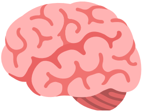
Learning context invariant representations for EEG data

Thibault de Surrel

LAMSADE Spring School - April 17th 2024

Motivation - Brain Computer Interfaces (BCI)



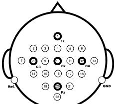
Translate the electrical signals

from the brain into commands

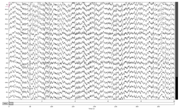
Applications

- Control prosthesis
- Write using a virtual keyboard
- Study the sleep level
- Many more applications...

The data - Electroencephalogram (EEG)



A cap equipped with sensors

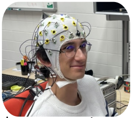


Electroencephalogram (EEG)

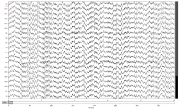
$$X = \begin{pmatrix} a_{1,1} & a_{1,2} & & & \\ a_{1,2} & a_{2,2} & & & \\ & & \ddots & & \\ & & & a_{n-1,n-1} & a_{n-1,n} \\ & & & a_{n-1,n} & a_{n,n} \end{pmatrix}$$

Covariance matrix

The data - Electroencephalogram (EEG)



A cap equipped with sensors



Electroencephalogram (EEG)

$$X = \begin{pmatrix} a_{1,1} & a_{1,2} & & & \\ a_{1,2} & a_{2,2} & & & \\ & & \ddots & & \\ & & & a_{n-1,n-1} & a_{n-1,n} \\ & & & a_{n-1,n} & a_{n,n} \end{pmatrix}$$

Covariance matrix

The data - From EEG to covariance matrices

$$\mathbf{E} \in \mathbb{R}^{c \times t} \longrightarrow \mathrm{Cov}(\mathbf{E}) = \frac{1}{1-t} \sum_{i=1}^{t} \mathbf{E}_i \mathbf{E}_i^\top$$

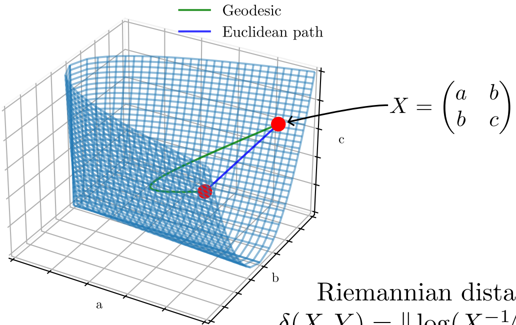
The data - From EEG to covariance matrices

$$\mathbf{E} \in \mathbb{R}^{c \times t} \longrightarrow \underbrace{\mathrm{Cov}(\mathbf{E})}_{\in \mathbb{S}_c^{++} = \left\{ \begin{array}{l} X \in \mathbb{R}^{c \times c} : X = X^\top, \\ \forall u \in \mathbb{R}^c \setminus \{0\},\ u^\top X u > 0 \end{array} \right\}} = \frac{1}{1-t} \sum_{i=1}^{t} \mathbf{E}_i \mathbf{E}_i^\top$$

The data - From EEG to covariance matrices

$$\mathbf{E} \in \mathbb{R}^{c \times t} \longrightarrow \underbrace{\mathrm{Cov}(\mathbf{E})}_{\in\, \mathbb{S}_c^{++}} = \frac{1}{1-t} \sum_{i=1}^{t} \mathbf{E}_i \mathbf{E}_i^\top$$

$$\in \mathbb{S}_c^{++} = \begin{Bmatrix} X \in \mathbb{R}^{c \times c} \colon X = X^\top, \\ \forall u \in \mathbb{R}^c \setminus \{0\},\ u^\top X u > 0 \end{Bmatrix}$$



Riemannian distance:

$$\delta(X, Y) = \| \log(X^{-1/2} Y X^{-1/2}) \|_F$$

The limitations

A lot of variabilities

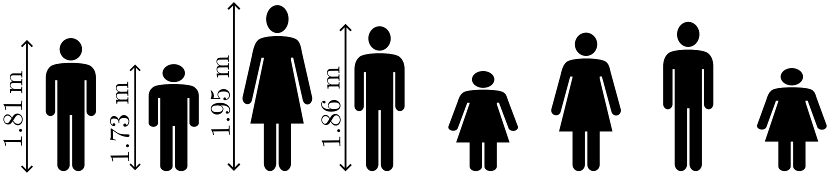
- Intrasubjet : noise, mental state, the subject's state of fatigue...
- Intersubjet

The problematic

How to build a context invariant representation for EEG data ?

Modelization using probability distributions

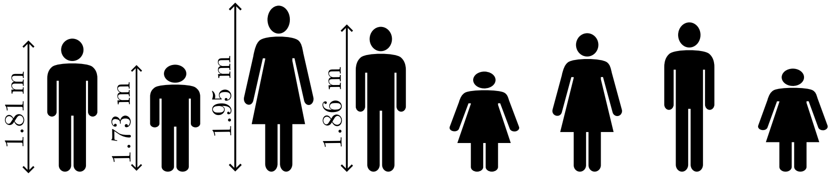
An example : Model the height of a population



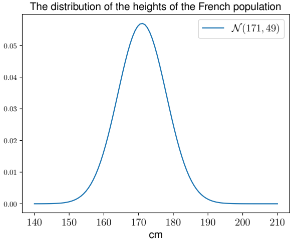
In France, the mean height of the population is 1.71 m. We can model the height of the french population using the *normal distribution* $\mathcal{N}(\mu, \sigma^2)$.

Modelization using probability distributions

An example : Model the height of a population

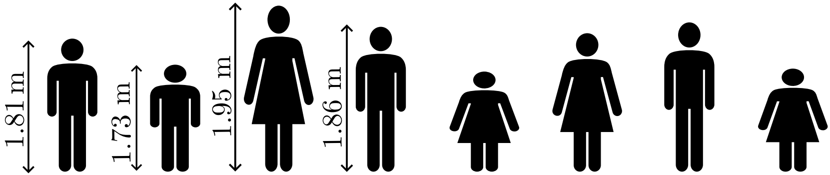


In France, the mean height of the population is 1.71 m. We can model the height of the french population using the *normal distribution* $\mathcal{N}(\mu, \sigma^2)$.

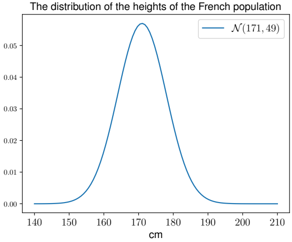


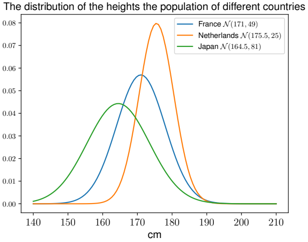
Modelization using probability distributions

An example : Model the height of a population



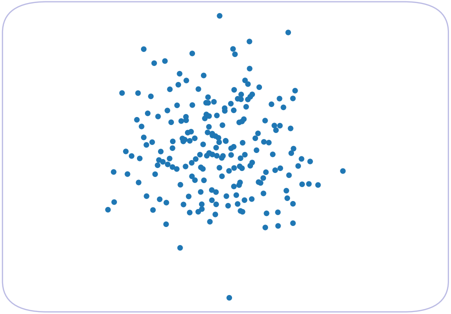
In France, the mean height of the population is 1.71 m. We can model the height of the french population using the *normal distribution* $\mathcal{N}(\mu, \sigma^2)$.





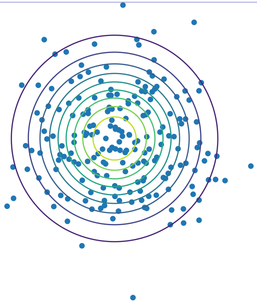
Modelization using probability distributions

Two examples in 2D



Modelization using probability distributions

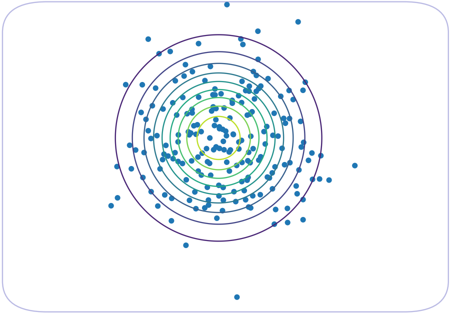
Two examples in 2D

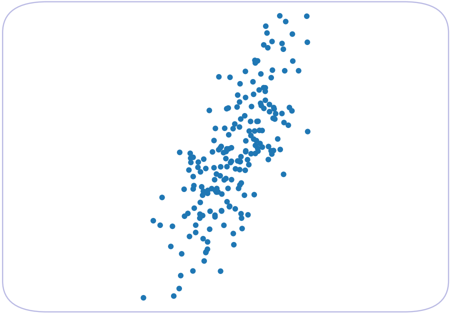


Isotropic gaussian : the spread is uniform.

Modelization using probability distributions

Two examples in 2D

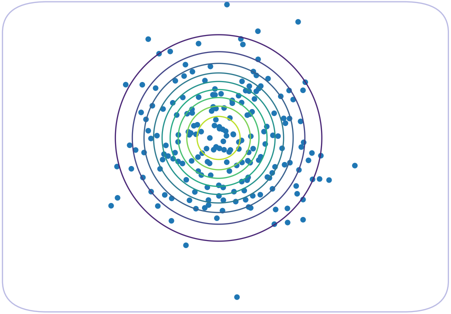


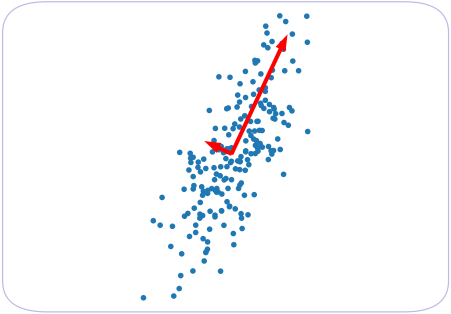


Isotropic gaussian : the spread is uniform.

Modelization using probability distributions

Two examples in 2D

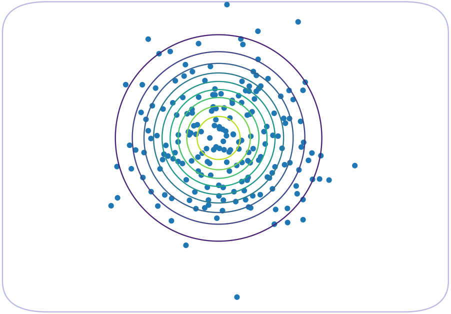




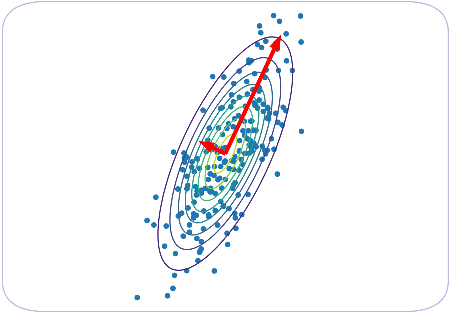
Isotropic gaussian : the spread is uniform.

Modelization using probability distributions

Two examples in 2D



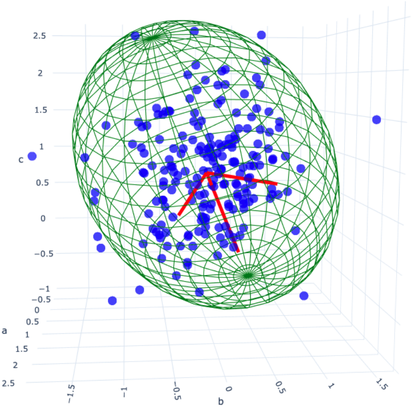
Isotropic gaussian : the spread is uniform.



Anisotropic gaussian : there are some preferred directions.

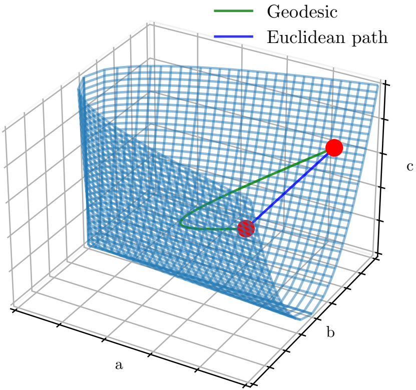
Modelization using probability distributions

One example in 3D



Modelization of the variabilities of EEG

Caution ! We need to take into account the Riemannian geometry of covariance matrices !



2 solutions to send an Euclidean Gaussian onto a Riemannian manifold

- First solution : an *isotropic* gaussian

$$p_{(\mu,\sigma^2)}(x) = \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{\|x-\mu\|^2}{2\sigma^2}\right)$$

2 solutions to send an Euclidean Gaussian onto a Riemannian manifold

- First solution : an *isotropic* gaussian

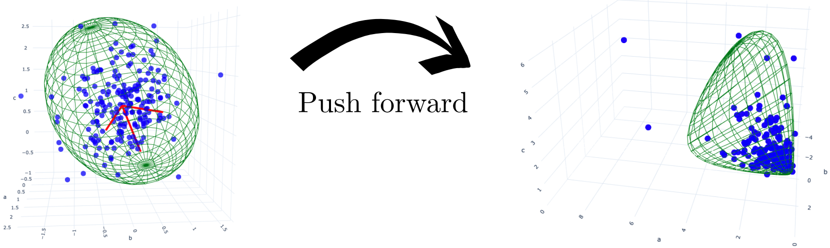
$$p_{(\bar{X},\sigma^2)}(X) = \zeta(\sigma)^{-1} \exp\left(-\frac{\delta(X,\bar{X})^2}{2\sigma^2}\right)$$

2 solutions to send an Euclidean Gaussian onto a Riemannian manifold

- First solution : an *isotropic* gaussian

$$p_{(\bar{X},\sigma^2)}(X) = \zeta(\sigma)^{-1} \exp\left(-\frac{\delta(X,\bar{X})^2}{2\sigma^2}\right)$$

- Second solution : an *anisotropic* gaussian

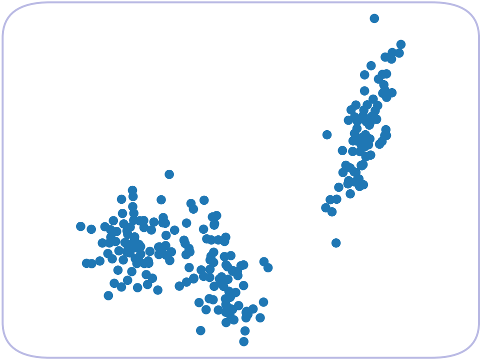


Euclidean anisotropic Gaussian

Riemannian anisotropic Gaussian

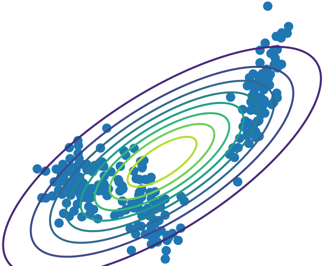
Mixture of gaussians

Sometime one gaussian fails at capturing the complexity of the dataset.



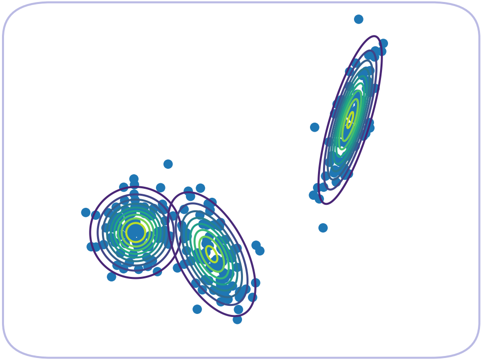
Mixture of gaussians

Sometime one gaussian fails at capturing the complexity of the dataset.



Mixture of gaussians

Sometime one gaussian fails at capturing the complexity of the dataset.



Estimating the parameters of a mixture of gaussians can be done using an *Expectation-Maximization* (EM) algorithm.

The applications

The applications of a Riemannian gaussian in BCI

- Better understand the variabilities
- Build an "informed" classifier
- Detect outliers
- Do some transfer learning
- Do some data augmentation

Conclusion

Motivations

Translate the electrical signals of the brain into commands.

Type of data

Electroencephalogram (multivariate time series) that are then converted into covariance matrices.

Modelization

A probability distribution.

The tools used

Riemannian geometry, probability theory, statistics...

Thank you !
Questions, remarks, comments... ?