# CSE 6363 - Machine Learning

## Gaussian Distribution

Alex Dillhoff

University of Texas at Arlington

# What is covered?

1. Univariate Gaussian
2. Multivariate Gaussian
3. Estimating parameters via MLE

# Univariate Gaussian Distribution

The univariate Gaussian is written as

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(x-\mu)^2),$$

where $\mu$ is the mean and $\sigma^2$ is the variance. The probability distribution function is plotted in the figure below.

# Univariate Gaussian Distribution

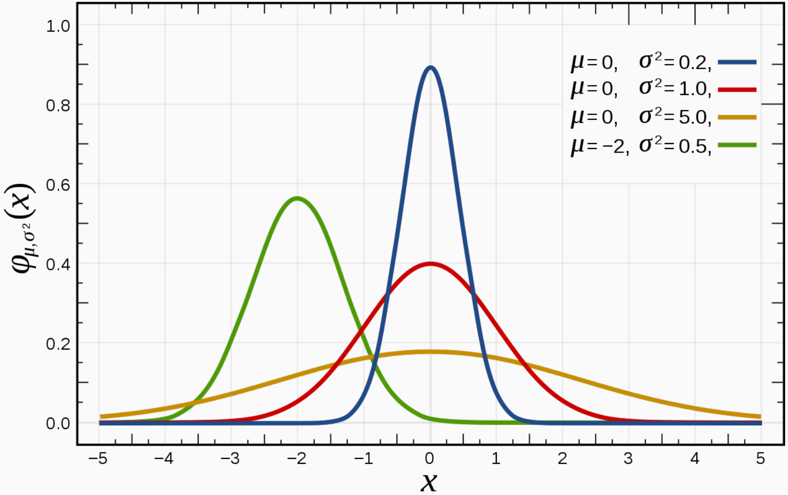


**Figure 1:** A plot of different Gaussians with varying mean and variance. Source: Wikipedia

# Univariate Gaussian Distribution

Changing the value of $\mu$ shifts the Gaussian either left or right.

As the variance increases, the pdf flattens. A small variance produces a pdf with a peak.

The pdf must integrate to 1.

# Univariate Gaussian Distribution

The **standard deviation** is simply the square of the variance: $\sigma = \sqrt{\sigma^2}$.

This quantity is useful as it represents the variance using the same units as the input $x$.

# Estimating Parameters

Given some dataset $X = x_1, \ldots, x_n$, we can estimate a Gaussian that best fits the data via Maximum Likelihood Estimation.

This requires maximizing the log-likelihood function.

# Estimating Parameters

The likelihood function is the probability of the data given the parameters, $p(X|\theta)$.

For a Gaussian, this is simply the pdf: $p(X|\theta) = \mathcal{N}(X|\mu, \sigma^2)$.

# Estimating Parameters

Taking the log of this function results in

$$\ln \mathcal{N}(X|\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

# Estimating Parameters

$$\mathcal{N}(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(x-\mu)^2),$$

$$\ln \mathcal{N}(X|\mu,\sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2.$$

# Estimating Parameters

Taking the derivative with respect to $\mu$ and $\sigma^2$ yields the maximum likelihood estimates

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2.$$

# Multivariate Gaussian Distribution

The multivariate Gaussian is written as

$$\mathcal{N}(\mathbf{x}|\mu, \Sigma) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \right\}$$

# Multivariate Gaussian Distribution

$\mu \in \mathbb{R}^D$ is the mean vector

$\Sigma \in \mathbb{R}^{D \times D}$ is the covariance matrix

$|\Sigma|$ is the determinant of $\Sigma$

# Multivariate Gaussian Distribution

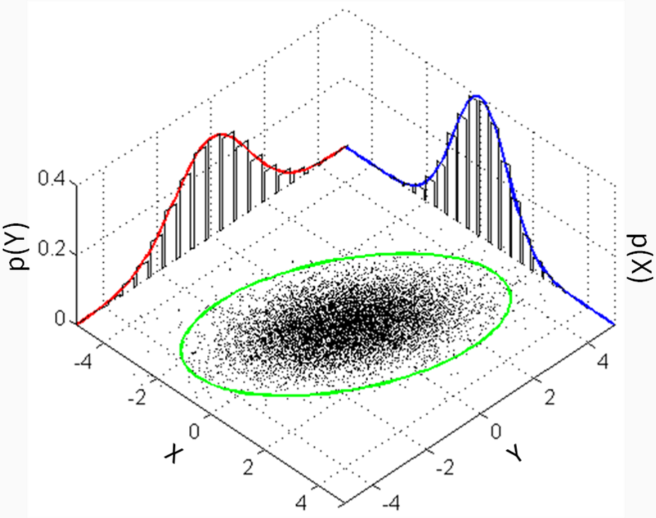


Figure 2: Multivariate normal with visualization of mean and covariance. Source: Wikipedia

# Estimating Parameters

Given a dataset $\mathsf{X} = \mathsf{x}_1, \ldots, \mathsf{x}_n$, we can again use MLE to estimate the parameters of the distribution.

For a multivariate distribution, the parameter estimates are given as follows

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \mathsf{x}_i$$

$$\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(\mathsf{x}_i - \hat{\mu})(\mathsf{x}_i - \hat{\mu})^T.$$