CS-E4895 Gaussian Processes
Lecture 0: Example

A.N. Onymous

Aalto University

Thursday 16.2.2023

Agenda for today

1. Motivation for Gaussian processes
2. Course content, format, and evaluation
3. Warm up for Gaussian processes: Review of the multivariate Gaussian distribution
4. First assignment

The multivariate Gaussian distribution

- **Definition** A random vector $\boldsymbol{x} = [x_1, x_2, \cdots, x_D]^\top$ is said to have the multivariate Gaussian distribution if all linear combinations of $\boldsymbol{x}$ are Gaussian distributed:

$$y = \boldsymbol{a}^\top \boldsymbol{x} = a_1 x_1 + a_2 x_2 + \cdots + a_D x_D \sim \mathcal{N}(m, v) \tag{1}$$

for all $\boldsymbol{a} \in \mathbb{R}^D$, where $\boldsymbol{a} \neq \mathbf{0}$

The multivariate Gaussian distribution

- **Definition** A random vector $\boldsymbol{x} = [x_1, x_2, \cdots, x_D]^\top$ is said to have the multivariate Gaussian distribution if all linear combinations of $\boldsymbol{x}$ are Gaussian distributed:

$$y = \boldsymbol{a}^\top \boldsymbol{x} = a_1 x_1 + a_2 x_2 + \cdots + a_D x_D \sim \mathcal{N}(m, v) \tag{1}$$

  for all $\boldsymbol{a} \in \mathbb{R}^D$, where $\boldsymbol{a} \neq \mathbf{0}$

- The multivariate Gaussian density for a variable $\boldsymbol{x} \in \mathbb{R}^D$:

$$N(x|\mu, S) = (2\pi)^{-\frac{D}{2}} |S|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(x-\mu)^\top S^{-1}(x-\mu)\right] \quad \in \mathbb{R}_{\geq 0} \tag{2}$$

$$\log N(x|\mu, S) = -\frac{D}{2}\log 2\pi - \frac{1}{2}\log|S| - \frac{1}{2}(x-\mu)^\top S^{-1}(x-\mu) \quad \in \mathbb{R} \tag{3}$$

The multivariate Gaussian distribution

- **Definition** A random vector $\boldsymbol{x} = [x_1, x_2, \cdots, x_D]^\top$ is said to have the multivariate Gaussian distribution if all linear combinations of $\boldsymbol{x}$ are Gaussian distributed:

$$y = \boldsymbol{a}^\top \boldsymbol{x} = a_1 x_1 + a_2 x_2 + \cdots + a_D x_D \sim \mathcal{N}(m, v) \tag{1}$$

  for all $\boldsymbol{a} \in \mathbb{R}^D$, where $\boldsymbol{a} \neq \mathbf{0}$

- The multivariate Gaussian density for a variable $\boldsymbol{x} \in \mathbb{R}^D$:

$$N(x|\mu, S) = (2\pi)^{-\frac{D}{2}} |S|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(x-\mu)^\top S^{-1}(x-\mu)\right] \quad \in \mathbb{R}_{\geq 0} \tag{2}$$

$$\log N(x|\mu, S) = -\frac{D}{2}\log 2\pi - \frac{1}{2}\log|S| - \frac{1}{2}(x-\mu)^\top S^{-1}(x-\mu) \quad \in \mathbb{R} \tag{3}$$

- Completely described by its parameters:
  - $\boldsymbol{\mu} \in \mathbb{R}^D$ is the mean vector
  - $\boldsymbol{\Sigma} \in \mathbb{R}^{D \times D}$ is the covariance matrix (positive definite)

The multivariate Gaussian distribution

- **Definition** A random vector $\boldsymbol{x} = [x_1, x_2, \cdots, x_D]^\top$ is said to have the multivariate Gaussian distribution if all linear combinations of $\boldsymbol{x}$ are Gaussian distributed:

$$y = \boldsymbol{a}^\top \boldsymbol{x} = a_1 x_1 + a_2 x_2 + \cdots + a_D x_D \sim \mathcal{N}(m, v) \tag{1}$$

for all $\boldsymbol{a} \in \mathbb{R}^D$, where $\boldsymbol{a} \neq \mathbf{0}$

- The multivariate Gaussian density for a variable $\boldsymbol{x} \in \mathbb{R}^D$:

$$N(x|\mu, S) = (2\pi)^{-\frac{D}{2}} |S|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(x-\mu)^\top S^{-1}(x-\mu)\right] \quad \in \mathbb{R}_{\geq 0} \tag{2}$$

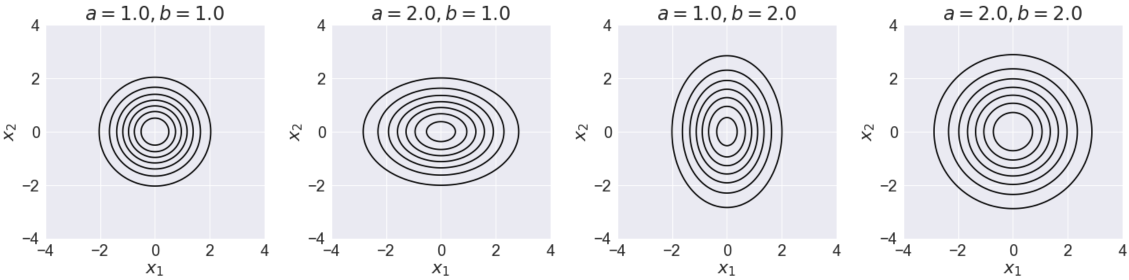
$$\log N(x|\mu, S) = -\frac{D}{2}\log 2\pi - \frac{1}{2}\log|S| - \frac{1}{2}(x-\mu)^\top S^{-1}(x-\mu) \quad \in \mathbb{R} \tag{3}$$

- Completely described by its parameters:
  - $\boldsymbol{\mu} \in \mathbb{R}^D$ is the mean vector
  - $\boldsymbol{\Sigma} \in \mathbb{R}^{D \times D}$ is the covariance matrix (positive definite)
- $(\boldsymbol{\Sigma})_{ij}$ is the covariance between the i'th and j'th elements x_i and x_j of $\boldsymbol{x}$

Interpretation of the covariance matrix - 2D examples

The diagonal of the covariance controls the scaling/marginal variances

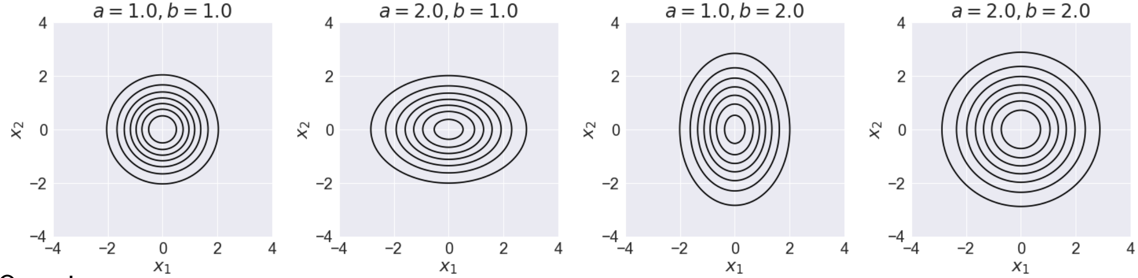
$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{\Sigma} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \tag{4}$$



Interpretation of the covariance matrix - 2D examples

The diagonal of the covariance controls the scaling/marginal variances

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{\Sigma} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \tag{4}$$



Questions:

1. If $\boldsymbol{\Sigma}$ is diagonal, then x_1 and x_2 are uncorrelated? True or false?
2. If $\boldsymbol{\Sigma}$ is diagonal, then x_1 and x_2 are independent? True or false?
3. What is the volume (integral) of density?
4. Which of the four densities has the highest peak and why?

More slides...

The end of today's lecture

- Next lecture: ...
  - We will introduce Gaussian processes more formally

- Time to work: first assignment

The end of today's lecture

- Next lecture: ...
  - We will introduce Gaussian processes more formally
  - Read Chapter 1 & 2 of the Gaussian process book `gaussianprocess.org/gpml`

- Time to work: first assignment

The end of today's lecture

- Next lecture: . . .
  - We will introduce Gaussian processes more formally
  - Read Chapter 1 & 2 of the Gaussian process book `gaussianprocess.org/gpml`

- Time to work: first assignment
  - . . .

The end of today's lecture

- Next lecture: ...
  - We will introduce Gaussian processes more formally
  - Read Chapter 1 & 2 of the Gaussian process book `gaussianprocess.org/gpml`

- Time to work: first assignment
  - ...
  - ...

The end of today's lecture

- Next lecture: ...
  - We will introduce Gaussian processes more formally
  - Read Chapter 1 & 2 of the Gaussian process book `gaussianprocess.org/gpml`

- Time to work: first assignment
  - ...
  - ...
  - ...