Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
IHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

# Principal Component Analysis

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

**Unsupervised learning**, so far, focused on clustering techniques, which seek to “discover” labels on feature data that has no historical labels.

Let’s now talk about unsupervised algorithms that focus on **dimension reduction**.

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

Motivation of Dimension Reduction:

- Imagine a dataset with 30+ features
- How would you understand the key features?
- Visualization and Data Analysis have limitations when the number of feature dimensions increases.

iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Dimensionality Reduction Outcomes:

- Understand which features describe the most variance in the data set.
- Aid human understanding of large feature sets, especially through visualization.

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

Dimensionality Reduction algorithms such as PCA **do not** simply choose a subset of the existing features.

iHUB DivyaSampark, IIT Roorkee
and

They create **new** dimensional components that are combinations of proportions of the existing features.

Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited

# Theory and Intuition - Part One

Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

Principal Component Analysis Outcomes:

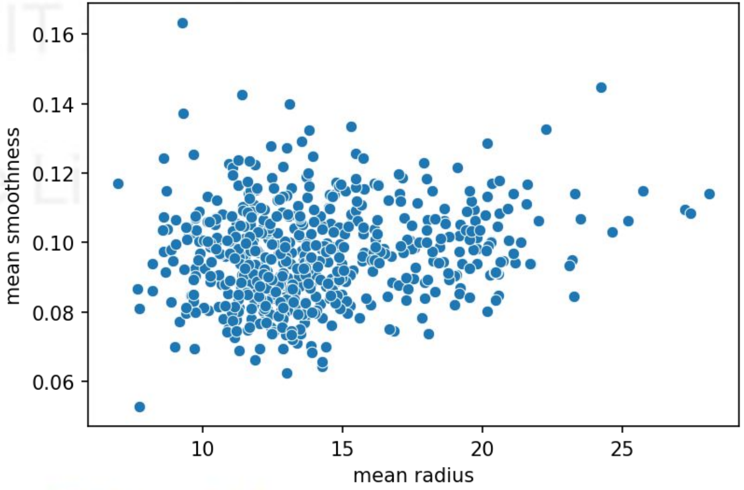
- Reduce number of dimensions in data.
- Show which features explain the most variance in the data.

iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

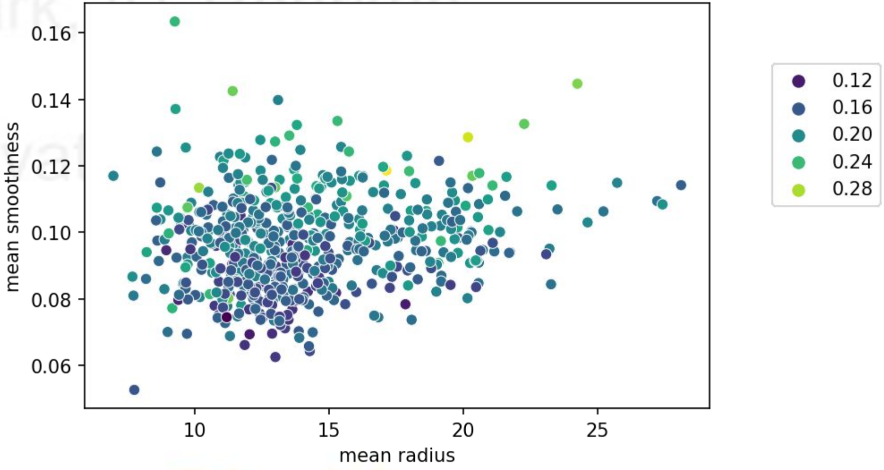
# Let’s talk about Dimension Reduction

| | mean radius | mean smoothness |
|---|---|---|
| **0** | 17.99 | 0.11840 |
| **1** | 20.57 | 0.08474 |
| **2** | 19.69 | 0.10960 |
| **3** | 11.42 | 0.14250 |
| **4** | 20.29 | 0.10030 |
| **...** | ... | ... |
| **564** | 21.56 | 0.11100 |
| **565** | 20.13 | 0.09780 |
| **566** | 16.60 | 0.08455 |
| **567** | 20.60 | 0.11780 |
| **568** | 7.76 | 0.05263 |



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

| | mean radius | mean smoothness | mean symmetry |
|---|---|---|---|
| **0** | 17.99 | 0.11840 | 0.2419 |
| **1** | 20.57 | 0.08474 | 0.1812 |
| **2** | 19.69 | 0.10960 | 0.2069 |
| **3** | 11.42 | 0.14250 | 0.2597 |
| **4** | 20.29 | 0.10030 | 0.1809 |
| **...** | ... | ... | ... |
| **564** | 21.56 | 0.11100 | 0.1726 |
| **565** | 20.13 | 0.09780 | 0.1752 |
| **566** | 16.60 | 0.08455 | 0.1590 |
| **567** | 20.60 | 0.11780 | 0.2397 |
| **568** | 7.76 | 0.05263 | 0.1587 |

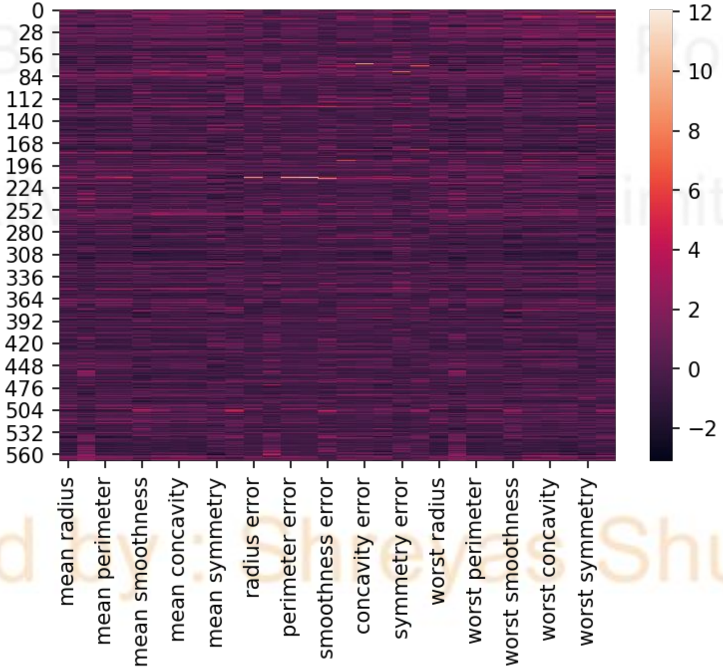



Sampark, IIT Roorkee
and
Bharat Private
Shreyas Shukla

| | mean radius | mean texture | mean perimeter | mean area | mean smoothness | mean compactness | mean concavity | mean concave points | mean symmetry | mean fractal dimension | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 17.99 | 10.38 | 122.80 | 1001.0 | 0.11840 | 0.27760 | 0.30010 | 0.14710 | 0.2419 | 0.07871 | ... |
| **1** | 20.57 | 17.77 | 132.90 | 1326.0 | 0.08474 | 0.07864 | 0.08690 | 0.07017 | 0.1812 | 0.05667 | ... |
| **2** | 19.69 | 21.25 | 130.00 | 1203.0 | 0.10960 | 0.15990 | 0.19740 | 0.12790 | 0.2069 | 0.05999 | ... |
| **3** | 11.42 | 20.38 | 77.58 | 386.1 | 0.14250 | 0.28390 | 0.24140 | 0.10520 | 0.2597 | 0.09744 | ... |
| **4** | 20.29 | 14.34 | 135.10 | 1297.0 | 0.10030 | 0.13280 | 0.19800 | 0.10430 | 0.1809 | 0.05883 | ... |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **564** | 21.56 | 22.39 | 142.00 | 1479.0 | 0.11100 | 0.11590 | 0.24390 | 0.13890 | 0.1726 | 0.05623 | ... |
| **565** | 20.13 | 28.25 | 131.20 | 1261.0 | 0.09780 | 0.10340 | 0.14400 | 0.09791 | 0.1752 | 0.05533 | ... |
| **566** | 16.60 | 28.08 | 108.30 | 858.1 | 0.08455 | 0.10230 | 0.09251 | 0.05302 | 0.1590 | 0.05648 | ... |
| **567** | 20.60 | 29.33 | 140.10 | 1265.0 | 0.11780 | 0.27700 | 0.35140 | 0.15200 | 0.2397 | 0.07016 | ... |
| **568** | 7.76 | 24.54 | 47.92 | 181.0 | 0.05263 | 0.04362 | 0.00000 | 0.00000 | 0.1587 | 0.05884 | ... |

569 rows × 30 columns

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

Dimension Reduction

- Helps visualize and understand complex data sets.
- Can also act as a simpler data set for training data for machine learning algorithms.
  - Reduce dimensions then train ML Algorithm on smaller data set.

iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Dimension Reduction

- Helps reduce N features to a desired smaller set of components through a **transformation**.
- It does **not** simply select a subset of features.

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

Variance Explained

- Often, certain features are more important or have more explanatory power than other features.
- Eg: Size of a house is probably much more important than the color of a house when explaining the price of a house for sale.

iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

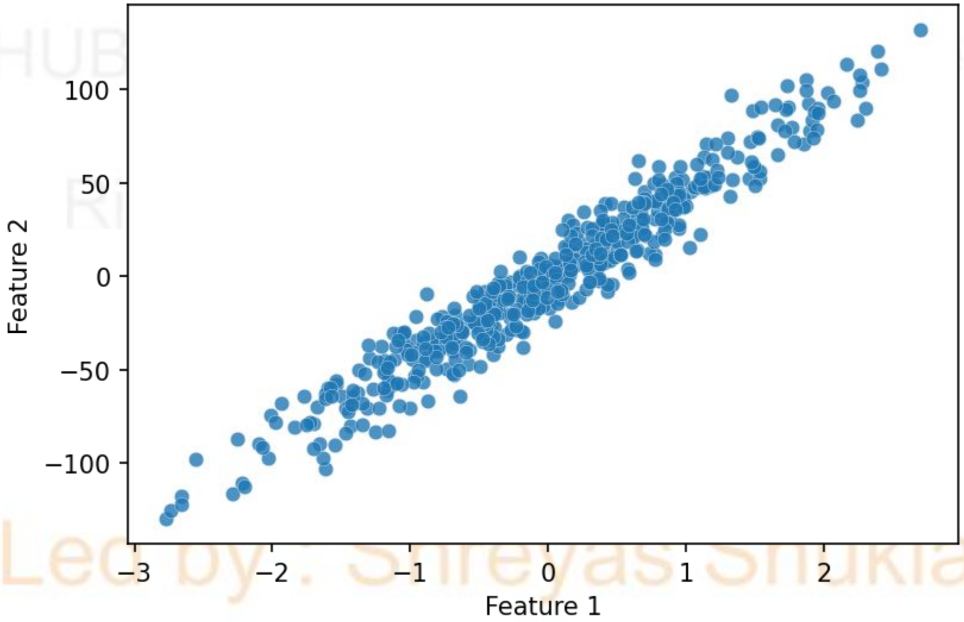
Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

- This idea of more important features is easy to understand when we can directly correlate features to a known label. But what about unlabeled data?
- What measurement can we use to determine feature “importance”?

Measure the proportion to which each feature accounts for dispersion in the data set.

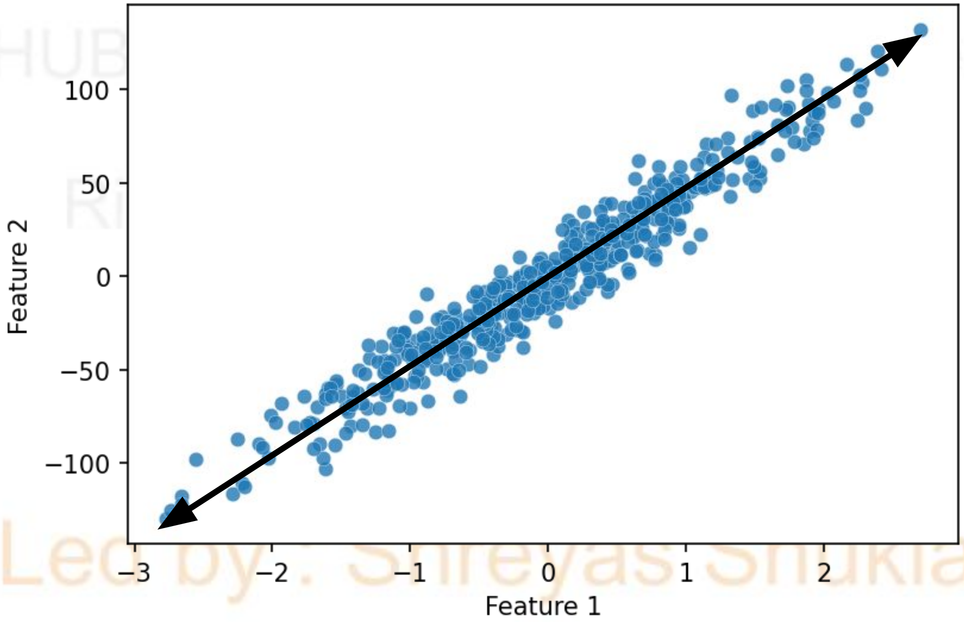
Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)



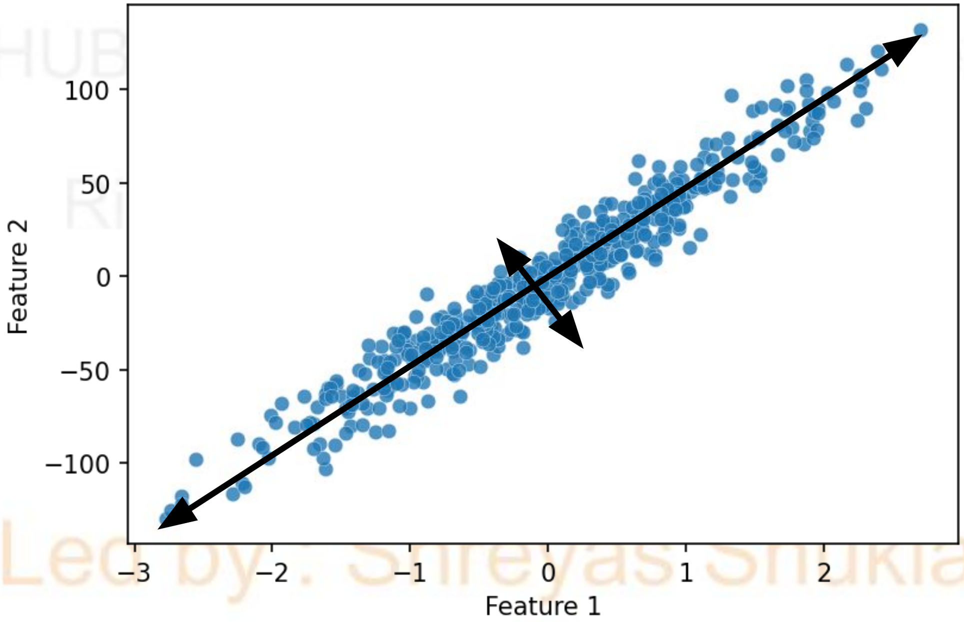

Led by: Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)



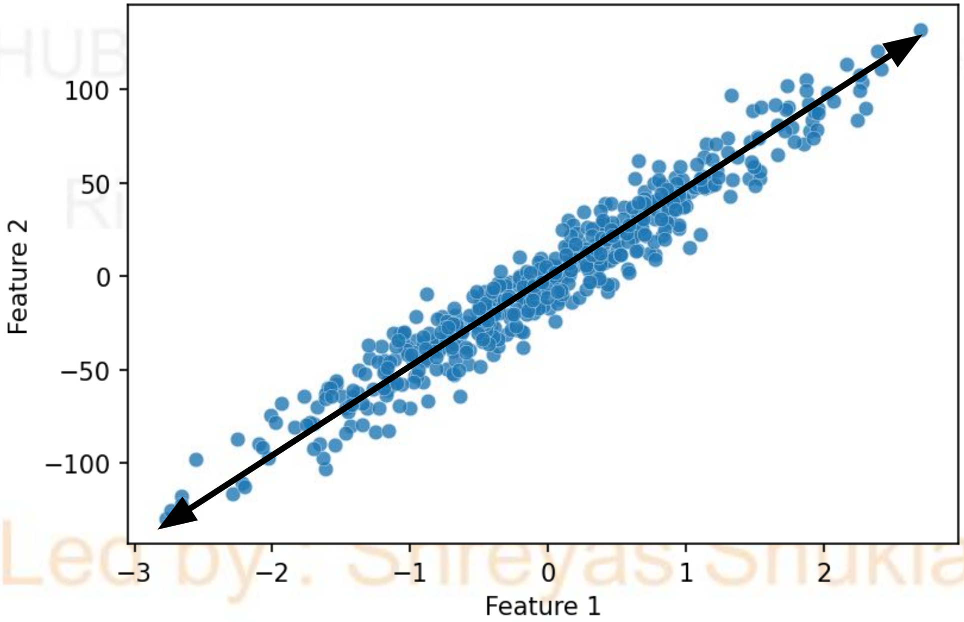

Led by: Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

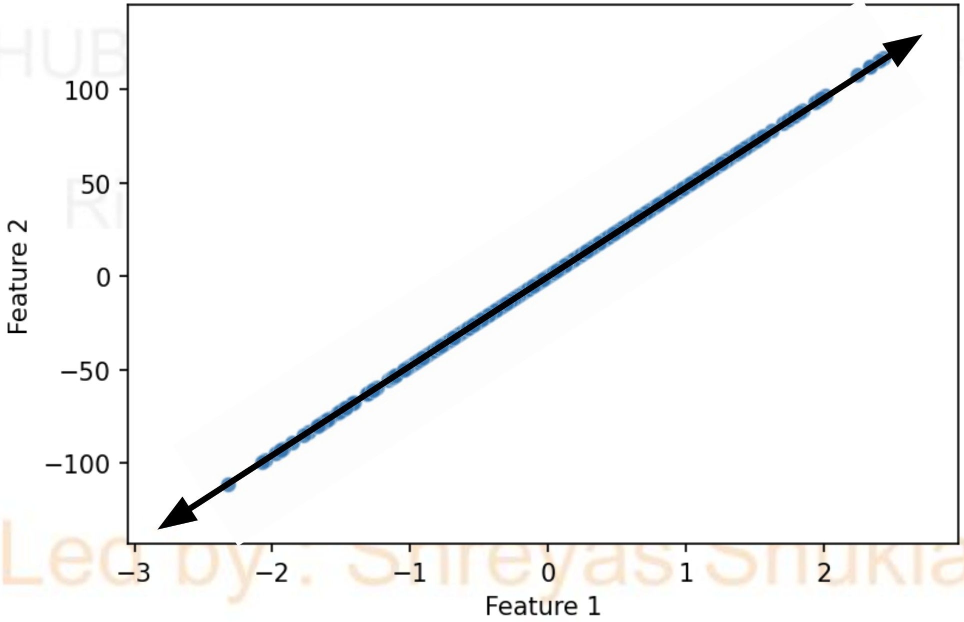



Led by: Shreyas Shukla

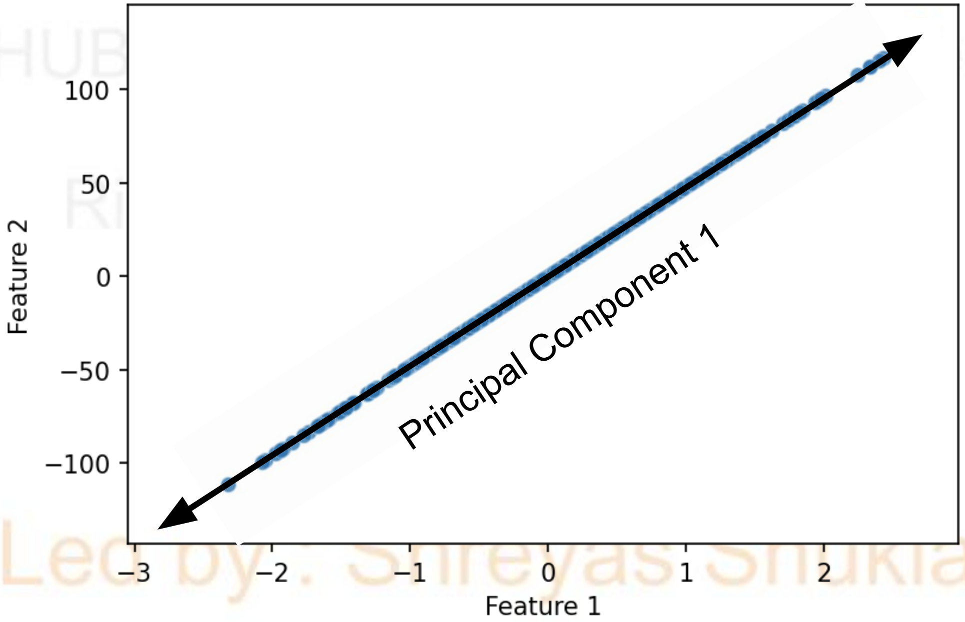
Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)



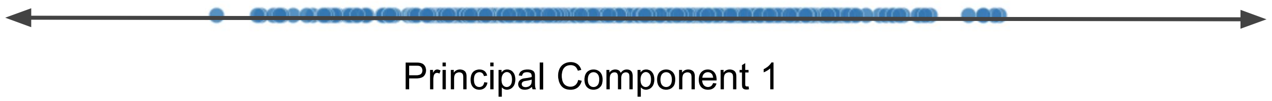
Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB
Led by : Shreyas Shukla



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

Variance Explained

- Principal Component is basically a linear combination of original features.
- The more variance the original feature accounts for, the more influence it has over the principal components.

IHUB DivyaSampark, IIT Roorkee
and
Ritwij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Rinvij Bharat Private Limited
Led by : Shreyas Shukla

Here we went from 2 features down to 1 principal component.

This single principal component can “explain” some percentage of the original data, for example 90% of variance explained by principal component.

100% of the variance in the data is explained by all the original features.

We trade off some of the explained variance for less dimensions.

This can be significant savings for data sets with many dimensions, but only a few strong features.

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

Let’s see how PCA actually works mathematically.

iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited

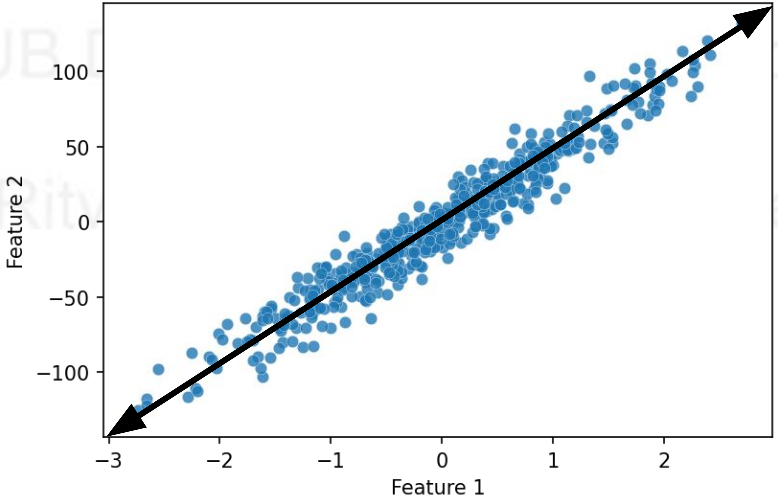
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
IHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

PCA operates by creating a new set of dimensions, that is the principal components, that are normalized linear combinations of the original features.

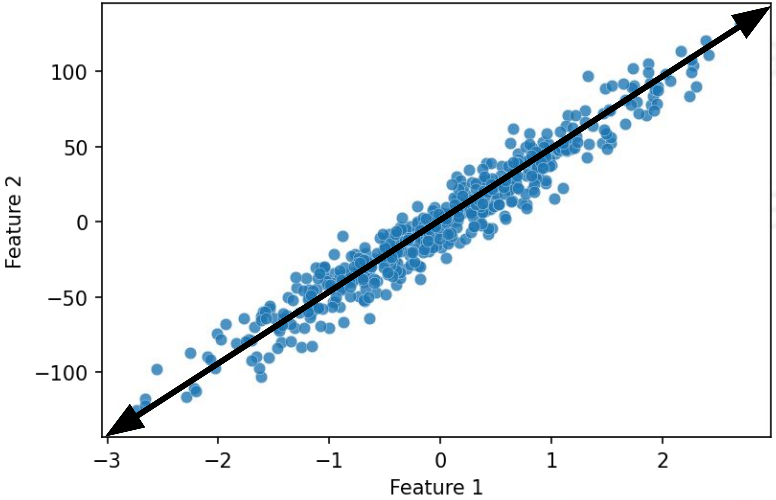
$$Z_1 = \phi_{11}X_1 + \phi_{21}X_2 + \ldots + \phi_{p1}X_p$$

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)



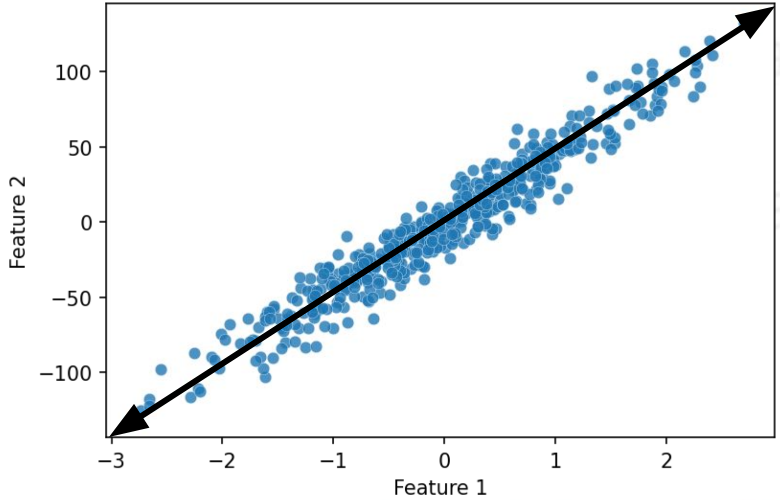

$$Z_1 = \phi_{11} X_1 + \phi_{21} X_2$$

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

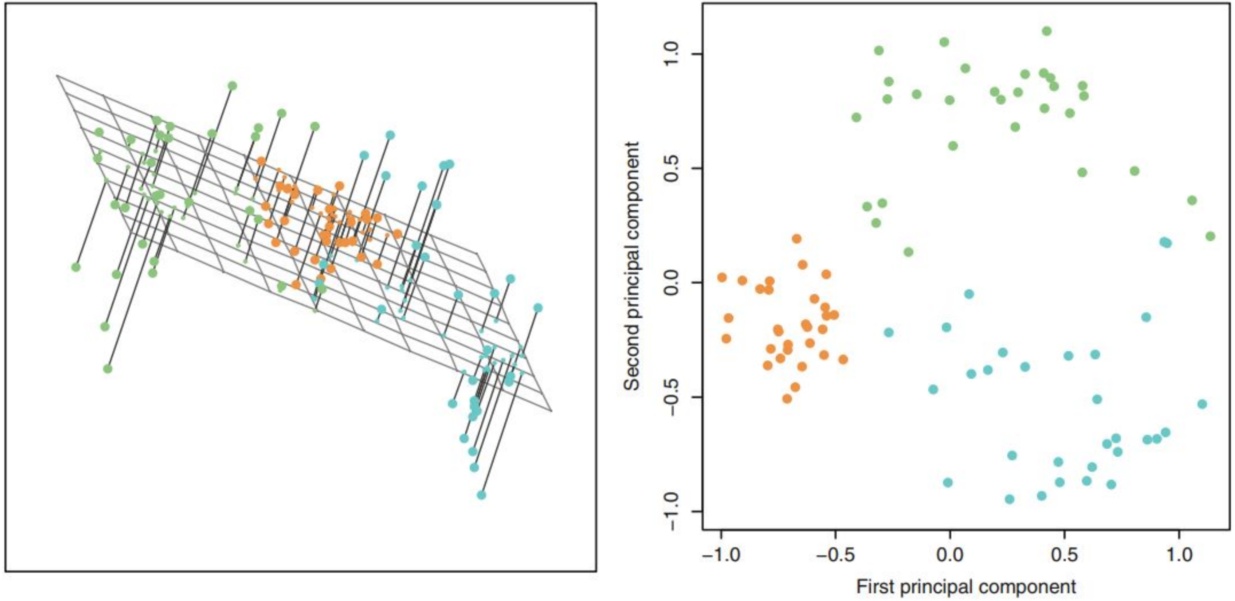



$$Z_1 = \phi_{11}X_1 + \phi_{21}X_2$$

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)




$$Z_1 = \phi_{11} X_1 + \phi_{21} X_2$$


Second principal component
First principal component
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

How do we calculate these components?

Let’s walk through the steps visually.

iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited

Led by : Shreyas Shukla

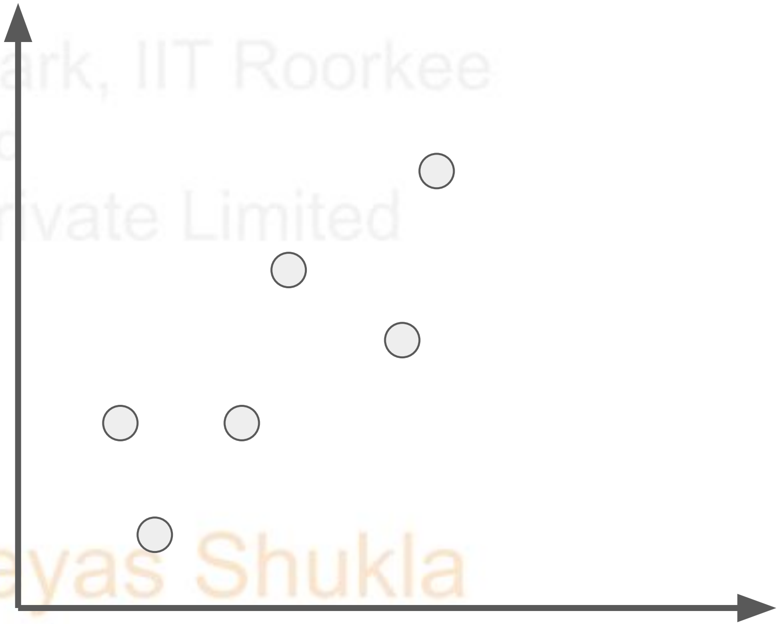
Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

| X1 | X2 |
|---|---|
| 1 | 2 |
| 2 | 1 |
| 3 | 2 |
| 4 | 4 |
| 5 | 3 |
| 6 | 5 |

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

| X1 | X2 |
|---|---|
| 1 | 2 |
| 2 | 1 |
| 3 | 2 |
| 4 | 4 |
| 5 | 3 |
| 6 | 5 |

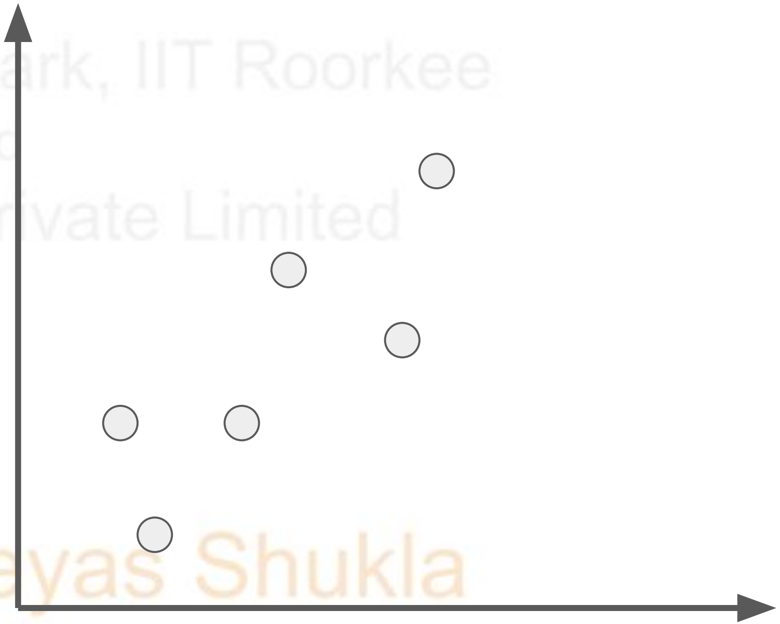
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Standardize the data:

| X1 | X2 |
|---|---|
| 1 | 2 |
| 2 | 1 |
| 3 | 2 |
| 4 | 4 |
| 5 | 3 |
| 6 | 5 |

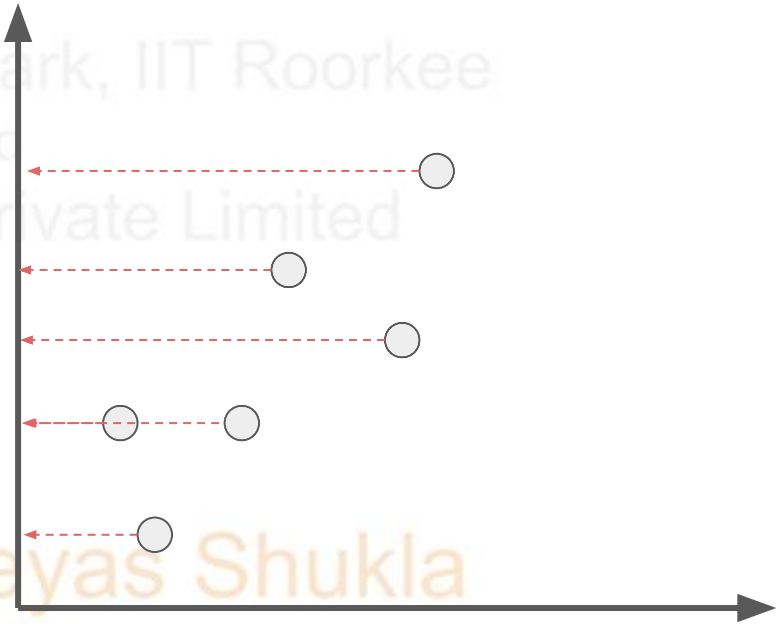


iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Standardize the data:

| X1 | X2 |
|---|---|
| 1 | 2 |
| 2 | 1 |
| 3 | 2 |
| 4 | 4 |
| 5 | 3 |
| 6 | 5 |

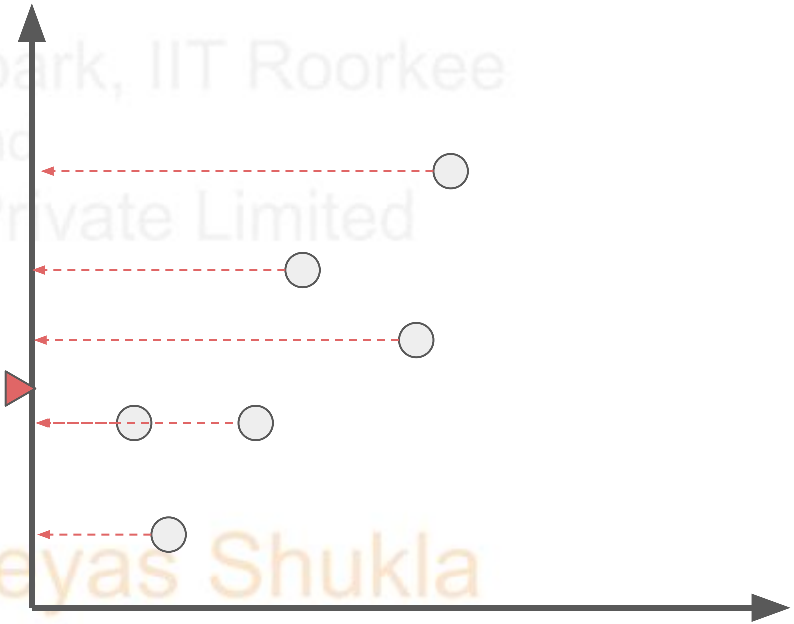


iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Standardize the data:

| X1 | X2 |
|---|---|
| 1 | 2 |
| 2 | 1 |
| 3 | 2 |
| 4 | 4 |
| 5 | 3 |
| 6 | 5 |

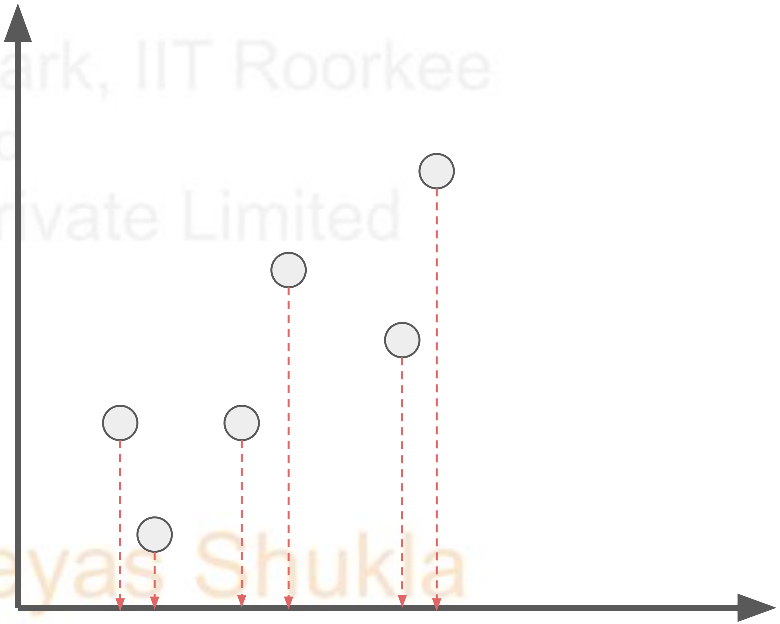


iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Standardize the data:

| X1 | X2 |
|---|---|
| 1 | 2 |
| 2 | 1 |
| 3 | 2 |
| 4 | 4 |
| 5 | 3 |
| 6 | 5 |



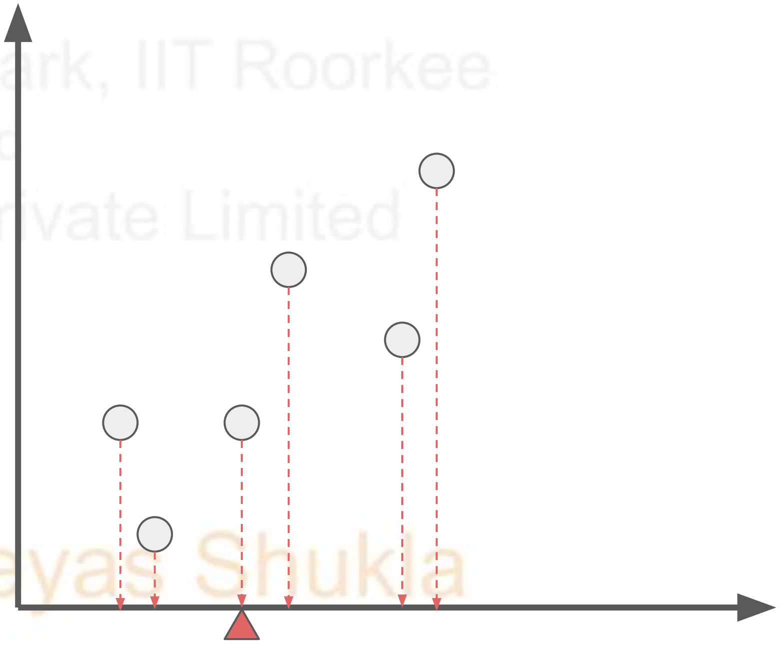
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Standardize the data:

| X1 | X2 |
|---|---|
| 1 | 2 |
| 2 | 1 |
| 3 | 2 |
| 4 | 4 |
| 5 | 3 |
| 6 | 5 |

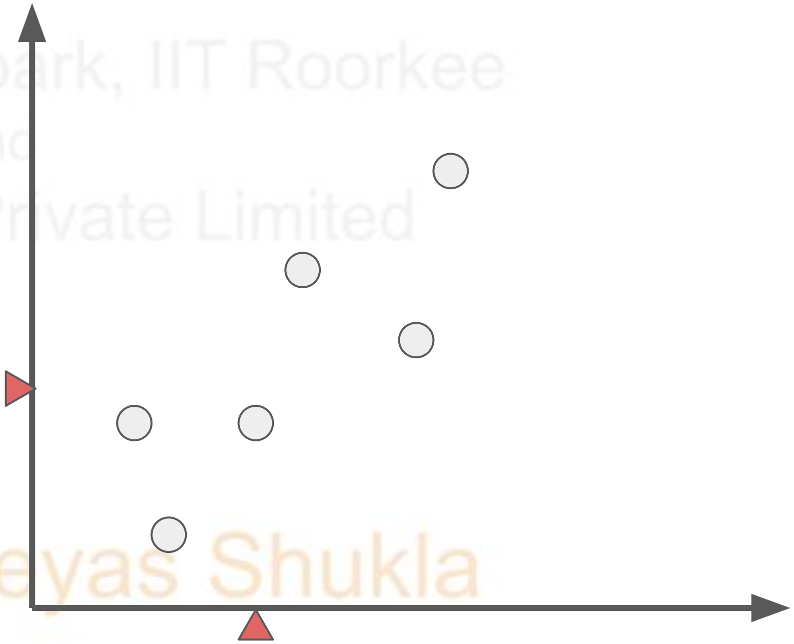
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Standardize the data:

| X1 | X2 |
|---|---|
| 1 | 2 |
| 2 | 1 |
| 3 | 2 |
| 4 | 4 |
| 5 | 3 |
| 6 | 5 |

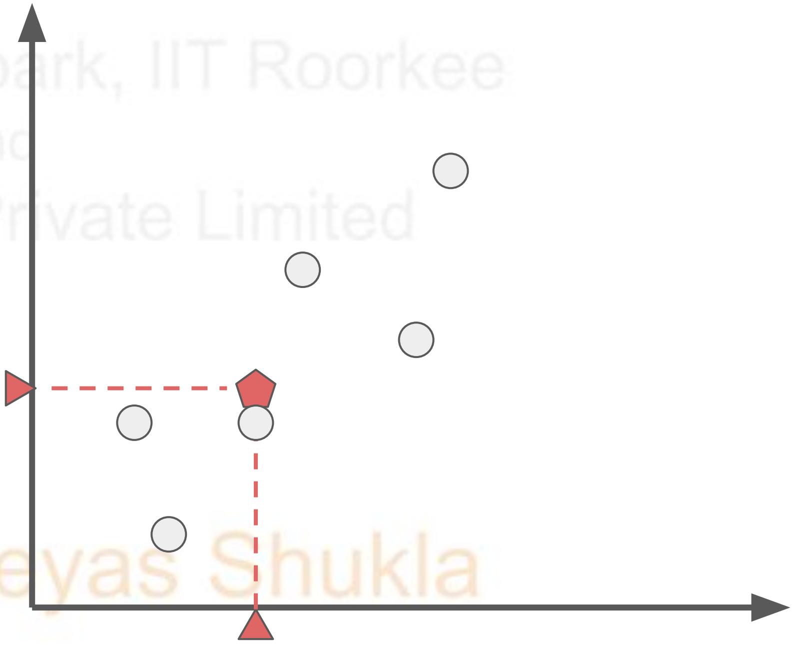


iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Standardize the data:

| X1 | X2 |
| --- | --- |
| 1 | 2 |
| 2 | 1 |
| 3 | 2 |
| 4 | 4 |
| 5 | 3 |
| 6 | 5 |

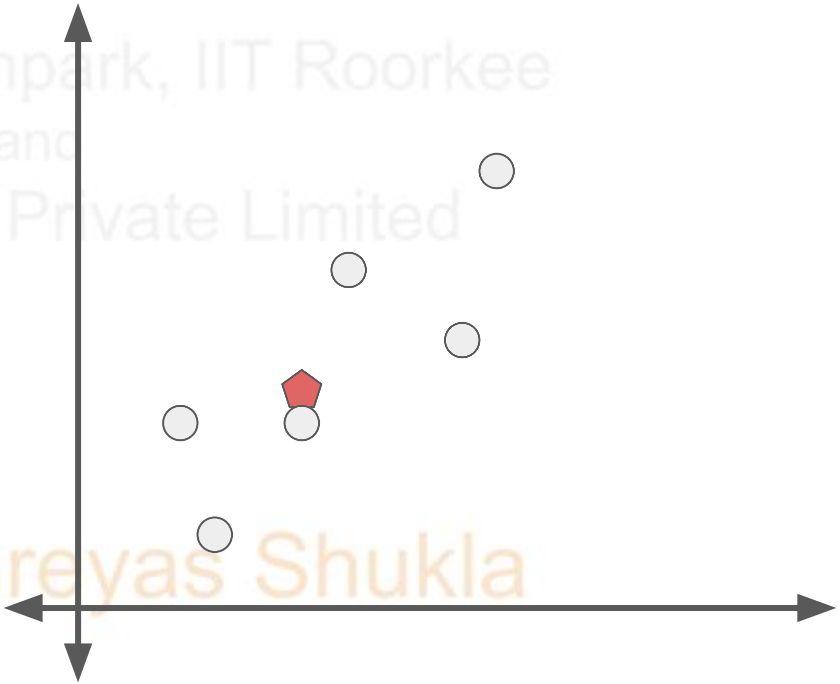


iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Standardize the data:

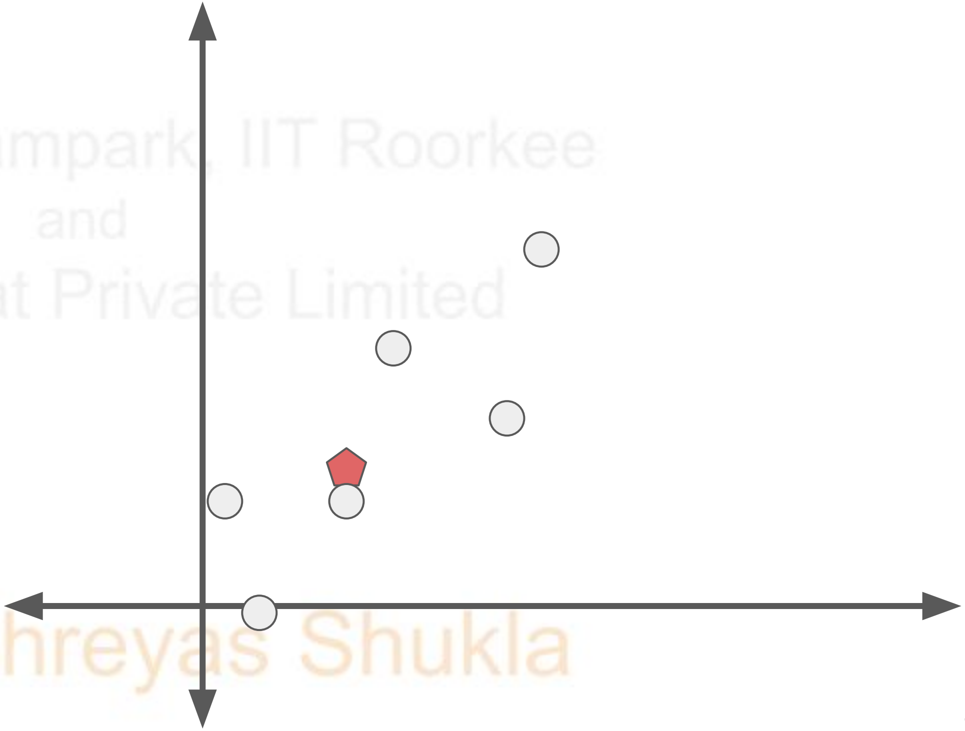
| X1 | X2 |
|---|---|
| 1 | 2 |
| 2 | 1 |
| 3 | 2 |
| 4 | 4 |
| 5 | 3 |
| 6 | 5 |



iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

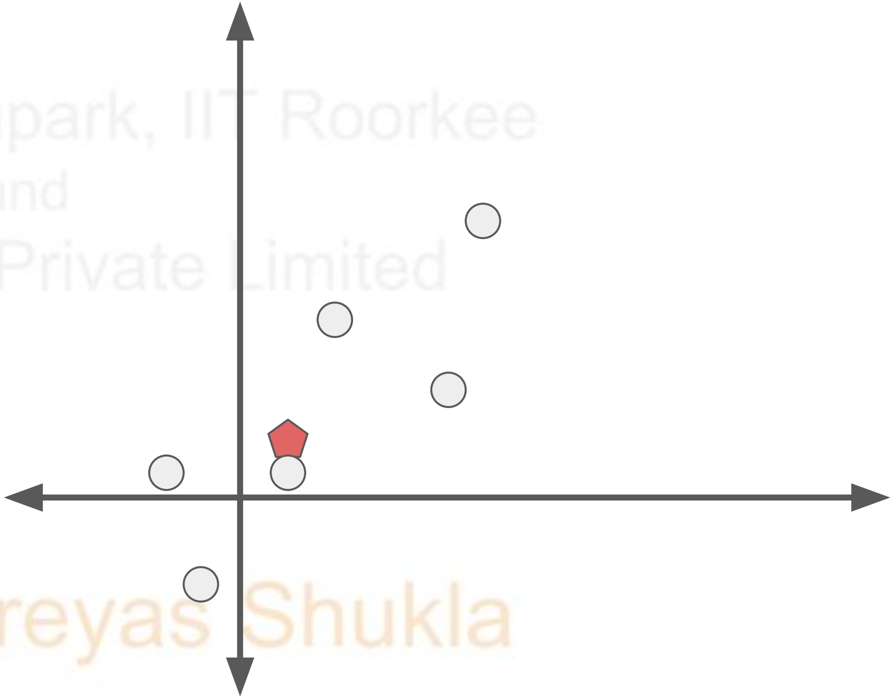
Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Standardize the data:



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

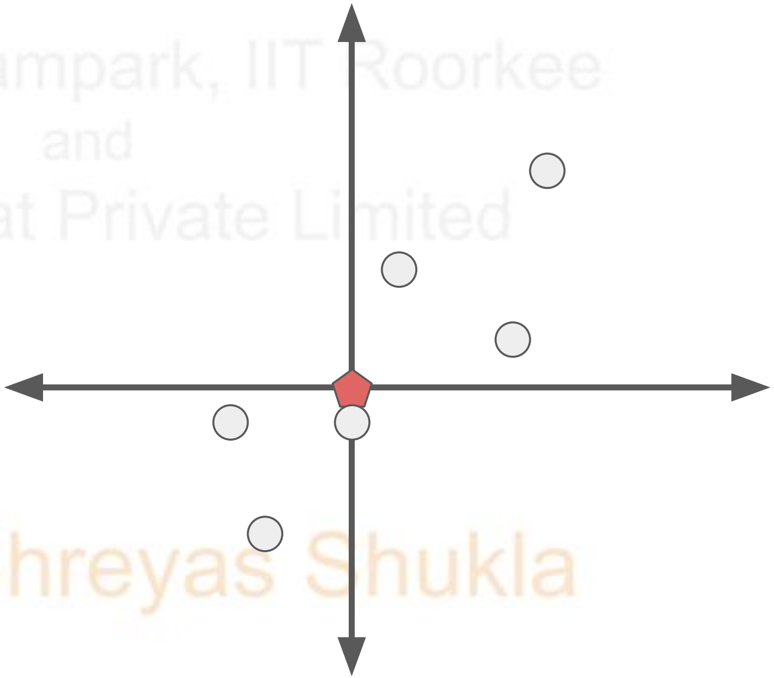
Standardize the data:



iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Standardize the data:

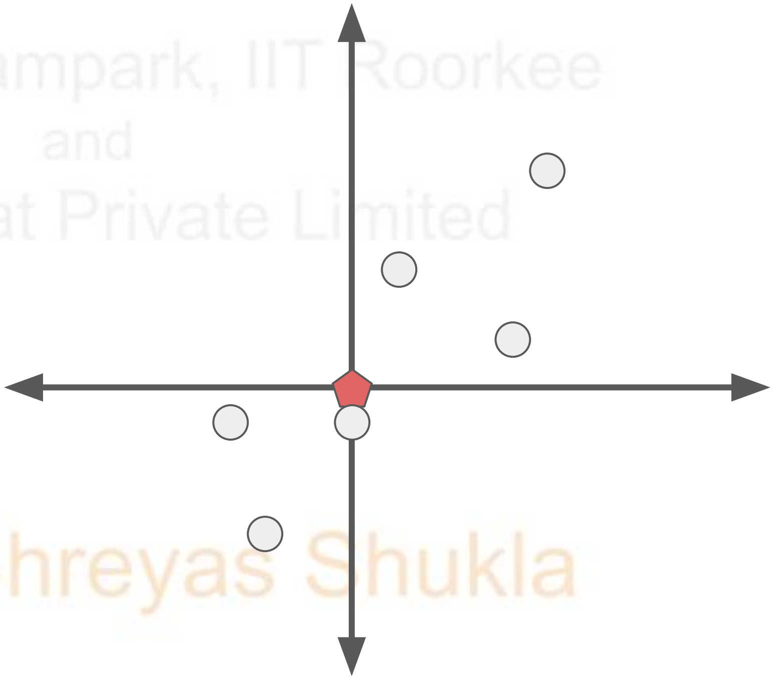


iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited

Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

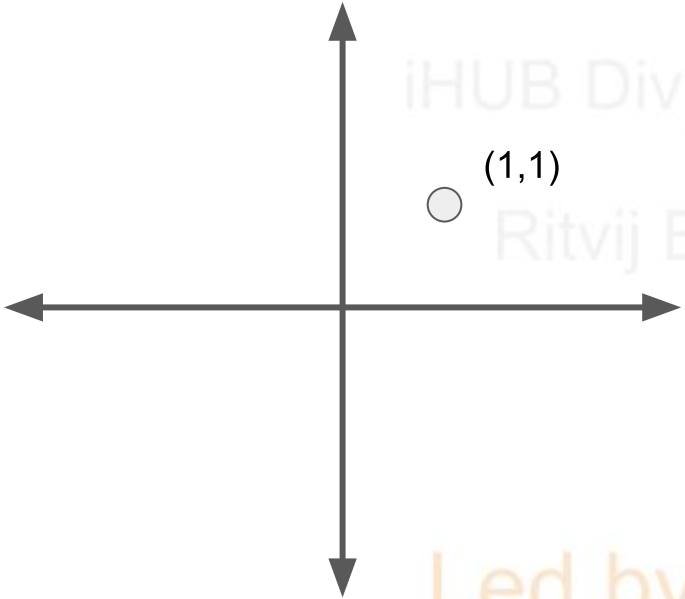
# Calculate covariance matrix for data:



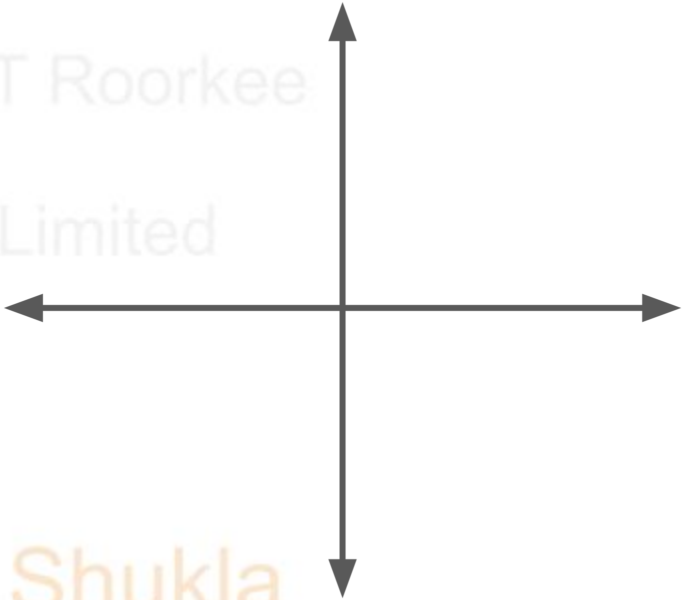
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Linear transformation of data:

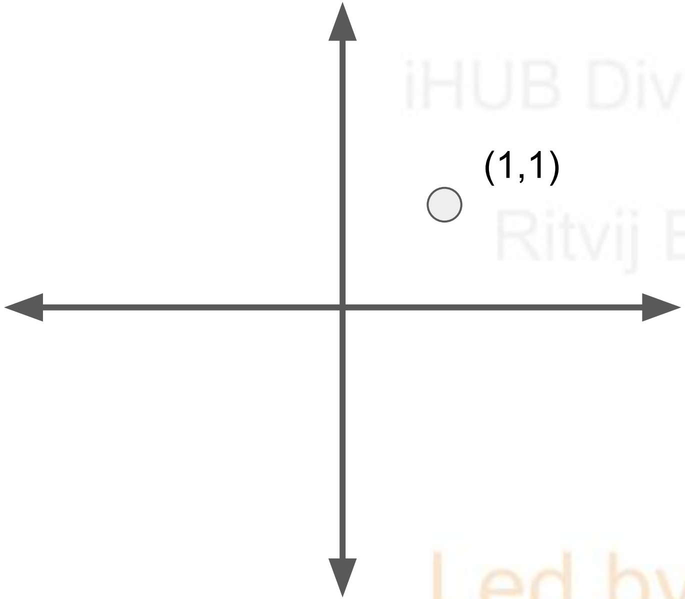



iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited

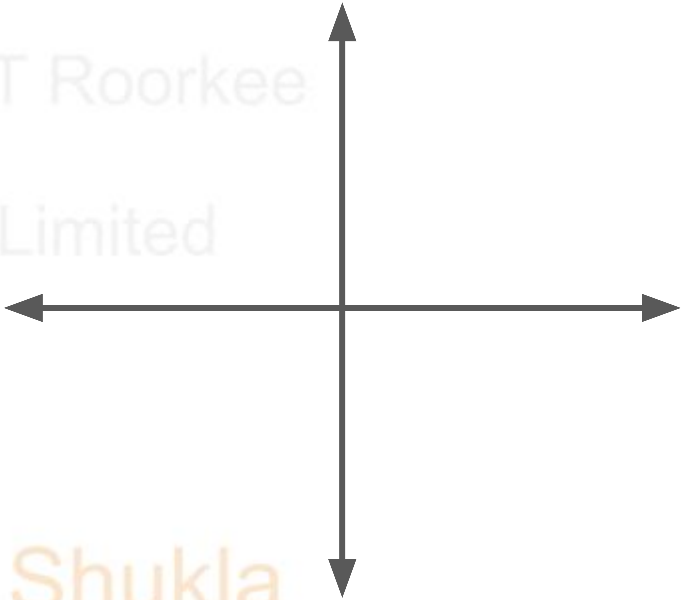


Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

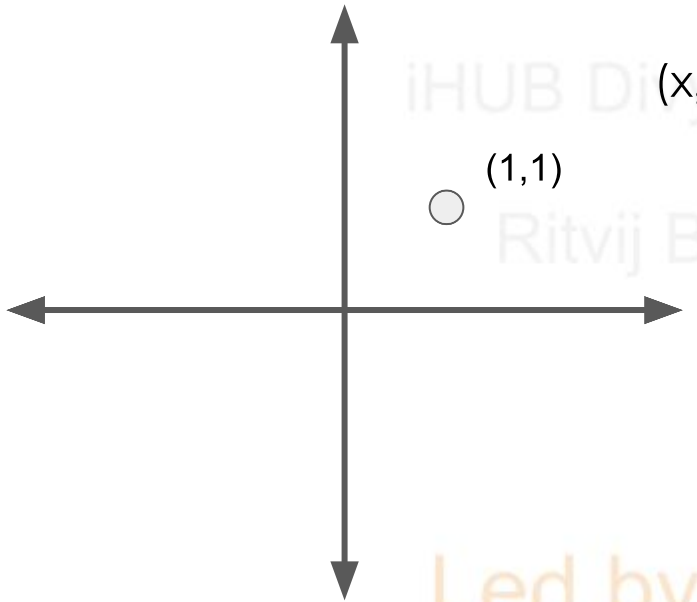



$$\begin{bmatrix} a1 & b1 \\ a2 & b2 \end{bmatrix}$$

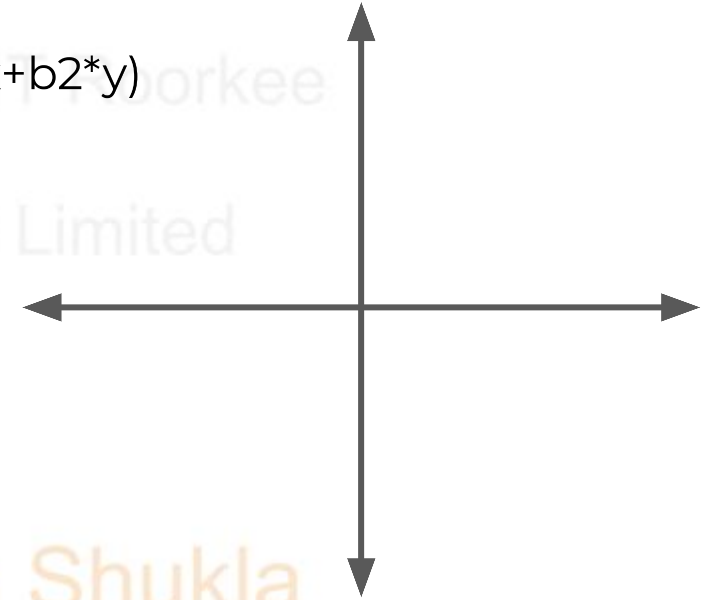


Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

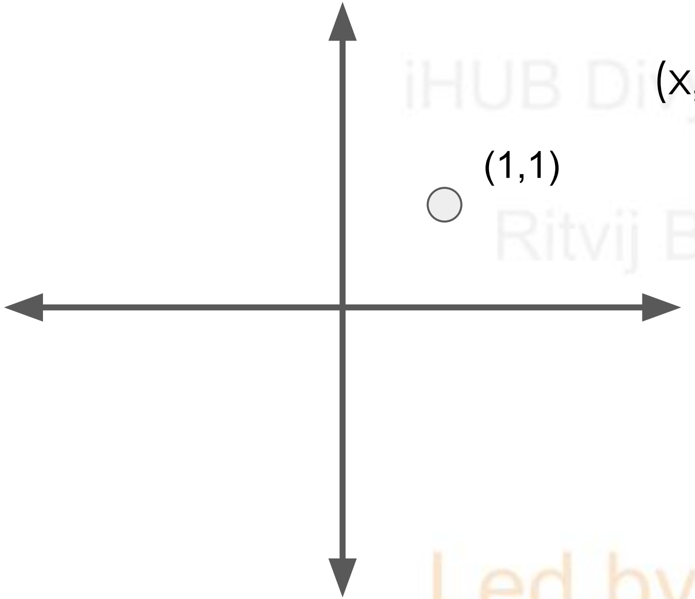
$(x,y) \rightarrow (a1*x+b1*y, a2*x+b2*y)$



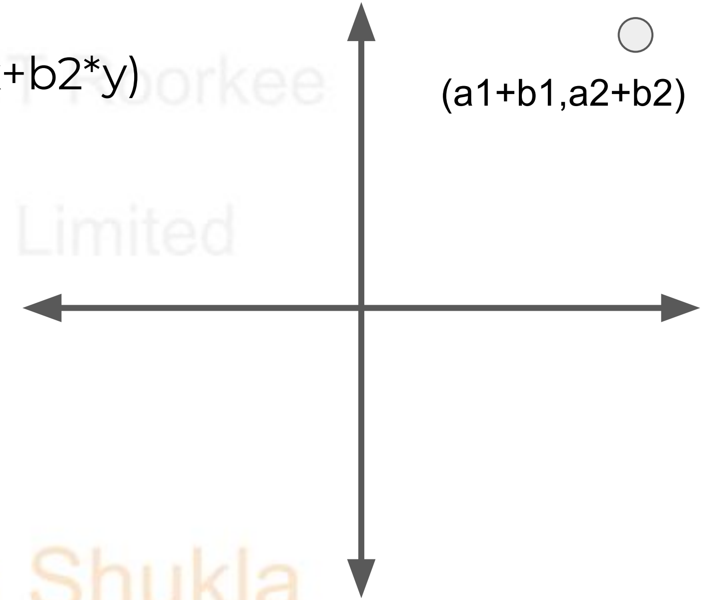

$$\begin{bmatrix} a1 & b1 \\ a2 & b2 \end{bmatrix}$$



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

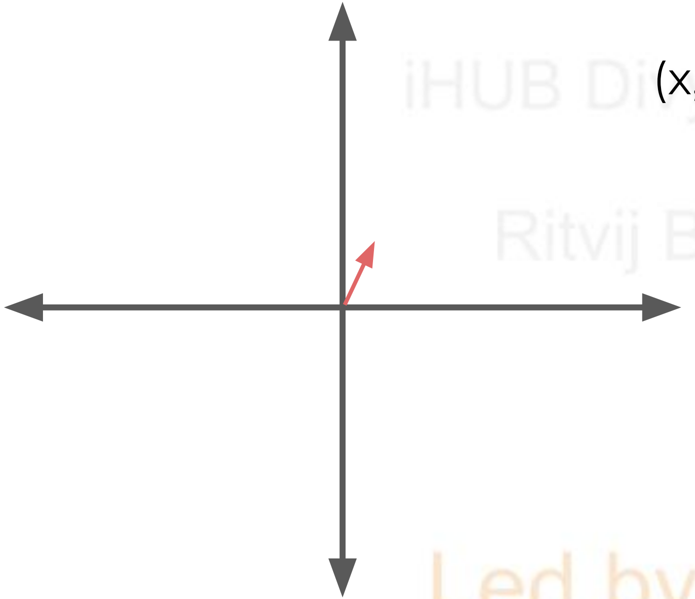



$(x,y) \rightarrow (a1*x+b1*y, a2*x+b2*y)$



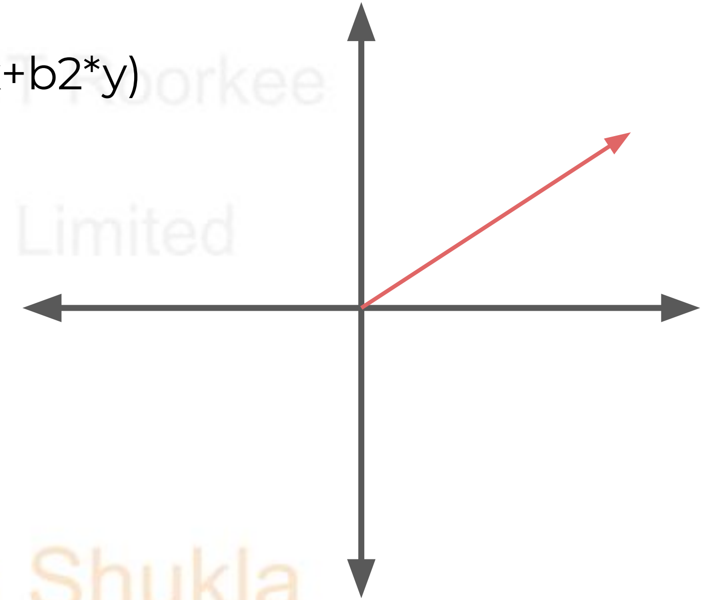

$$\begin{bmatrix} a1 & b1 \\ a2 & b2 \end{bmatrix}$$

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

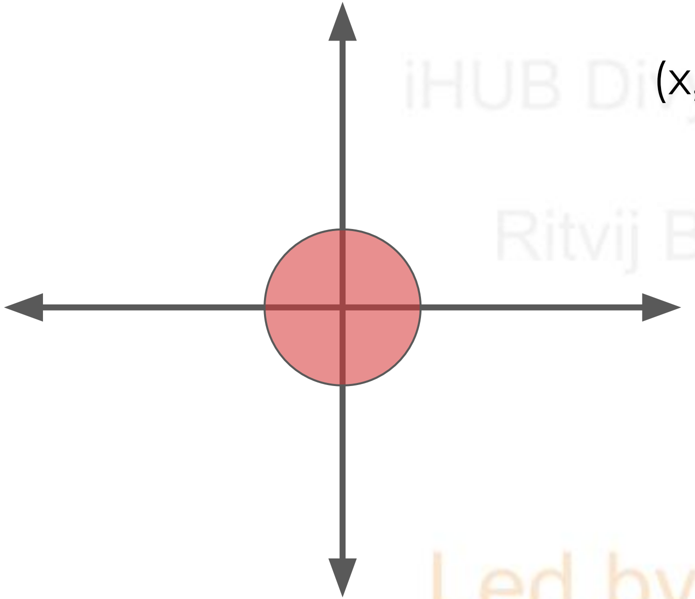


$$(x,y) \rightarrow (a1*x+b1*y, a2*x+b2*y)$$

$$\begin{bmatrix} a1 & b1 \\ a2 & b2 \end{bmatrix}$$

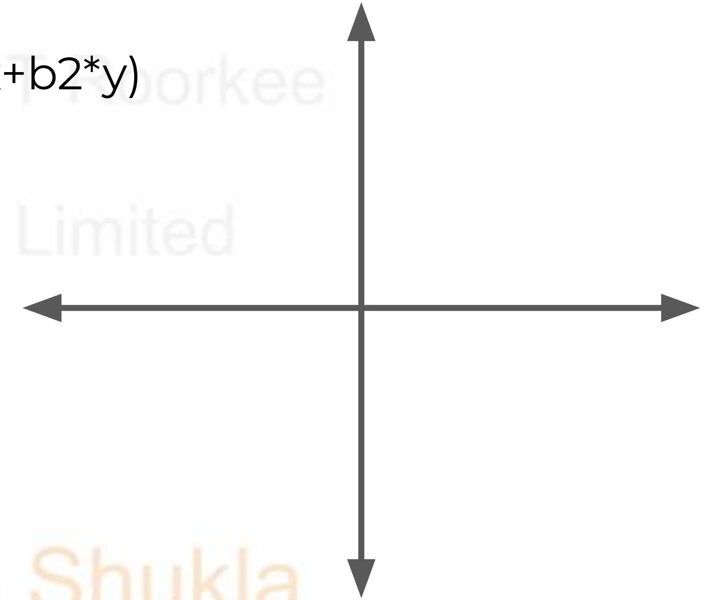


Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla



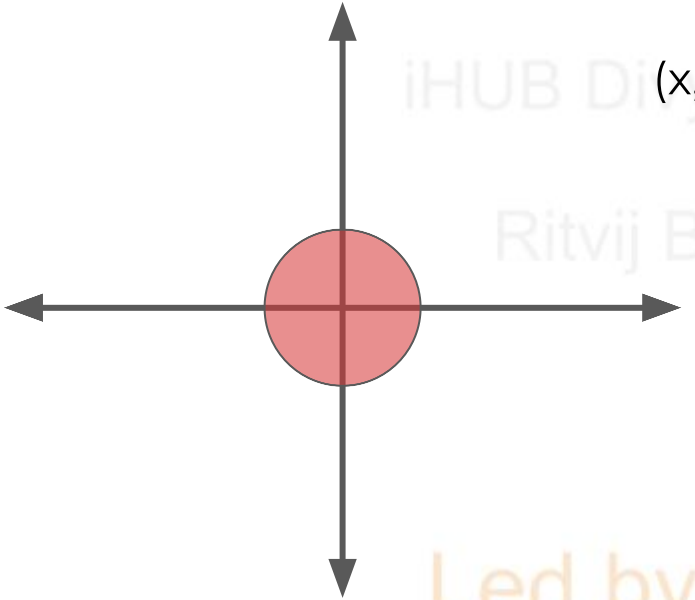
$$(x,y) \rightarrow (a1*x+b1*y, a2*x+b2*y)$$

$$\begin{bmatrix} a1 & b1 \\ a2 & b2 \end{bmatrix}$$

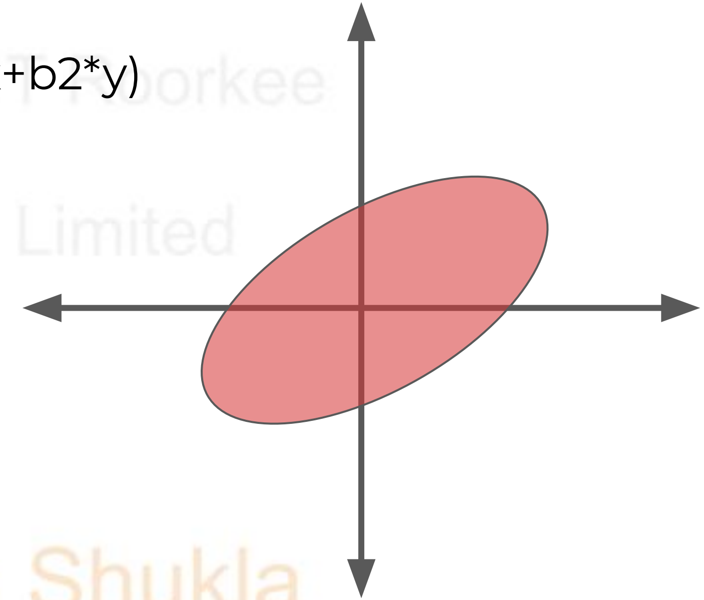


Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

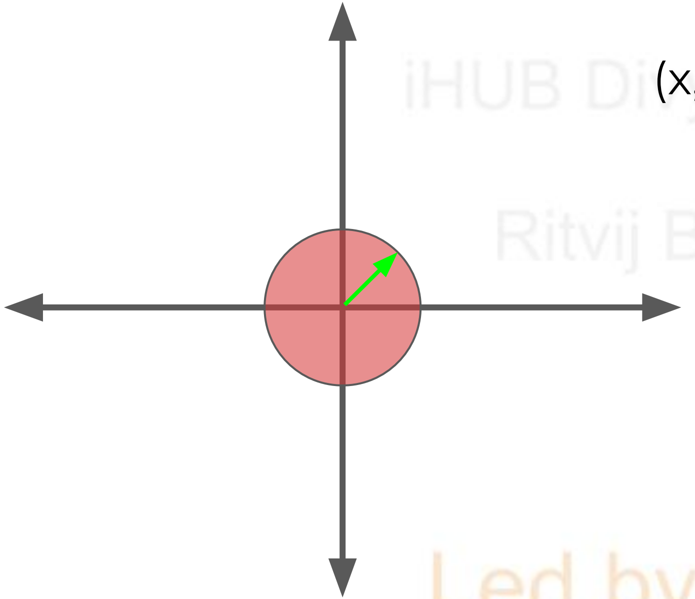
$$(x,y) \rightarrow (a1*x+b1*y, a2*x+b2*y)$$



$$\begin{bmatrix} a1 & b1 \\ a2 & b2 \end{bmatrix}$$

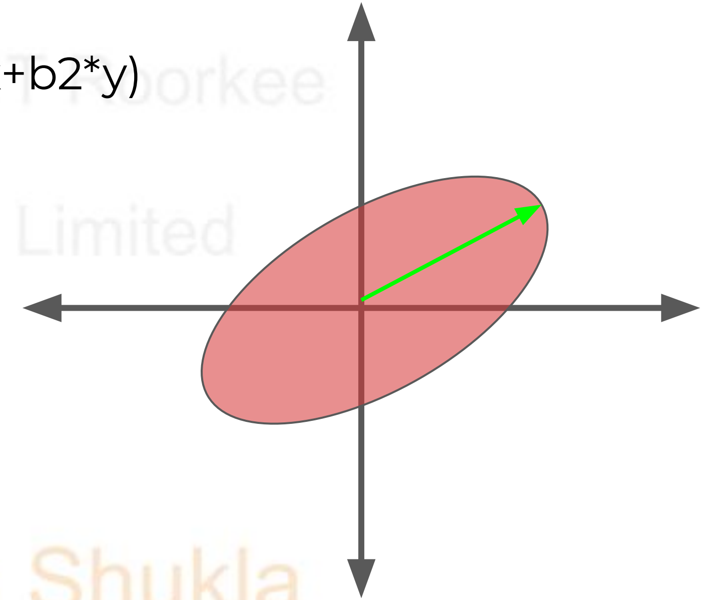


Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla



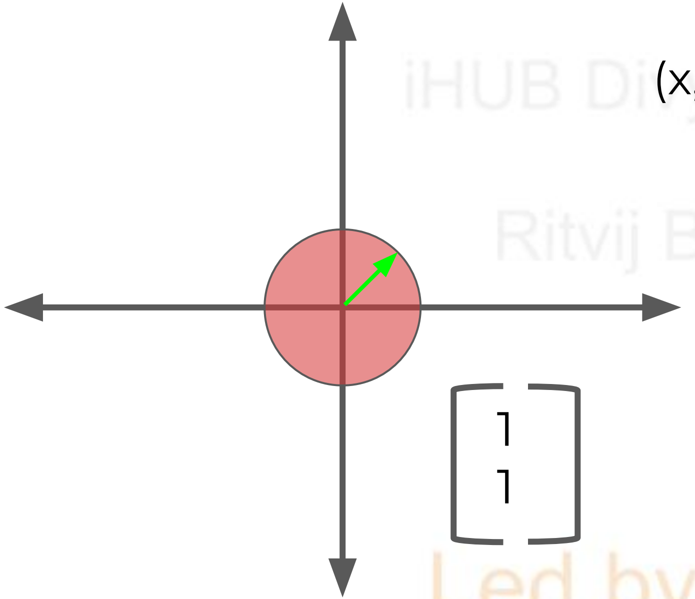
$(x,y) \rightarrow (a1*x+b1*y, a2*x+b2*y)$

$$\begin{bmatrix} a1 & b1 \\ a2 & b2 \end{bmatrix}$$



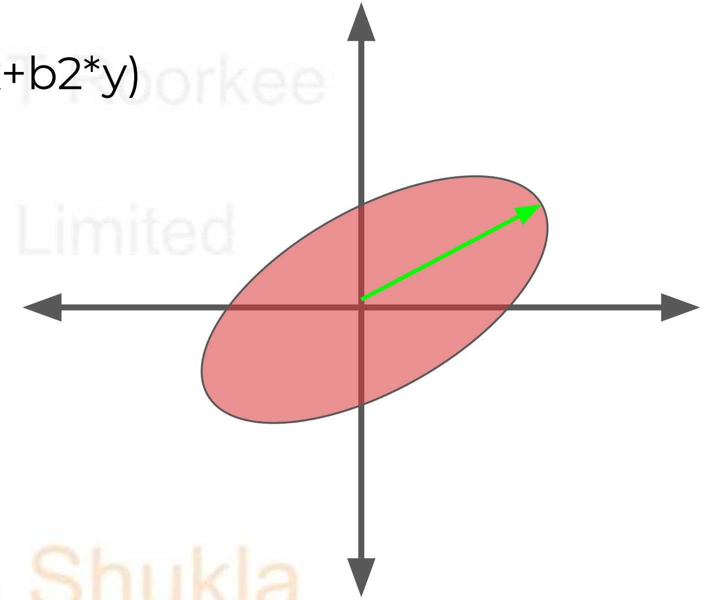
Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

# EigenVector: Directional Information



$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

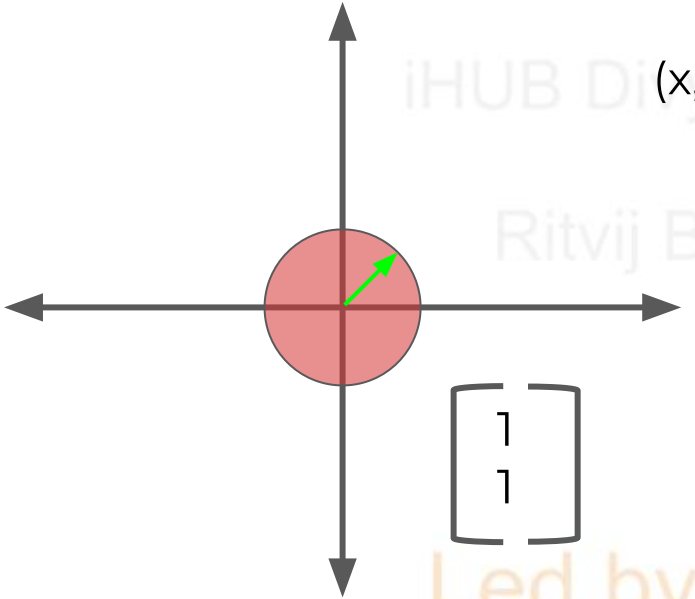
(x,y) → (a1*x+b1*y,a2*x+b2*y)



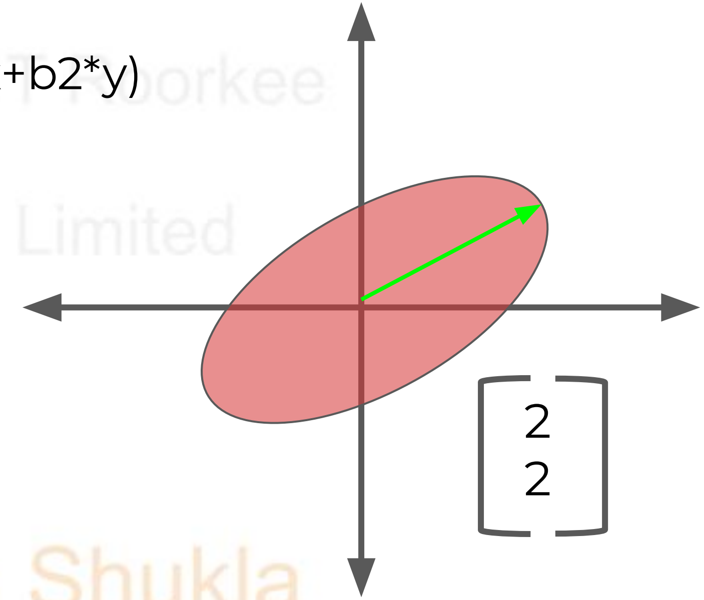
Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

# EigenVector: Directional Information

(x,y) → (a1*x+b1*y,a2*x+b2*y)



$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

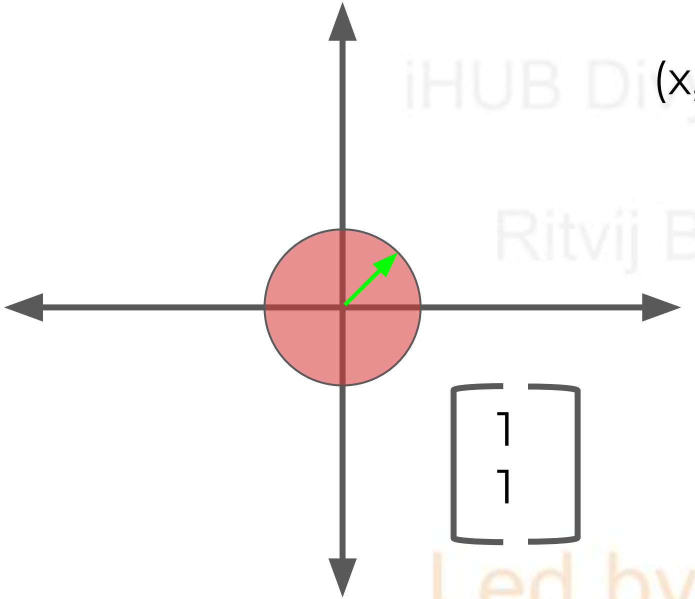


$$\begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

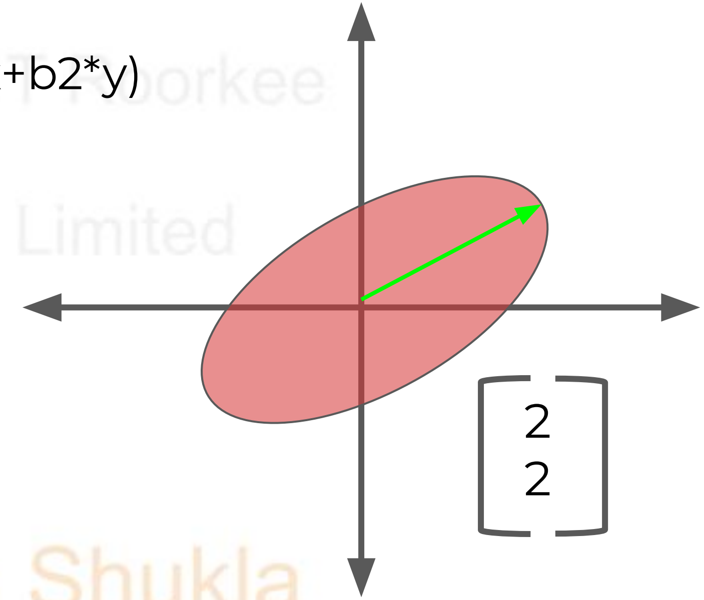
# EigenValue: Magnitude Information

(x,y) → (a1*x+b1*y,a2*x+b2*y)



$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} a1 & b1 \\ a2 & b2 \end{bmatrix}$$



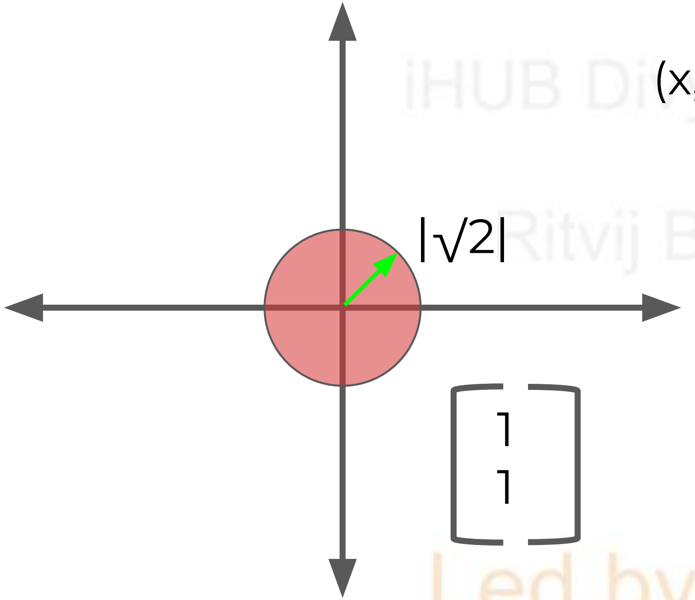
$$\begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

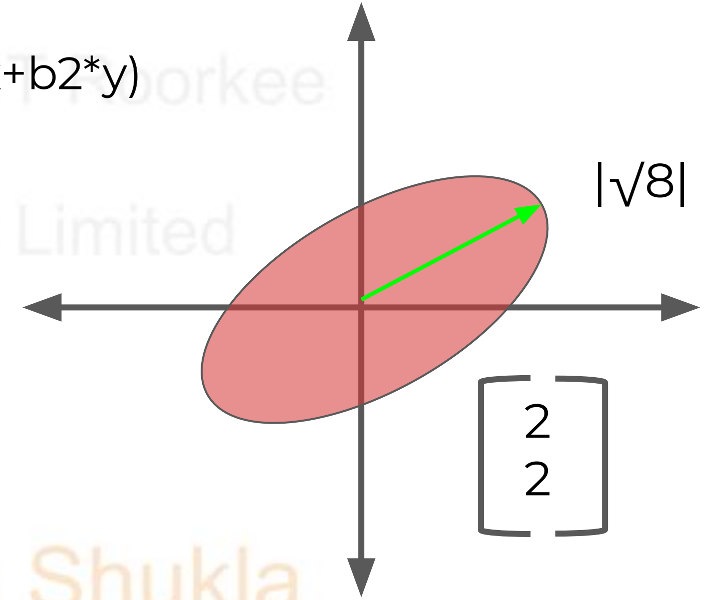
# EigenValue: Magnitude Information

iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited

$(x,y) \rightarrow (a1*x+b1*y, a2*x+b2*y)$



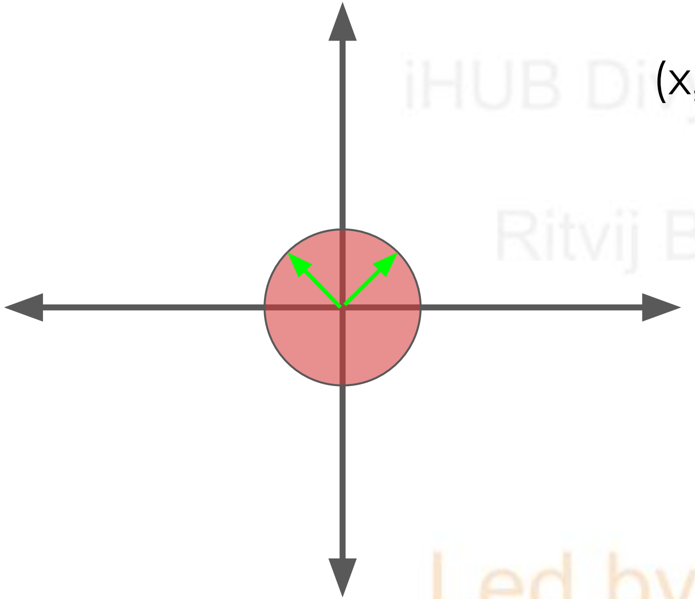

$$\begin{bmatrix} a1 & b1 \\ a2 & b2 \end{bmatrix}$$



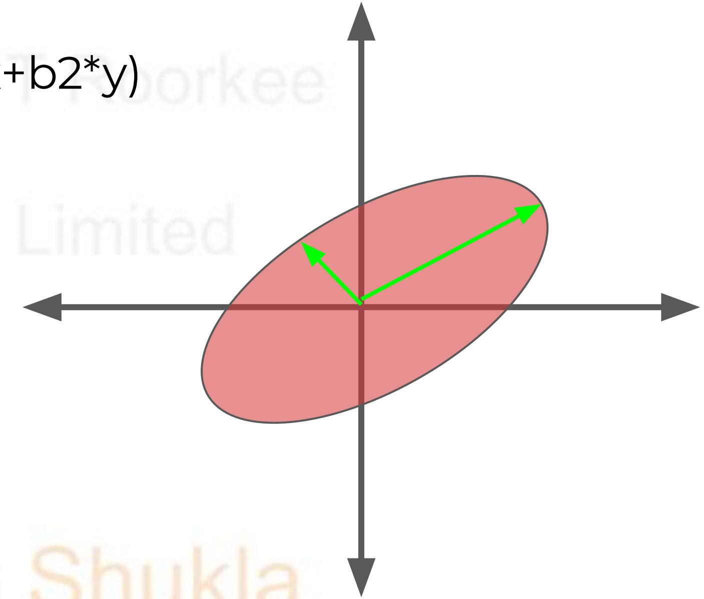

Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Orthogonal EigenVector



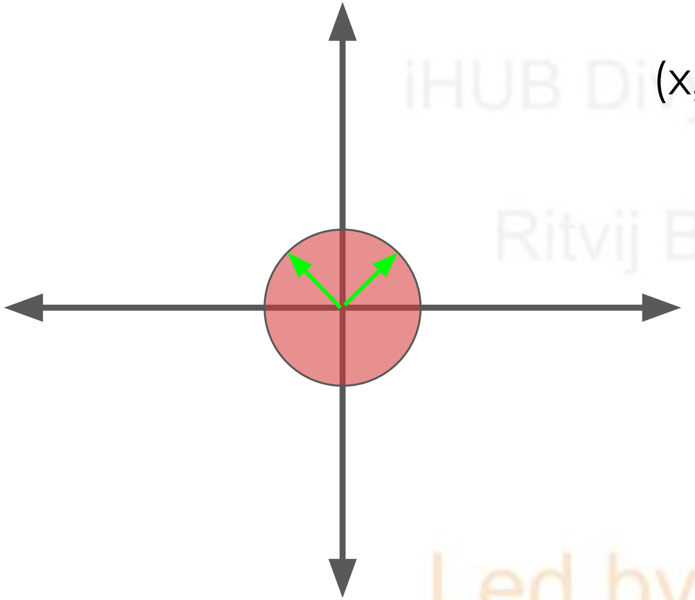
(x,y) → (a1*x+b1*y,a2*x+b2*y)



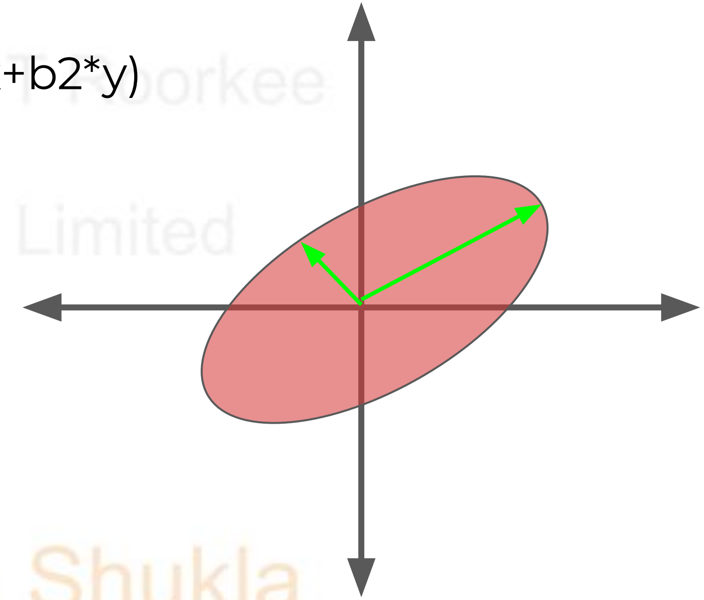
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

# EigenVector is just a linear transformation



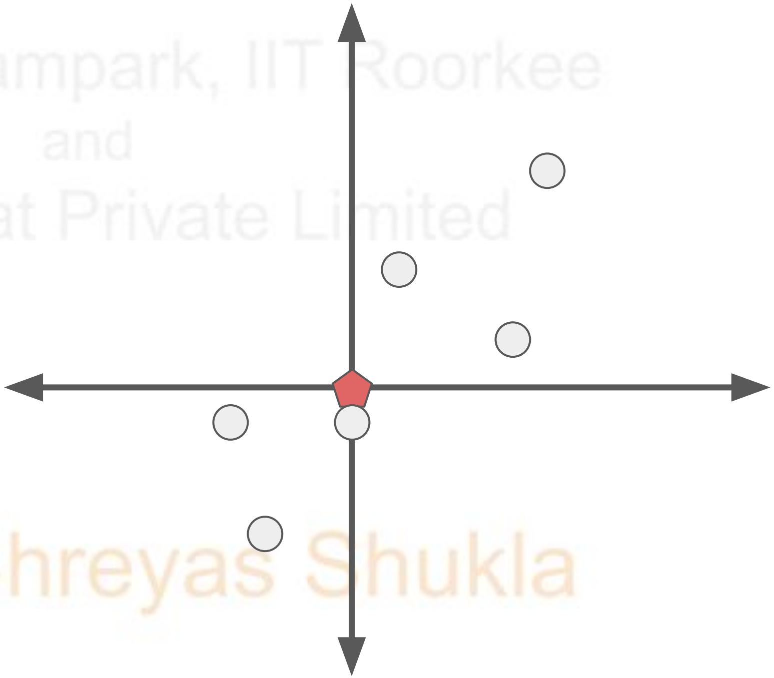
(x,y) → (a1*x+b1*y,a2*x+b2*y)



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Rivij Bharat Private Limited
Led by : Shreyas Shukla

# Apply Linear Transformation:

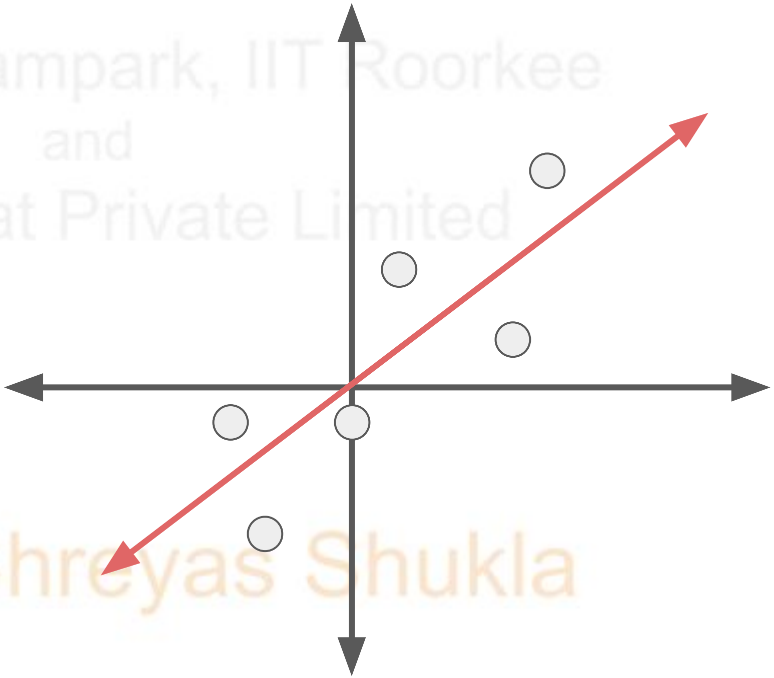
$$\begin{bmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{bmatrix}$$



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Kirvij Bharat Private Limited
Led by : Shreyas Shukla

# Apply Linear Transformation:

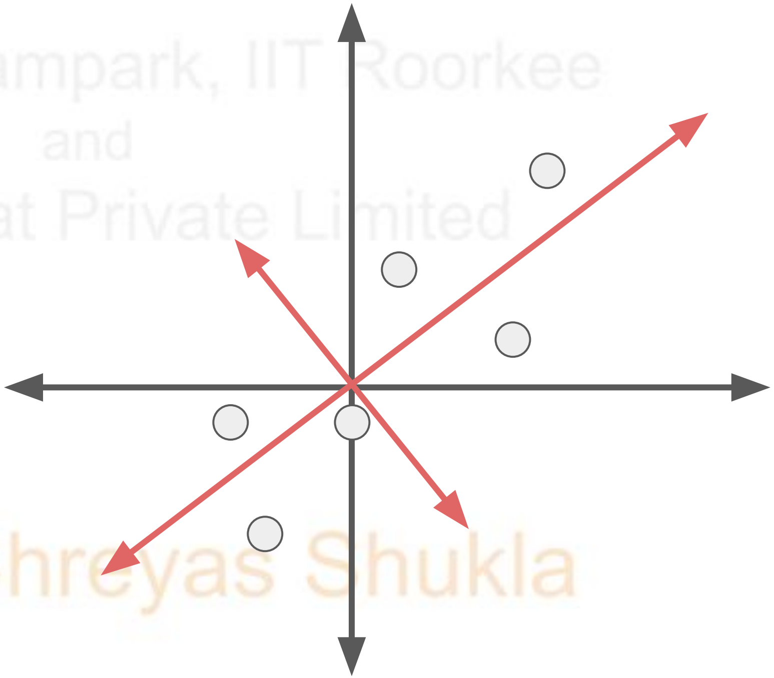
$$\begin{bmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{bmatrix}$$



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# Apply Linear Transformation:

$$\begin{bmatrix} \text{Var}(X) & \text{Cov}(X,Y) \\ \text{Cov}(X,Y) & \text{Var}(Y) \end{bmatrix}$$

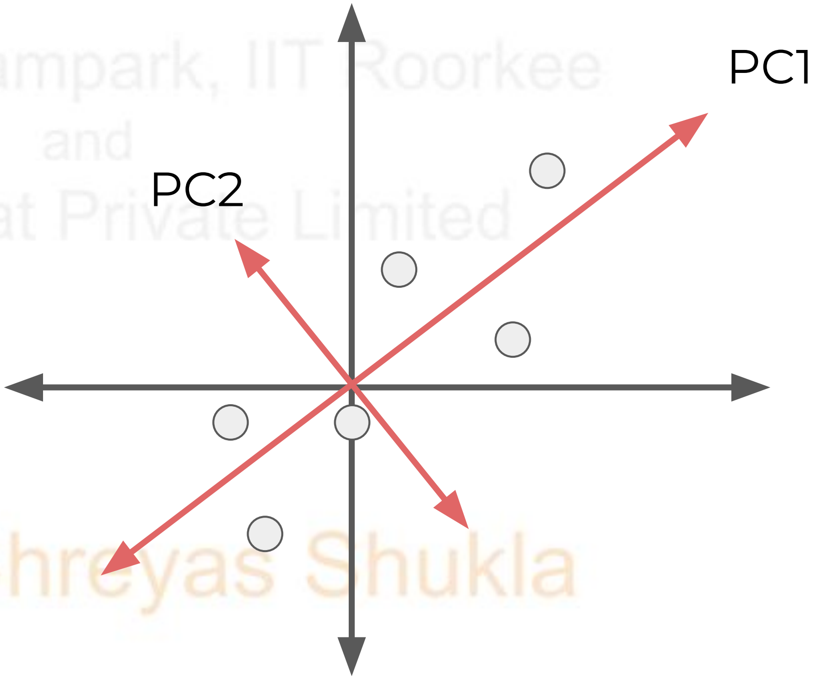


iHUB DivyaSampark, IIT Roorkee
and
Rivij Bharat Private Limited
Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla

# EigenValue measures variance explained:

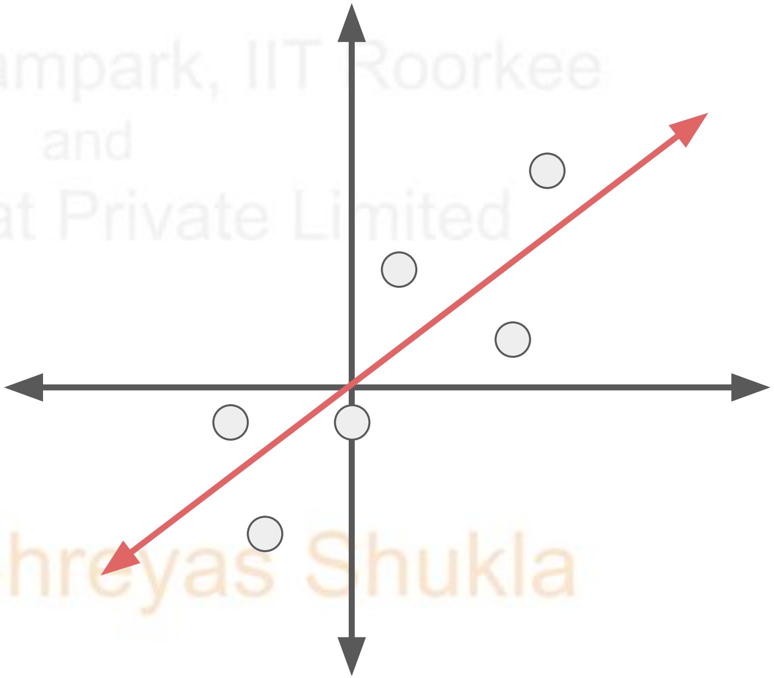
$$\begin{bmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{bmatrix}$$



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Kirvij Bharat Private Limited
Led by : Shreyas Shukla

# EigenValue measures variance explained:

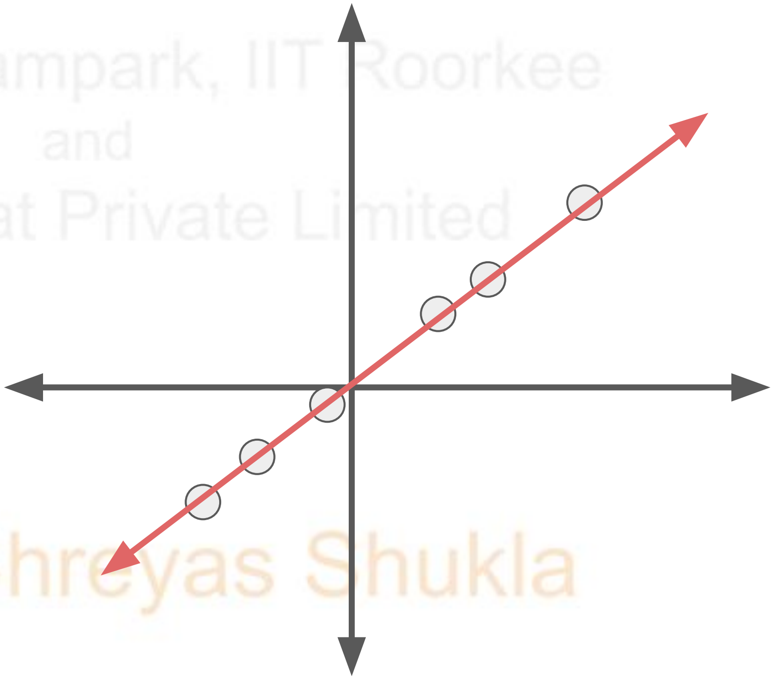
$$\begin{bmatrix} \text{Var(X)} & \text{Cov(X,Y)} \\ \text{Cov(X,Y)} & \text{Var(Y)} \end{bmatrix}$$



Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)
iHUB DivyaSampark, IIT Roorkee
and
Rivij Bharat Private Limited
Led by : Shreyas Shukla

# EigenValue measures variance explained:

$$\begin{bmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \mathrm{Var}(Y) \end{bmatrix}$$

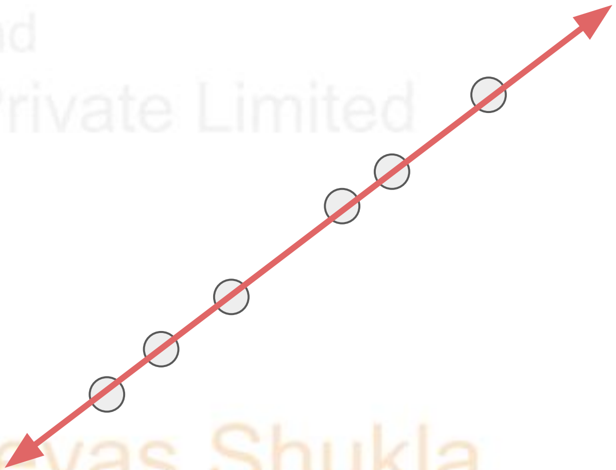


Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# EigenValue measures variance explained:

iHUB DivyaSampark, IIT Roorkee
and
Bharat Private Limited

$$\begin{bmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \mathrm{Var}(Y) \end{bmatrix}$$



Led by : Shreyas Shukla

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

# EigenValue measures variance explained:

iHUB DivyaSampark, IIT Roorkee
and
[illegible] Bharat Private Limited

$$\begin{bmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \mathrm{Var}(Y) \end{bmatrix}$$

Principal Component 1

Led by : Shreyas Shukla

# PCA Steps

Mastering Machine Learning with Python
(27th Aug 2024 - 18th Oct 2024)

1. Get original data
2. Calculate Covariance Matrix
3. Calculate EigenVectors
4. Sort EigenVectors by EigenValues
5. Choose N largest EigenValues
6. Project original data onto EigenVectors

iHUB DivyaSampark, IIT Roorkee
and
Ritvij Bharat Private Limited
Led by : Shreyas Shukla