

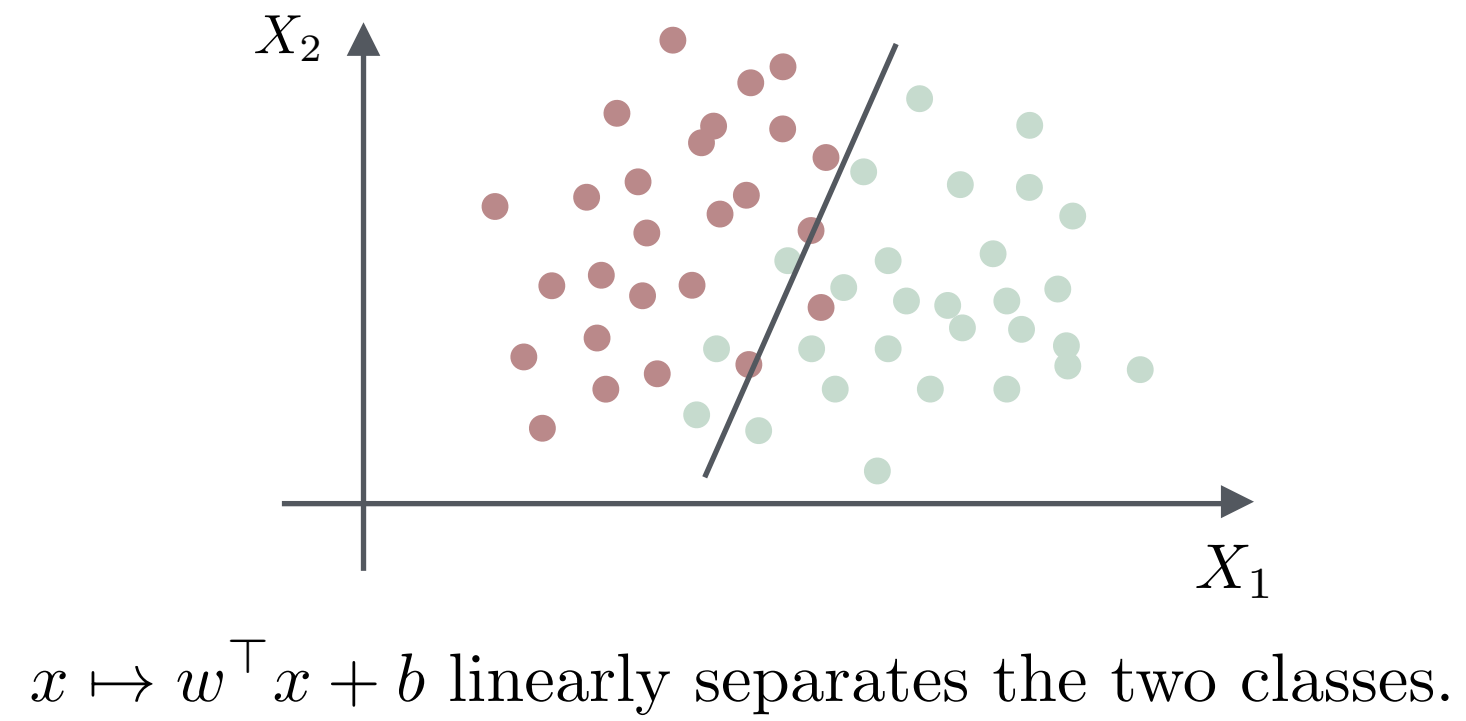
# Fundamentals of Machine Learning

Yassine Laguel

Mail : [yassine.laguel@univ-cotedazur.fr](mailto:yassine.laguel@univ-cotedazur.fr)

# Limitations of linear models

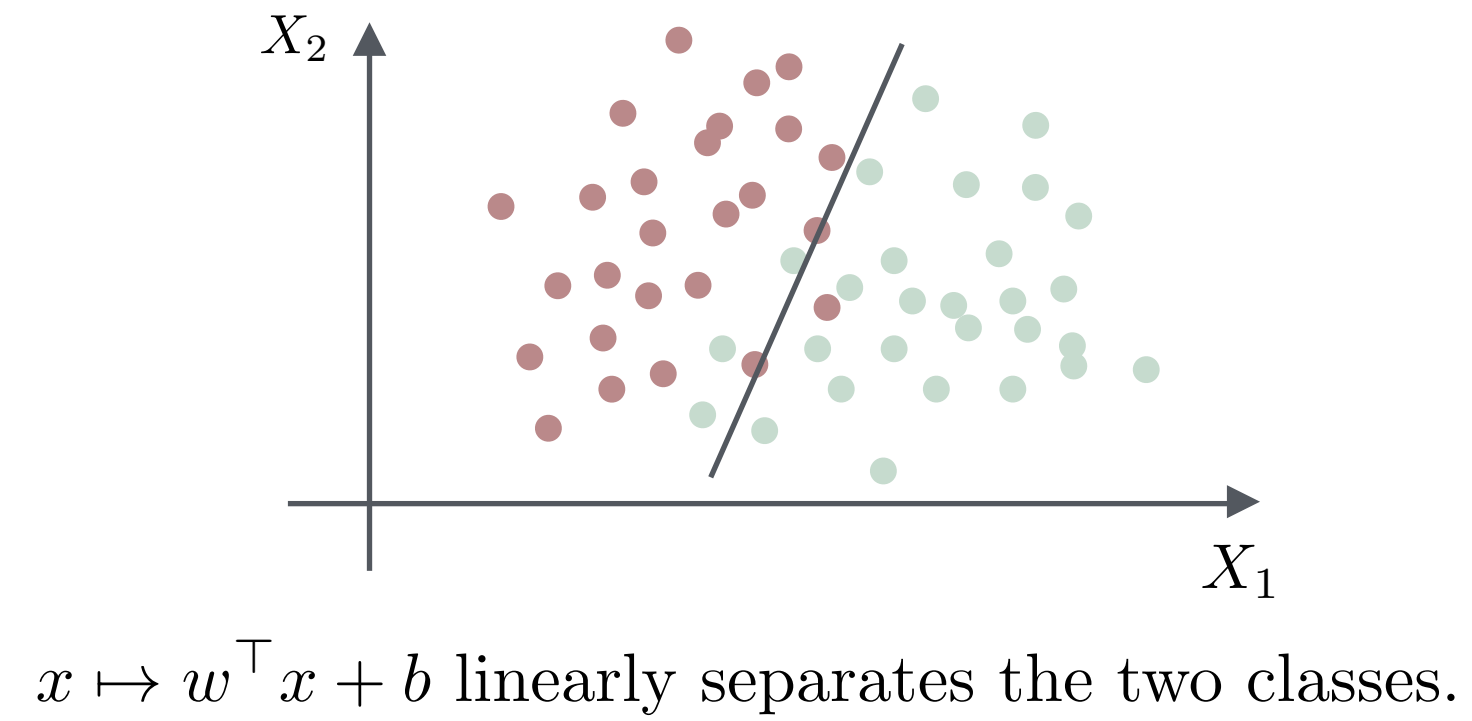
- Linear methods are great for linear tasks
  - In linear classification



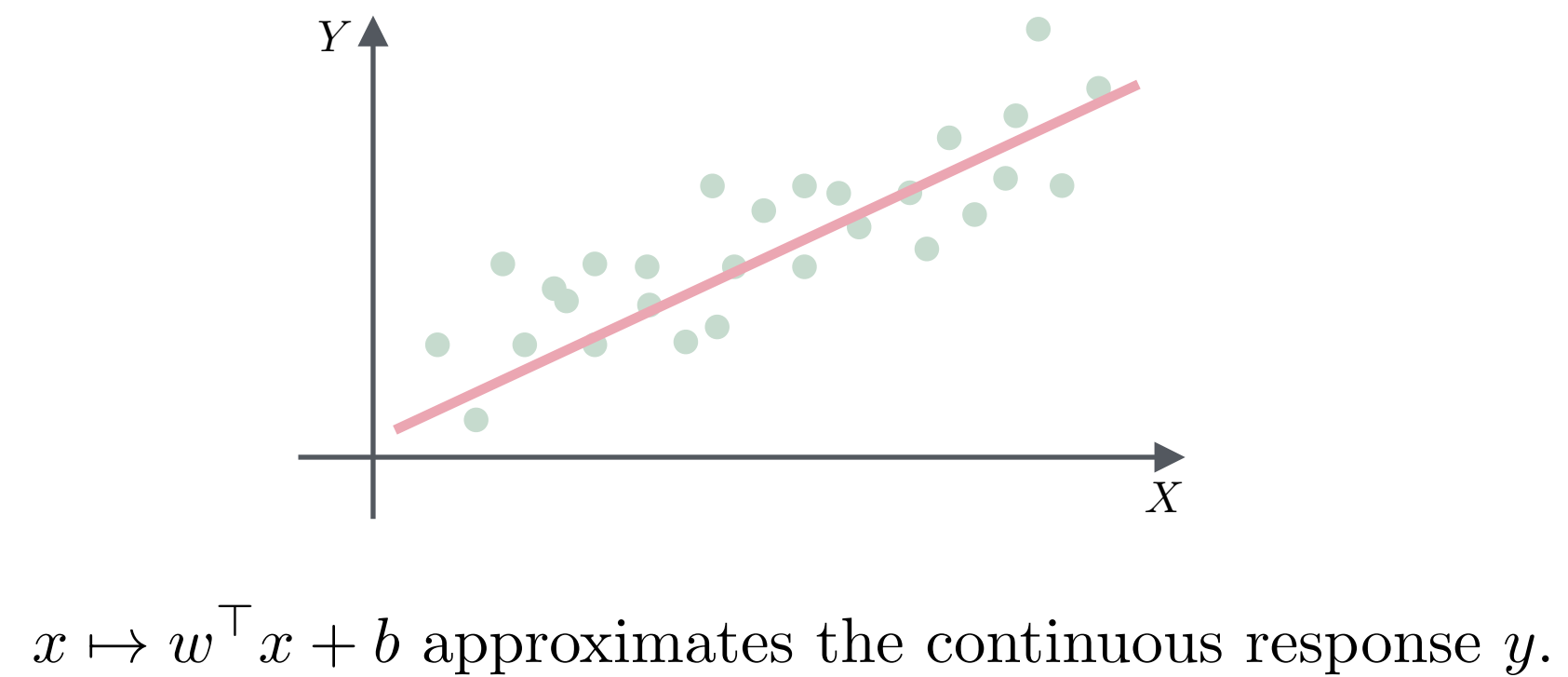
# Limitations of linear models

- Linear methods are great for linear tasks

- In linear classification



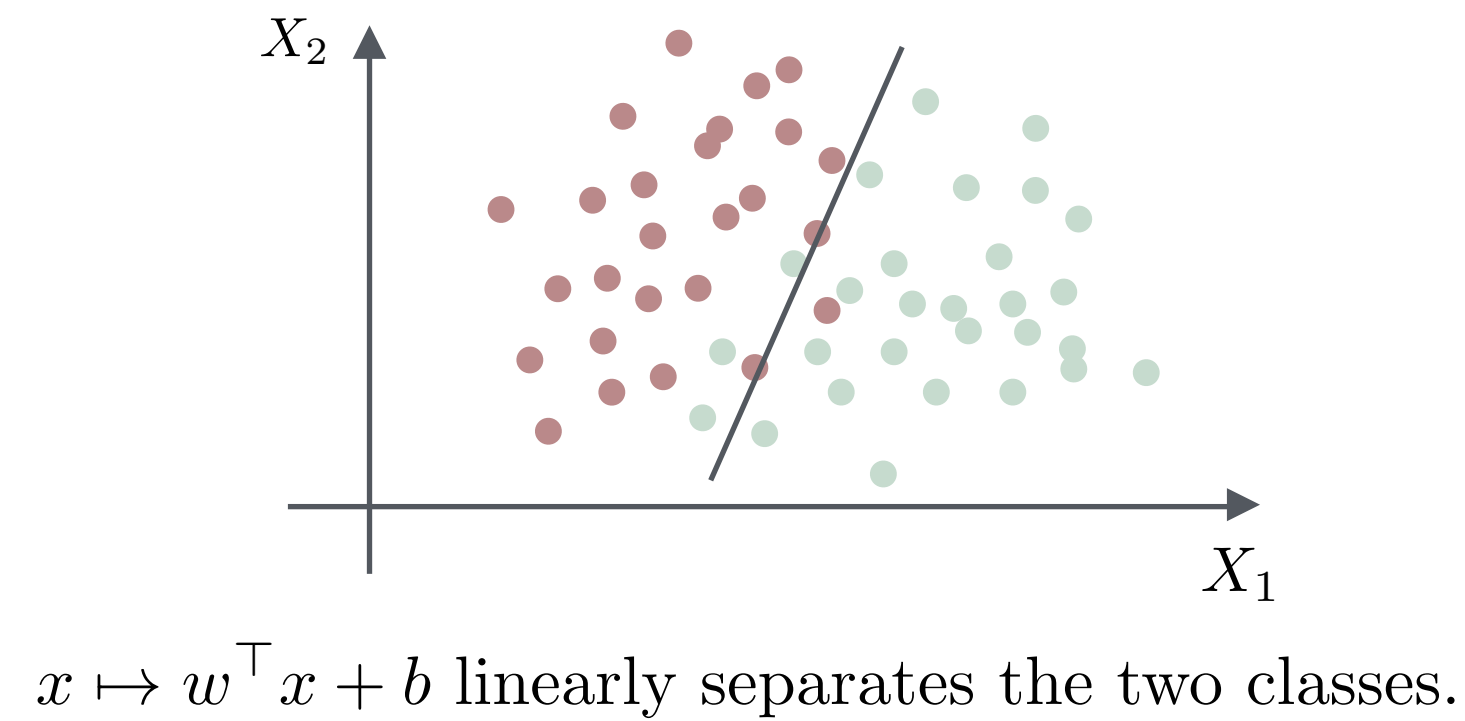
- In linear regression



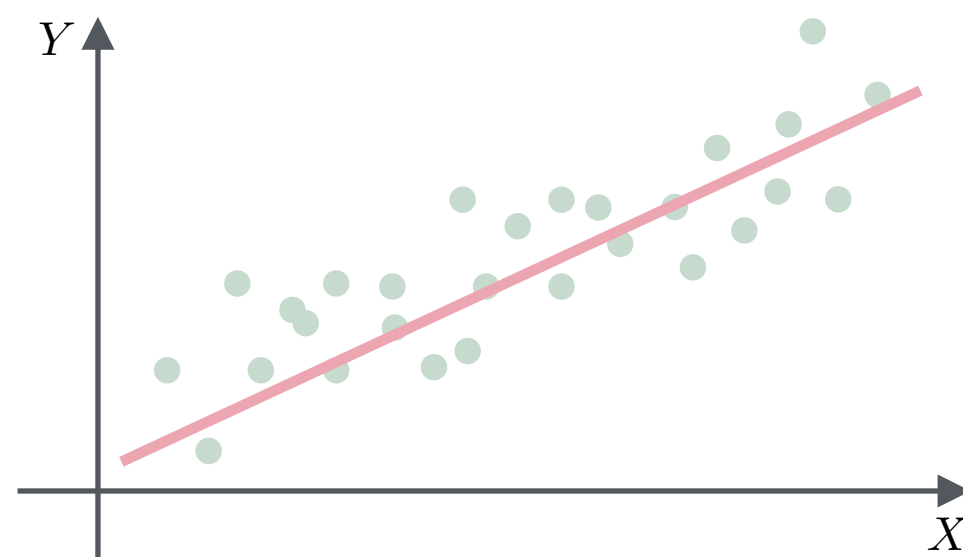
# Limitations of linear models

- Linear methods are great for linear tasks

- In linear classification

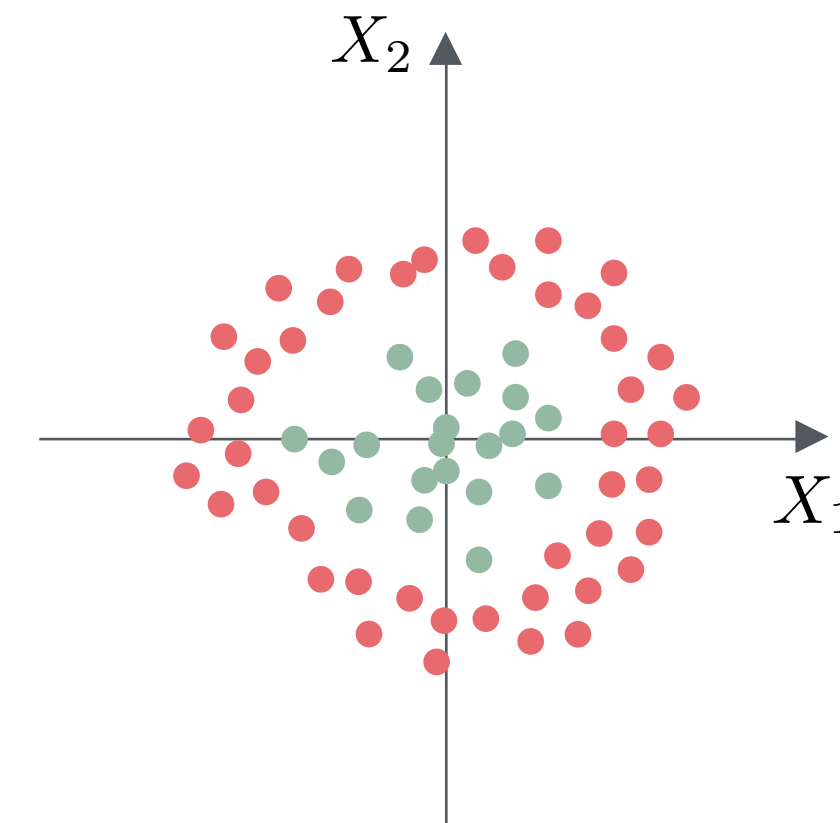


- In linear regression



- What can we do when for non-linear tasks ?

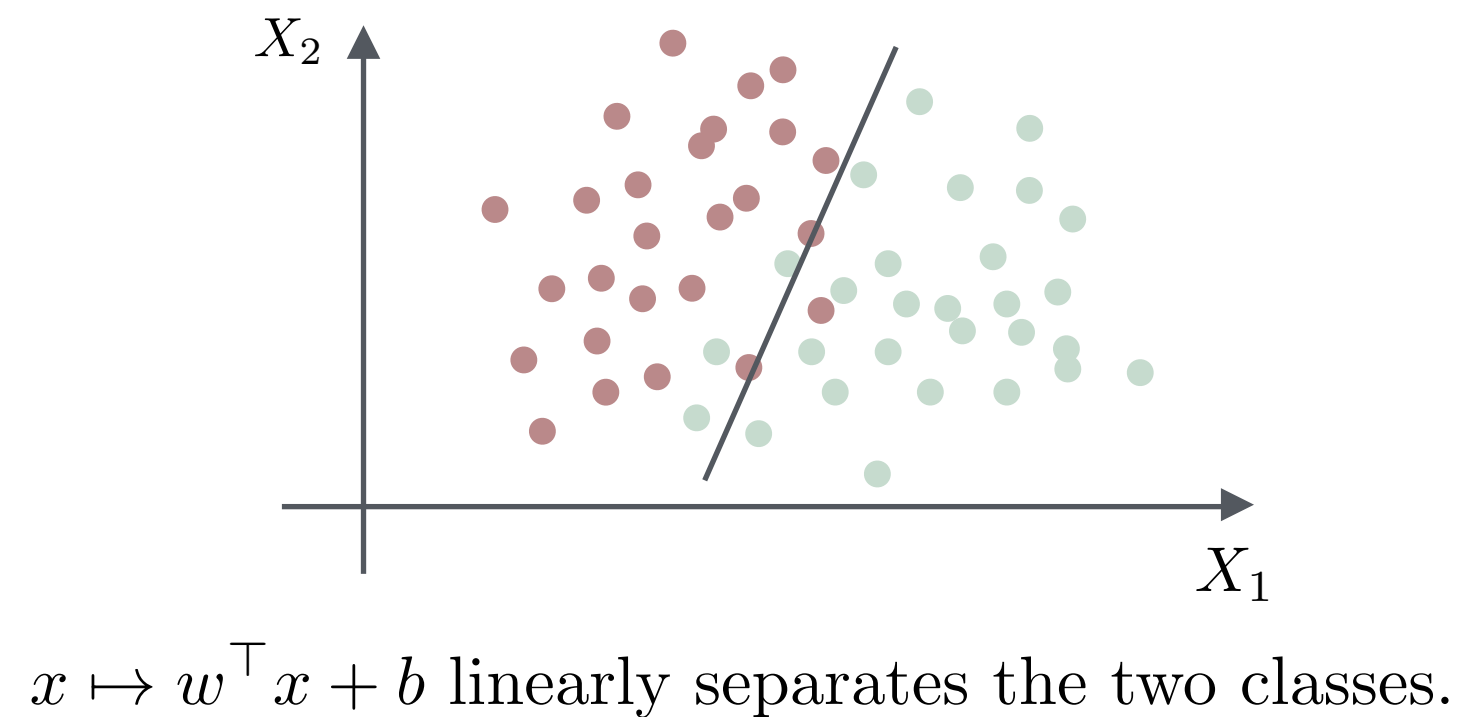
- In  $\mathbb{R}^2$  no hyperplane separates the following two classes.



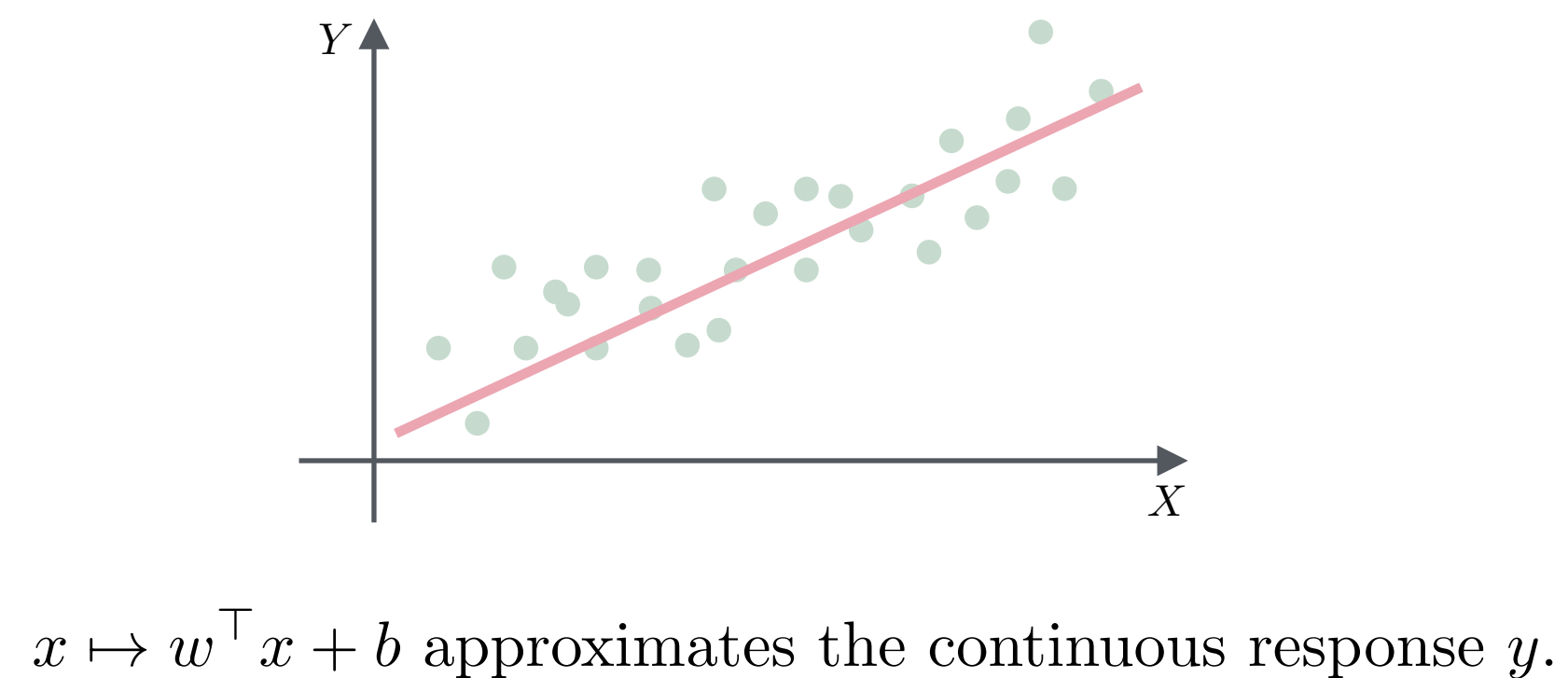
# Limitations of linear models

- Linear methods are great for linear tasks

- In linear classification

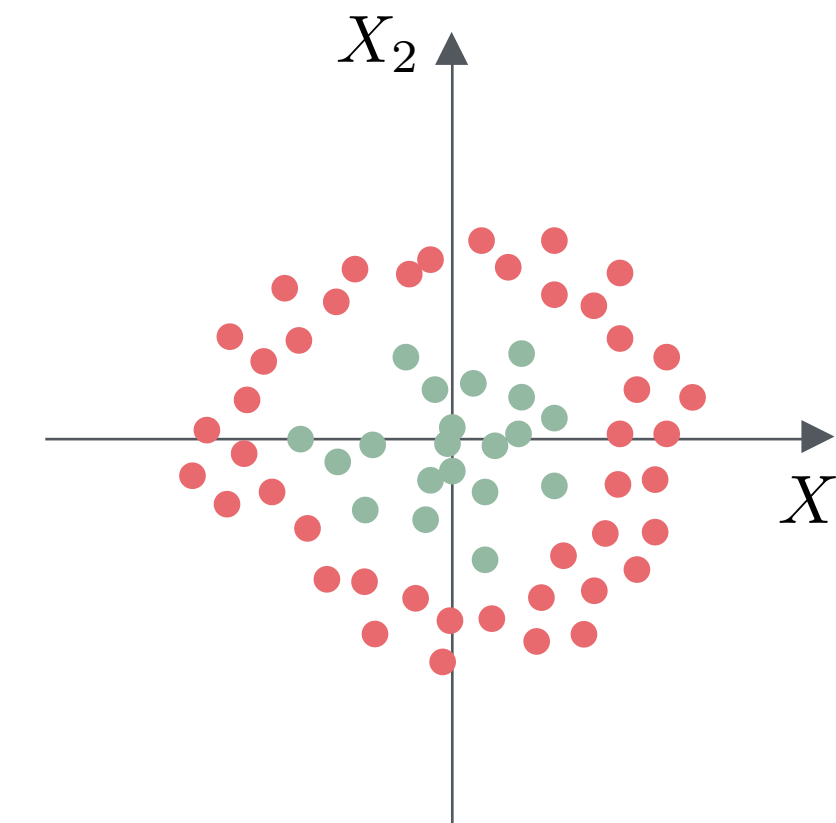


- In linear regression

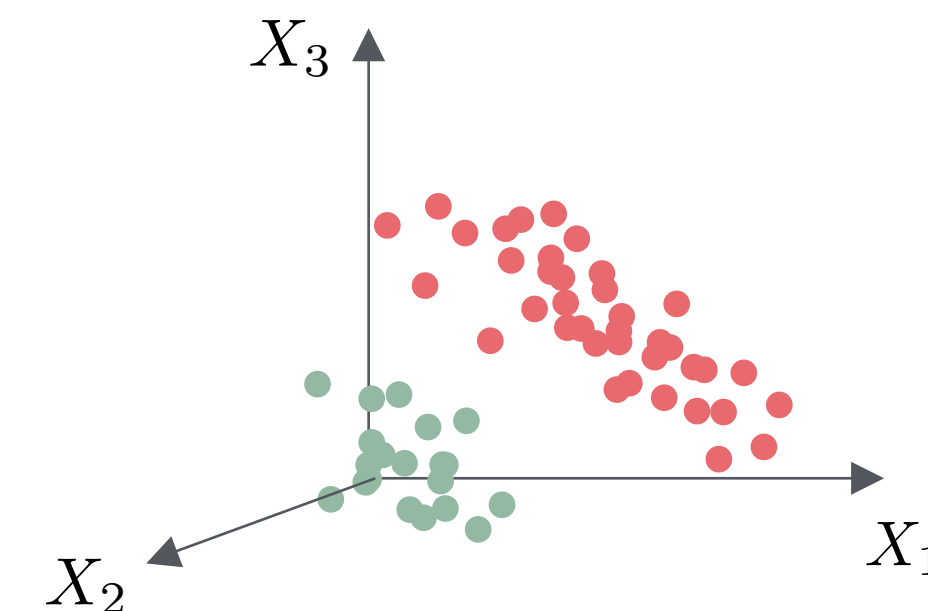


- What can we do when for non-linear tasks ?

- In  $\mathbb{R}^2$  no hyperplane separates the following two classes.



- Yet, if we apply the transformation  $x \mapsto \varphi(x) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)^\top$ .  
We obtain linearly separable data points in  $\mathbb{R}^3$ .



# Positive Definite Kernels

## ■ A First definition

### Definition : Positive definite Kernel

A positive definite kernel on a set  $\mathcal{X}$  is a function  $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  which satisfies

$$\text{(symmetry) (i) } K(x, x') = K(x', x), \quad \forall x, x' \in \mathcal{X}^2$$

$$\begin{aligned} \text{(positive-definiteness) (ii) } \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) &\geq 0 && \forall (x_1, \dots, x_n) \in \mathcal{X}^n \\ &&& \forall (a_1, \dots, a_n) \in \mathbb{R}^n \end{aligned}$$

# Positive Definite Kernels

## ■ A First definition

### Definition : Positive definite Kernel

A positive definite kernel on a set  $\mathcal{X}$  is a function  $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  which satisfies

$$\text{(symmetry) (i) } K(x, x') = K(x', x), \quad \forall x, x' \in \mathcal{X}^2$$

$$\begin{aligned} \text{(positive-definiteness) (ii) } \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) &\geq 0 & \forall (x_1, \dots, x_n) \in \mathcal{X}^n \\ & & \forall (a_1, \dots, a_n) \in \mathbb{R}^n \end{aligned}$$

### Remarks :

- In other words,  $K$  is a p.d. kernel if and only if, for any  $N \in \mathbb{N}$ , and any set of points  $(x_1, \dots, x_n) \in \mathcal{X}^N$ , the similarity matrix  $\mathbf{K}_{i,j} = K(x_i, x_j)$  is positive definite.

# Positive Definite Kernels

## ■ A First definition

### Definition : Positive definite Kernel

A positive definite kernel on a set  $\mathcal{X}$  is a function  $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  which satisfies

$$\text{(symmetry) (i) } K(x, x') = K(x', x), \quad \forall x, x' \in \mathcal{X}^2$$

$$\text{(positive-definiteness) (ii) } \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0 \quad \begin{array}{l} \forall (x_1, \dots, x_n) \in \mathcal{X}^n \\ \forall (a_1, \dots, a_n) \in \mathbb{R}^n \end{array}$$

### Remarks :

- In other words,  $K$  is a p.d. kernel if and only if, for any  $N \in \mathbb{N}$ , and any set of points  $(x_1, \dots, x_n) \in \mathcal{X}^N$ , the similarity matrix  $\mathbf{K}_{i,j} = K(x_i, x_j)$  is positive definite.
- We call **kernel methods**, algorithm that apply kernel transformations to input data.



# Positive Definite Kernels

## ■ A First definition

### Definition : Positive definite Kernel

A positive definite kernel on a set  $\mathcal{X}$  is a function  $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  which satisfies

$$\text{(symmetry) (i) } K(x, x') = K(x', x), \quad \forall x, x' \in \mathcal{X}^2$$

$$\text{(positive-definiteness) (ii) } \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0 \quad \begin{array}{l} \forall (x_1, \dots, x_n) \in \mathcal{X}^n \\ \forall (a_1, \dots, a_n) \in \mathbb{R}^n \end{array}$$

### Remarks :

- In other words,  $K$  is a p.d. kernel if and only if, for any  $N \in \mathbb{N}$ , and any set of points  $(x_1, \dots, x_n) \in \mathcal{X}^N$ , the similarity matrix  $\mathbf{K}_{i,j} = K(x_i, x_j)$  is positive definite.
- We call **kernel methods**, algorithm that apply kernel transformations to input data.

## ■ Stability properties

### Proposition : Stability through elementary operations

Let  $K_1, K_2$  be two positive definite kernels. Let  $c > 0$ .

Then

- (i)  $K_1 + K_2$  is a positive definite kernel.
- (ii)  $c \cdot K_1$  is a positive definite kernel.
- (iii)  $K_1 \cdot K_2$  is a positive definite kernel.

**Proof :** on the blackboard.

# Positive Definite Kernels

## ■ A First definition

### Definition : Positive definite Kernel

A positive definite kernel on a set  $\mathcal{X}$  is a function  $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  which satisfies

$$\text{(symmetry) (i) } K(x, x') = K(x', x), \quad \forall x, x' \in \mathcal{X}^2$$

$$\text{(positive-definiteness) (ii) } \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0 \quad \begin{array}{l} \forall (x_1, \dots, x_n) \in \mathcal{X}^n \\ \forall (a_1, \dots, a_n) \in \mathbb{R}^n \end{array}$$

### Remarks :

- In other words,  $K$  is a p.d. kernel if and only if, for any  $N \in \mathbb{N}$ , and any set of points  $(x_1, \dots, x_n) \in \mathcal{X}^N$ , the similarity matrix  $\mathbf{K}_{i,j} = K(x_i, x_j)$  is positive definite.
- We call **kernel methods**, algorithm that apply kernel transformations to input data.

## ■ Stability properties

### Proposition : Stability through elementary operations

Let  $K_1, K_2$  be two positive definite kernels. Let  $c > 0$ .

Then

- (i)  $K_1 + K_2$  is a positive definite kernel.
- (ii)  $c \cdot K_1$  is a positive definite kernel.
- (iii)  $K_1 \cdot K_2$  is a positive definite kernel.

**Proof :** on the blackboard.

### Proposition : Pointwise limit of p.d. kernels

Let  $(K_i)_{i \geq 0}$  be a sequence of p.d. kernels that converges pointwise to a function  $K$ . Then  $K$  is also a p.d. kernel.

**Proof :** on the blackboard.

# Positive Definite Kernels

## ■ A First definition

### Definition : Positive definite Kernel

A positive definite kernel on a set  $\mathcal{X}$  is a function  $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  which satisfies

$$\text{(symmetry) (i) } K(x, x') = K(x', x), \quad \forall x, x' \in \mathcal{X}^2$$

$$\text{(positive-definiteness) (ii) } \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0 \quad \begin{array}{l} \forall (x_1, \dots, x_n) \in \mathcal{X}^n \\ \forall (a_1, \dots, a_n) \in \mathbb{R}^n \end{array}$$

### Remarks :

- In other words,  $K$  is a p.d. kernel if and only if, for any  $N \in \mathbb{N}$ , and any set of points  $(x_1, \dots, x_n) \in \mathcal{X}^N$ , the similarity matrix  $\mathbf{K}_{i,j} = K(x_i, x_j)$  is positive definite.
- We call **kernel methods**, algorithm that apply kernel transformations to input data.

## ■ Stability properties

### Proposition : Stability through elementary operations

Let  $K_1, K_2$  be two positive definite kernels. Let  $c > 0$ .

Then

- (i)  $K_1 + K_2$  is a positive definite kernel.
- (ii)  $c \cdot K_1$  is a positive definite kernel.
- (iii)  $K_1 \cdot K_2$  is a positive definite kernel.

**Proof :** on the blackboard.

### Proposition : Pointwise limit of p.d. kernels

Let  $(K_i)_{i \geq 0}$  be a sequence of p.d. kernels that converges pointwise to a function  $K$ . Then  $K$  is also a p.d. kernel.

**Proof :** on the blackboard.

### Remark :

- We deduce from the above properties that if  $K$  is a p.d. kernel, then
  - $K^n$  is a p.d. kernel for all  $n \in \mathbb{N}, n \geq 1$ .

# Positive Definite Kernels

## ■ A First definition

### Definition : Positive definite Kernel

A positive definite kernel on a set  $\mathcal{X}$  is a function  $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  which satisfies

$$\text{(symmetry) (i) } K(x, x') = K(x', x), \quad \forall x, x' \in \mathcal{X}^2$$

$$\text{(positive-definiteness) (ii) } \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0 \quad \begin{array}{l} \forall (x_1, \dots, x_n) \in \mathcal{X}^n \\ \forall (a_1, \dots, a_n) \in \mathbb{R}^n \end{array}$$

### Remarks :

- In other words,  $K$  is a p.d. kernel if and only if, for any  $N \in \mathbb{N}$ , and any set of points  $(x_1, \dots, x_n) \in \mathcal{X}^N$ , the similarity matrix  $\mathbf{K}_{i,j} = K(x_i, x_j)$  is positive definite.
- We call **kernel methods**, algorithm that apply kernel transformations to input data.

## ■ Stability properties

### Proposition : Stability through elementary operations

Let  $K_1, K_2$  be two positive definite kernels. Let  $c > 0$ .

Then

- (i)  $K_1 + K_2$  is a positive definite kernel.
- (ii)  $c \cdot K_1$  is a positive definite kernel.
- (iii)  $K_1 \cdot K_2$  is a positive definite kernel.

**Proof :** on the blackboard.

### Proposition : Pointwise limit of p.d. kernels

Let  $(K_i)_{i \geq 0}$  be a sequence of p.d. kernels that converges pointwise to a function  $K$ . Then  $K$  is also a p.d. kernel.

**Proof :** on the blackboard.

### Remark :

- We deduce from the above properties that if  $K$  is a p.d. kernel, then
  - $K^n$  is a p.d. kernel for all  $n \in \mathbb{N}, n \geq 1$ .
  - $e^K$  is a p.d. kernel.

# Implementation



# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.



# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- Property (ii) is often called the reproducing property.



# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- Property (ii) is often called the reproducing property.
- For  $\mathcal{H}$  given, if a reproducing kernel exists, then  $\mathcal{H}$  is called a **reproducing kernel Hilbert space**.

# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- Property (ii) is often called the reproducing property.
- For  $\mathcal{H}$  given, if a reproducing kernel exists, then  $\mathcal{H}$  is called a **reproducing kernel Hilbert space**.

### Proposition : Reproducing kernels are p.d. kernels

Reproducing kernels are positive definite kernels.

# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- Property (ii) is often called the reproducing property.
- For  $\mathcal{H}$  given, if a reproducing kernel exists, then  $\mathcal{H}$  is called a **reproducing kernel Hilbert space**.

### Proposition : Reproducing kernels are p.d. kernels

Reproducing kernels are positive definite kernels.

**Proof :** On the black board

# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- Property (ii) is often called the reproducing property.
- For  $\mathcal{H}$  given, if a reproducing kernel exists, then  $\mathcal{H}$  is called a **reproducing kernel Hilbert space**.

### Proposition : Reproducing kernels are p.d. kernels

Reproducing kernels are positive definite kernels.

**Proof :** On the black board

## ■ Unicity results for the kernel and the RKHS

### Proposition : Continuity. of the evaluation map

A Hilbert space of functions  $H \subset \mathbb{R}^{\mathcal{X}}$  is a RKHS if and only if for any  $x \in \mathcal{X}$ ,  $f \mapsto f(x)$  from  $\mathcal{H}$  to  $\mathbb{R}$  is continuous.

# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- Property (ii) is often called the reproducing property.
- For  $\mathcal{H}$  given, if a reproducing kernel exists, then  $\mathcal{H}$  is called a **reproducing kernel Hilbert space**.

### Proposition : Reproducing kernels are p.d. kernels

Reproducing kernels are positive definite kernels.

**Proof :** On the black board

## ■ Unicity results for the kernel and the RKHS

### Proposition : Continuity. of the evaluation map

A Hilbert space of functions  $H \subset \mathbb{R}^{\mathcal{X}}$  is a RKHS if and only if for any  $x \in \mathcal{X}$ ,  $f \mapsto f(x)$  from  $\mathcal{H}$  to  $\mathbb{R}$  is continuous.

**Proof :** On the blackboard.

# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- Property (ii) is often called the reproducing property.
- For  $\mathcal{H}$  given, if a reproducing kernel exists, then  $\mathcal{H}$  is called a **reproducing kernel Hilbert space**.

### Proposition : Reproducing kernels are p.d. kernels

Reproducing kernels are positive definite kernels.

**Proof :** On the black board

## ■ Unicity results for the kernel and the RKHS

### Proposition : Continuity. of the evaluation map

A Hilbert space of functions  $H \subset \mathbb{R}^{\mathcal{X}}$  is a RKHS if and only if for any  $x \in \mathcal{X}$ ,  $f \mapsto f(x)$  from  $\mathcal{H}$  to  $\mathbb{R}$  is continuous.

**Proof :** On the blackboard.

### Proposition : Unicity of the reproducing kernel

If two reproducing kernel exist for a Hilbert space, then they are equal.



# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- Property (ii) is often called the reproducing property.
- For  $\mathcal{H}$  given, if a reproducing kernel exists, then  $\mathcal{H}$  is called a **reproducing kernel Hilbert space**.

### Proposition : Reproducing kernels are p.d. kernels

Reproducing kernels are positive definite kernels.

**Proof :** On the black board

## ■ Unicity results for the kernel and the RKHS

### Proposition : Continuity. of the evaluation map

A Hilbert space of functions  $H \subset \mathbb{R}^{\mathcal{X}}$  is a RKHS if and only if for any  $x \in \mathcal{X}$ ,  $f \mapsto f(x)$  from  $\mathcal{H}$  to  $\mathbb{R}$  is continuous.

**Proof :** On the blackboard.

### Proposition : Unicity of the reproducing kernel

If two reproducing kernel exist for a Hilbert space, then they are equal.

**Proof :** On the blackboard.

# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- Property (ii) is often called the reproducing property.
- For  $\mathcal{H}$  given, if a reproducing kernel exists, then  $\mathcal{H}$  is called a **reproducing kernel Hilbert space**.

### Proposition : Reproducing kernels are p.d. kernels

Reproducing kernels are positive definite kernels.

**Proof :** On the black board

## ■ Unicity results for the kernel and the RKHS

### Proposition : Continuity. of the evaluation map

A Hilbert space of functions  $H \subset \mathbb{R}^{\mathcal{X}}$  is a RKHS if and only if for any  $x \in \mathcal{X}$ ,  $f \mapsto f(x)$  from  $\mathcal{H}$  to  $\mathbb{R}$  is continuous.

**Proof :** On the blackboard.

### Proposition : Unicity of the reproducing kernel

If two reproducing kernel exist for a Hilbert space, then they are equal.

**Proof :** On the blackboard.

### Proposition : Unicity of the. RKHS

If two RKHS have the same reproducing kernel, then they are equal.



# Reproducing Kernel Hilbert space

## ■ Kernels as inner products

### Definition : Reproducing Kernel Hilbert space

Let  $\mathcal{X}$  be a set and  $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$  be a class of functions forming a (real) Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ .

The function  $K : \mathcal{X}^2 \mapsto \mathbb{R}$  is called a **reproducing kernel** of  $\mathcal{H}$  if

- (i)  $\mathcal{H}$  contains all functions of the form  $K_x : u \mapsto K(x, u)$ , for  $x \in \mathcal{X}$ .
- (ii) For every  $\mathbf{x} \in \mathcal{X}$  and  $f \in \mathcal{H}$ , we have  $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$ .

### Remarks :

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- Property (ii) is often called the reproducing property.
- For  $\mathcal{H}$  given, if a reproducing kernel exists, then  $\mathcal{H}$  is called a **reproducing kernel Hilbert space**.

### Proposition : Reproducing kernels are p.d. kernels

Reproducing kernels are positive definite kernels.

**Proof :** On the black board

## ■ Unicity results for the kernel and the RKHS

### Proposition : Continuity. of the evaluation map

A Hilbert space of functions  $H \subset \mathbb{R}^{\mathcal{X}}$  is a RKHS if and only if for any  $x \in \mathcal{X}$ ,  $f \mapsto f(x)$  from  $\mathcal{H}$  to  $\mathbb{R}$  is continuous.

**Proof :** On the blackboard.

### Proposition : Unicity of the reproducing kernel

If two reproducing kernel exist for a Hilbert space, then they are equal.

**Proof :** On the blackboard.

### Proposition : Unicity of the. RKHS

If two RKHS have the same reproducing kernel, then they are equal.

**Proof :** On the blackboard.

# Reproducing Kernel Hilbert space

## ■ Aronszajn

### Proposition : Aronszajn theorem

$K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  is a p.d. kernel on  $\mathcal{X}$  if and only if there exists a Hilbert space  $\mathcal{H}$  and a mapping  $\varphi : \mathcal{X} \rightarrow \mathcal{H}$  such that

$$K(\mathbf{x}, \mathbf{x}') = \langle \Phi(\mathbf{x}), \Phi(\mathbf{x}') \rangle_{\mathcal{H}}, \quad \forall x, x' \in \mathcal{X}.$$

**Proof :** Delayed for a bit.

**Remarks :**

- As a recall, a Hilbert space is a complete vector space endowed with an inner product.
- In other words, the kernel operates a measure of similarity on data, made explicit as an inner product after application of  $\varphi$ .
- We a good choice for  $\varphi$ , one may expect to transform their nonlinear prediction task into a linear one in the space  $\mathcal{H}$ .