

IFT3395/6390

Fondements de l'apprentissage machine

**Dépasser les limitations
des classifieurs linéaires**

Première solution:

Mapping non-linéaire explicite

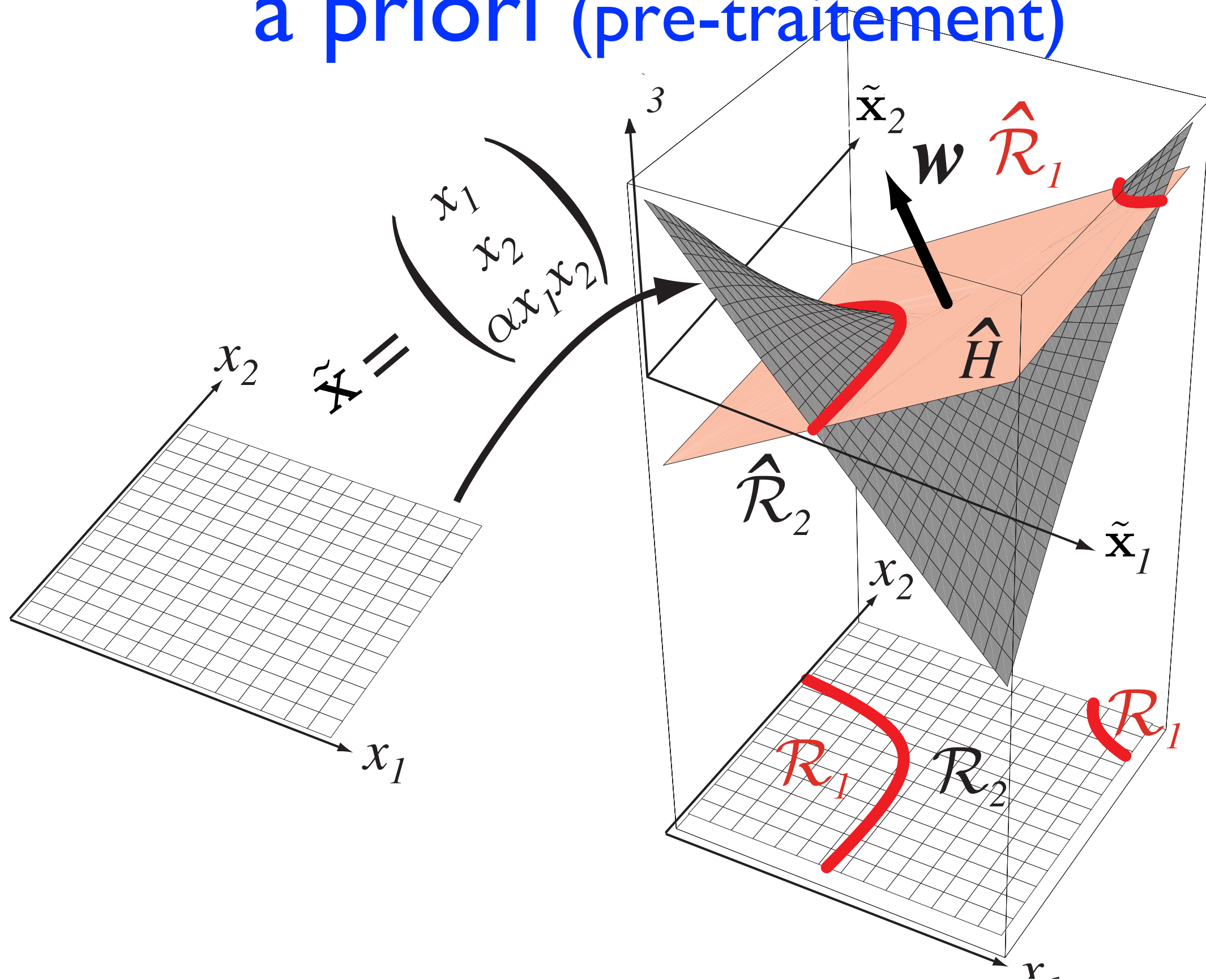
Professeur: Pascal Vincent

Comment obtenir un classifieur non linéaire à partir d'un classifieur linéaire?

Une vieille technique...

- D'abord appliquer une transformation non-linéaire à \mathbf{x}
(projection dans un **nouvel** espace de **traits caractéristiques utiles**)
$$\tilde{\mathbf{x}} = \phi(\mathbf{x})$$
- Apprendre une fonction discriminante linéaire dans ce **nouvel espace** (avec l'une des nombreuses techniques de classifieur linéaire)
- L'hyperplan séparateur dans ce nouvel espace correspond à une **frontière de décision non linéaire dans l'espace d'origine**.
- Et voilà! \Rightarrow **Classifieur non-linéaire de plus forte capacité.**

Ex. transformation non-linéaire à priori (pré-traitement)



Ex. transformation non-linéaire à priori (pré-traitement)

Comment obtenir un classifieur (ou régresseur) polynômial à partir d'un classifieur (ou régresseur) linéaire?

Qu'est-ce qu'une fonction linéaire (affine) de x (ex en dimension 2)?

$$y = f_{\theta}([x_1, x_2]) = w_1 x_1 + w_2 x_2 + b$$

Qu'est-ce qu'une fonction polynômiale de degré 2 de x ?

$$poly_2([x_1, x_2]) = a_{11}x_1^2 + a_{22}x_2^2 + a_{12}x_1x_2 + a_1x_1 + a_2x_2 + b$$

Qu'est-ce qu'une fonction polynômiale de degré 3 de x ?

Ex. transformation non-linéaire à priori (pré-traitement)

Comment obtenir un classifieur (ou régresseur) polynômial à partir d'un classifieur (ou régresseur) linéaire?

- D'abord **appliquer une transformation non-linéaire à x**
(projection dans un nouvel espace de traits caractéristiques)

$$\tilde{x} = \varphi(x) = \varphi([x_1, x_2]) = [x_1^2, x_2^2, x_1 x_2, x_1, x_2]$$

- Apprendre une fonction discriminante linéaire dans ce nouvel **espace** (avec l'une des nombreuses techniques de classifieur linéaire)

$$f(x) = w^T \tilde{x} + b = w_1 \underbrace{\tilde{x}_1}_{x_1^2} + w_2 \underbrace{\tilde{x}_2}_{x_2^2} + w_3 \underbrace{\tilde{x}_3}_{x_1 x_2} + w_4 \underbrace{\tilde{x}_4}_{x_1} + w_5 \underbrace{\tilde{x}_5}_{x_2} + b$$

- L'hyperplan séparateur dans ce nouvel espace correspond à une **frontière de décision non linéaire dans l'espace d'origine.**
- Et voilà! \Rightarrow **Classifieur non-linéaire de plus forte capacité.**

Mais quelle transformation?

Trois manières d'avoir une transformation $\tilde{\mathbf{x}} = \phi(\mathbf{x})$

- Utiliser une transformation fixée à priori explicitement
 - ➡ Exemple précédent
- Utiliser une transformation fixée à priori implicitement
 - ➡ Méthodes à noyau (SVM à noyau, Régression Logistique à noyau, ...)
Prochain cours...
- Apprendre les paramètres d'une transformation paramétrée:
 - ➡ Réseaux de neurones
de type Perceptron Multicouche ($M_{ulti}L_{ayer}P_{erceptron}$)