CS 224 - Multivariate Gaussian Distribution

Fall 2024

Goal

- Basics about multivariate Gaussian distribution.
- Visualizations for better understanding.

Univariate Gaussian Distribution

$x \in \mathbb{R}$, the density function is given by

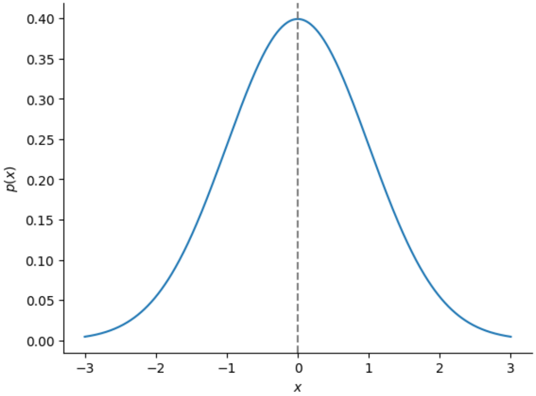
$$p(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right).$$

- μ: mean; σ^2: variance

Examples: Effect of μ

μ represents the **mean** of the data.

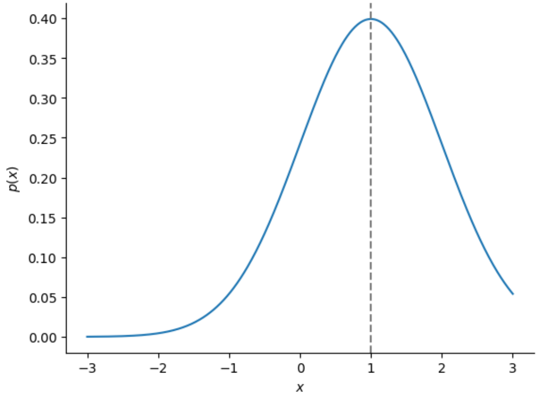
$$\mu = 0, \ \sigma = 1$$



Examples: Effect of μ

μ represents the **mean** of the data.

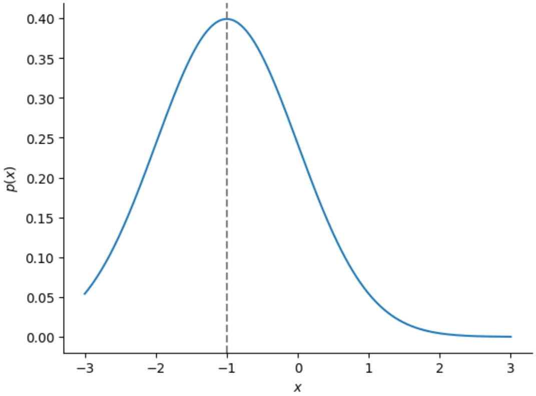
$$\mu = 1, \ \sigma = 1$$



Examples: Effect of μ

μ represents the **mean** of the data.

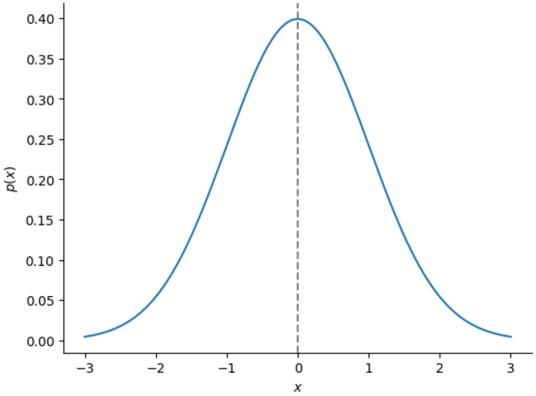
$$\mu = -1,\ \sigma = 1$$



Examples: Effect of σ

σ represents the **spread** of data from the mean.

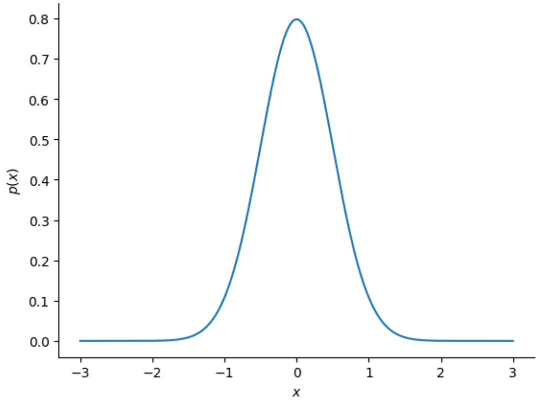
$$\mu = 0, \;\; \sigma = 1$$



Examples: Effect of σ

σ represents the **spread** of data from the mean.

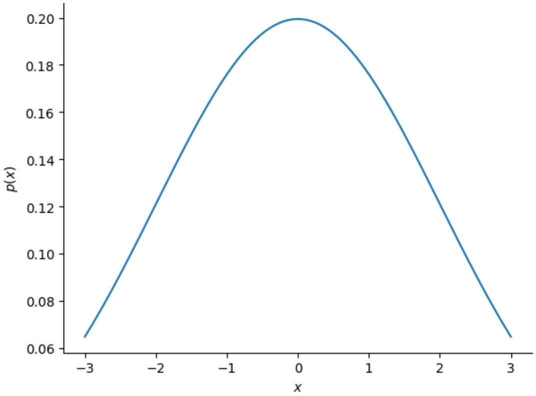
$$\mu = 0, \;\; \sigma = 0.5$$



Examples: Effect of σ

σ represents the **spread** of data from the mean.

$$\mu = 0, \ \sigma = 2$$



Covariance

The covariance of two random variable X and Y is defined as

$$\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y].$$

- It measures the degree to which X and Y are (linearly) related.

Covariance Matrix

If **x** is a D-dimensional random vector, its covariance matrix is defined as

$$\mathrm{Cov}[\mathbf{x}] = \Sigma = \begin{pmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] & \cdots & \mathrm{Cov}[X_1, X_D] \\ \mathrm{Cov}[X_2, X_1] & \mathrm{Var}[X_2] & \cdots & \mathrm{Cov}[X_2, X_D] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}[X_D, X_1] & \mathrm{Cov}[X_D, X_2] & \cdots & \mathrm{Var}[X_D] \end{pmatrix}$$

Alternatively, For any random vector **x** with mean μ and covariance matrix Σ,

$$\Sigma = E[(\mathbf{x} - \mu)(\mathbf{x} - \mu)] = E[(\mathbf{x}\mathbf{x}^\top] - \mu\mu^\top)$$

- Σ is symmetric, positive semi definite matrix.

Multivariate Gaussian Distribution

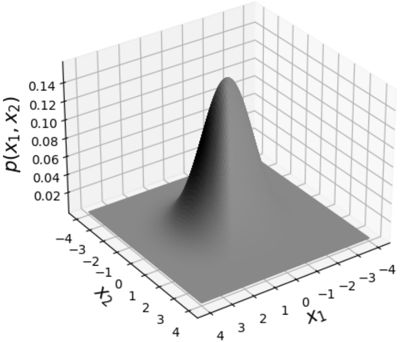
$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{k}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

- $\boldsymbol{\mu}$: mean; Σ^2: covariance matrix
- Restriction: Σ must be symmetric positive definite (in order for Σ^{-1} to exist).

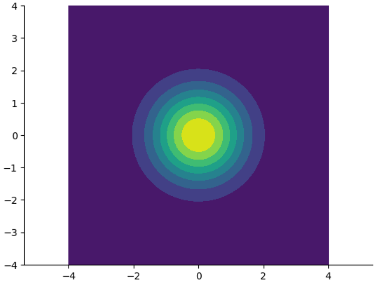
Examples: Effect of μ

μ: the **mean** of the data.

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



Surface

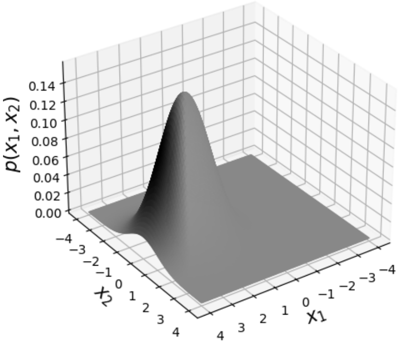


Contour

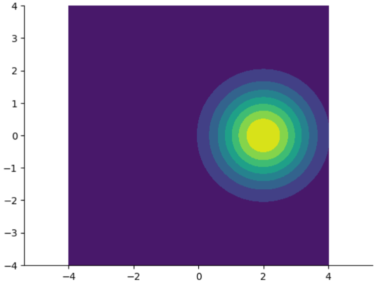
Examples: Effect of μ

μ: the **mean** of the data.

$$\mu = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \; \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



Surface

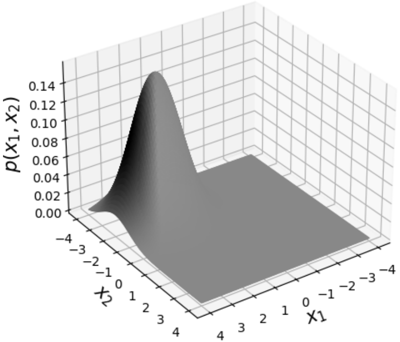


Contour

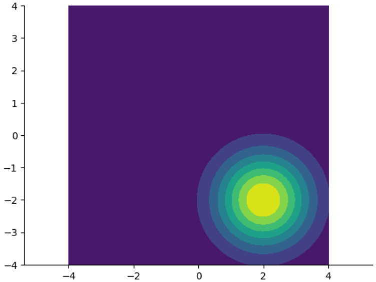
Examples: Effect of μ

μ: the **mean** of the data.

$$\mu = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \; \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



Surface

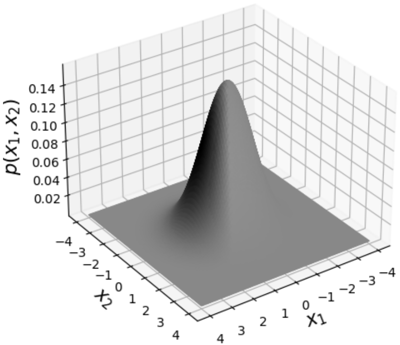


Contour

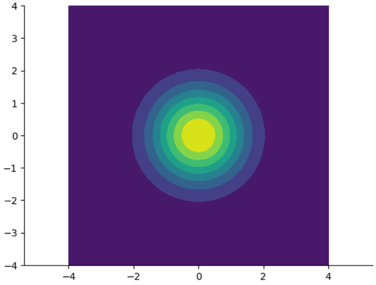
Examples: Effect of Σ

Σ represents the **spread** of data from the mean.

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



Surface

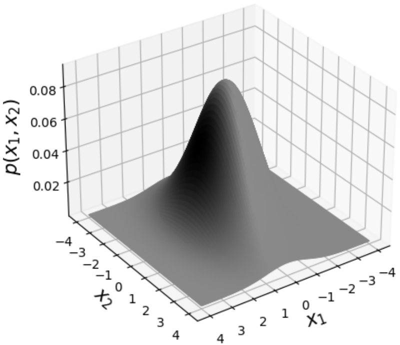


Contour

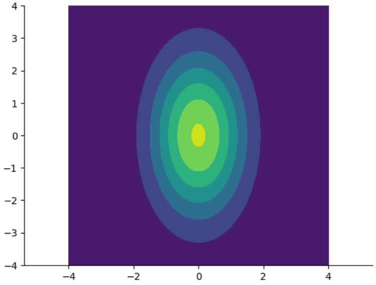
Examples: Effect of Σ

Σ represents the **spread** of data from the mean.

$$\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$



Surface

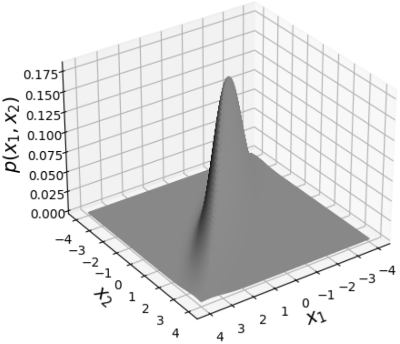


Contour

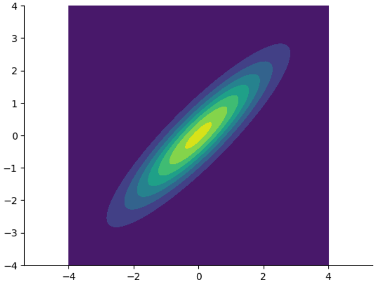
Examples: Effect of Σ

Σ represents the **spread** of data from the mean.

$$\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ \Sigma = \begin{bmatrix} 2 & 1.8 \\ 1.8 & 2 \end{bmatrix}$$



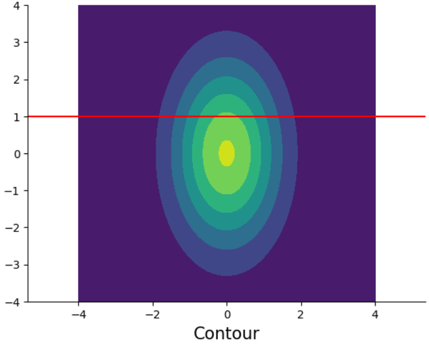
Surface

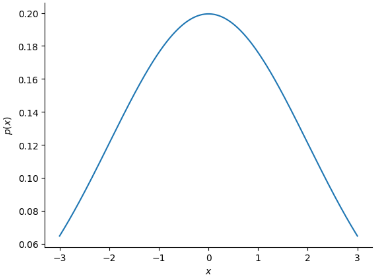


Contour

Conditioning a 2d Gaussian

The conditional $p(x_1|x_2)$ is obtained by "slicing" the joint distribution through the $X_2 = x_2$ line.





Resources

- Chuong B. Do. *The Multivariate Gaussian Distribution*. Lecture Notes. 2008
- Kevin P. Murphy. *Probabilistic Machine Learning: An introduction*. MIT Press, 2022