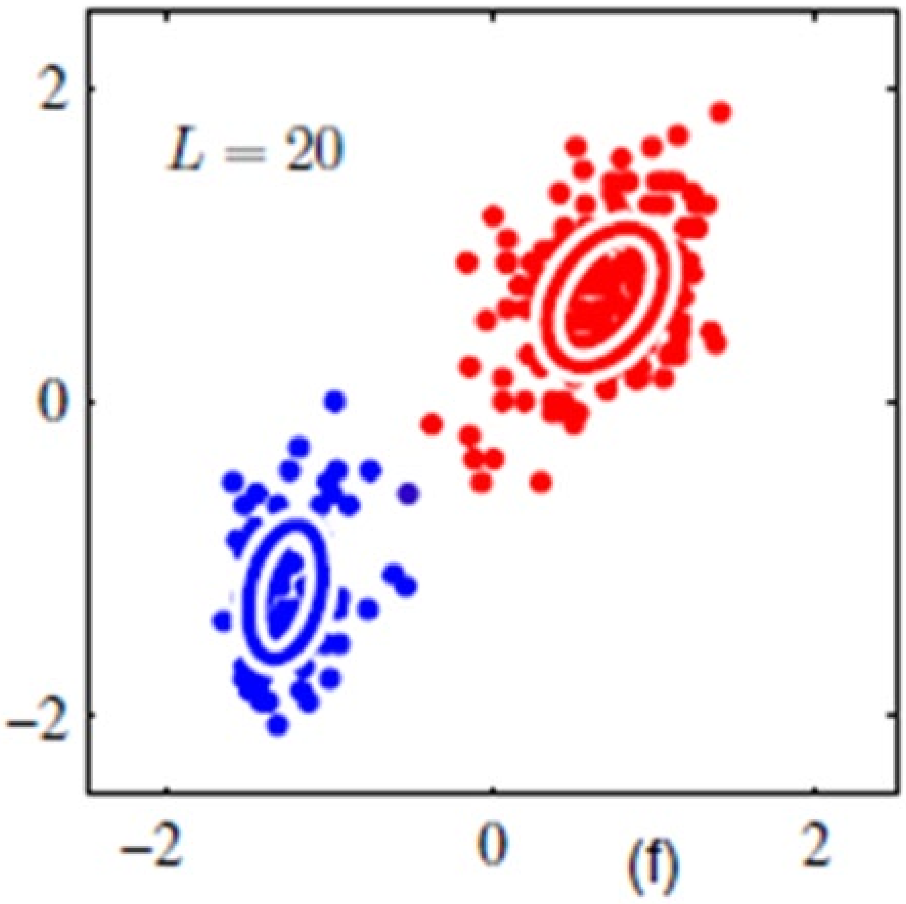
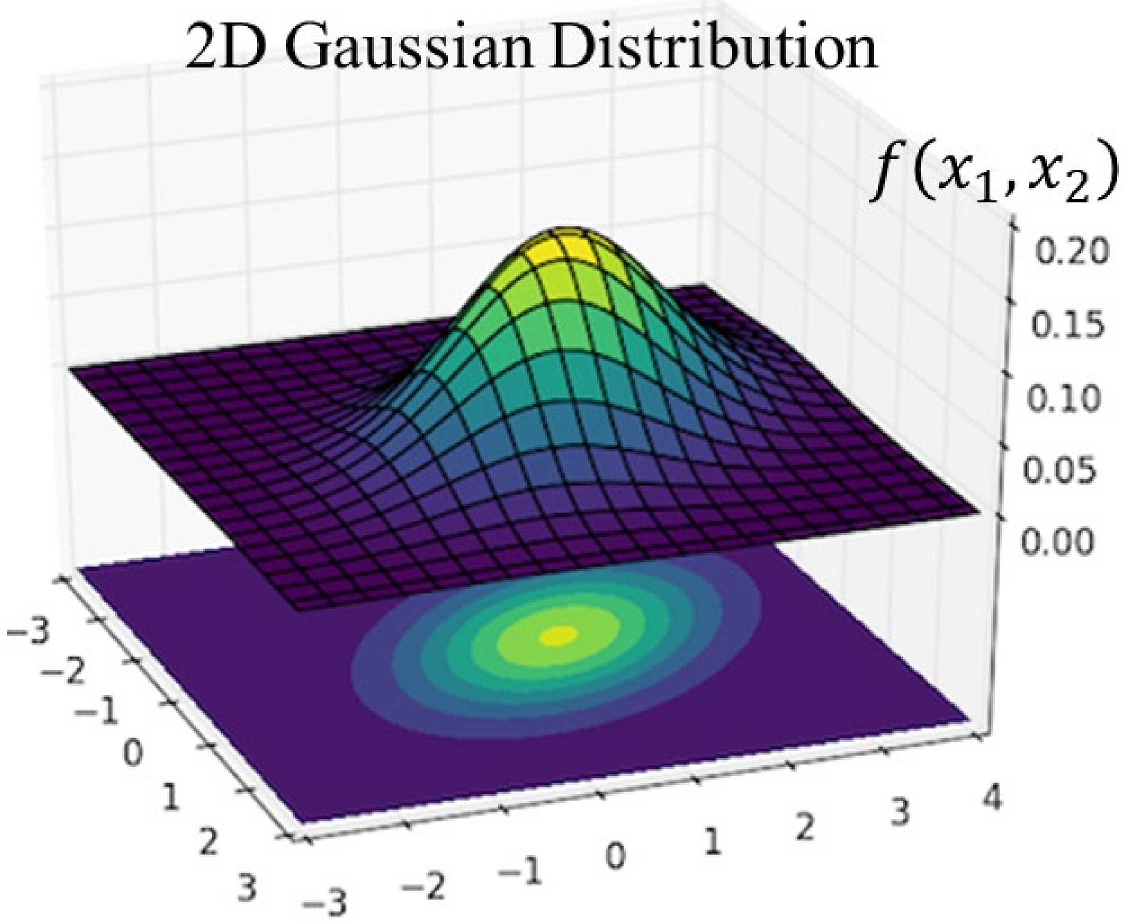
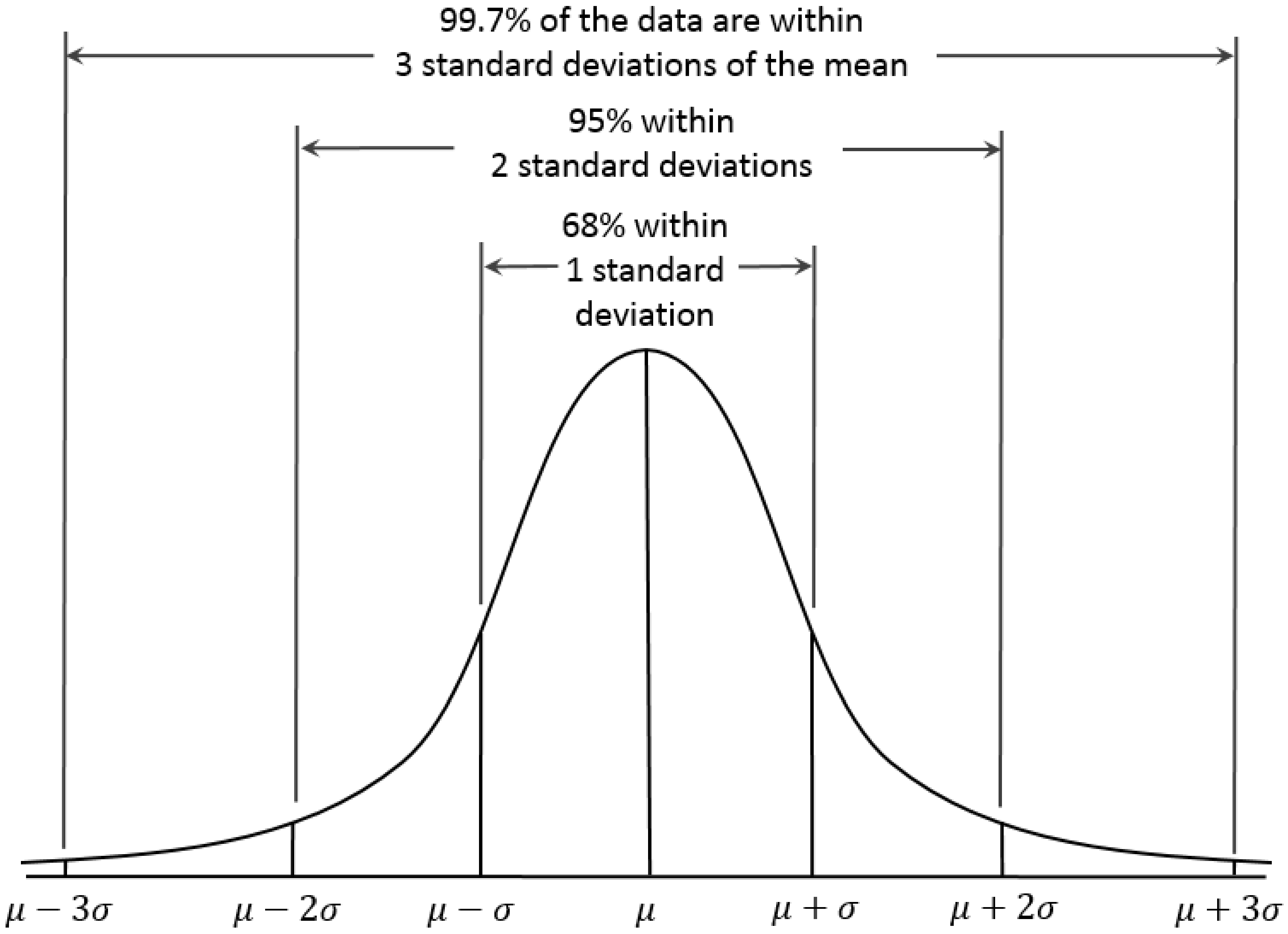
Visualize 2D Gaussian PDF

Why and how do we use a circle or ellipse to represent a 2D Gaussian function/PDF ?





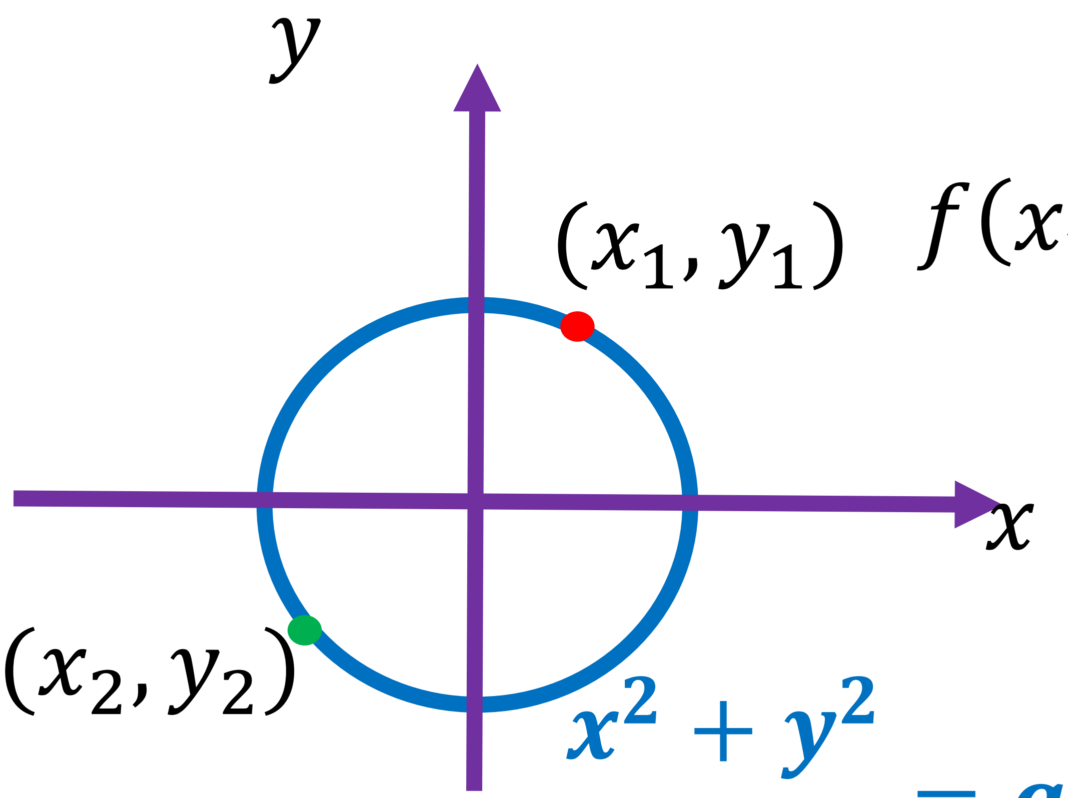
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



https://en.wikipedia.org/wiki/Normal_distribution#/media/File:Empirical_Rule.PNG

A 2D Gaussian PDF

$$f(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$



$$f(x_1,y_1) = \frac{1}{2\pi\sigma^2} e^{-\boldsymbol{a}}$$

$$f(x_2,y_2) = \frac{1}{2\pi\sigma^2} e^{-\boldsymbol{a}}$$

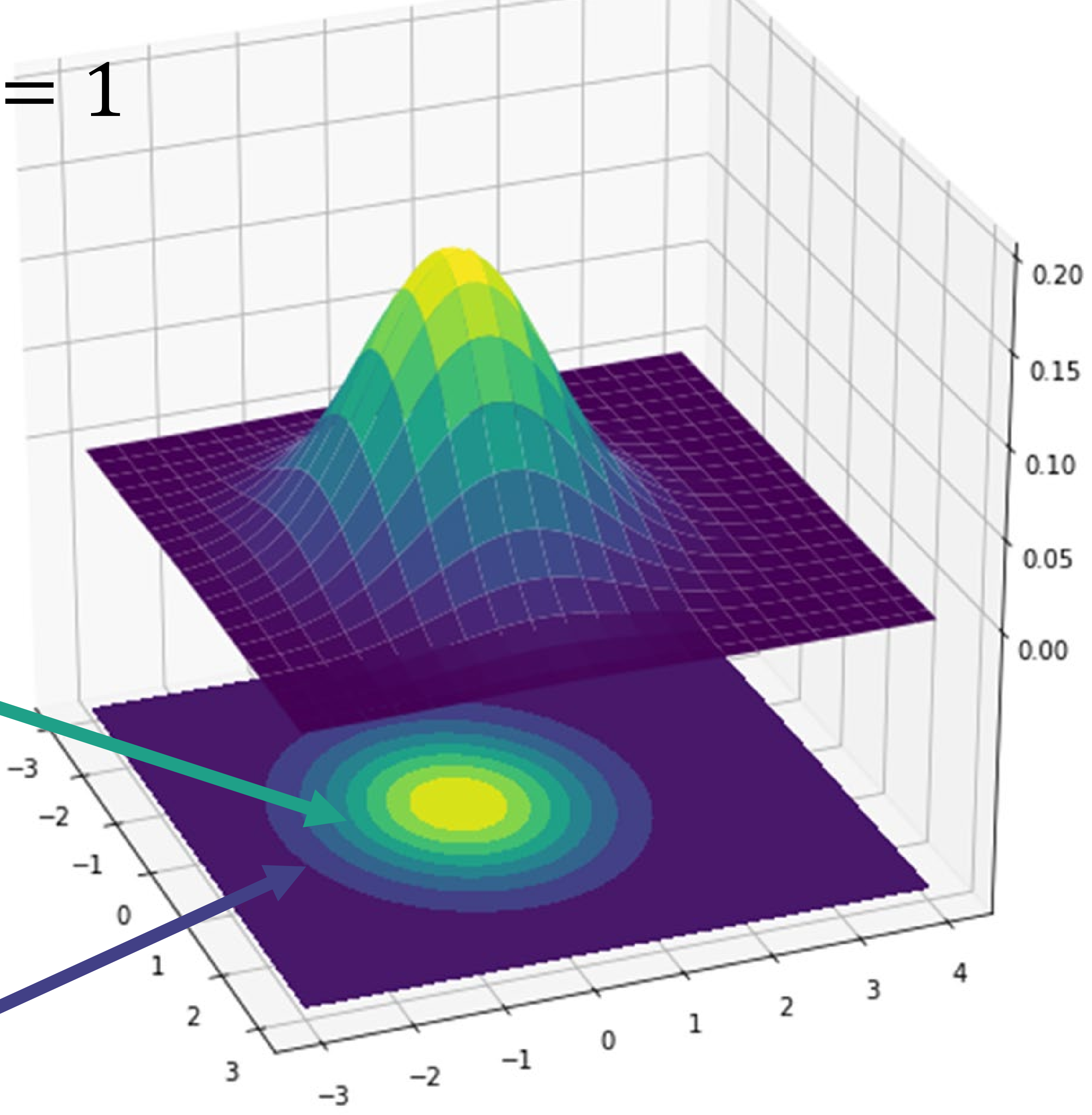
$$\frac{\boldsymbol{x^2+y^2}}{\boldsymbol{2\sigma^2}} = \boldsymbol{a}$$

$\boldsymbol{a}$ is a number (positive)

PDF values at all points on the circle are the same (level curve)

A 2D Gaussian PDF $\sigma = 1$

$$f(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$$

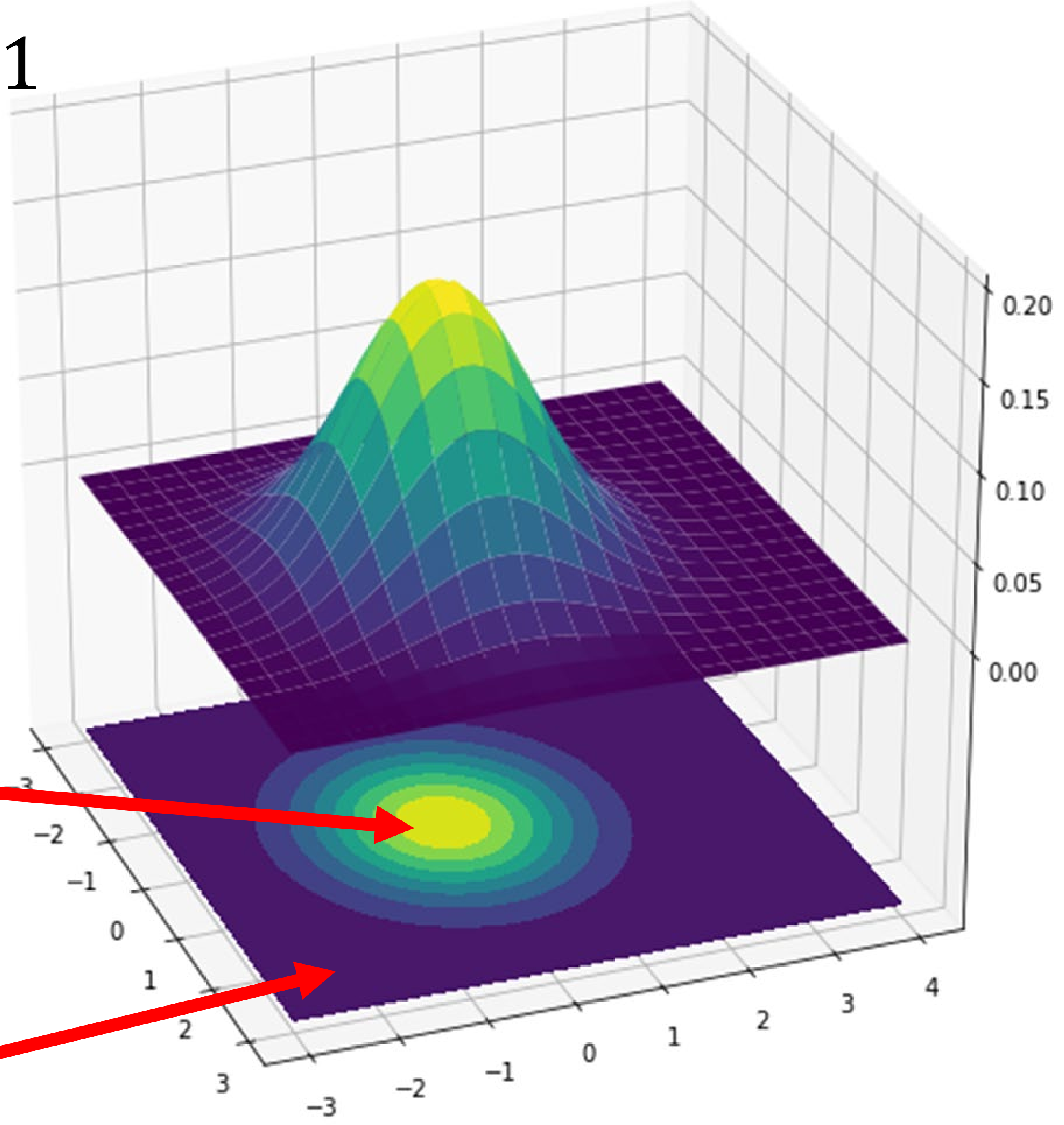


$$\boxed{\frac{x^2 + y^2}{2} = 1}$$

$$\boxed{\frac{x^2 + y^2}{2} = 4}$$

A 2D Gaussian PDF $\sigma = 1$

$$f(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$$

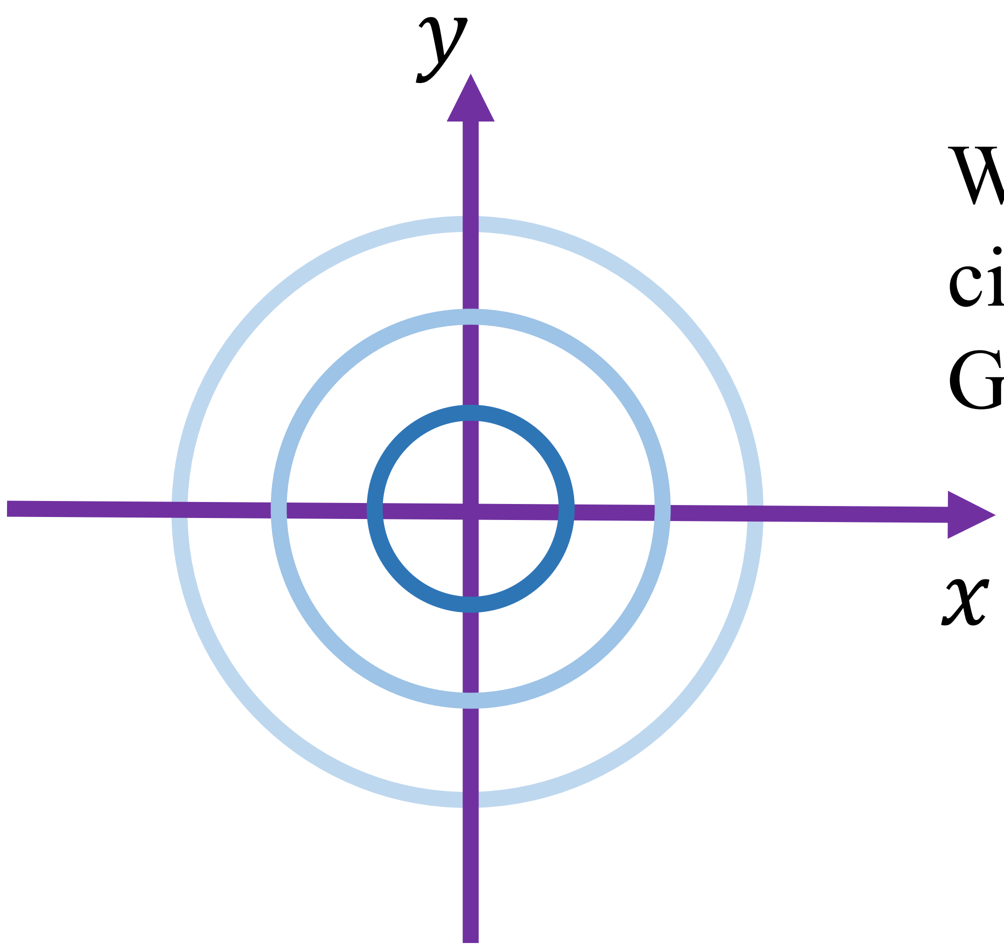


High (probability) density

Low (probability) density

A 2D Gaussian PDF

$$f(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

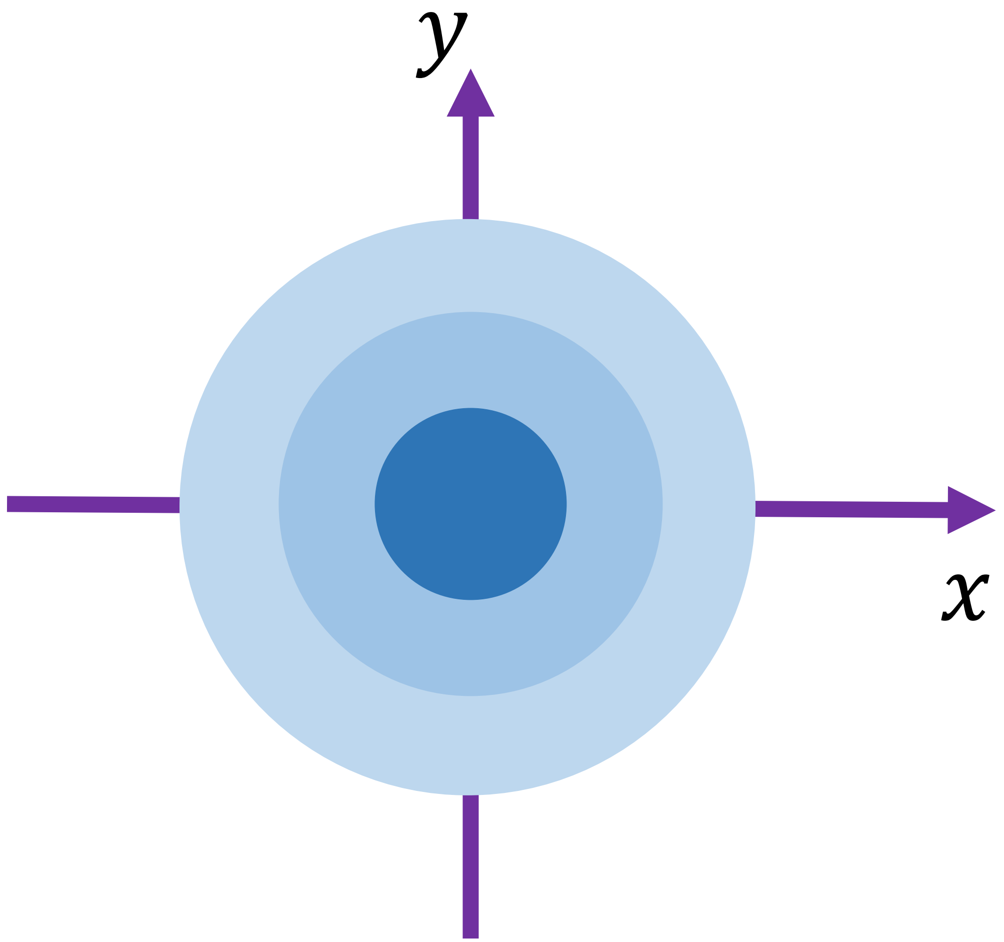


We can use several circles to represent a Gaussian PDF

PDF values at all points on a circle are the same (level curve)

A 2D Gaussian PDF

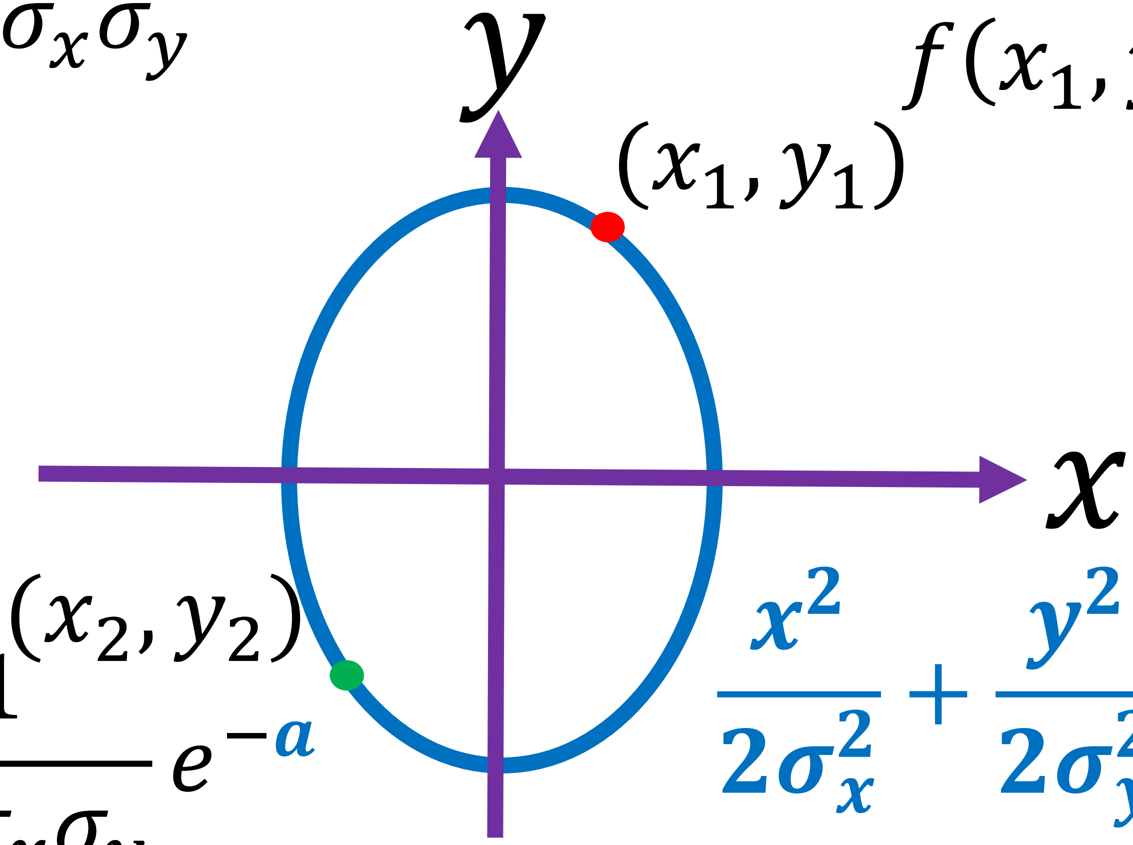
$$f(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$



A 2D Gaussian PDF – ellipse-shaped

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{x^2}{2\sigma_x^2}+\frac{y^2}{2\sigma_y^2}\right)}$$

$$f(x_1,y_1) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\boldsymbol{a}}$$



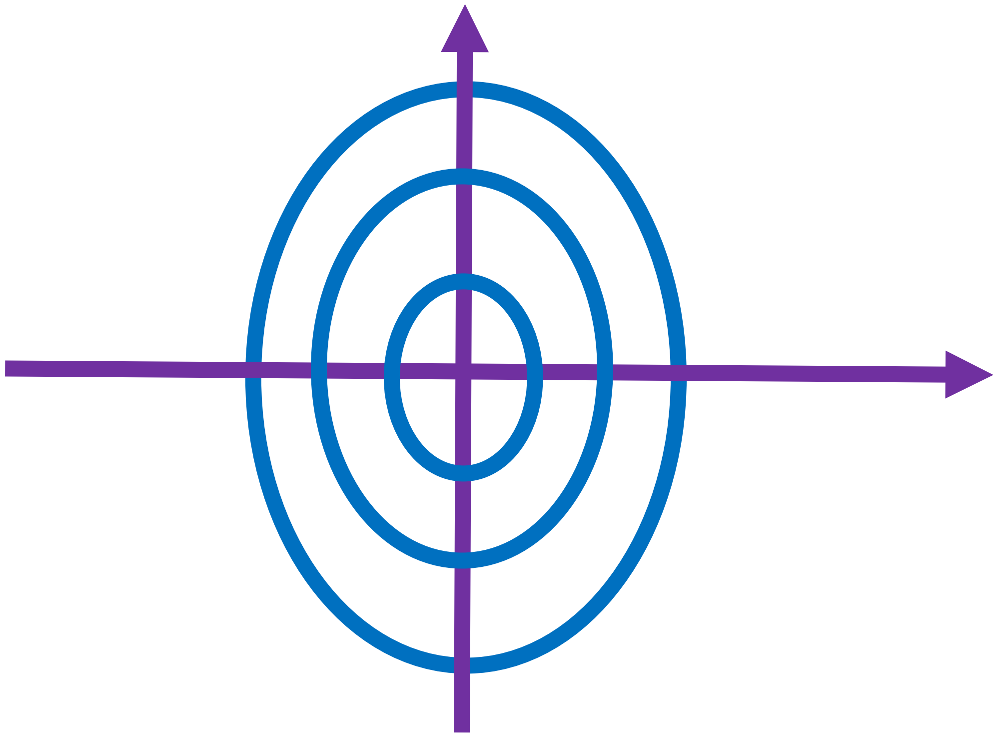
$$f(x_2,y_2) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\boldsymbol{a}}$$

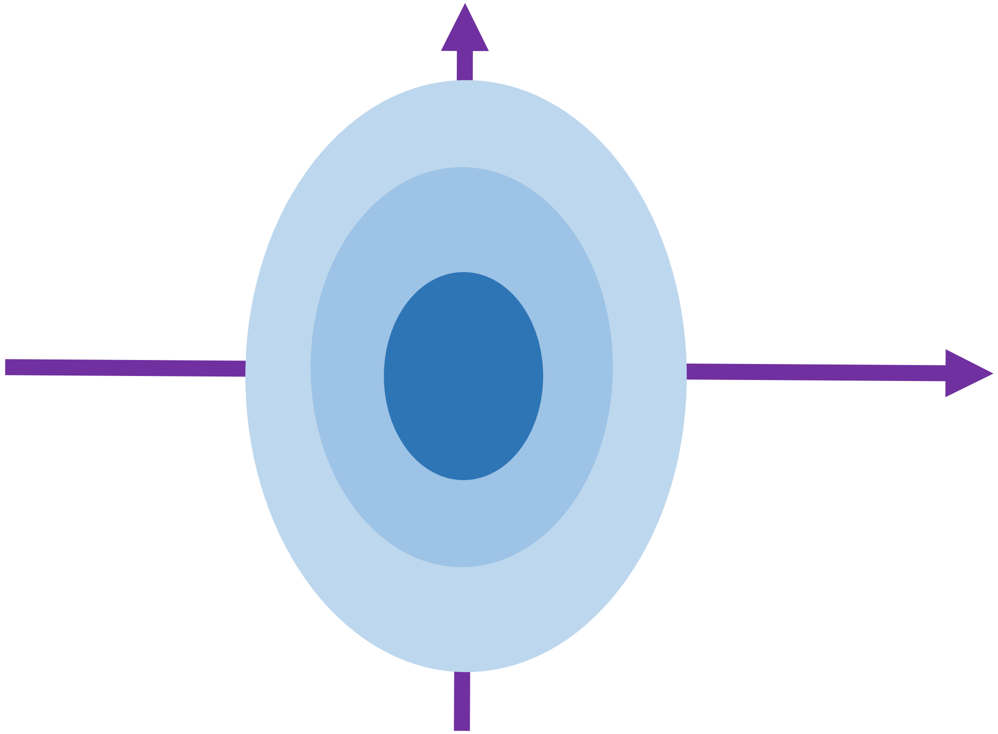
$$\boldsymbol{\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2} = \mathrm{a}}$$

PDF values at all points on the ellipse are the same (level curve)

A 2D Gaussian PDF – ellipse-shaped

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{x^2}{2\sigma_x^2}+\frac{y^2}{2\sigma_y^2}\right)}$$





$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\left(\frac{x_1^2}{2\sigma_1^2} + \frac{x_2^2}{2\sigma_2^2}\right)}$$

The general form of a M-D Gaussian

$$f(x) = \frac{1}{(2\pi)^{M/2}|\Sigma|^{1/2}} exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$

$$M{=}2,\ x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}, |\Sigma|^{1/2} = \sqrt{\sigma_1^2\sigma_2^2} = \sigma_1\sigma_2$$

$$\frac{1}{2}x^T\Sigma^{-1}x = \frac{1}{2}\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \sigma_1^{-2} & 0 \\ 0 & \sigma_2^{-2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{x_1^2}{2\sigma_1^2} + \frac{x_2^2}{2\sigma_2^2}$$

“Rotate” an ellipse-shaped 2D Gaussian Function

$$f(x_1, x_2) = \frac{1}{2\pi|\Sigma|^{1/2}} exp\left\{-\frac{1}{2}[x_1, x_2]\Sigma^{-1}\begin{bmatrix}x_1\\x_2\end{bmatrix}\right\}$$

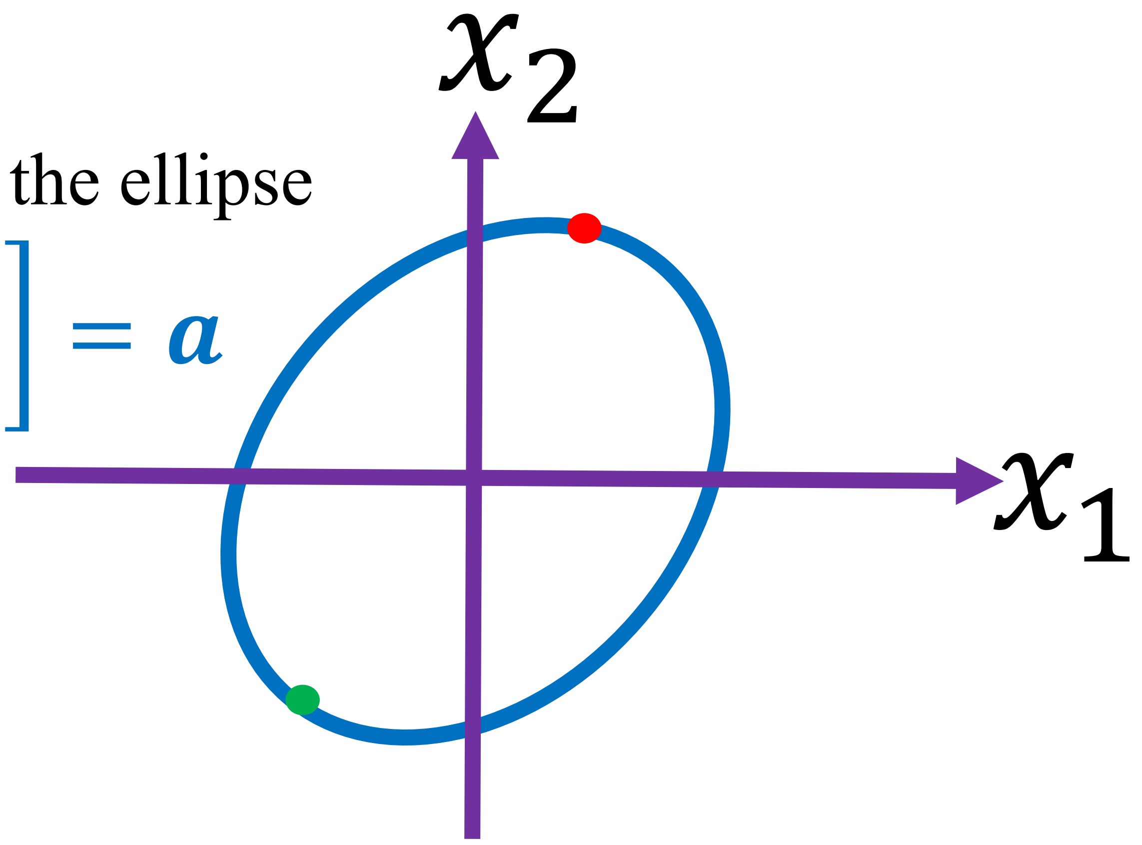
Note:

x_1 refers to the 1st feature component

x_2 refers to the 2nd feature component

This equation defines the ellipse

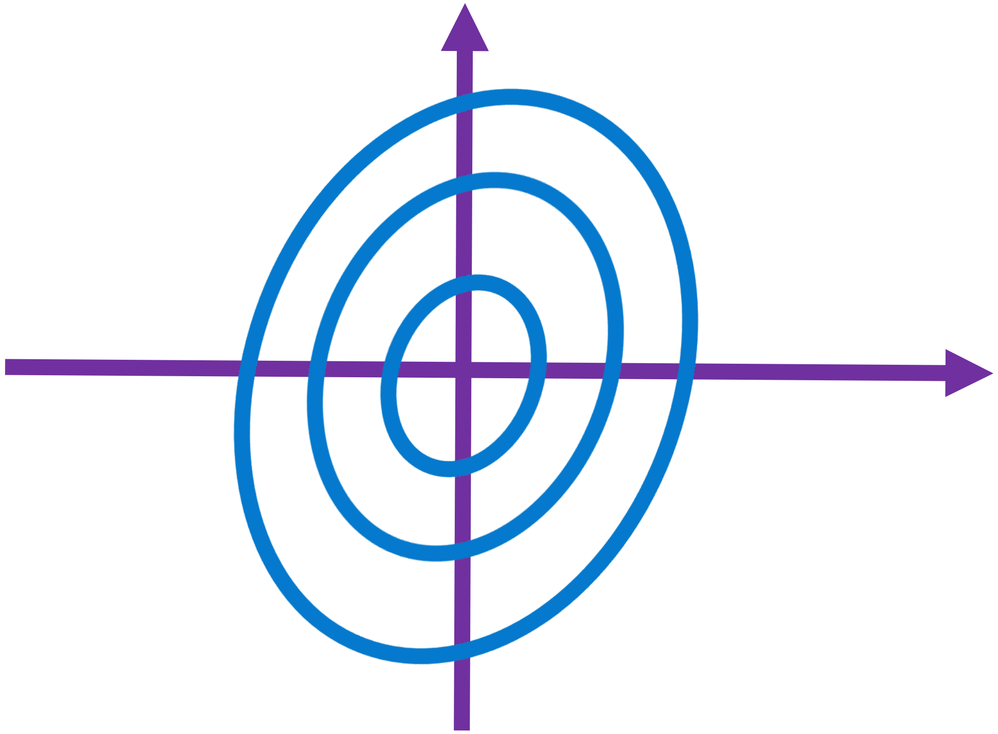
$$[\boldsymbol{x_1}, \boldsymbol{x_2}]\boldsymbol{\Sigma}^{-1}\begin{bmatrix}\boldsymbol{x_1}\\\boldsymbol{x_2}\end{bmatrix} = \boldsymbol{a}$$

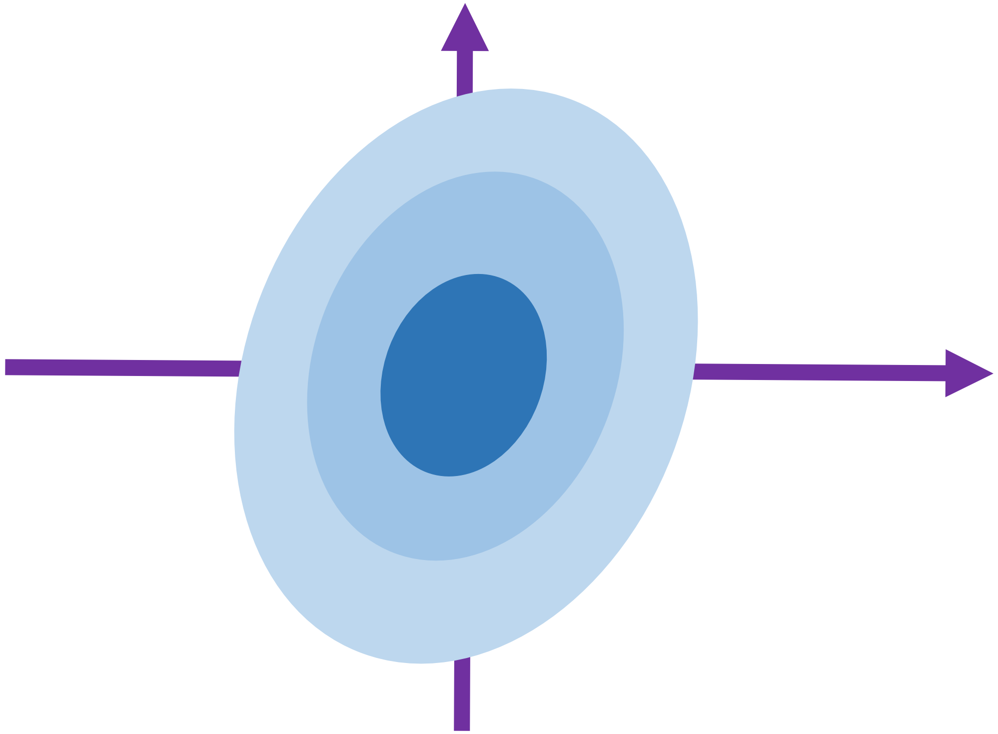


PDF values at all points on the ellipse are the same (level curve)

"Rotate" an ellipse-shaped 2D Gaussian Function

$$f(x_1, x_2) = \frac{1}{2\pi|\Sigma|^{1/2}} exp\left\{-\frac{1}{2}[x_1, x_2]\Sigma^{-1}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right\}$$

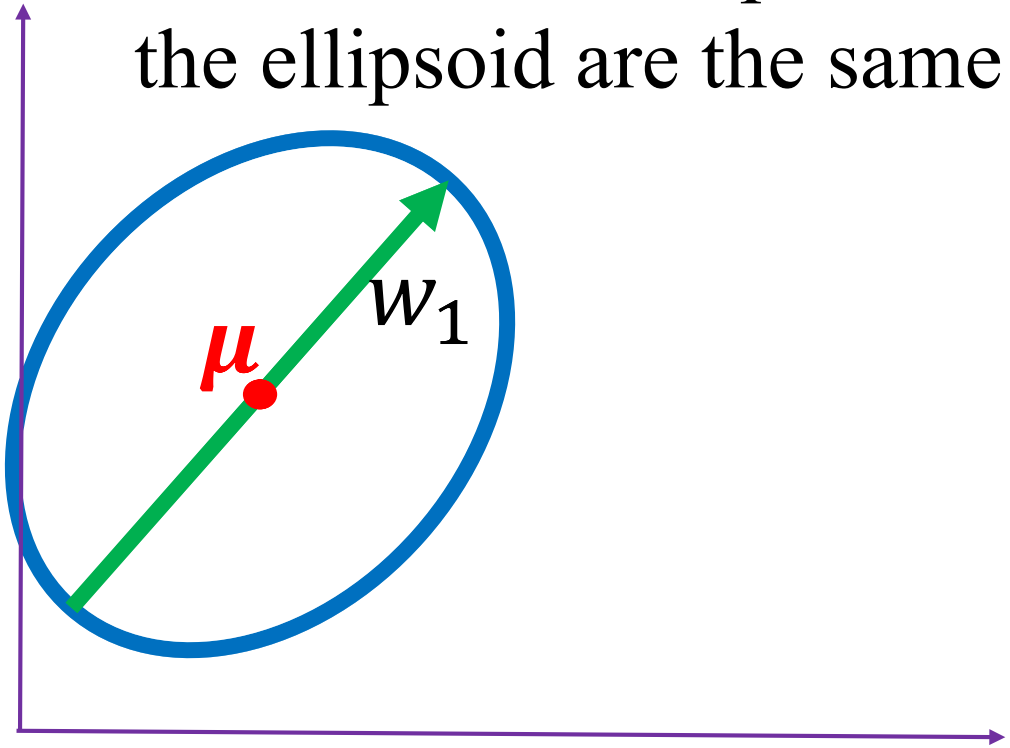




This equation defines a high dimensional ellipsoid (surface)

$$(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})=\boldsymbol{a}$$

PDF values at all points on the ellipsoid are the same



$\boldsymbol{x}$ is a high-dimensional vector

$\boldsymbol{\mu}$ is the center

$\boldsymbol{a}$ is a number (positive)

$\boldsymbol{\Sigma}$ is the covariance matrix

λ_1 is the largest eigenvalue of $\boldsymbol{\Sigma}$

$\boldsymbol{w}_1$ is the corresponding eigenvector, direction of the major axis

$\sqrt{\lambda_1}$ is proportional to the length of the major axis