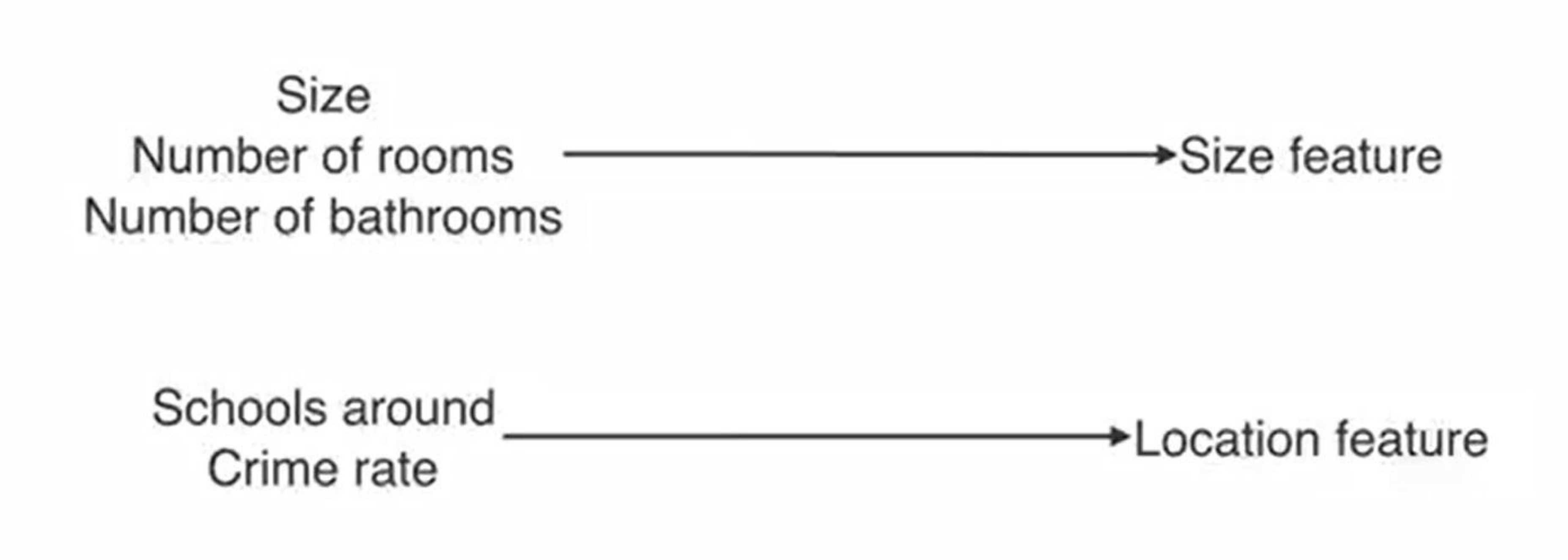
Principal Component Analysis (PCA)

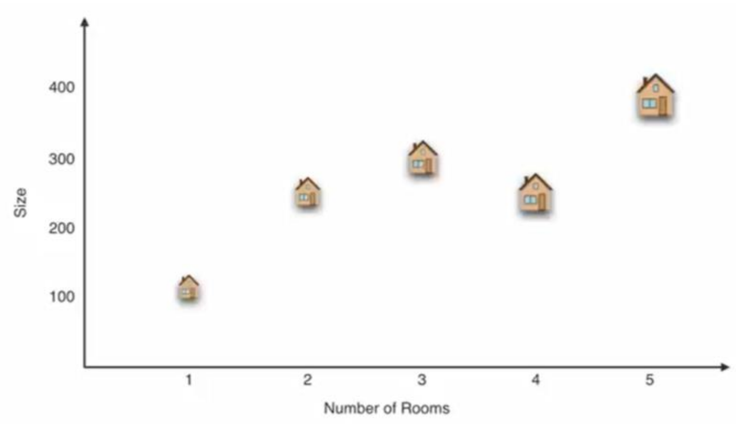
Gundimeda Venugopal

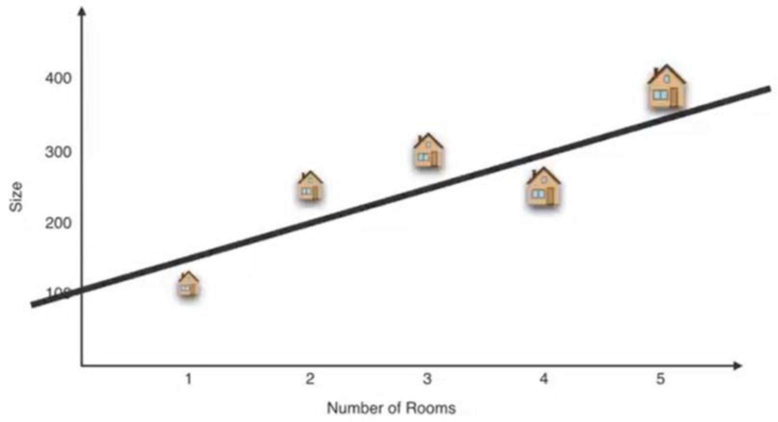
Dimensionality Reduction: Taking a picture of the Data

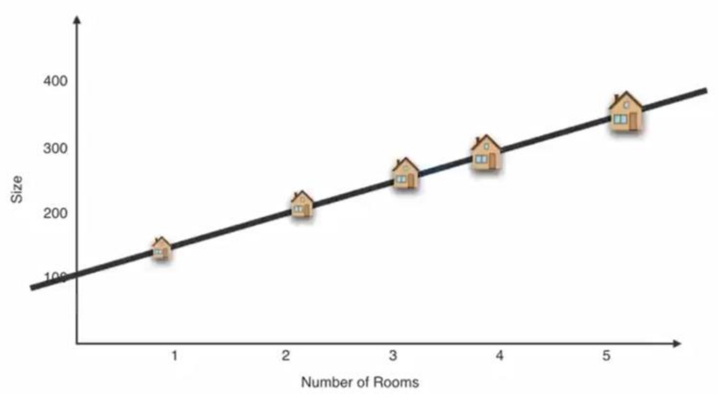
Housing Data



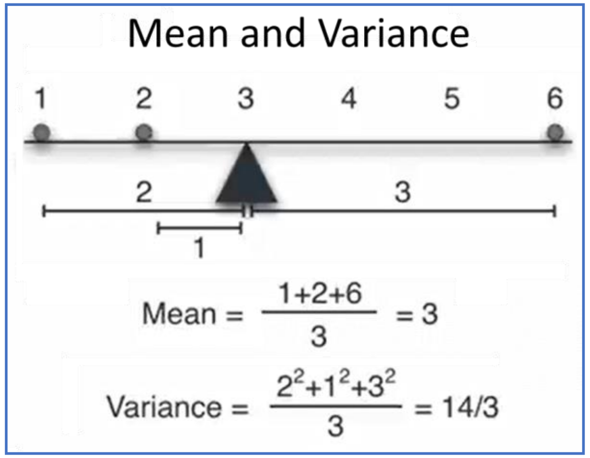
Dimensionality Reduction (2d to 1d example)

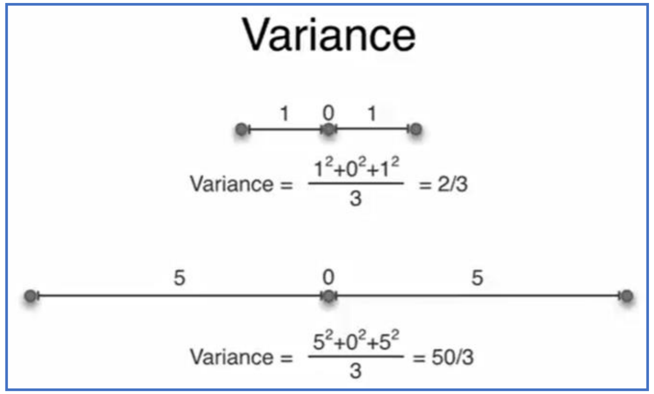


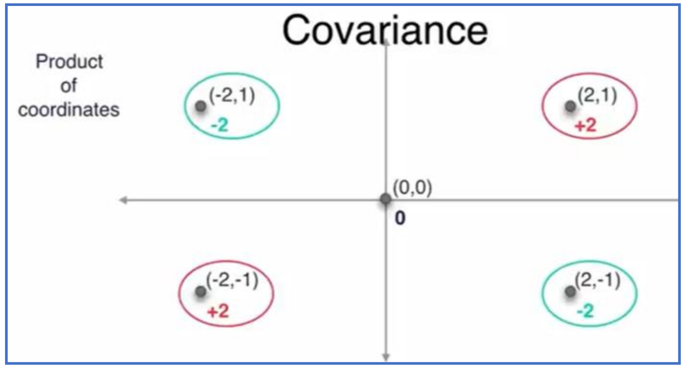


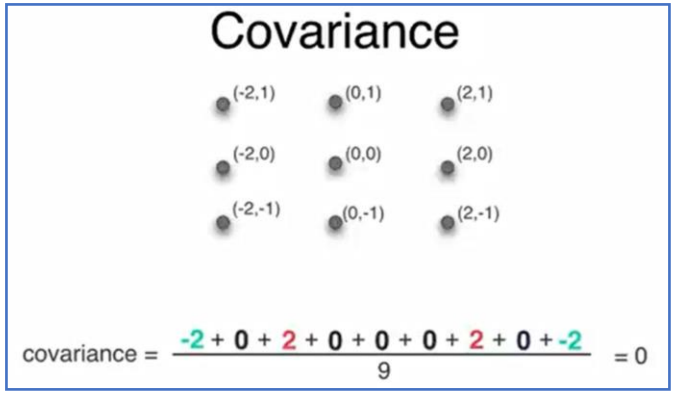


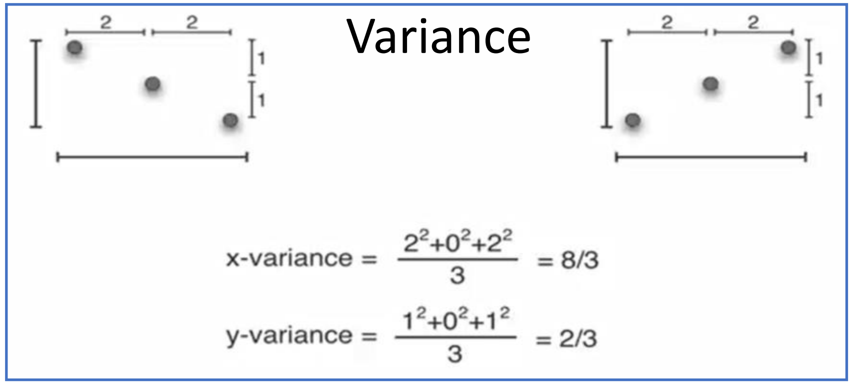
Mean Variance and Covariance

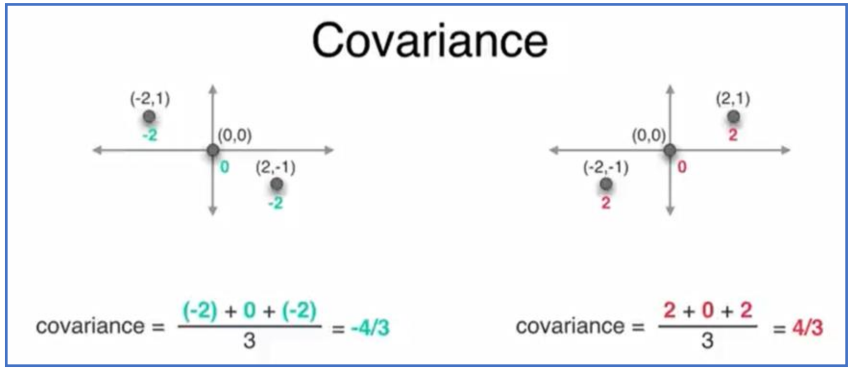


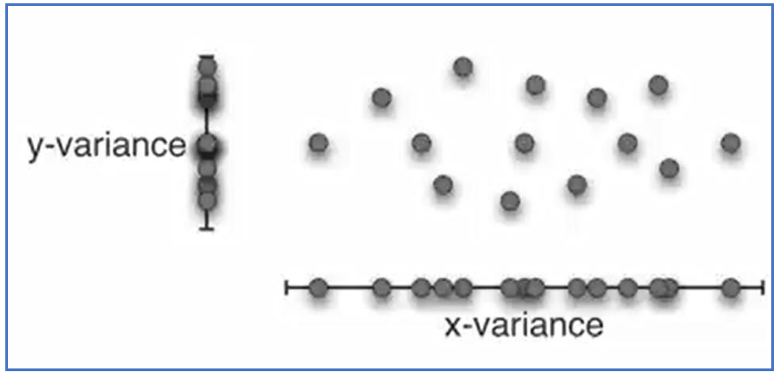


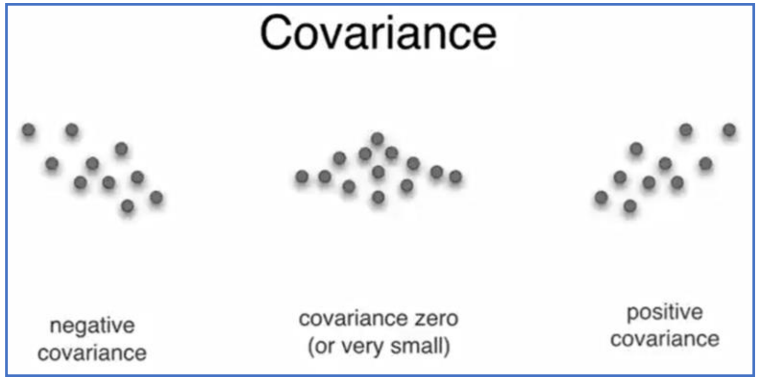


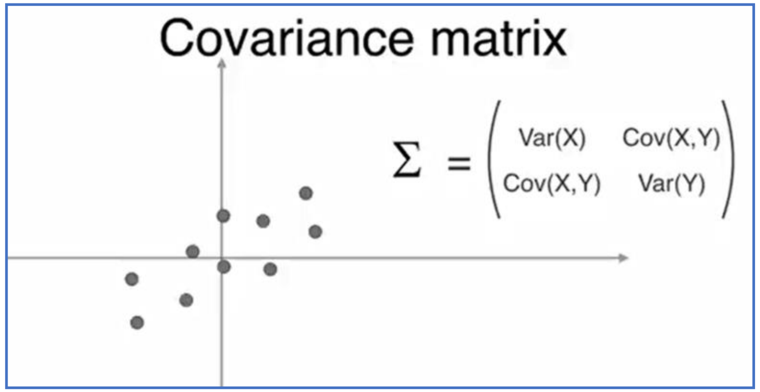




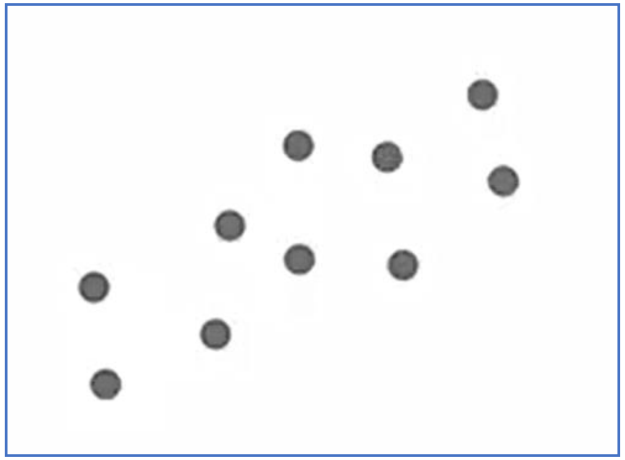


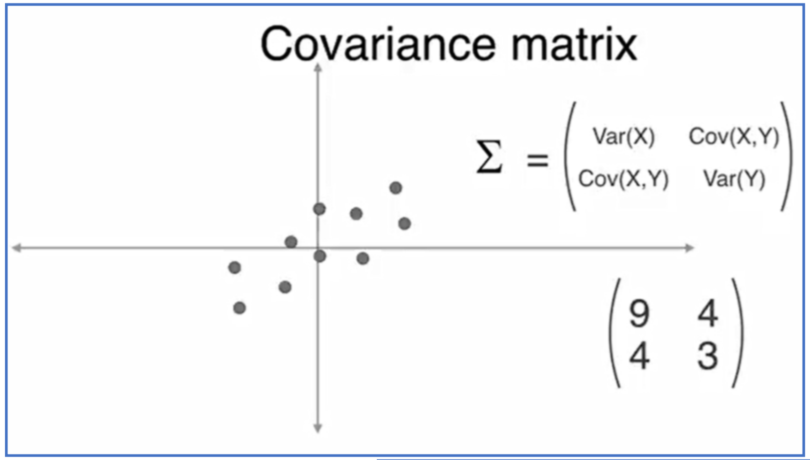


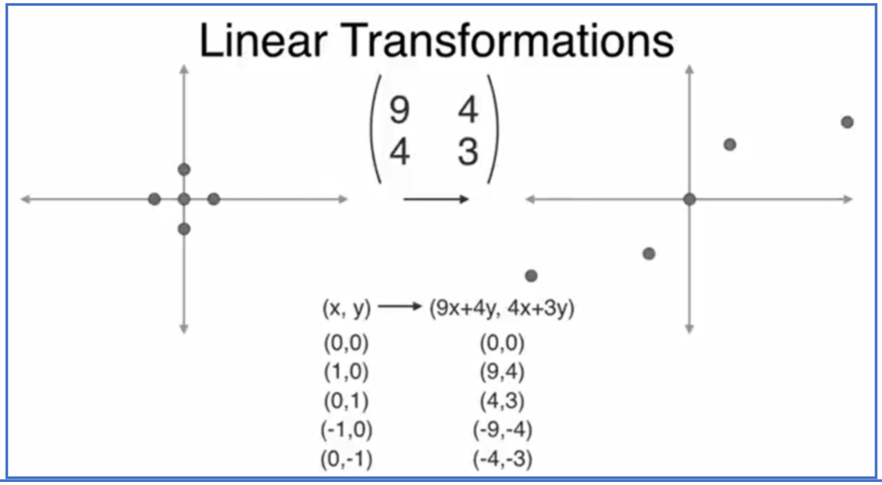


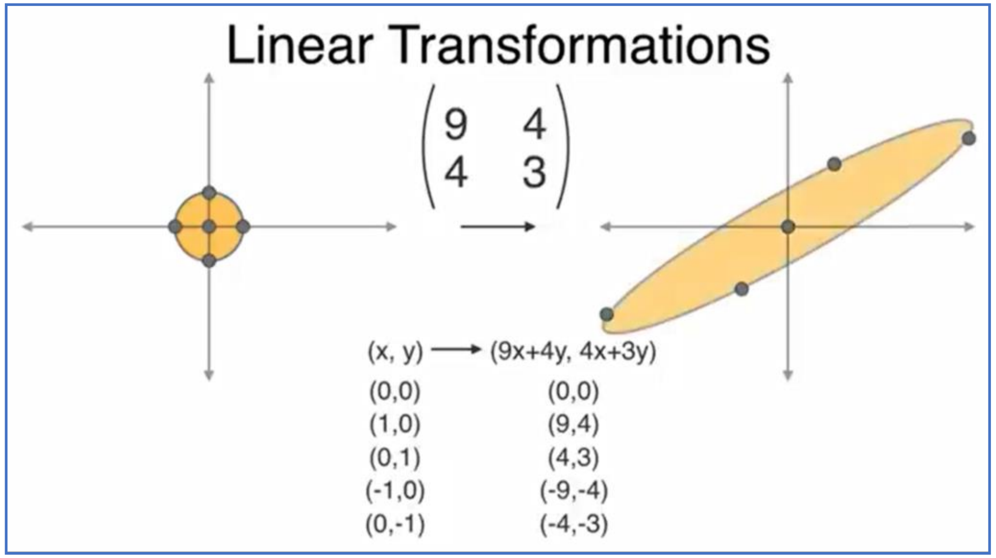


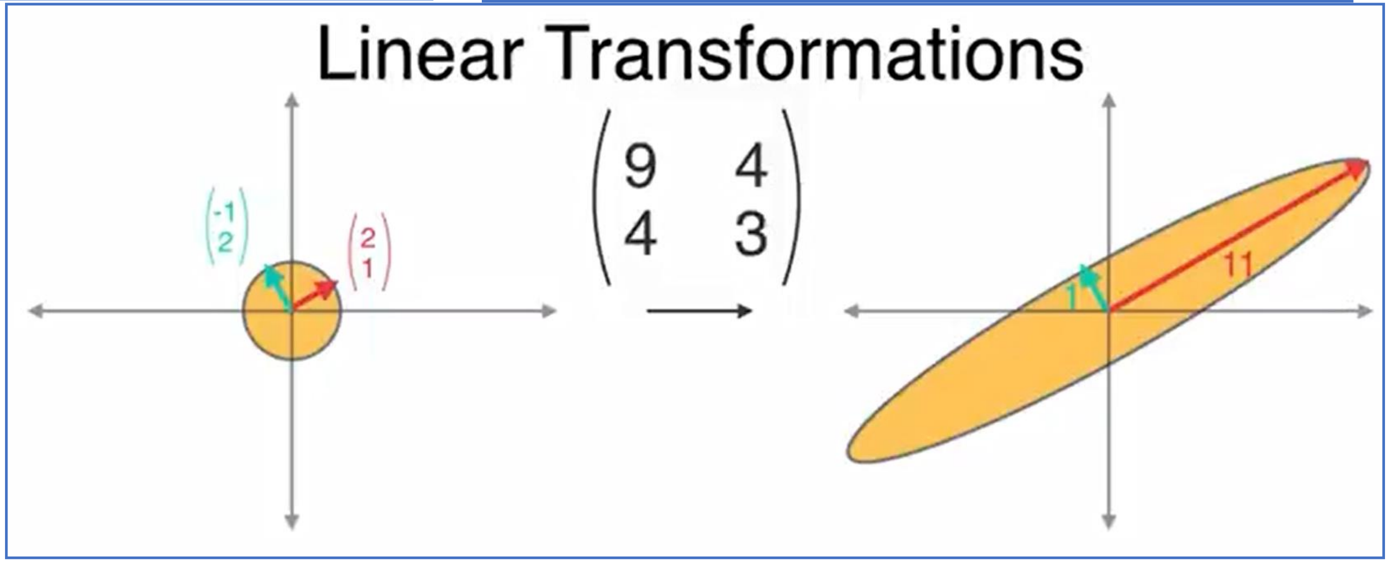
Covariance Matrix + Linear Transformations











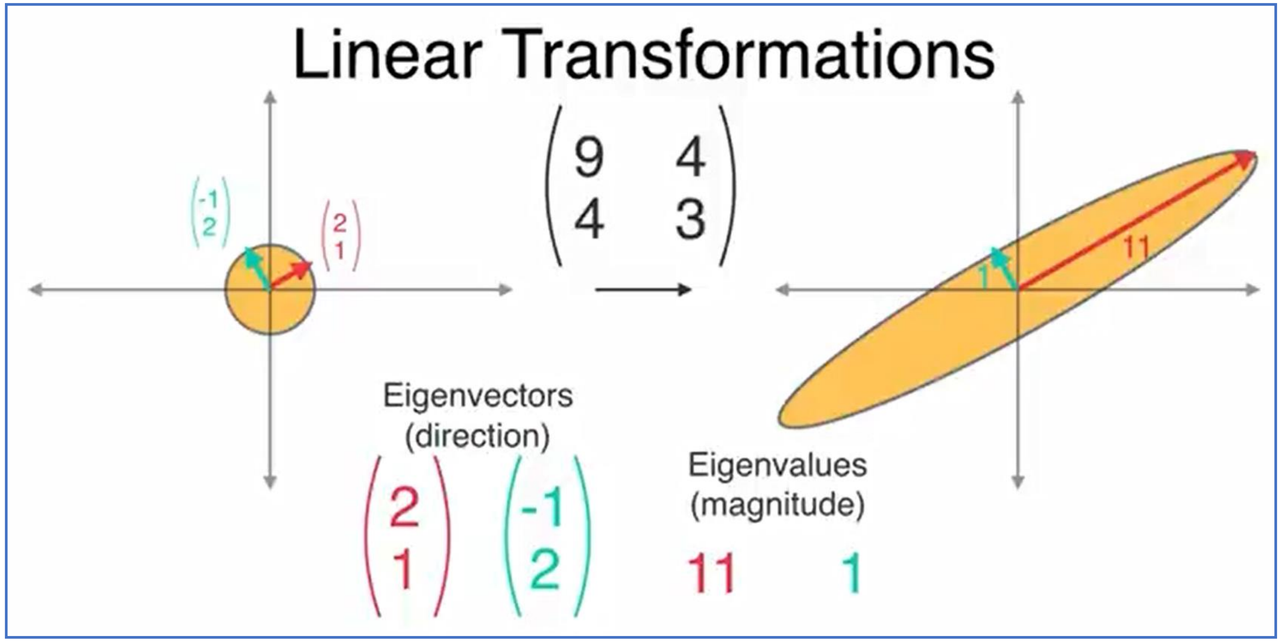
Linear transformations are operations that map one vector to another vector

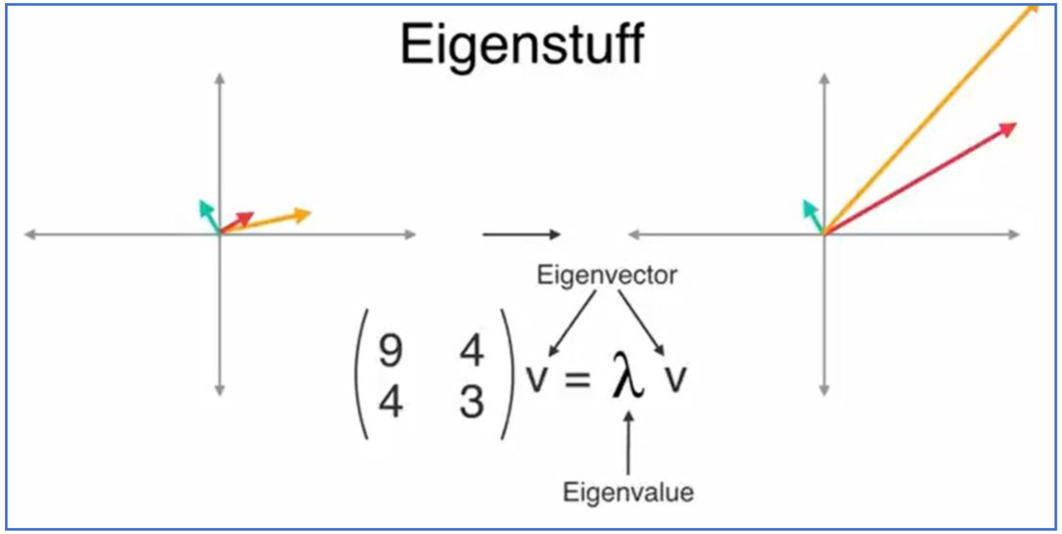
A matrix can be seen as a representation of a linear transformation, and a vector can be seen as a representation of a point in an image.

Eigen Values and Eigen Vectors

- An eigenvalue is a scalar value that indicates how much a vector is stretched or shrunk by a linear transformation.

- An eigenvector is a vector that does not change its direction when it is transformed by a matrix, only its magnitude.





- The eigenvalues can tell you how much variance or diversity there is in the image.
- The eigenvectors can tell you the directions or patterns that are most prominent in the image.
- By using the eigenvalues and eigenvectors, you can perform various tasks such as image compression, segmentation & recognition.

Eigen Value of a matrix

Eigenvalues

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

Characteristic Polynomial

$$\begin{vmatrix} x-9 & -4 \\ -4 & x-3 \end{vmatrix} = (x-9)(x-3) - (-4)(-4) = x^2 - 12x + 11$$

$$= (x-11)(x-1)$$

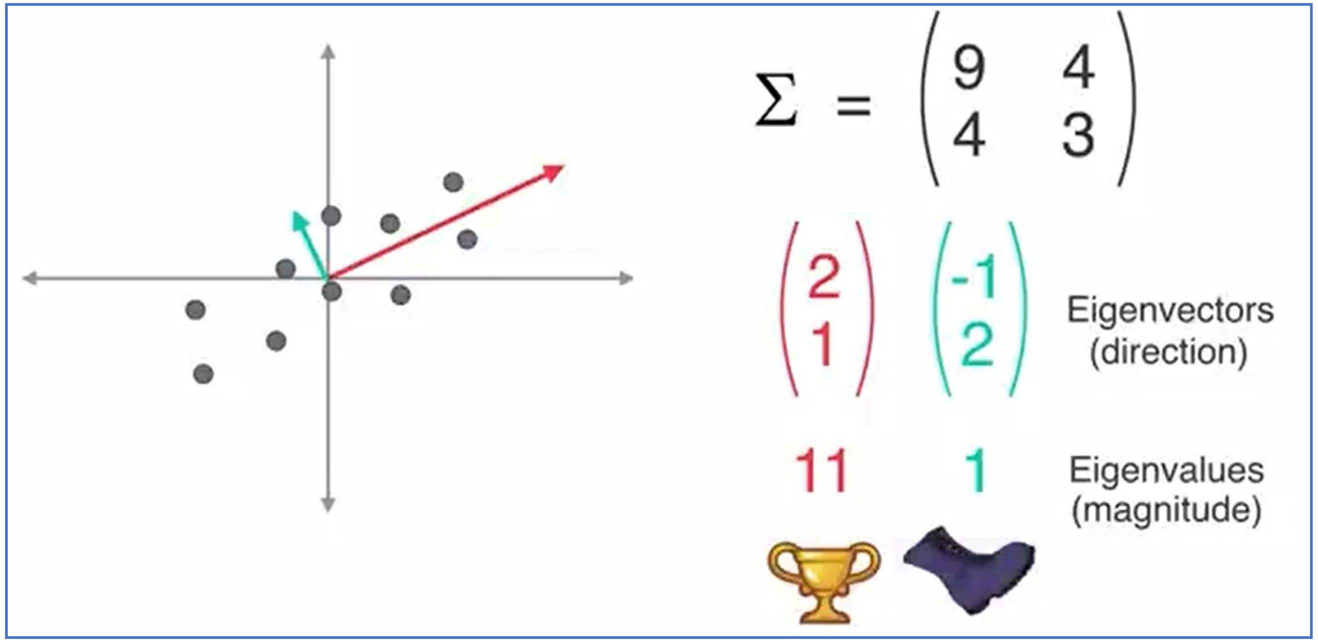
Eigenvalues 11 and 1

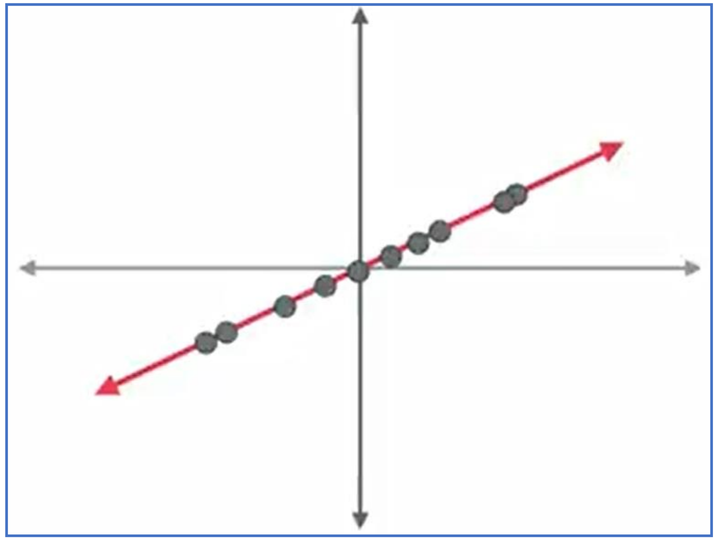
Eigenvectors

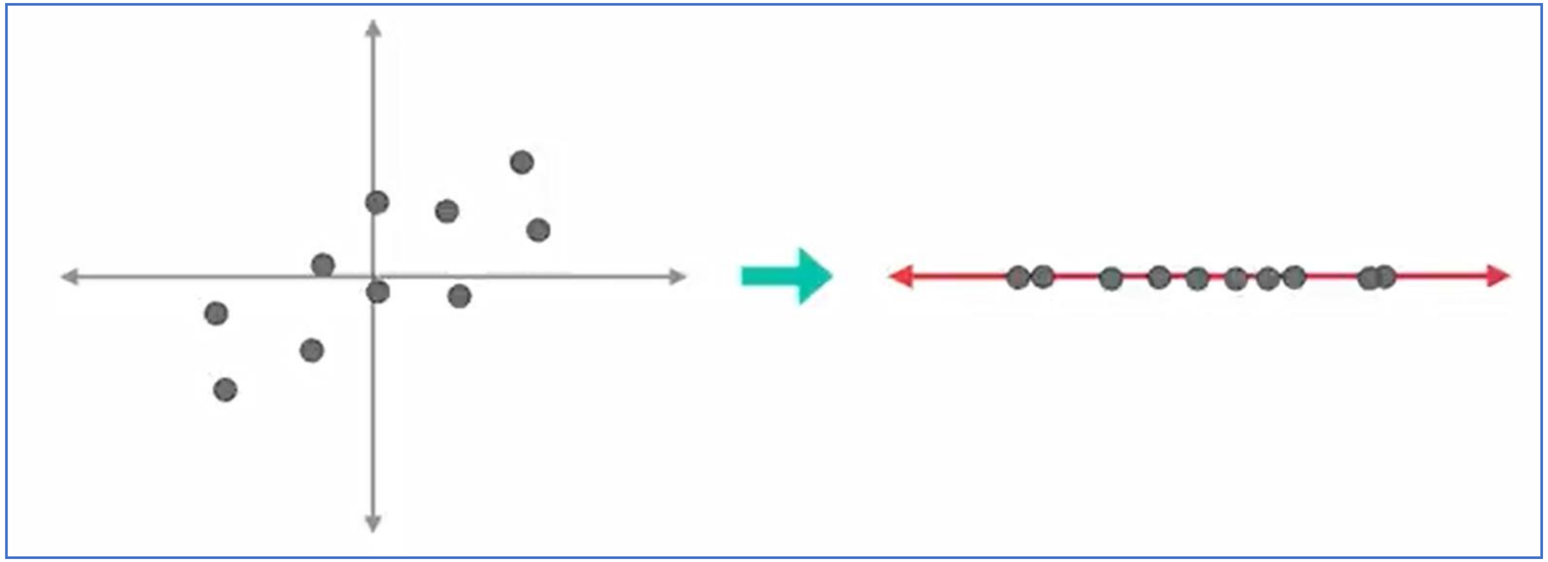
$$\begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = 11\begin{pmatrix} u \\ v \end{pmatrix} \qquad \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = 1\begin{pmatrix} u \\ v \end{pmatrix}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Principal Component Analysis (PCA)

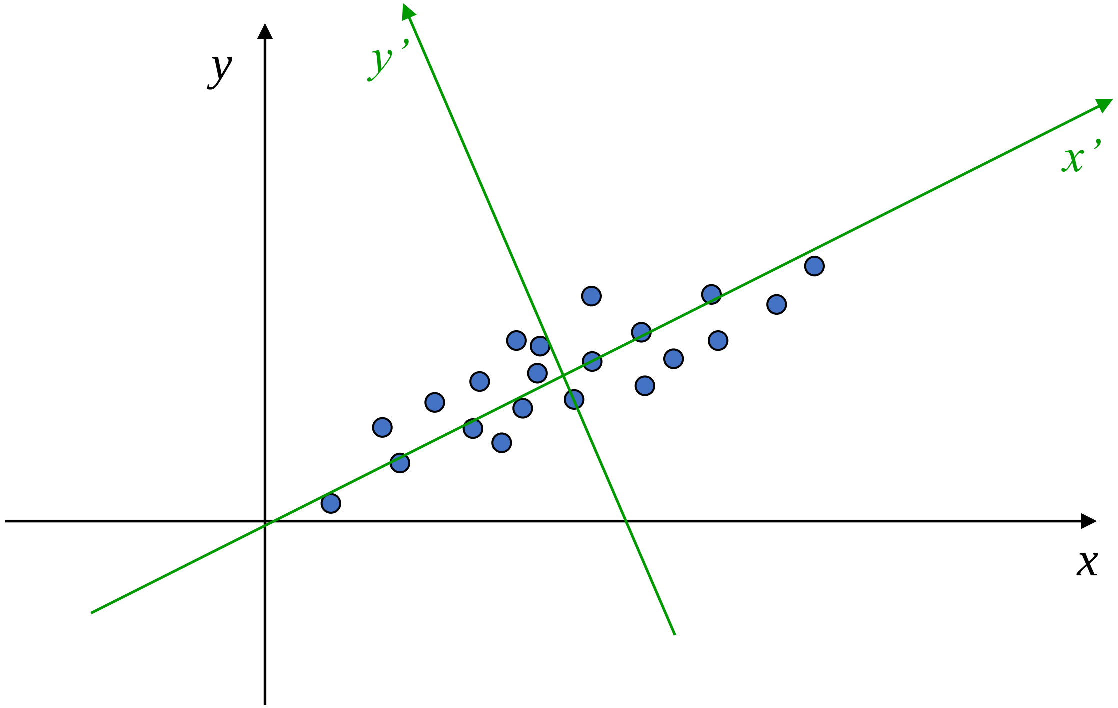






Principal Component Analysis – the general idea

❖PCA finds an orthogonal basis that best represents given data set.

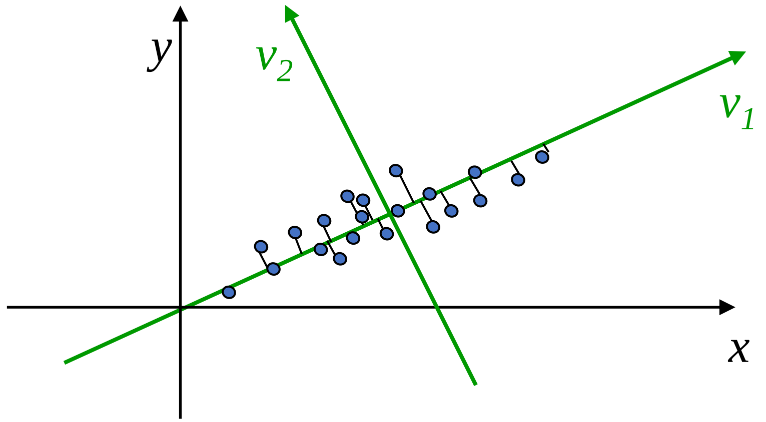


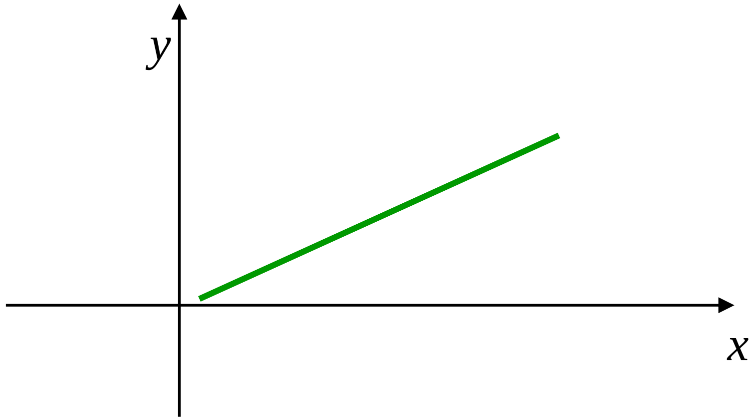
❖ Eigenvectors that correspond to big eigenvalues are the directions in which the data has strong components (= large variance).

❖ If the eigenvalues are more or less the same – there is no preferable direction.

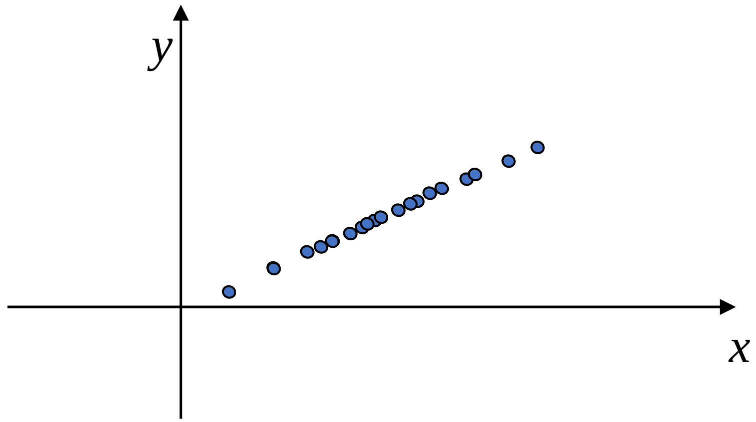
❖The sum of distances^2 from the x' axis is minimized.

PCA – the general idea





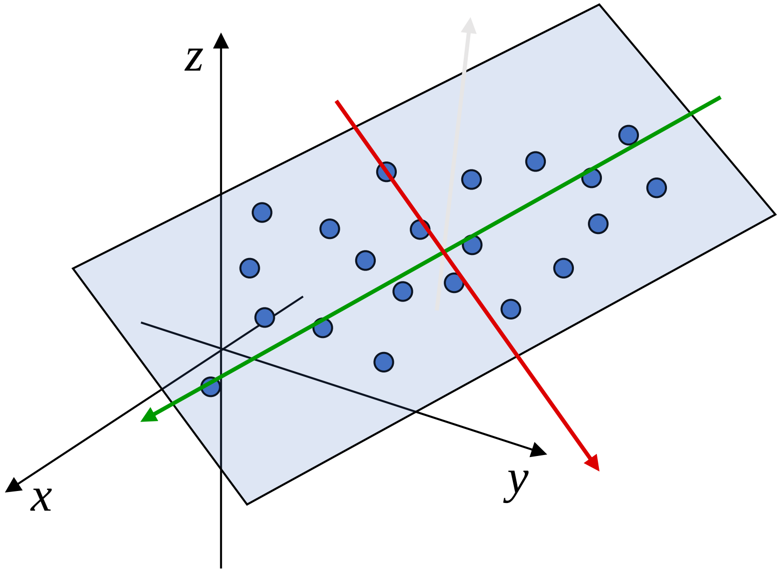
This line segment approximates the original data set



The projected data set approximates the original data set

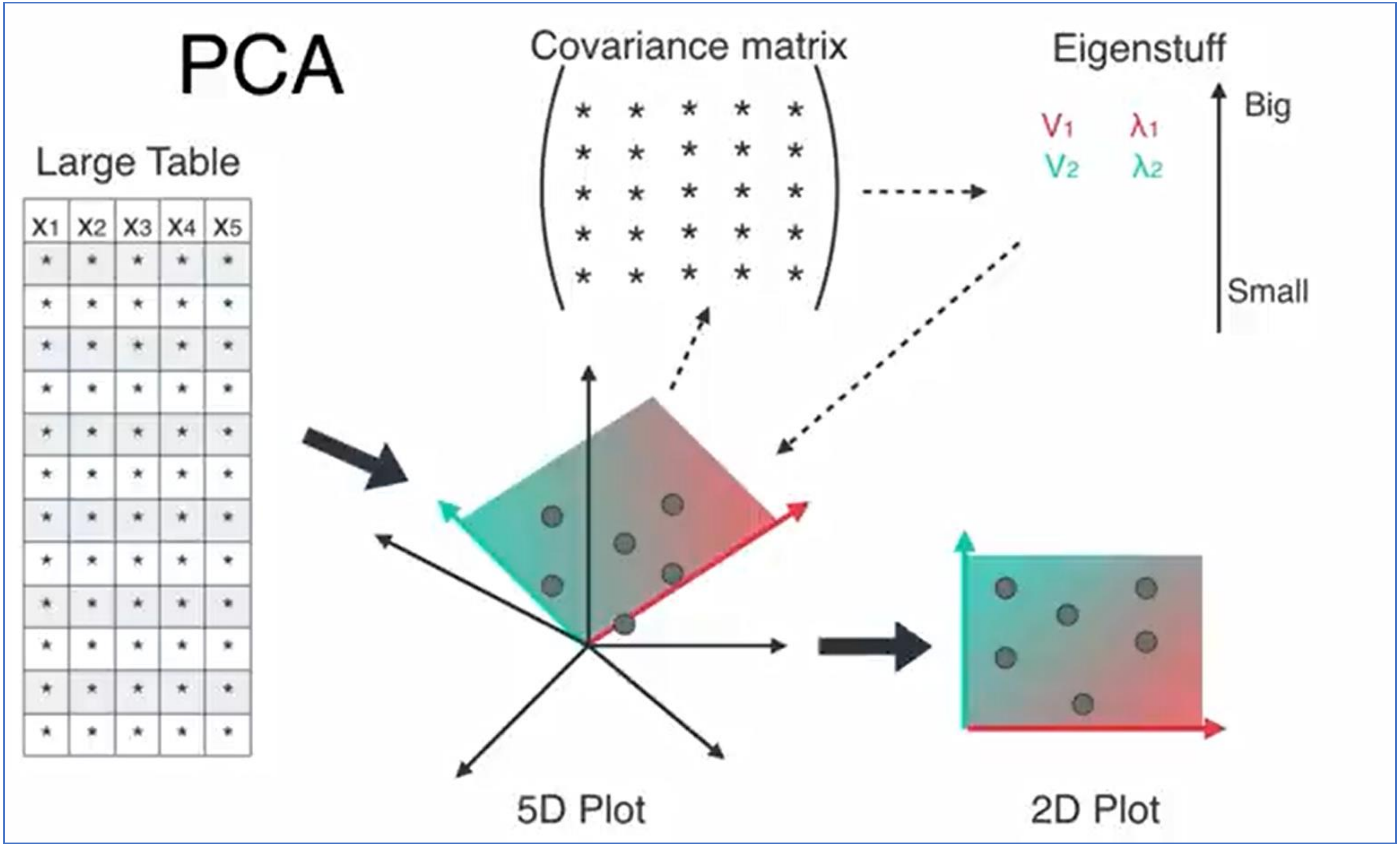
PCA – the general idea

❖PCA finds an orthogonal basis that best represents given data set.



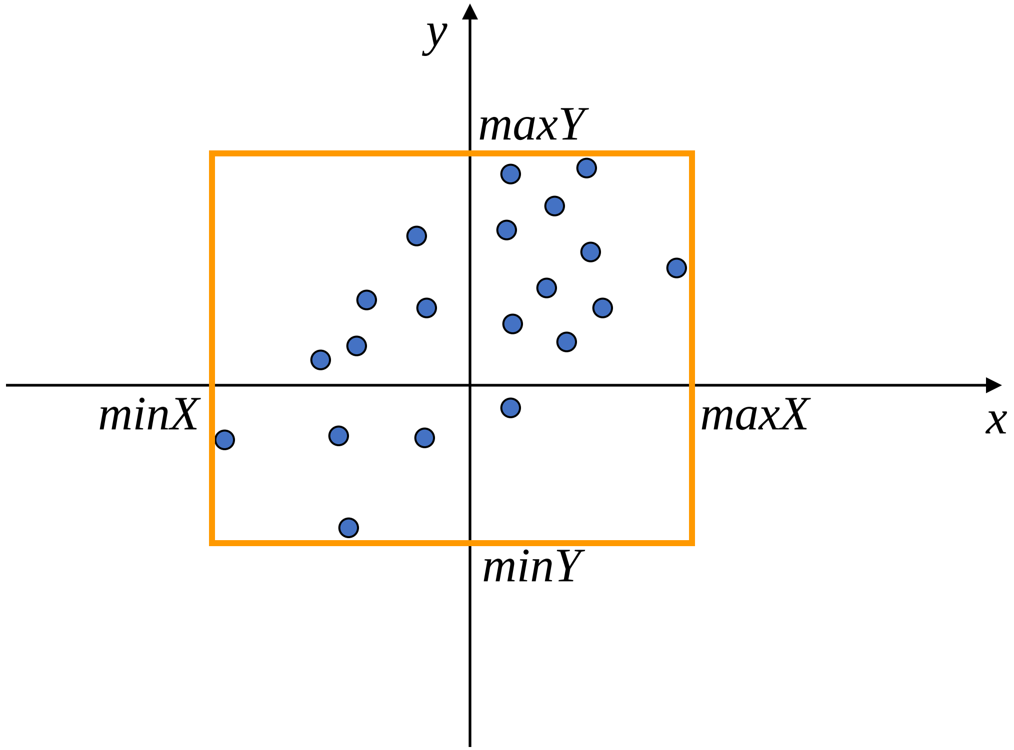
3D point set in standard basis

❖PCA finds a best approximating plane (again, in terms of $\Sigma distances^2$)


PCA
Large Table
X1 X2 X3 X4 X5
Covariance matrix
Eigenstuff
V1 λ1
V2 λ2
Big
Small
5D Plot
2D Plot

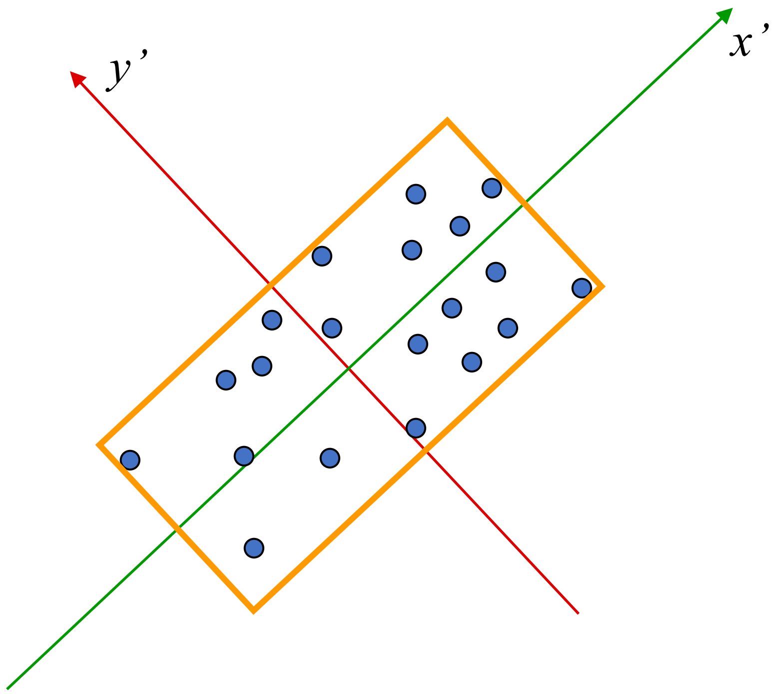
Application: finding tight bounding box

❖ An axis-aligned bounding box: agrees with the axes



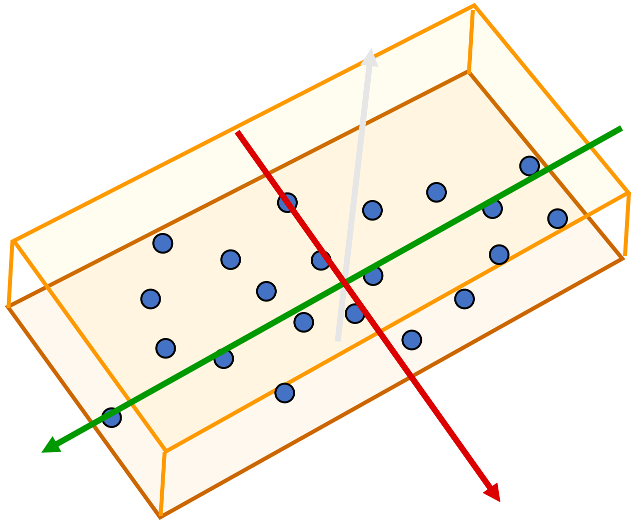
Application: finding tight bounding box

❖ Oriented bounding box: we find better axes!



Application: finding tight bounding box

- ❖ Oriented bounding box: we find better axes!



References and Credits

- Principal Component Analysis (PCA) by Luis Serrano
- StatQuest: Principal Component Analysis (PCA), Step-by-Step
- Principal Component Analysis (PCA) - easy and practical explanation
- https://www.linkedin.com/advice/0/how-can-you-use-eigenvalues-eigenvectors-improve-image-eid9f
- MATH 3191: Example Singular Value Decomposition for 3 x 2 Matrix
- https://byjus.com/maths/singular-value-decomposition/