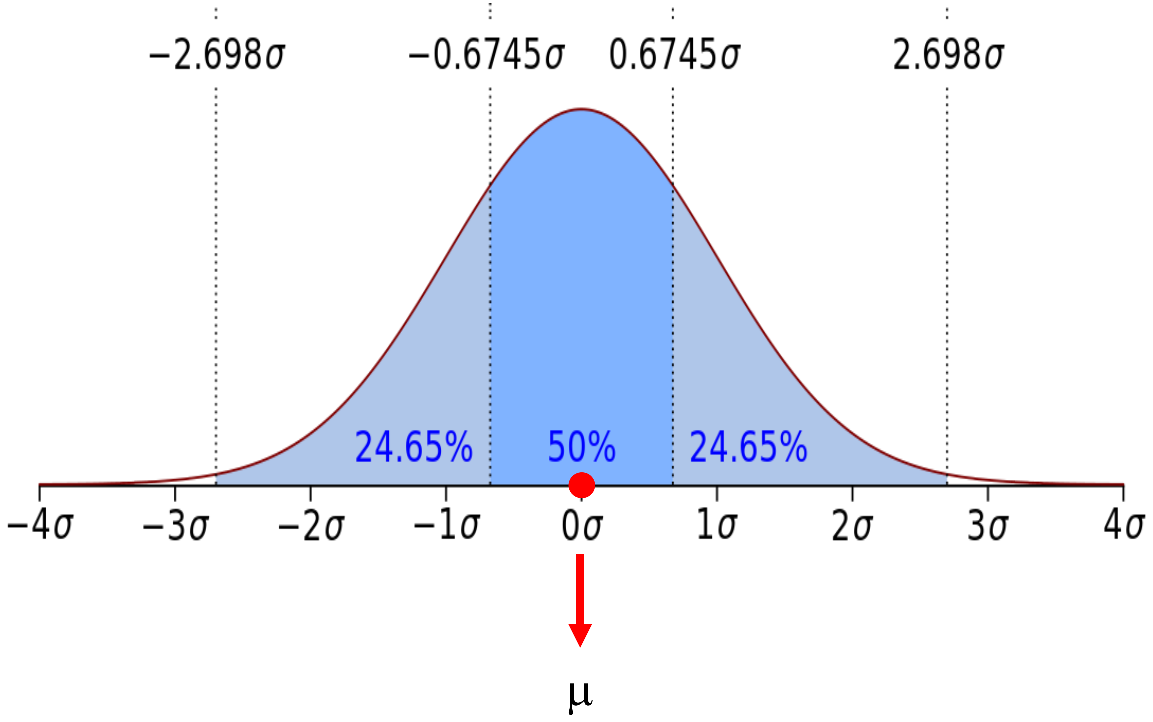
Train In Data

Multivariate Gaussian Distribution

Gaussian distribution

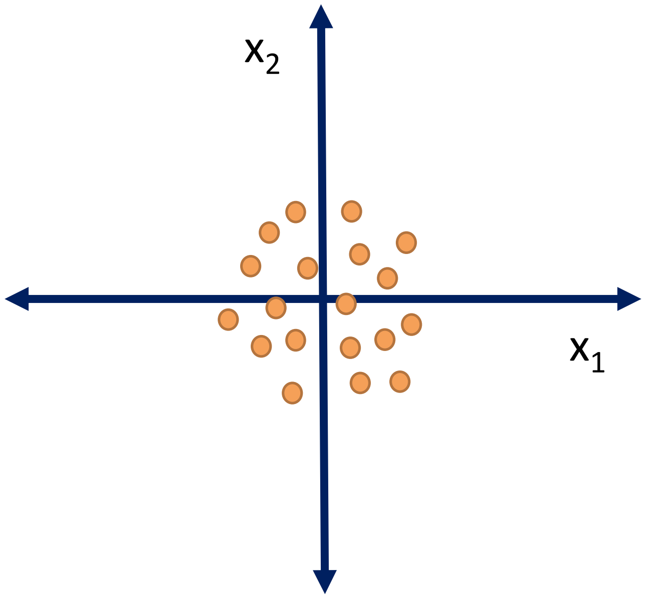


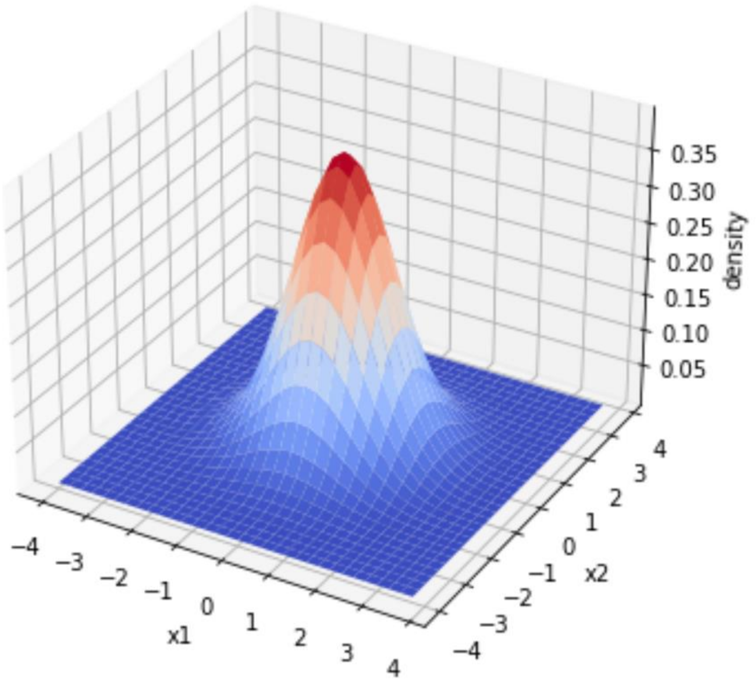
- Univariate Gaussian distributions are determined by **μ** and **σ**

- μ = Mean value
  ➔ centre of distribution

- σ = standard deviation
  ➔ measure of dispersion

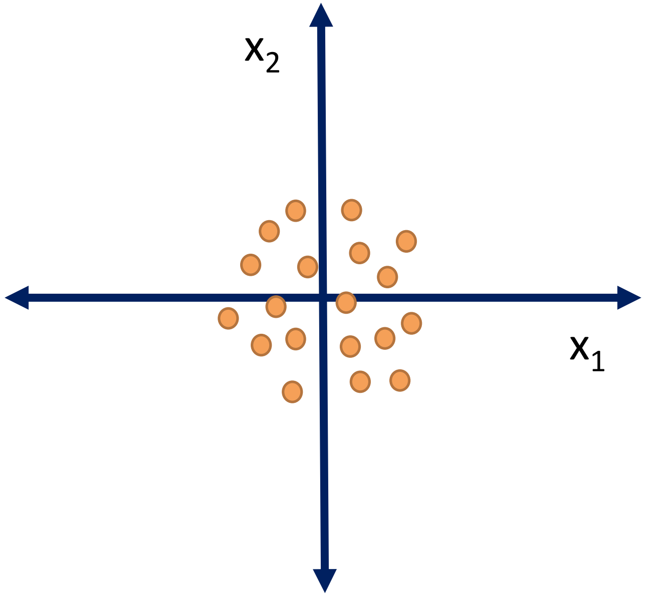
Multivariate Gaussian Distribution





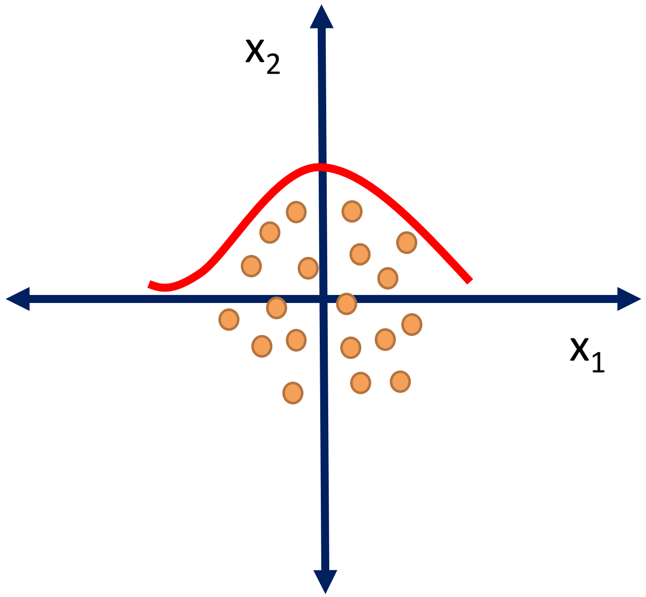
The probability of a value x occurring is given by the joint probability of x1 and x2

Multivariate Gaussian Distribution



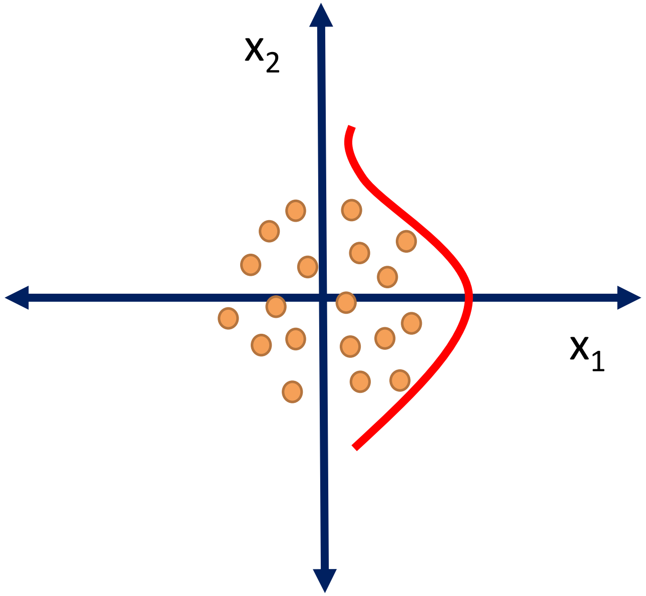
- $\mu 1$ and $\mu 2$.
- $\sigma^2 1$ and $\sigma^2 2$.

Multivariate Gaussian Distribution



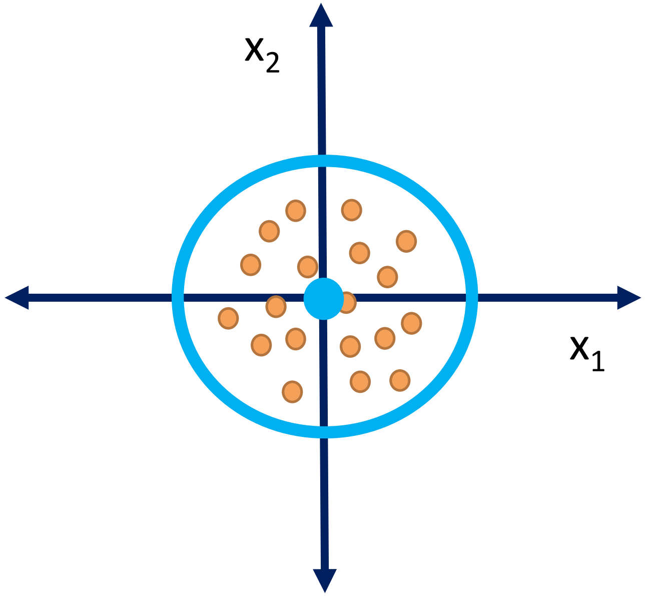
- $\mu 1$ and $\mu 2$.
- $\sigma^2 1$ and $\sigma^2 2$.
- $X1 \sim N(\mu 1=0, \sigma^2 1=1)$

Multivariate Gaussian Distribution



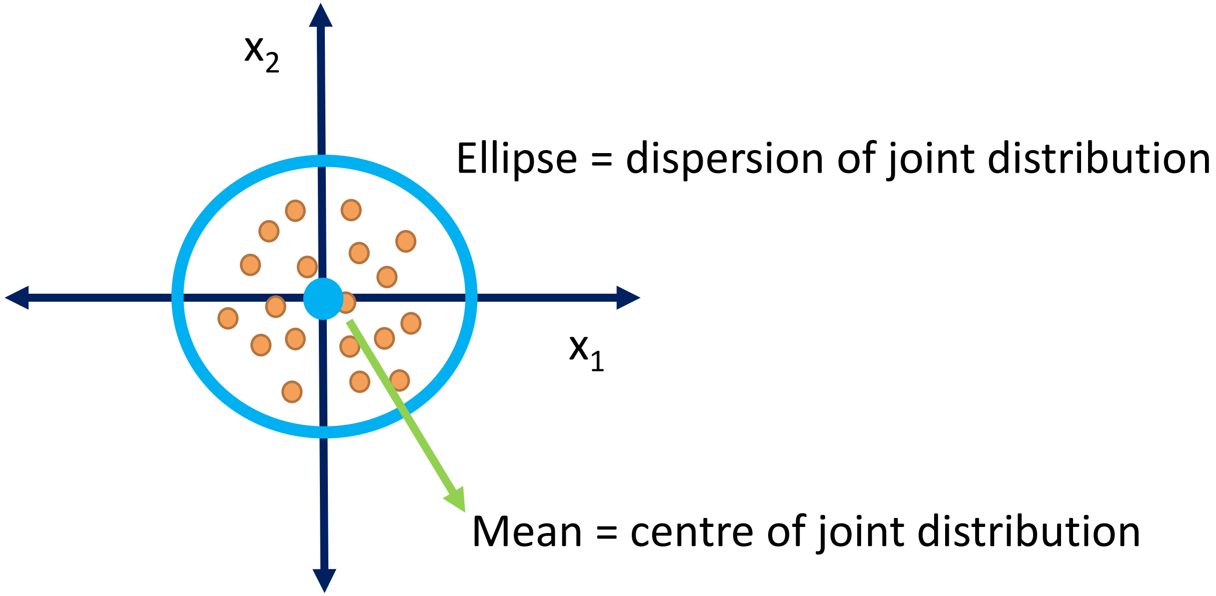
- $\mu 1$ and $\mu 2$.
- $\sigma^2 1$ and $\sigma^2 2$.
- $X1 \sim N(\mu 1=0, \sigma^2 1=1)$
- $X2 \sim N(\mu 2=0, \sigma^2 2=1)$

Multivariate Gaussian Distribution



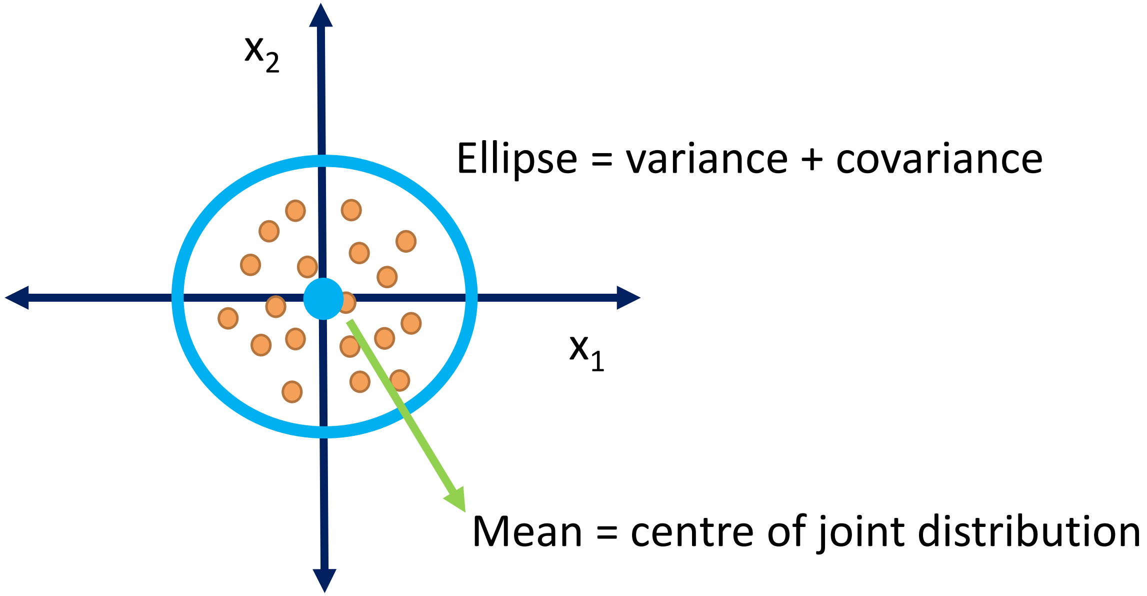
- $\mu 1$ and $\mu 2$.
- $\sigma^2 1$ and $\sigma^2 2$.
- $X1 \sim N(\mu 1=0, \sigma^2 1=1)$
- $X2 \sim N(\mu 2=0, \sigma^2 2=1)$

Multivariate Gaussian Distribution



- $\mu1$ and $\mu2$.
- $\sigma^2 1$ and $\sigma^2 2$.
- $X1 \sim N(\mu1=0, \sigma^2 1=1)$
- $X2 \sim N(\mu2=0, \sigma^2 2=1)$

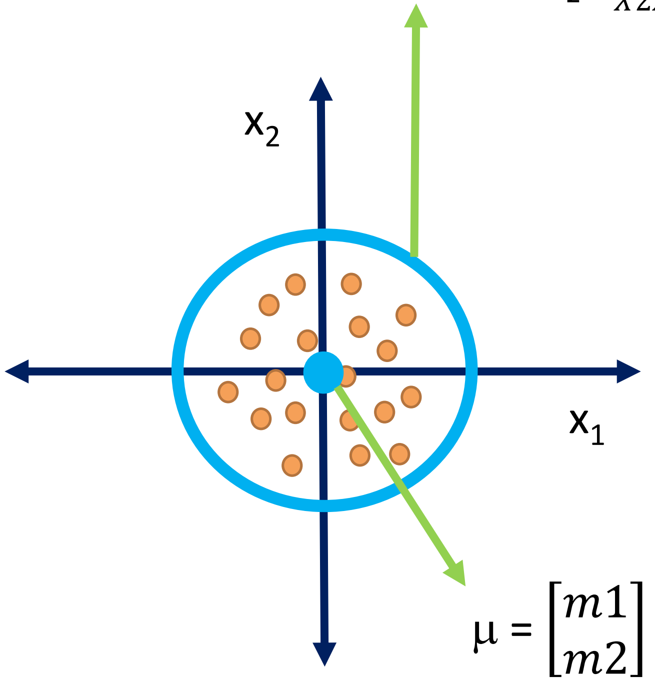
Multivariate Gaussian Distribution



- $\mu 1$ and $\mu 2$.
- $\sigma^2 1$ and $\sigma^2 2$.
- $X1 \sim N(\mu 1=0, \sigma^2 1=1)$
- $X2 \sim N(\mu 2=0, \sigma^2 2=1)$

Multivariate Gaussian Distribution

Covariance matrix = $\sum = \begin{bmatrix} K_{X1X1} & K_{X1X2} \\ K_{X2X1} & K_{X2X2} \end{bmatrix}$



$\mu = \begin{bmatrix} m1 \\ m2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

- $\mu1$ and $\mu2$.
- $\sigma^2 1$ and $\sigma^2 2$.
- X1 ~N($\mu1$=0, $\sigma^2 1$=1)
- X2 ~N($\mu2$=0, $\sigma^2 2$=1)
- $X = \begin{bmatrix} x1 \\ x2 \end{bmatrix} \sim N(\mu = \begin{bmatrix} m1 \\ m2 \end{bmatrix}, \sum = \begin{bmatrix} K_{X1X1} & K_{X1X2} \\ K_{X2X1} & K_{X2X2} \end{bmatrix})$

Multivariate Gaussian Distribution

- Generalizes the univariate Gaussian distribution to higher dimensions
  - ✓ More than 1 variable
  - ✓ Instead of values, we now have <u>vectors</u>

- <u>Multivariate</u> Gaussian distributions need $\boldsymbol{\mu}$, $\boldsymbol{\sigma}^2$ and the **covariance** $\sum$

- **Covariance matrix:** captures $\boldsymbol{\sigma}^2$ and $\sum$

Covariance

- Measure of joint probability of 2 random variables.

- Cov(X1, X2) = $\frac{\sum\left(\left(x_{ij} - x_{jmean}\right) \times \left(x_{ik} - x_{kmean}\right)\right)}{n}$

Covariance

- Measure of joint probability of 2 random variables.

- $\text{Cov(X1, X2)} = \frac{\sum(x_{ij} - x_{jmean}) \times (x_{ik} - x_{kmean}))}{n}$

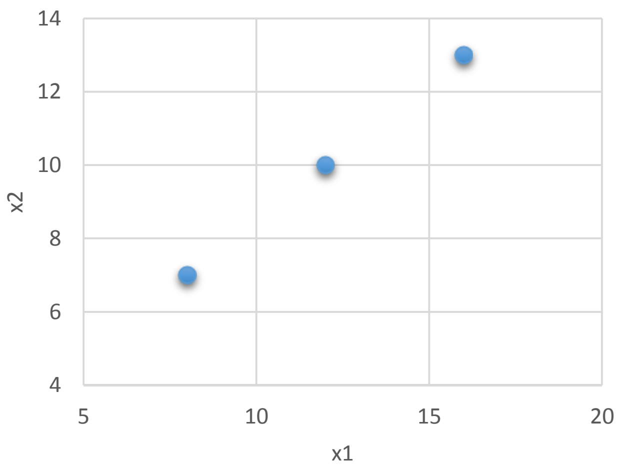
	j	k	
	X1	X2	
	12	10	
	8	7	i
	16	13	
	mean = 12	mean = 10	

$$\text{Cov(X1, X2)} = \frac{(12 - 12) * (10 - 10) + (8 - 12) * (7 - 10) + (16 - 12) * (13 - 10)}{3} = 8$$

Covariance

x1	x2
12	10
8	7
16	13

12 10



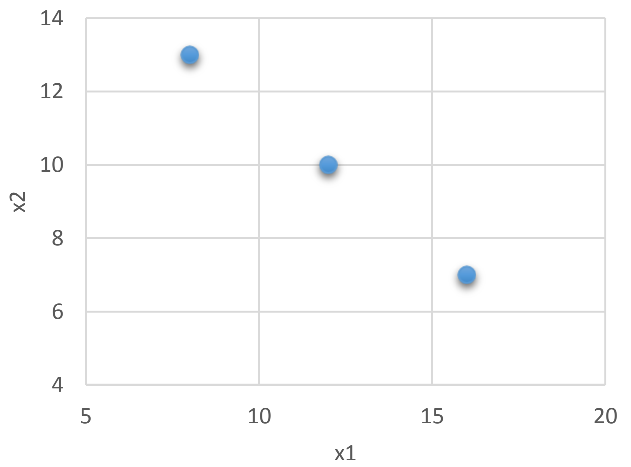
Cov(X1, X2) =

$$\frac{(12 - 12) * (10 - 10) + (8 - 12) * (7 - 10) + (16 - 12) * (13 - 10)}{3}$$

= 8

x1	x2
12	10
8	13
16	7

12 10



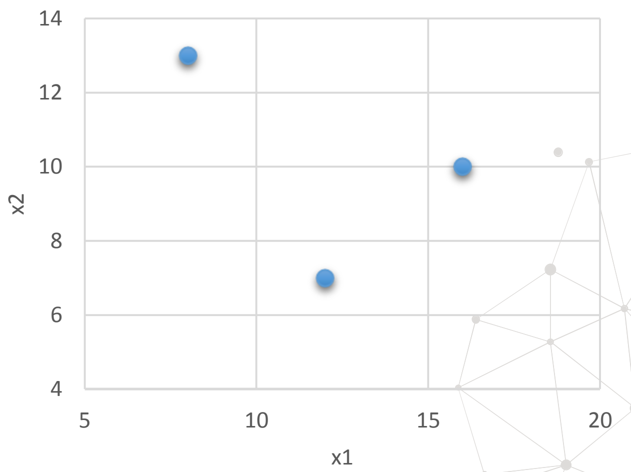
Cov(X1, X2) =

$$\frac{(12 - 12) * (10 - 10) + (8 - 12) * (13 - 10) + (16 - 12) * (7 - 10)}{3}$$

= - 8

x1	x2
12	7
8	13
16	10

12 10

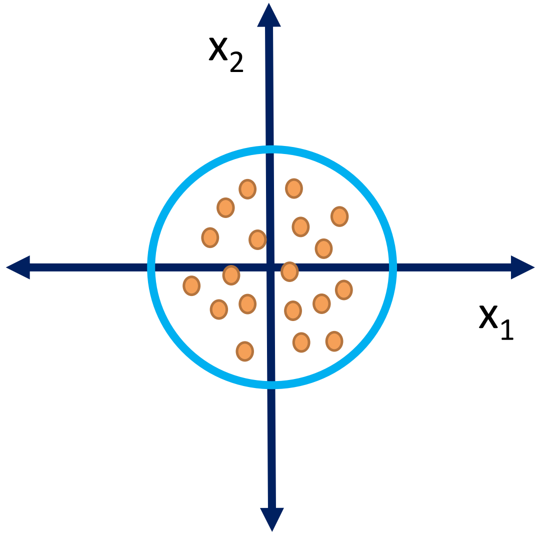


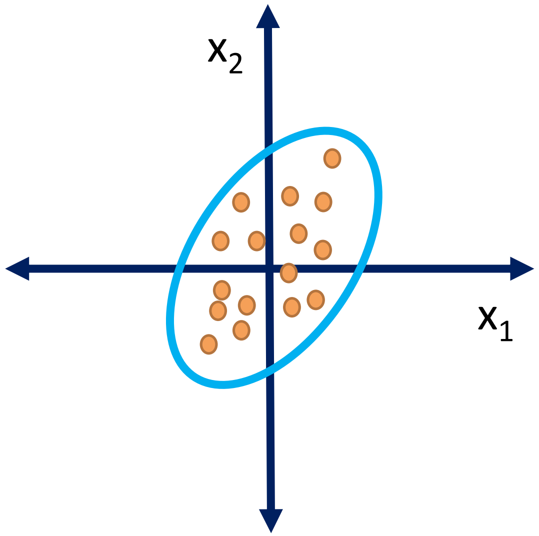
Cov(X1, X2) =

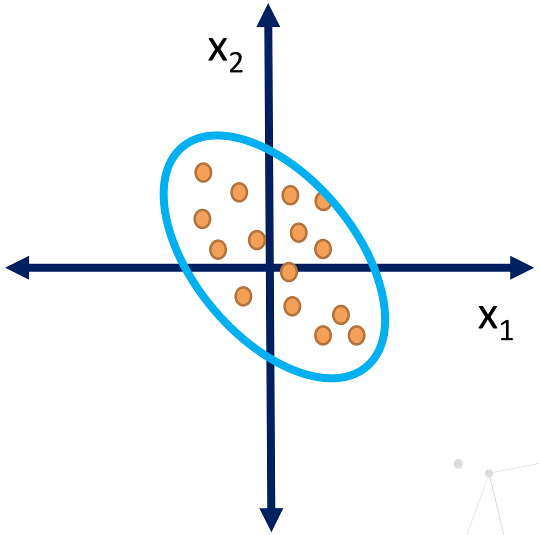
$$\frac{(12 - 12) * (7 - 10) + (8 - 12) * (13 - 10) + (16 - 12) * (10 - 10)}{3}$$

= -4

Covariance





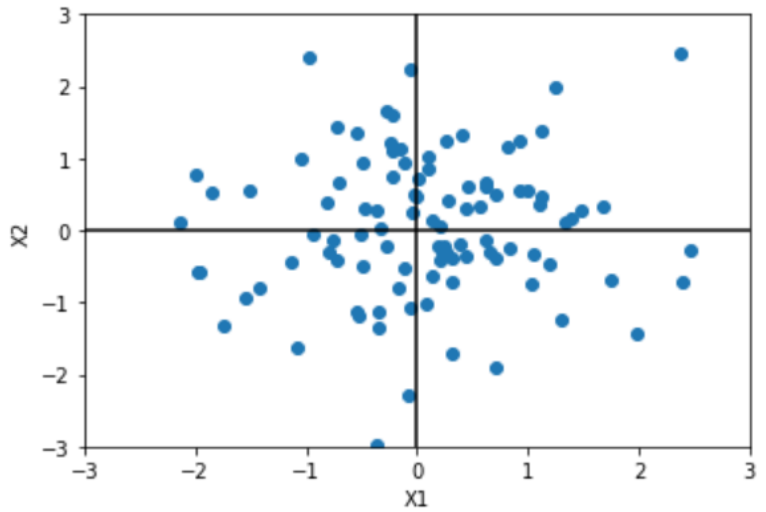


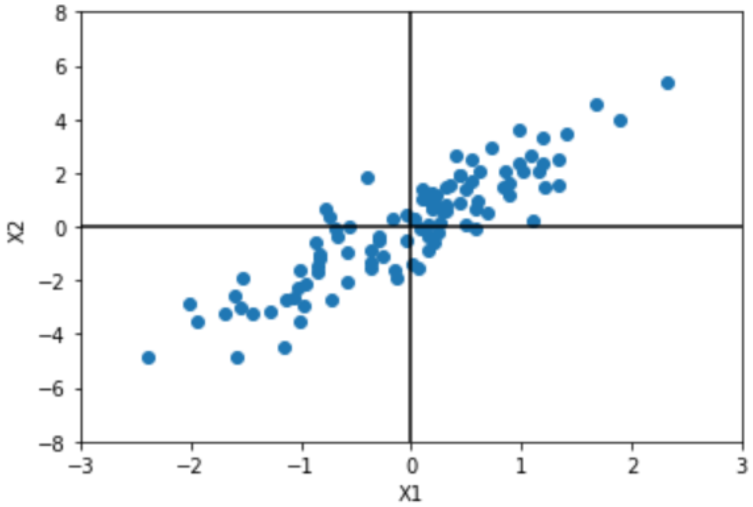
- Cov(x1,x2) = 0
- X1 and x2 are not correlated

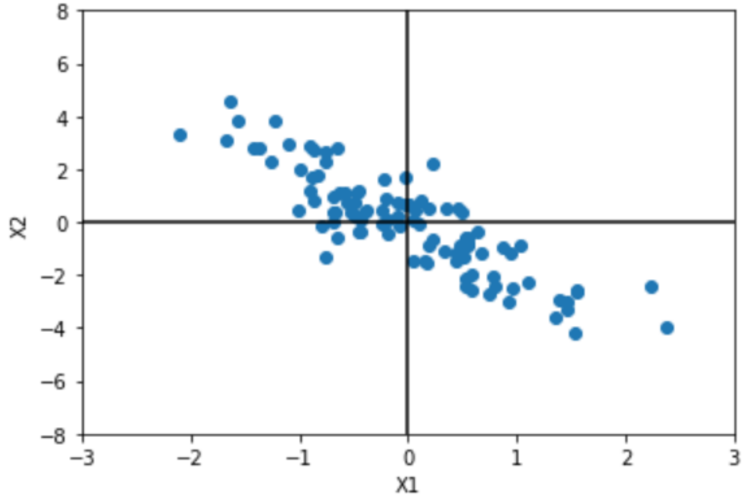
- Cov(x1,x2) > 0
- The bigger x1, the bigger x2

- Cov(x1,x2) < 0
- The bigger x1, the smaller x2

Covariance








```
np.sum( (x1 - np.mean(x1)) * (x2 - np.mean(x2))) / n
```

```
0.062008247055892605
```

```
np.sum( (x1 - np.mean(x1)) * (x2 - np.mean(x2))) / n
```

```
-1.574496561355141
```

```
np.sum( (x1 - np.mean(x1)) * (x2 - np.mean(x2))) / n
```

```
1.7016981044990922
```

Covariance Matrix

- Square matrix with the covariance of each pair of variables.

- Symmetric

- The diagonal contains the variances, i.e., the covariance of each variable with itself

- The covariance matrix provides a succinct way to summarize the covariance of all pairs of variables

$$\Sigma = \begin{bmatrix} K_{X1X1} & K_{X1X2} \\ K_{X2X1} & K_{X2X2} \end{bmatrix}$$

Where:

- Kx1x1 = var(x1)

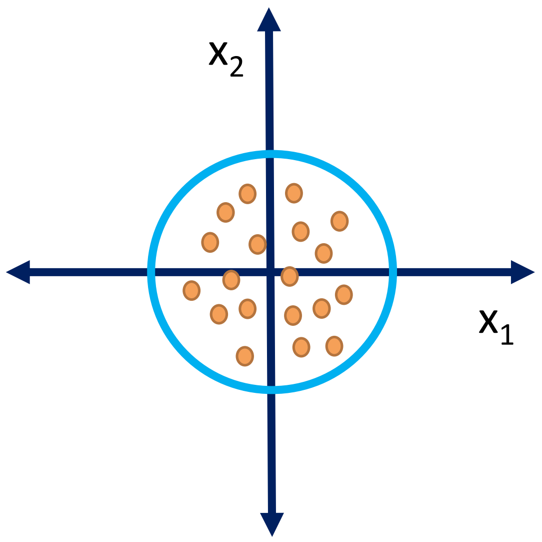
- Kx2x2 = var(x2)

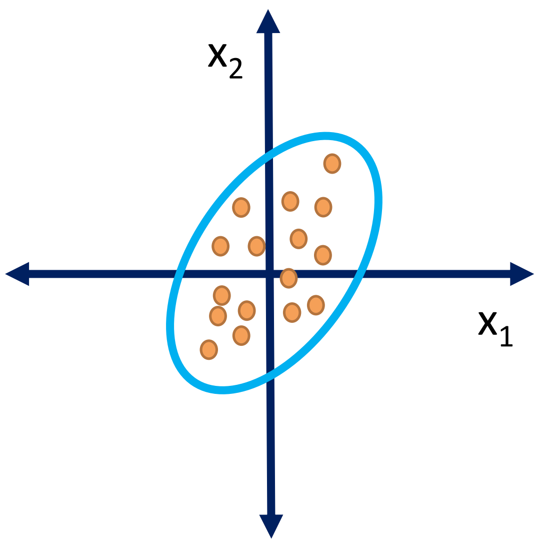
- Kx1x2 = K^Tx2x1 = cov(x1, x2)

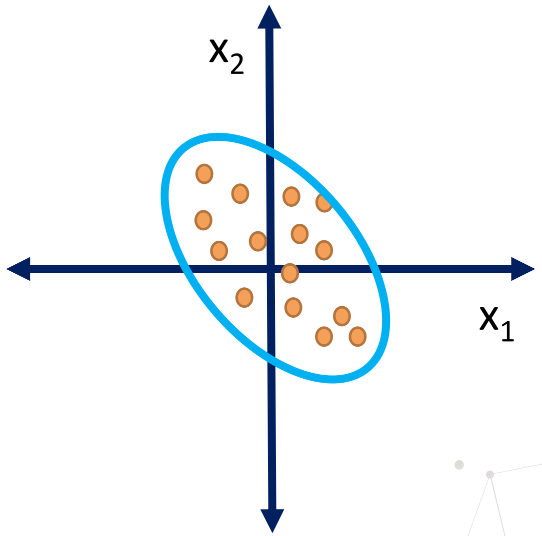
Covariance Matrix - General

$$\Sigma = \begin{bmatrix} Kx_1x_1 & \cdots & Kx_1x_n \\ \vdots & \ddots & \vdots \\ Kx_nx_1 & \cdots & Kx_nx_n \end{bmatrix}$$

Covariance Matrix





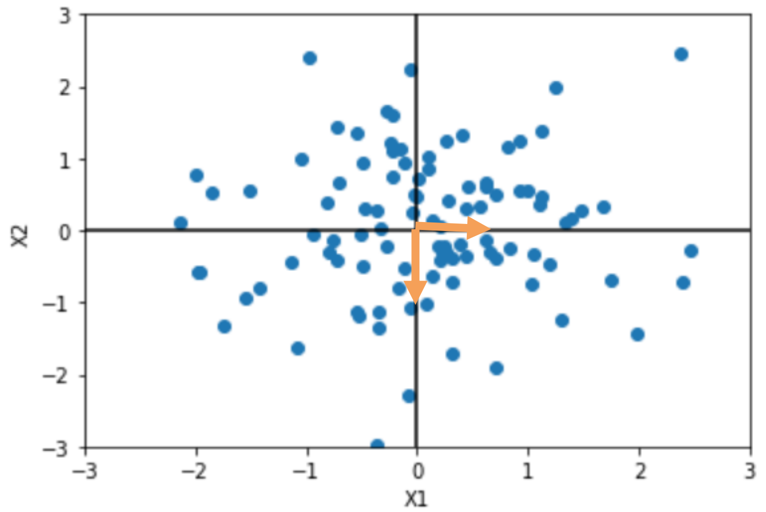


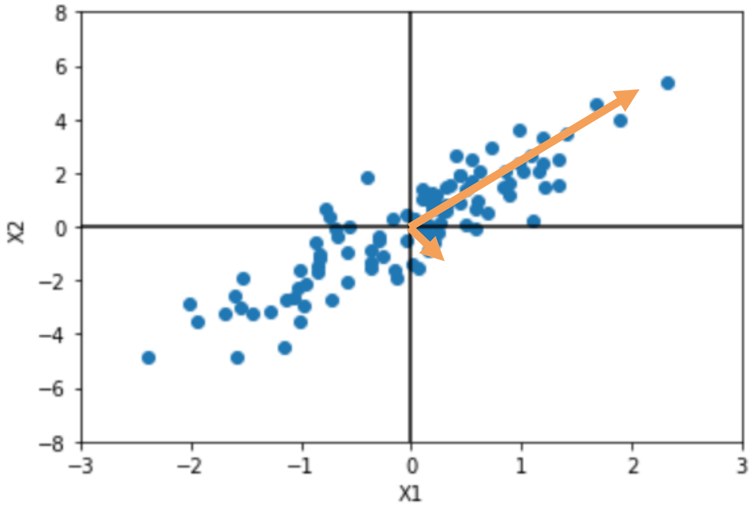
$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

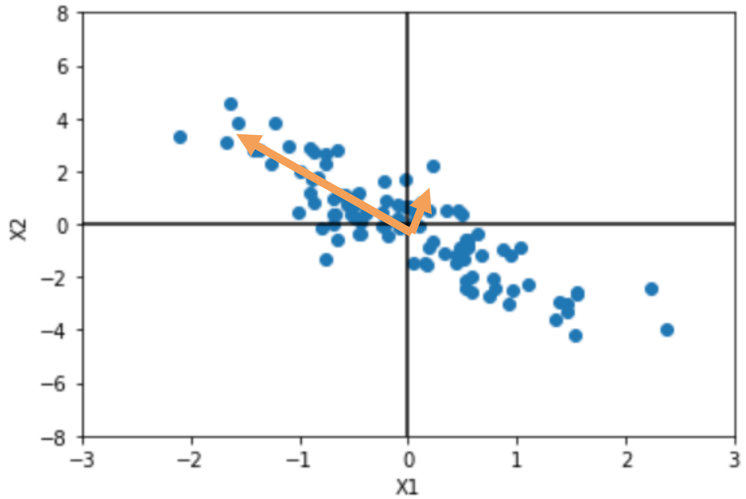
$$\Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$$

Covariance Matrix: simulation








```
np.cov(x1, x2)
```

```
array([[0.94461146, 0.06263459],
       [0.06263459, 0.98069174]])
```

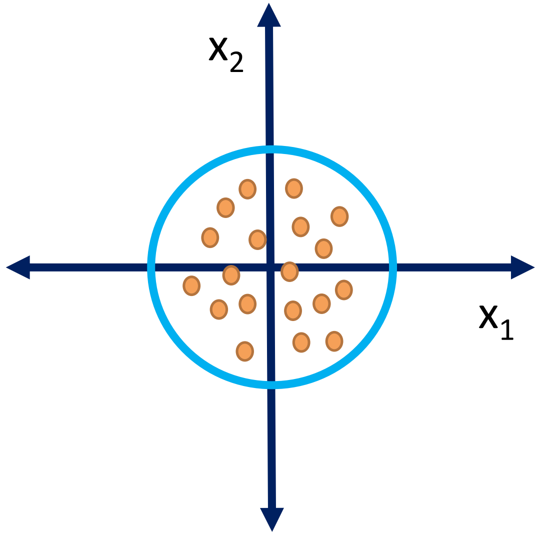
```
np.cov(x1, x2)
```

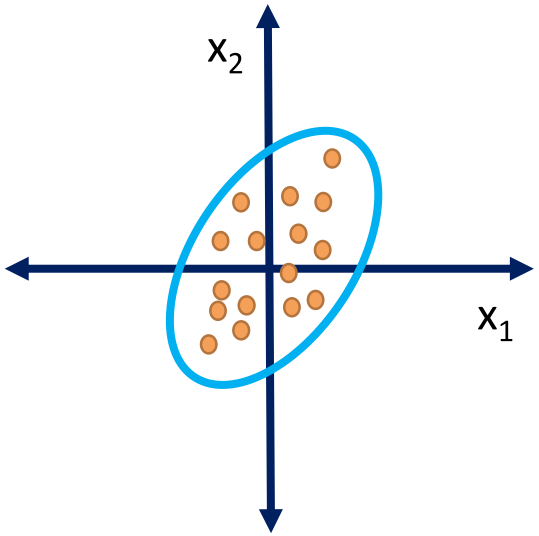
```
array([[0.84262922, 1.71888697],
       [1.71888697, 4.38490628]])
```

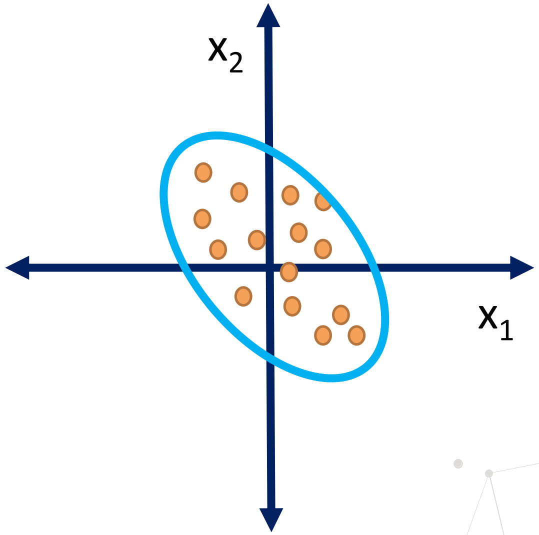
```
np.cov(x1, x2)
```

```
array([[ 0.85181132, -1.59040057],
       [-1.59040057,  3.84101983]])
```

Multivariate Gaussian Distribution







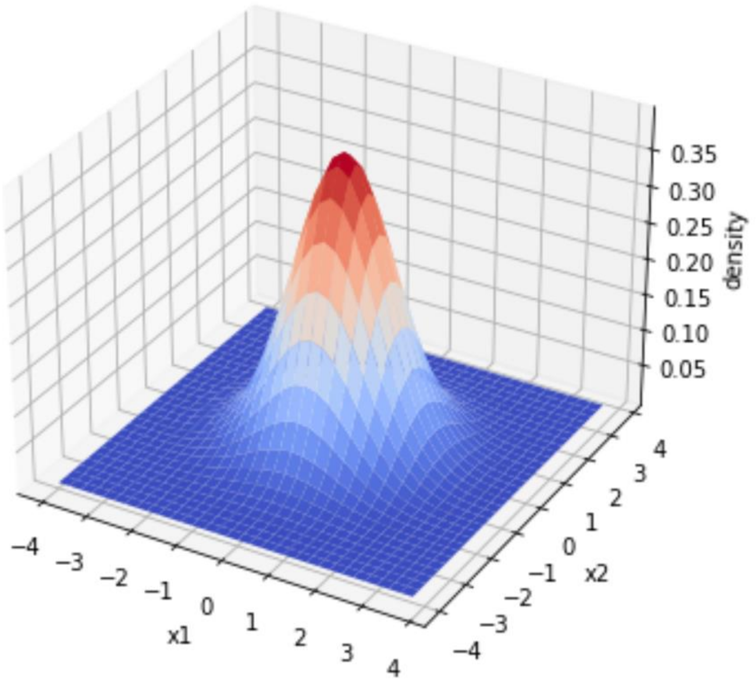
$$X = \begin{bmatrix} x1 \\ x2 \end{bmatrix} \sim N(\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix})$$

$$\begin{bmatrix} x1 \\ x2 \end{bmatrix} \sim N(\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix})$$

$$\begin{bmatrix} x1 \\ x2 \end{bmatrix} \sim N(\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix})$$

Multivariate Gaussian Distribution

$$f_{\mathbf{X}}(x_1,\ldots,x_k) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k|\boldsymbol{\Sigma}|}}$$



The probability of a value x occurring is given by the joint probability of x1 and x2

THANK YOU

www.trainindata.com

Train In Data