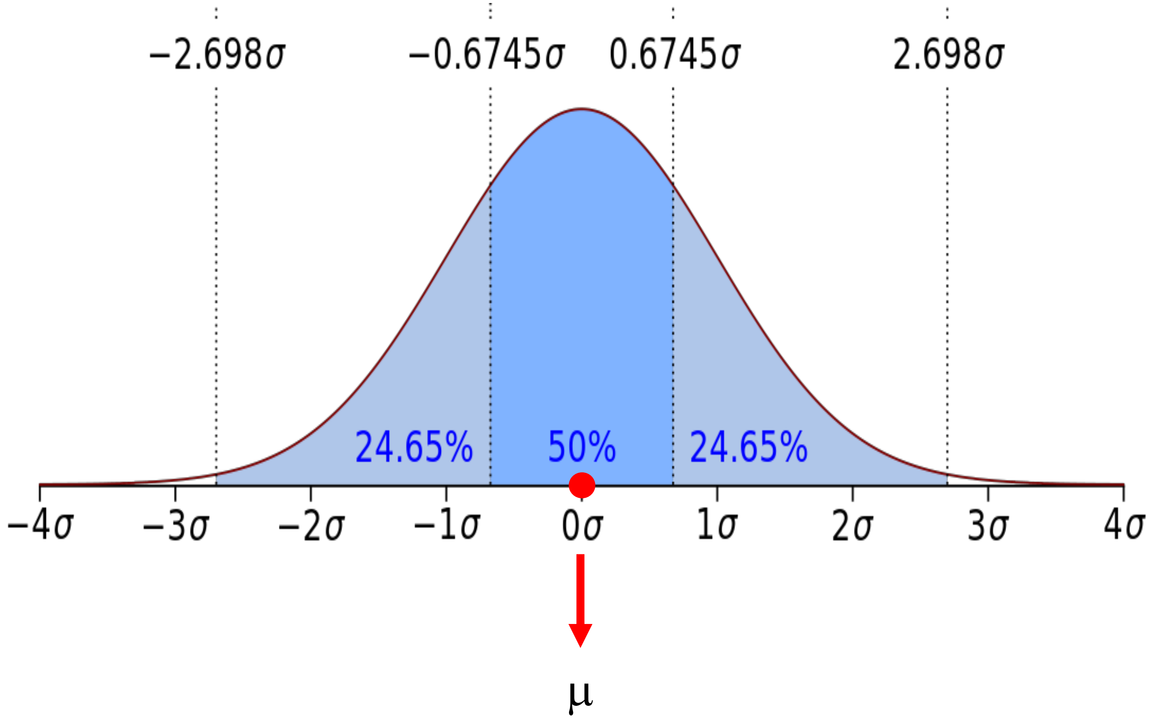
Train In Data

Multivariate Gaussian Distribution

Gaussian distribution

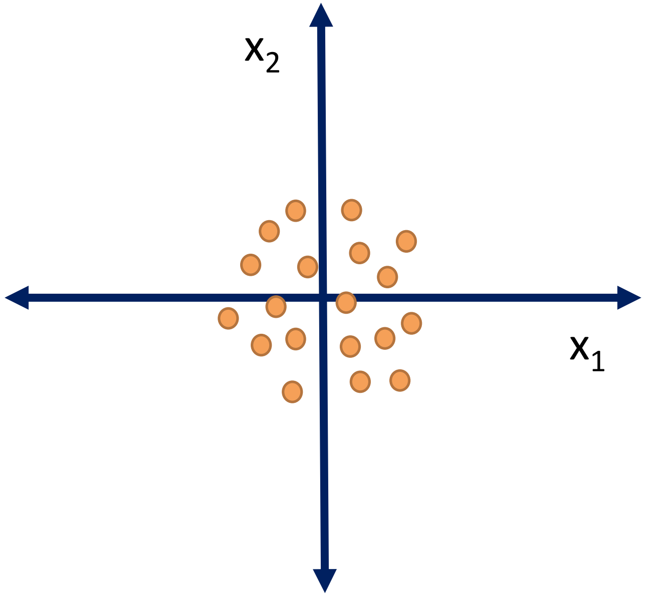


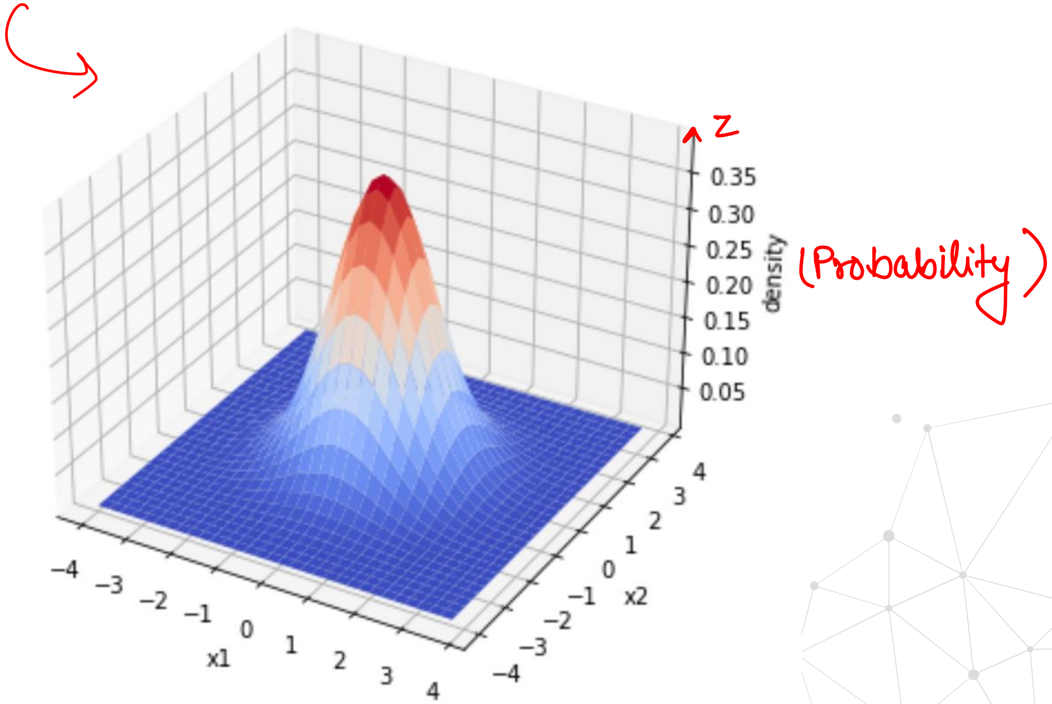
- Univariate Gaussian distributions are determined by **μ** and **σ**
- μ = Mean value
  - ➔ centre of distribution
- σ = standard deviation
  - ➔ measure of dispersion

How far from the mean can the values take place.

Multivariate Gaussian Distribution

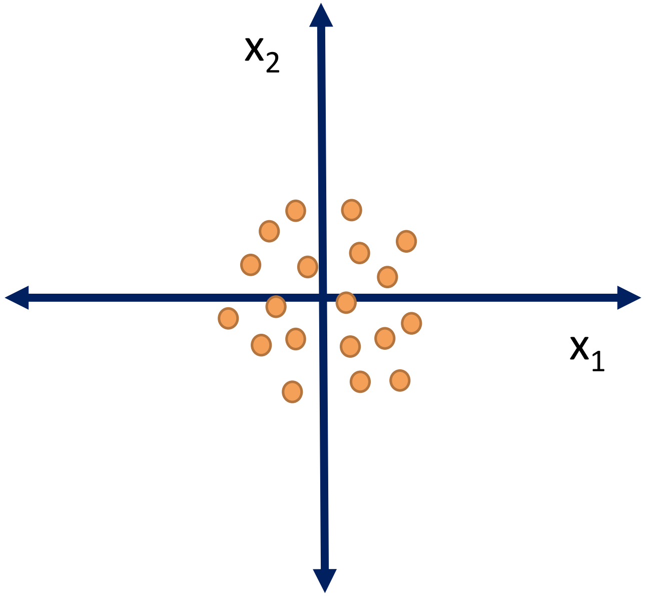
The prob. of x occuring, now x is a vector that contains values x_1 & x_2, is given by the joint probability of (x_1, x_2). Now, estimating this distribution.





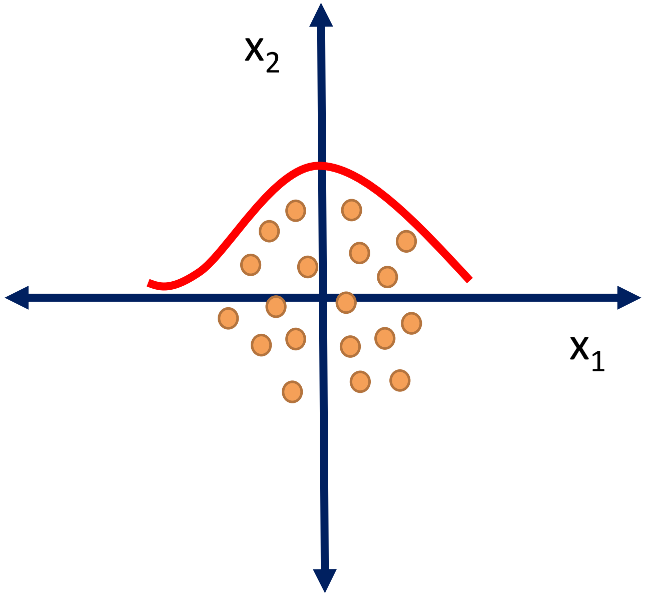
The probability of a value x occurring is given by the joint probability of x1 and x2

Multivariate Gaussian Distribution



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

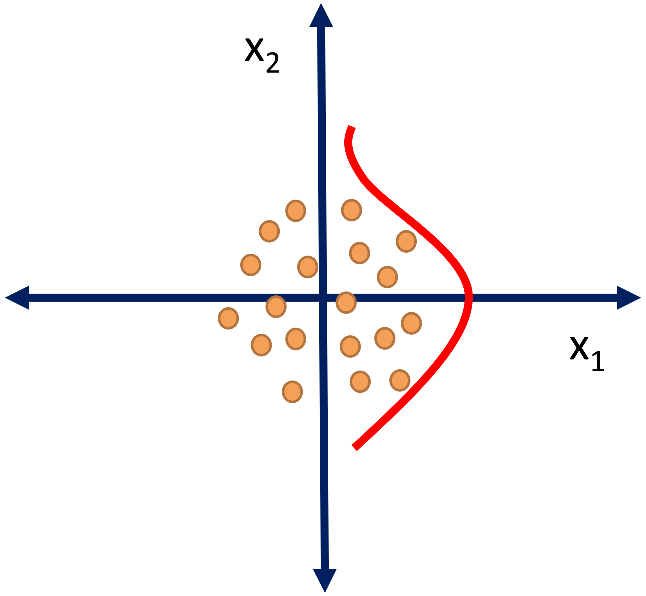
Multivariate Gaussian Distribution



From given values of x_1 & x_2 we can calculate the mean and variance.

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

Multivariate Gaussian Distribution

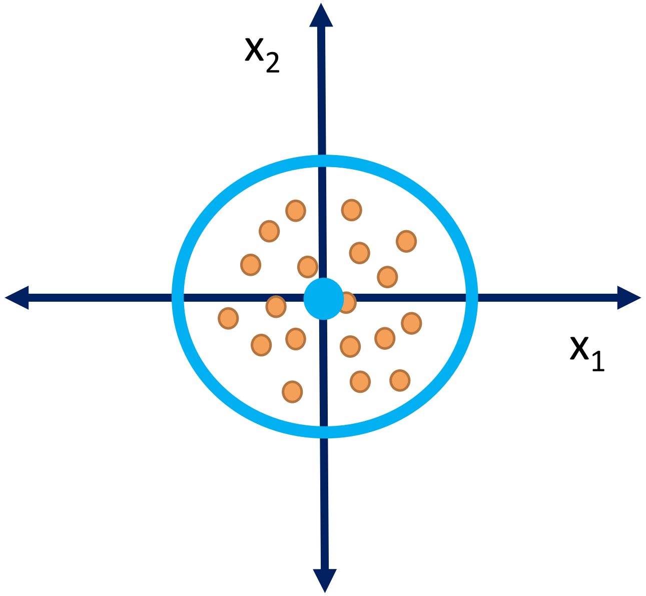


- $\mu 1$ and $\mu 2$.
- $\sigma^2 1$ and $\sigma^2 2$.
- X1 ~N($\mu 1$=0, $\sigma^2 1$=1)
- X2 ~N($\mu 2$=0, $\sigma^2 2$=1)

In this case, the $\mu = 0$, $\sigma^2 = 1$ for both

But the params. of the respective variables are not enough to describe the dispersion or joint occurance of each one of these values. We want to find a function that allows us to estimate the prob. of each of these vectors occuring.

Multivariate Gaussian Distribution

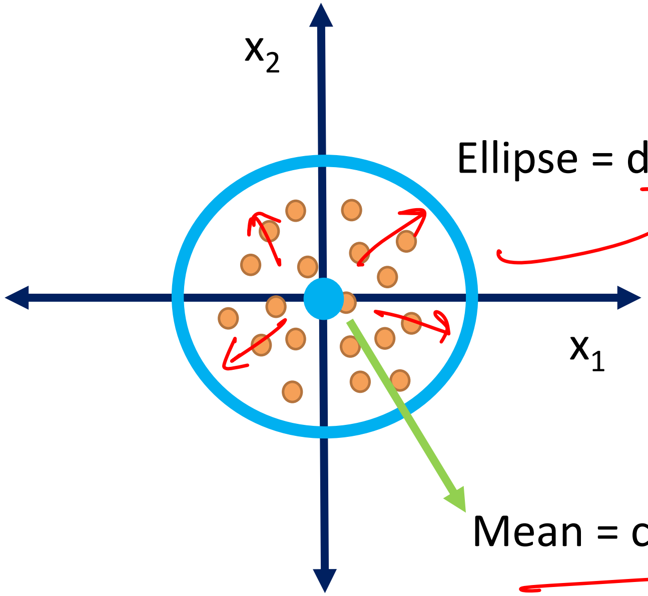


- $\mu 1$ and $\mu 2$.
- $\sigma^2 1$ and $\sigma^2 2$.
- X1 ~N($\mu 1$=0, $\sigma^2 1$=1)
- X2 ~N($\mu 2$=0, $\sigma^2 2$=1)

Multivariate Gaussian Distribution

This joint distribution has a center, denoting the mean of values. in this case it's a vector.

And a measure of dispersion, denoted by the Ellipse (here circle), shows the dispersion that the values can take away from the mean in any direction.



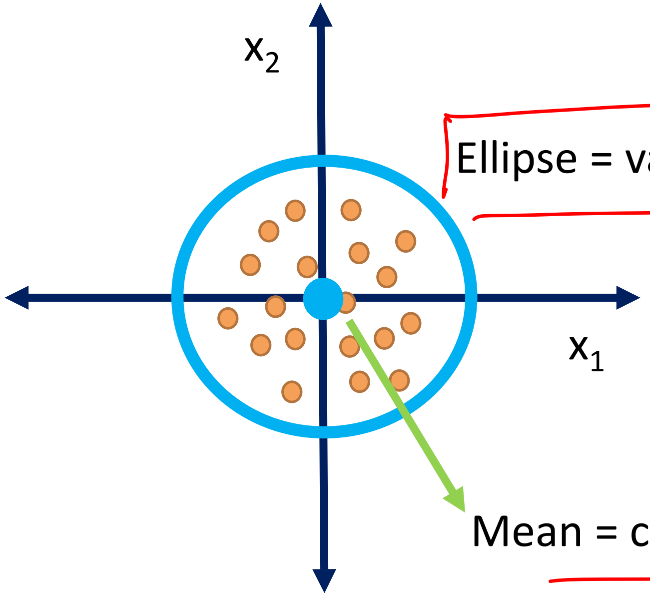
Ellipse = dispersion of joint distribution

Mean = centre of joint 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

Measure of dispersion is given by var. & co-var.



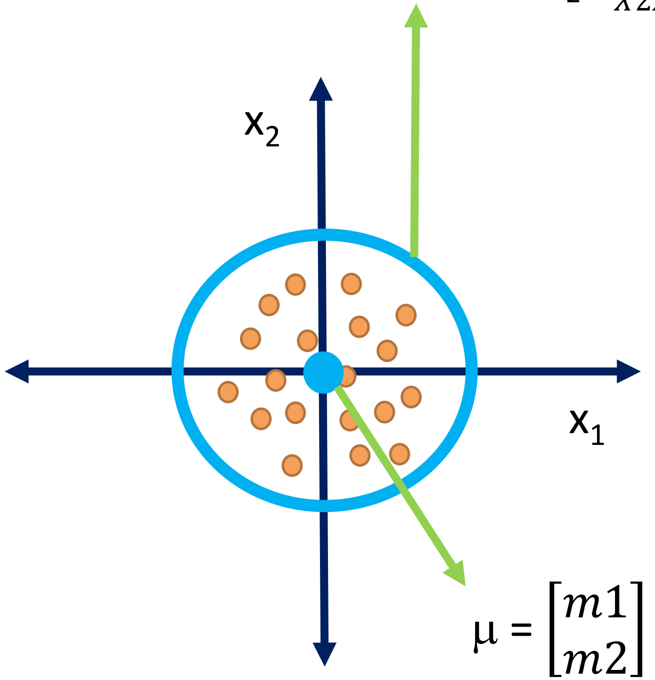
Ellipse = variance + covariance

Mean = centre of joint distribution

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

Multivariate Gaussian Distribution

Covariance matrix = $\Sigma = \begin{bmatrix} K_{X1X1} & K_{X1X2} \\ K_{X2X1} & K_{X2X2} \end{bmatrix} = \begin{bmatrix} \sigma^2_{x_1x_1} & \sigma^2_{x_1x_2} \\ \sigma^2_{x_2x_1} & \sigma^2_{x_2x_2} \end{bmatrix} \rightarrow$ Captures variance of variables tog.



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

Center i.e. Mean (μ) is now a vector. $m_1 = \sum x^1_i$; $m_2 = \sum x^2_i$

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

$X \sim N(\mu, \Sigma)$

Multivariate Gaussian Distribution

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

- Multivariate Gaussian distributions need **μ**, **σ**^{2} and the **covariance** ∑

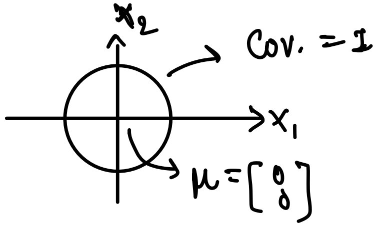
- **Covariance matrix:** captures **σ**^{2} and ∑

Covariance

- Measure of joint probability of 2 random variables. → Measure of correlation.

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

$$\Sigma = \begin{bmatrix} \sigma^2_{x_1x_1} & \sigma_{x_1x_2} \\ \sigma^2_{x_2x_1} & \sigma^2_{x_2x_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

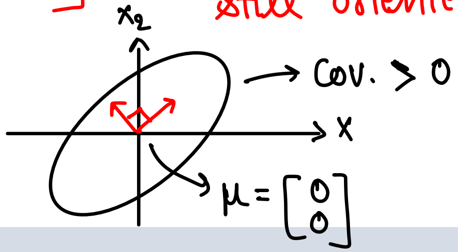


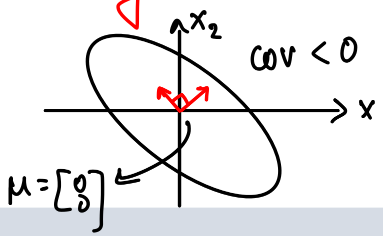
The covariance matrix (Σ) describes the shape of multivariate Gaussian dist.

The simplest covariance matrix is an identity matrix. This yields a circular Gaussian dist. in 2 dims. or a hypersphere in 3 dims. where each component has var. = 1

Eg: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$\begin{bmatrix} 1.5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ → If you scale the individual components, this will cause the dist. to be ellipsoid, but still oriented along the axes.





⟹ If some of the off-diagonal values are non-zero, this covariance b/w components skews the orientation of the Gaussian so that it no longer oriented along the axes. The off-diagonal cov. values tells that some components are non-independent, they vary w.r.to each other. The mean (μ) defines the offset of the whole dist. shifting the whole thing in space.

Eg: $\begin{bmatrix} 1.5 & 0 & 0.3 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ ⟹ $\mu \neq 0$

Covariance

- Measure of joint probability of 2 random variables.

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

j	k	
X1	X2	
12	10	
8	7	i
16	13	

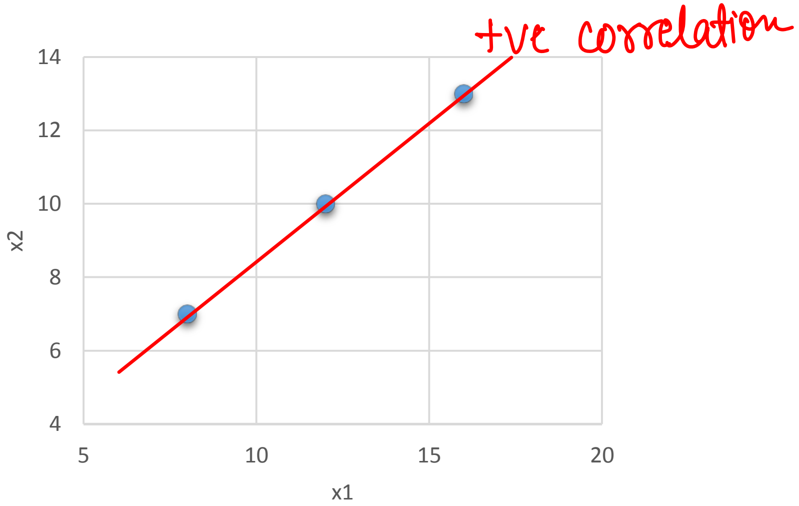
mean = 12 mean = 10

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

Covariance

x1	x2
12	10
8	7
16	13

12 10

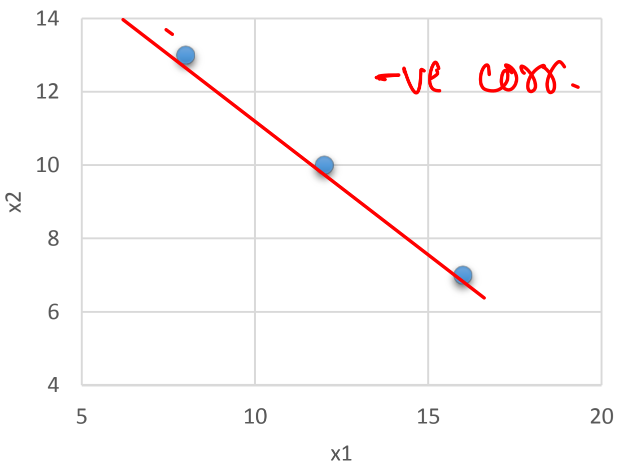


Cov(X1, X2) = 8

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

x1	x2
12	10
8	13
16	7

12 10

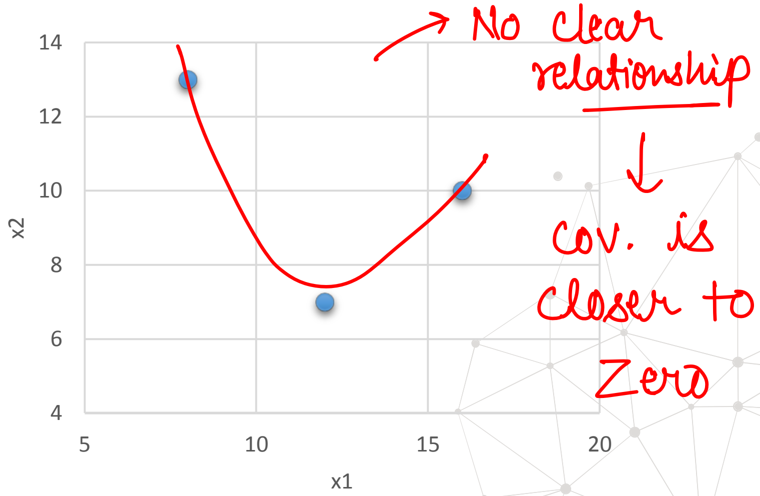


Cov(X1, X2) = -8

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

x1	x2
12	7
8	13
16	10

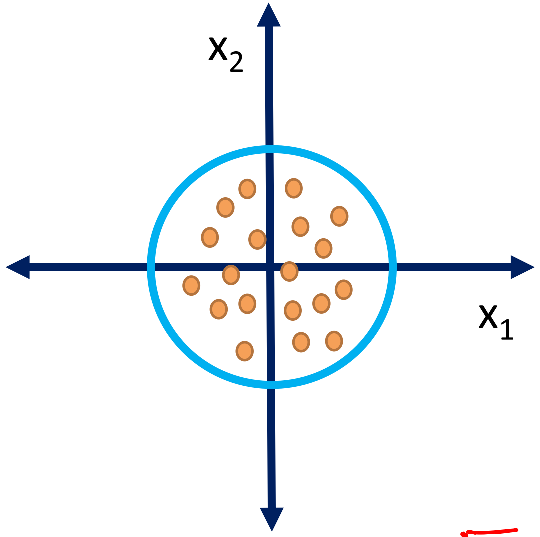
12 10



Cov(X1, X2) = -4

$$\frac{(12 - 12) * (7 - 10) + (8 - 12) * (13 - 10) + (16 - 12) * (10 - 10)}{3} = -4$$

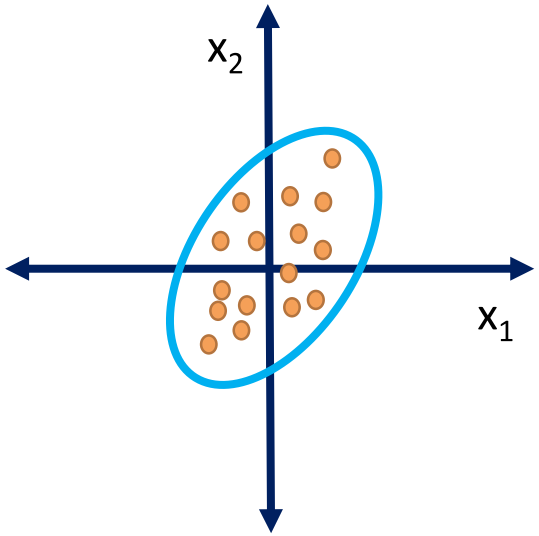
Covariance



- Cov(x1,x2) = 0

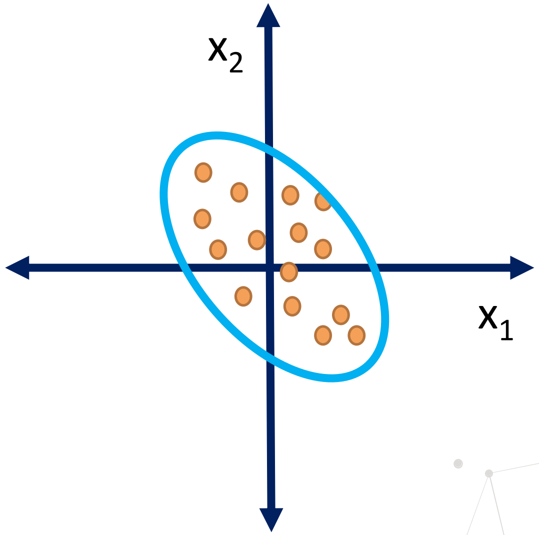
$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

- X1 and x2 are not correlated



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

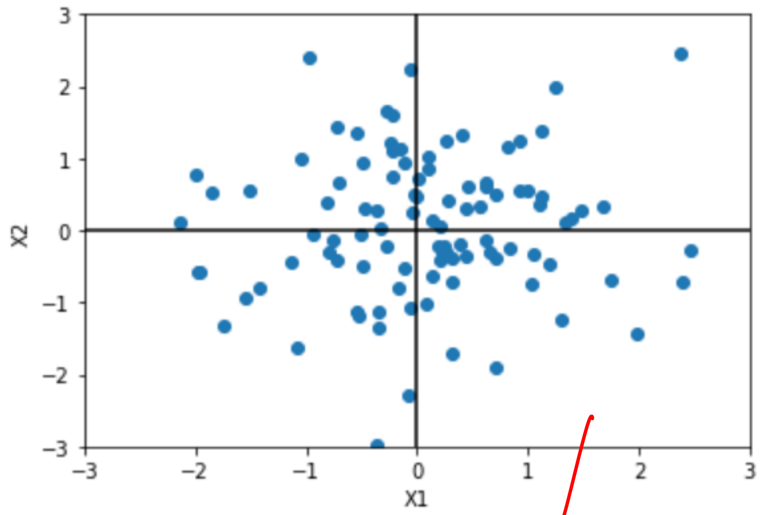
$x_1 \propto x_2$

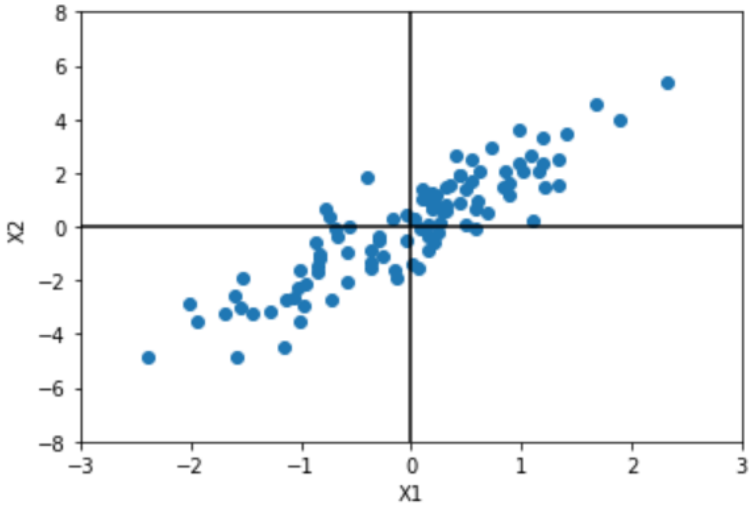


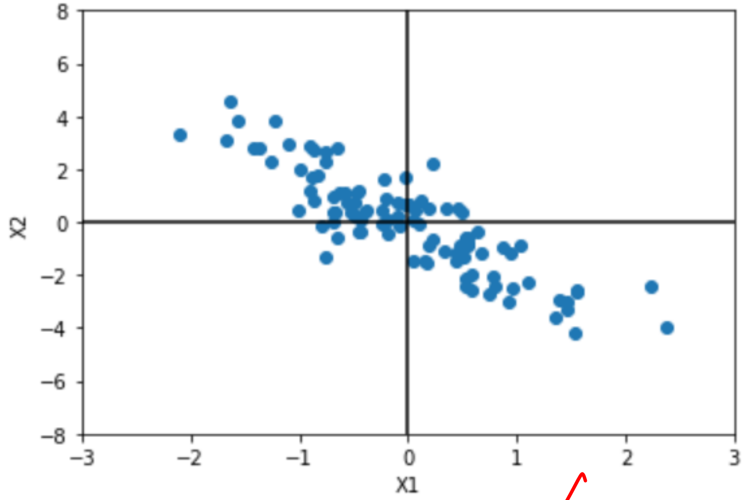
- Cov(x1,x2) < 0
- The bigger x1, the smaller x2

$x_1 \propto \frac{1}{x_2}$

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 = cov.

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

1.7016981044990922

Covariance Matrix

$$\Sigma = \begin{bmatrix} \sigma^2_{x_1 x_1} & \sigma^2_{x_1 x_2} \\ \sigma^2_{x_2 x_1} & \sigma^2_{x_2 x_2} \end{bmatrix}$$

- 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

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

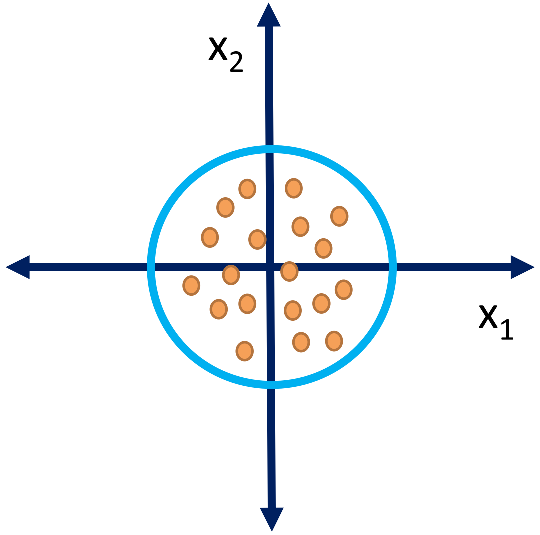
Where:

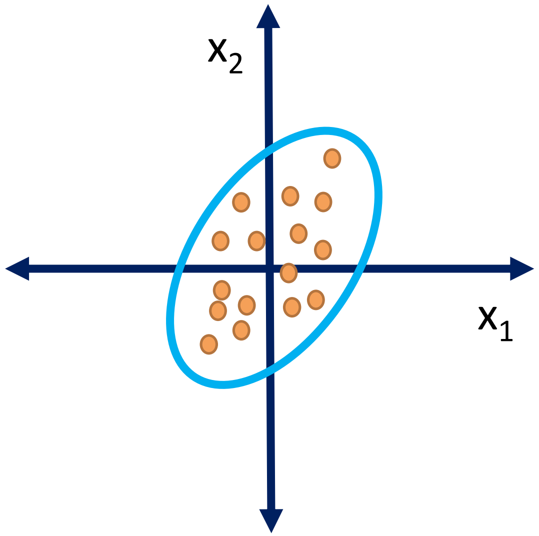
- Kx1x1 = var(x1)
- Kx2x2 = var(x2)
- Kx1x2 = K^Tx2x1 = cov(x1, x2)

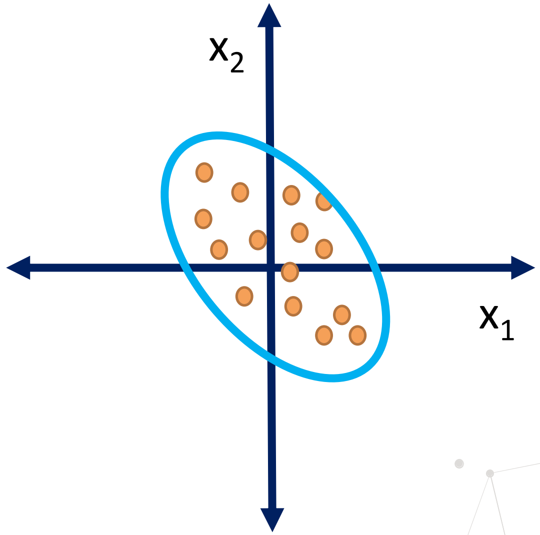
Covariance Matrix - General

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

Covariance Matrix





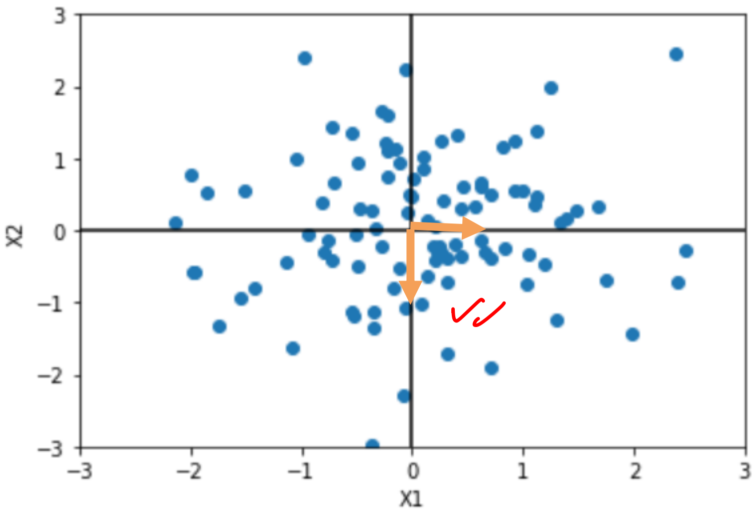


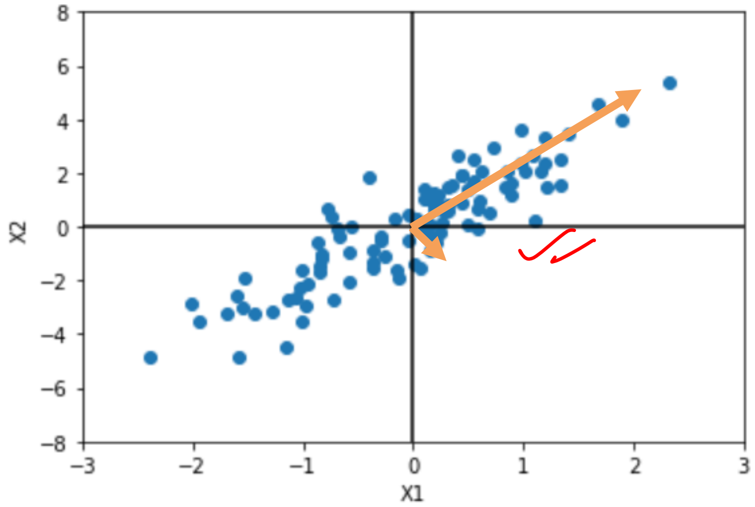
$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

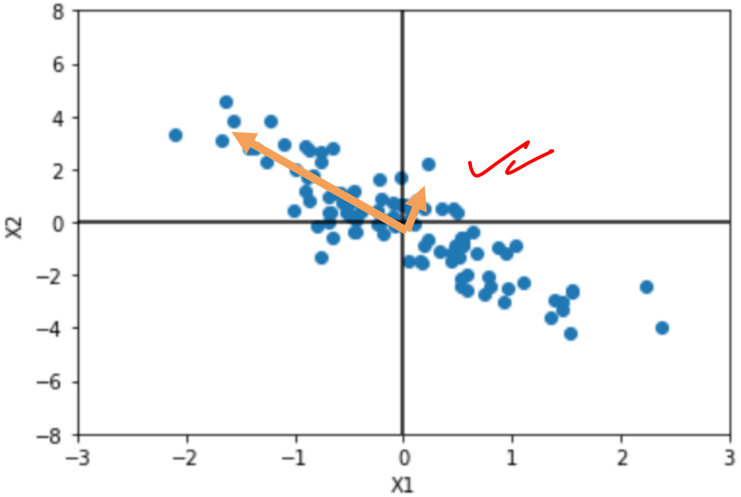
$$\Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

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

Covariance Matrix: simulation








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

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

$$= \begin{bmatrix} \sigma^2_{x_1 x_1} & \sigma^2_{x_1 x_2} \\ \sigma^2_{x_2 x_1} & \sigma^2_{x_2 x_2} \end{bmatrix}$$

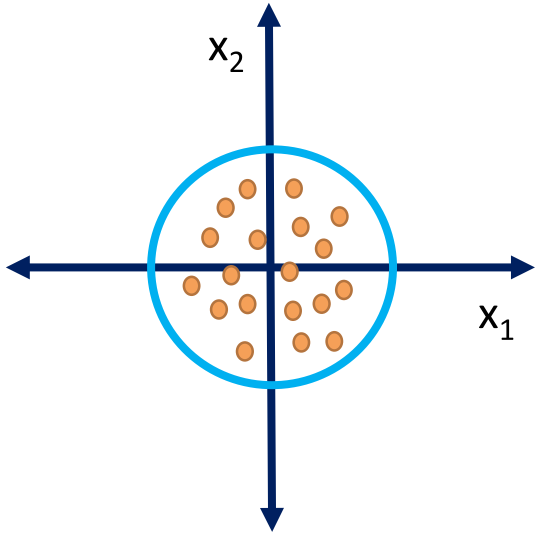
```
np.cov(x1, x2)
```

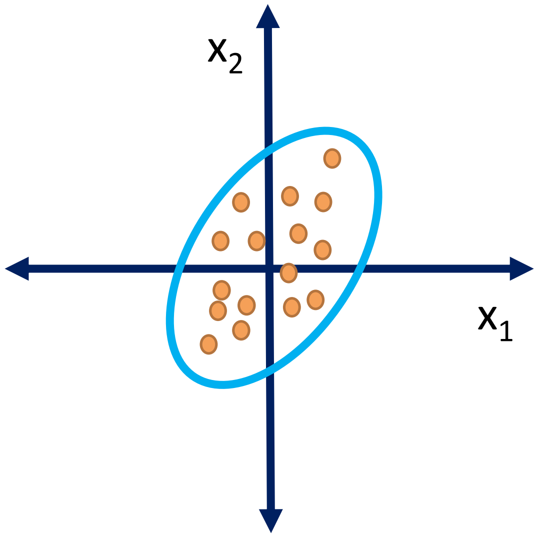
```
array([[0.84262922, 1.71888697],
       [1.71888697, 4.38490628]])
```

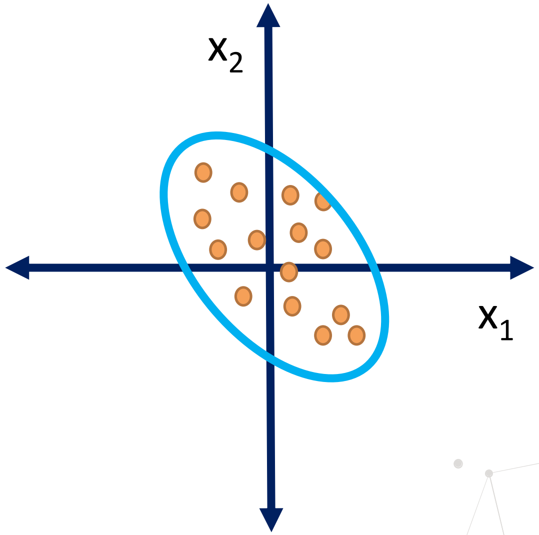
```
np.cov(x1, x2)
```

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

Multivariate Gaussian Distribution







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

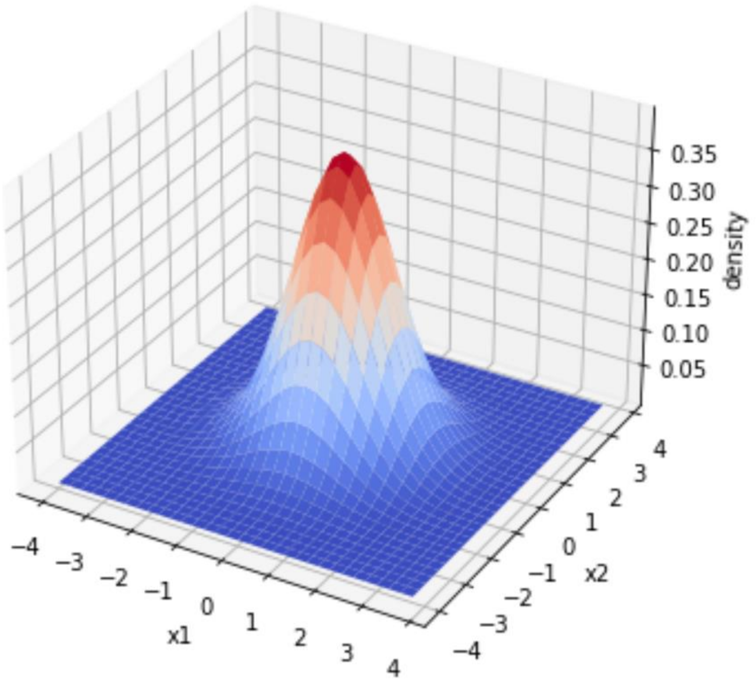
$X \sim N(\mu, \Sigma)$

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

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

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