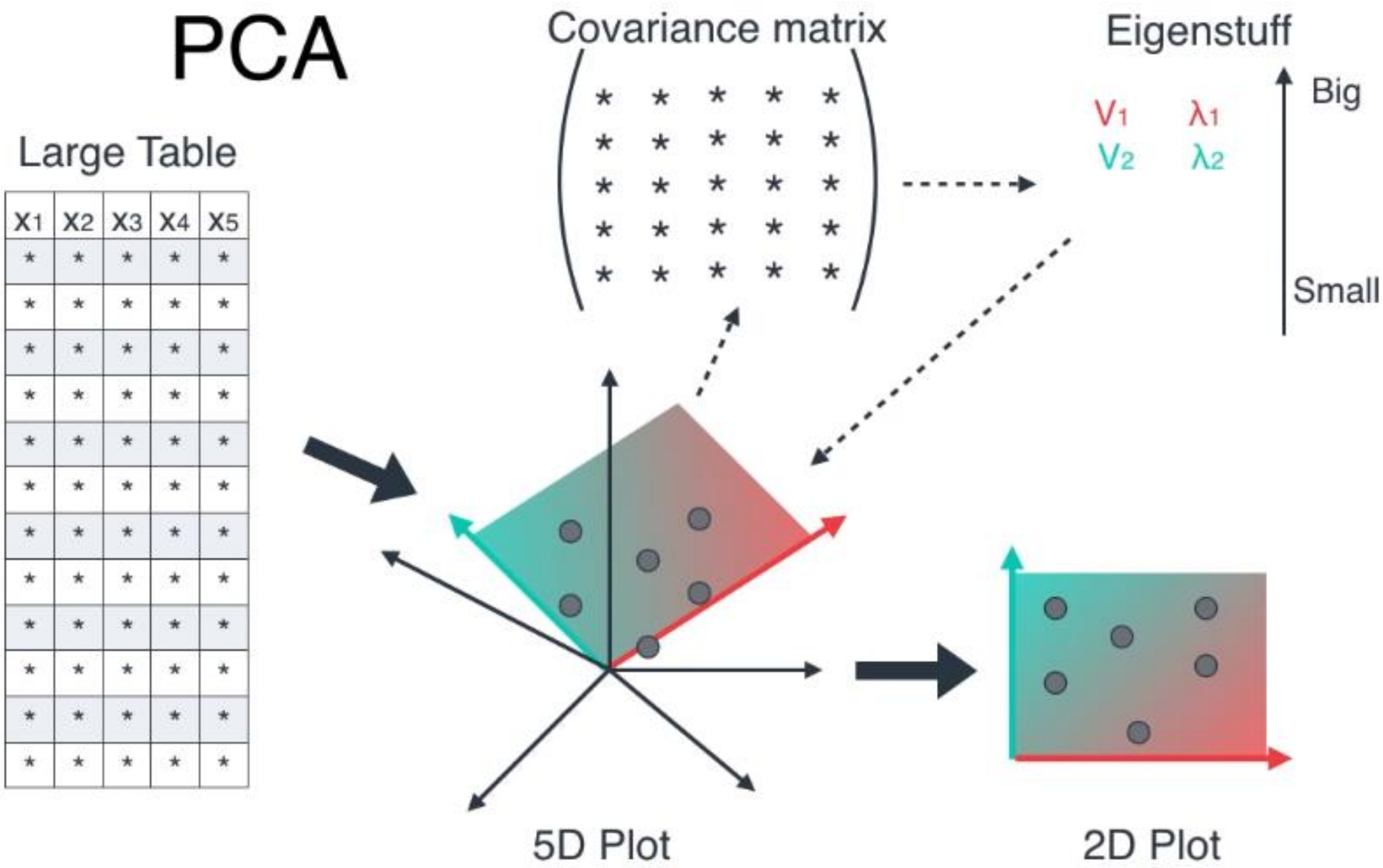

PCA
Large Table
x1 x2 x3 x4 x5
Covariance matrix
Eigenstuff
V1 λ1
V2 λ2
V3 λ3
V4 λ4
V5 λ5
Big
Small
5D Plot

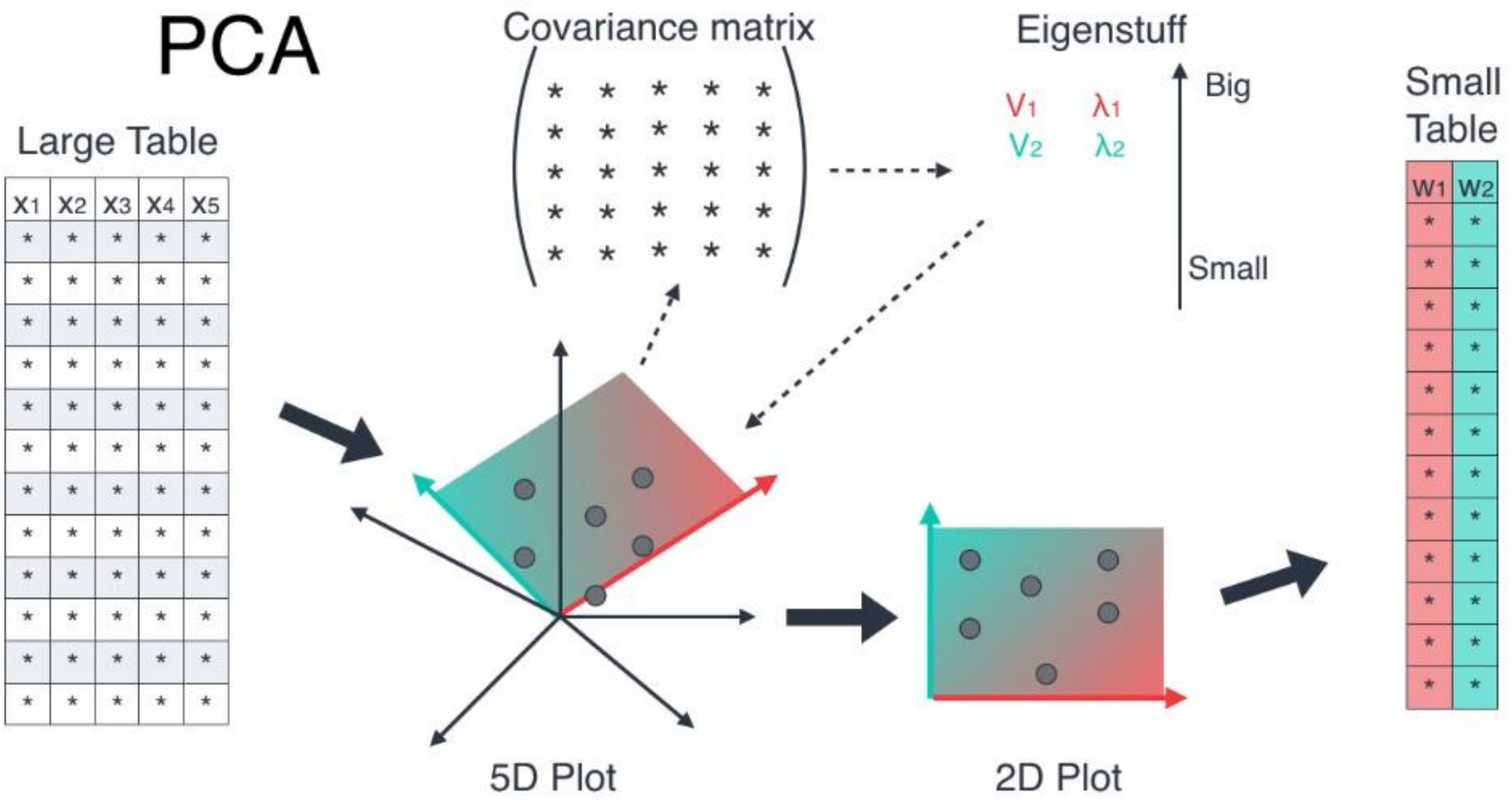
PCA


Covariance matrix
Eigenstuff
Large Table
x1 x2 x3 x4 x5
V1 λ1
V2 λ2
Big
Small
5D Plot

PCA


Covariance matrix
Eigenstuff
Big
Small
V_1 λ_1
V_2 λ_2
Large Table
x_1 x_2 x_3 x_4 x_5
5D Plot


PCA
Covariance matrix
Eigenstuff
Big
V_1 λ_1
V_2 λ_2
Small
Large Table
x_1 x_2 x_3 x_4 x_5
5D Plot
2D Plot


PCA
Large Table
x1 x2 x3 x4 x5
Covariance matrix
Eigenstuff
V1 λ1
V2 λ2
Big
Small
Small Table
w1 w2
5D Plot
2D Plot

Credits: The PCA content is taken from amazing explanation of Luis Serrano