University of Amsterdam
Informatics Institute

AMLAB
Amsterdam
Machine Learning Lab

Machine Learning 1

Lecture 2.2 - Gaussian Distribution

Erik Bekkers

(Bishop 1.2.4)

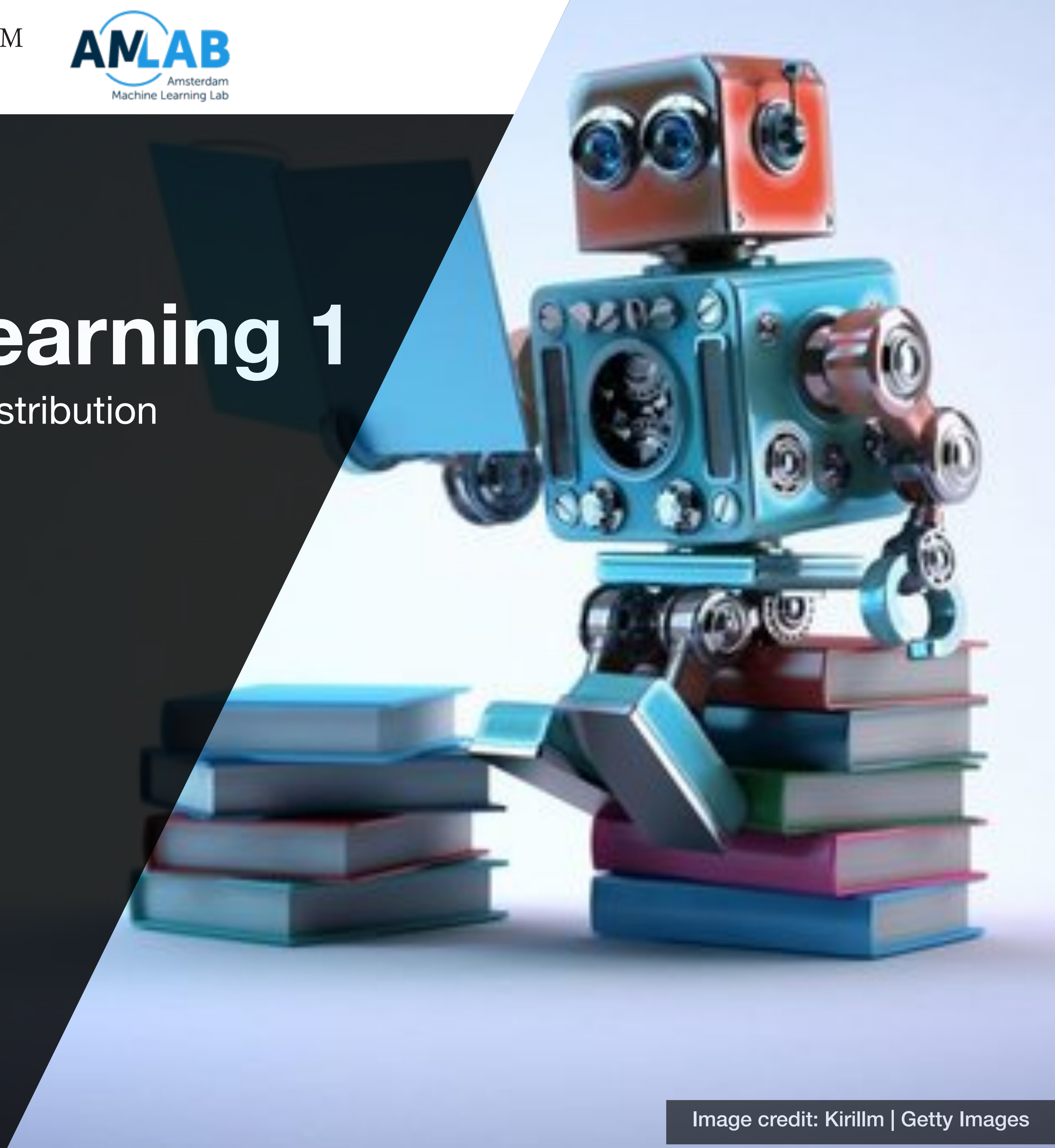
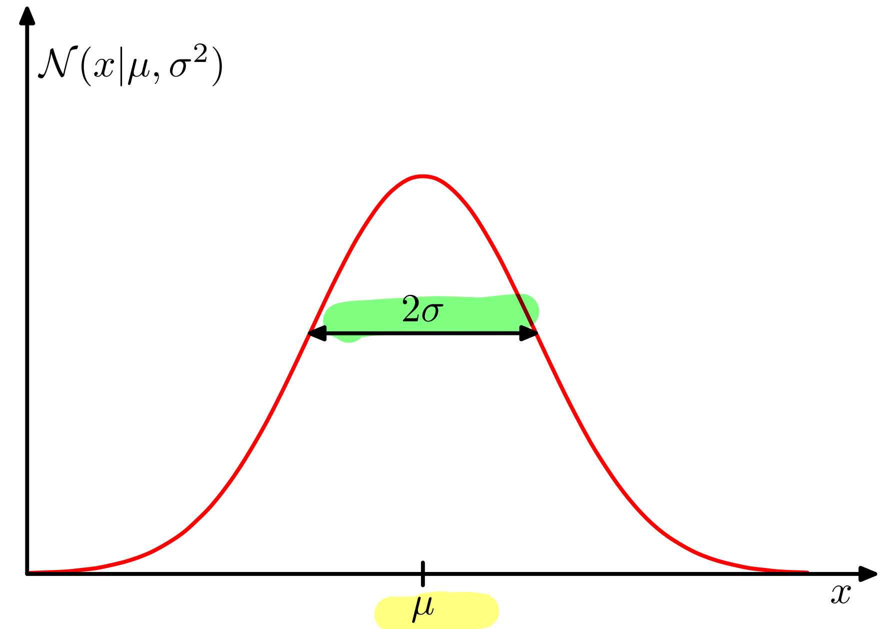


Image credit: Kirillm | Getty Images

Slide credits: Patrick Forré and Rianne van den Berg

Gaussian Distribution



- Real valued stochastic variable *X*

mean   variance

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

- Mean:

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} \mathcal{N}(x|\mu,\sigma^2)x \, \mathrm{d}x =$$

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

$$= \int_{-\infty}^{\infty} \frac{\sqrt{2\sigma^2}\,y + \mu}{\sqrt{\cancel{2}\pi\cancel{\sigma^2}}} e^{-y^2} \cancel{\sqrt{2\sigma^2}} \, dy$$

$$= \int_{-\infty}^{\infty} \left(\underbrace{\frac{\cancel{\sqrt{2\sigma^2}\,y}}{\cancel{\sqrt{\pi}}}}_{=0} + \frac{\mu}{\sqrt{\pi}} \right) e^{-y^2} dy = \frac{\mu}{\sqrt{\pi}} \cdot \sqrt{\pi} = \mu$$

Change of variables

$$y = \frac{1}{\sqrt{2\sigma^2}}(x-\mu) \rightarrow x = \sqrt{2\sigma^2}\cdot y + \mu$$

$$\frac{dy}{dx} = \frac{1}{\sqrt{2\sigma^2}} \rightarrow dx = \sqrt{2\sigma^2}\, dy$$

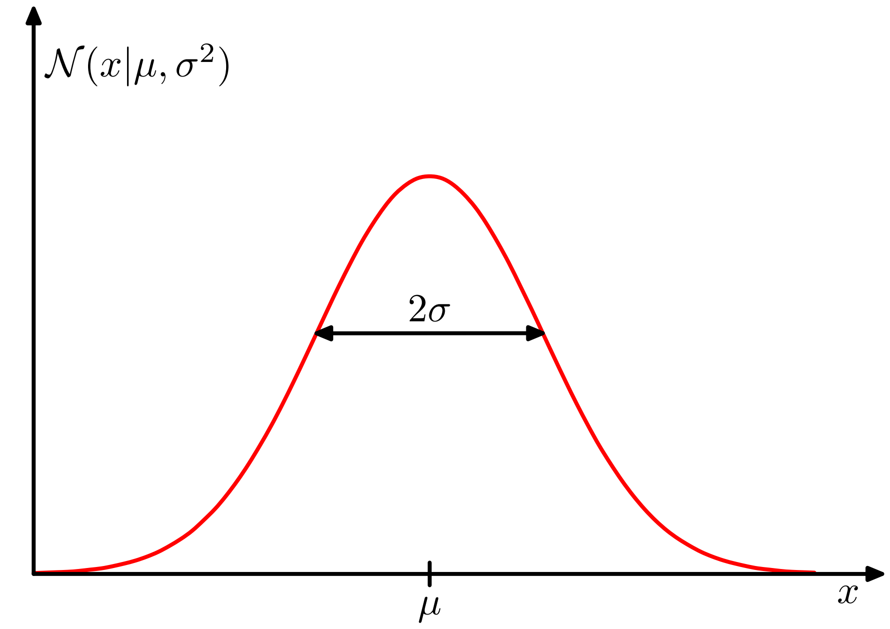
Integration of odd funcs

$$\int_{-\infty}^{\infty} y\, e^{-y^2} \mathrm{d}y = 0$$

Useful property:

$$\int_{-\infty}^{\infty} e^{-y^2} \mathrm{d}y = \sqrt{\pi}$$

Gaussian Distribution



- Real valued stochastic variable *X*

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

- Variance: $\mathrm{var}[x] = \mathbb{E}[(x-\mathbb{E}[x])^2]$

$$\mathrm{var}[x] = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

$$= \int_{-\infty}^{\infty} \frac{2\sigma^2 y^2}{\sqrt{\cancel{2}\pi\cancel{\sigma^2}}} e^{-y^2} \sqrt{\cancel{2\sigma^2}}\, dy$$

$$= \int_{-\infty}^{\infty} \frac{2\sigma^2}{\sqrt{\pi}} y^2 e^{-y^2} dy = \frac{2\sigma^2}{\sqrt{\pi}} \frac{1}{2}\sqrt{\pi} = \sigma^2$$

Change of variables

$$y = \frac{1}{\sqrt{2\sigma^2}}(x-\mu) \rightarrow x = \sqrt{2\sigma^2}y + \mu$$

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\sqrt{2\sigma^2}} \rightarrow \mathrm{d}x = \sqrt{2\sigma^2}\mathrm{d}y$$

Useful property:

$$\int_{-\infty}^{\infty} e^{-y^2}\mathrm{d}y = \sqrt{\pi}$$

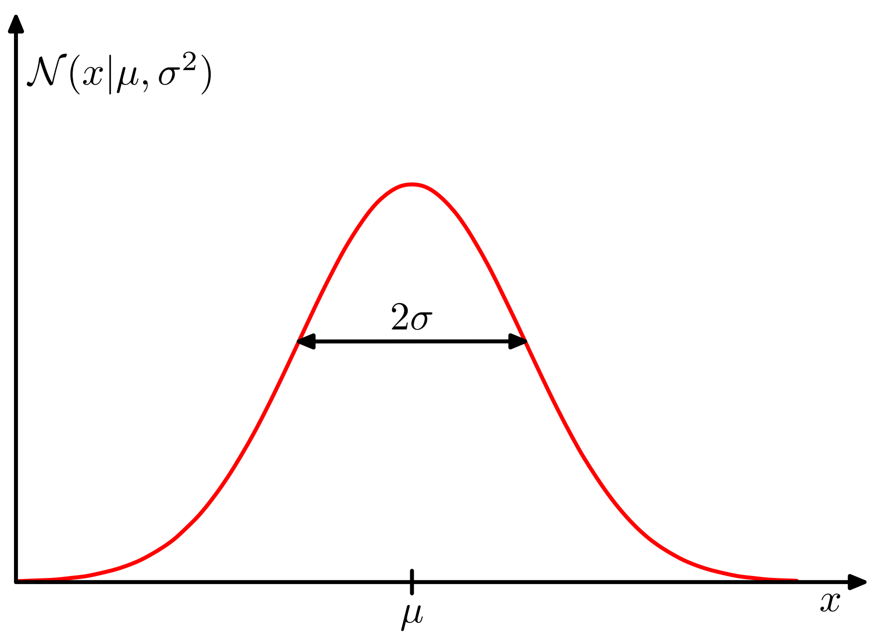
Convenient trick: $a=1$

$$\int_{-\infty}^{\infty} x^2 e^{-ax^2}\mathrm{d}x = -\frac{\partial}{\partial a}\int_{-\infty}^{\infty} e^{-ax^2}\mathrm{d}x$$

$$= \frac{\partial}{\partial a}\sqrt{\frac{\pi}{a}} = \frac{1}{2}\sqrt{\frac{\pi}{a^3}}$$

Gaussian Distribution

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



$$x \sim \mathcal{N}(x \mid \mu, \sigma^2) : \qquad \mathbb{E}[x] = \mu$$

$$\mathrm{Var}[x] = \sigma^2$$

Multivariate Gaussian Distribution

- D-dimensional vector $\mathbf{x} = (x_1, x_2, ..., x_D)^T$

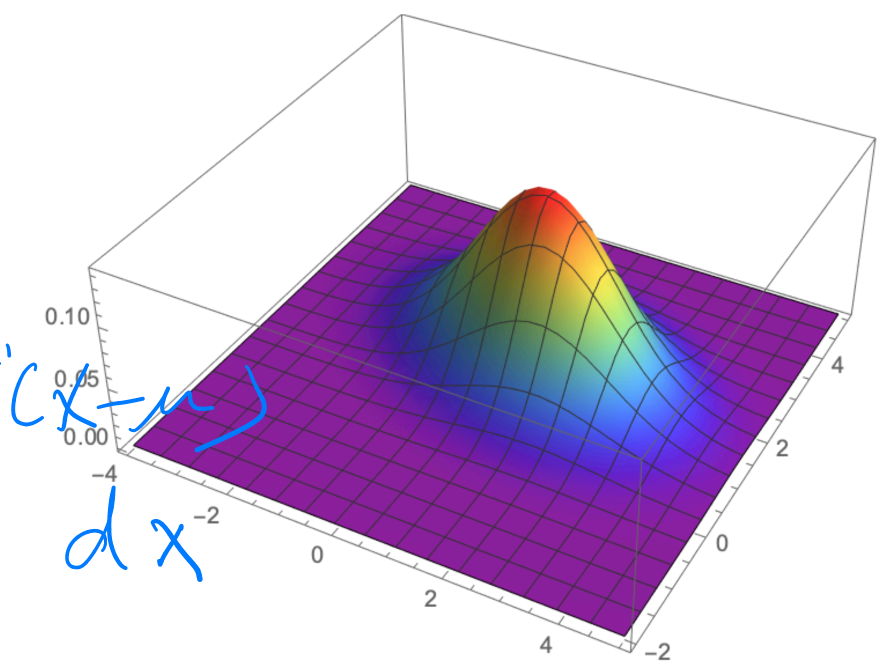
mean ↘ covariance matrix

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

$$|\mathbf{\Sigma}| = \det \mathbf{\Sigma}$$

Ch 2.3

- $\mathbf{\Sigma} = \mathrm{cov}[x, x]$ $\quad D \times D$



- $\mathbb{E}[\mathbf{x}] = \int_{\mathbb{R}^D} x \frac{1}{(2\pi)^{D/2} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)} dx$

$y = x - \mu$

"because odd"

$$= \int_{\mathbb{R}^D} (\cancel{y} + \mu) \cdot \underbrace{\frac{1}{(2\pi)^{D/2} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2} y^T \Sigma^{-1} y} dy}_{=1}$$

"because $\int_{\mathbb{R}^D} \mathcal{N}(y \mid 0, \Sigma) dy$ " = 1"

$= \mu$

Normalization factor:

$$\int_{\mathbb{R}^D} e^{-\frac{1}{2}\mathbf{x}^T \mathbf{A} \mathbf{x}} = \frac{2\pi^{D/2}}{|\mathbf{A}|^{1/2}}$$