SVM

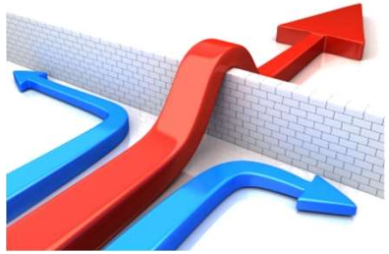
SUPPORT VECTOR MACHINE
SÉPARATEUR À VASTE MARGE

Mohamed Heny SELMI

Ecole Supérieure Privée
d'Ingénierie et de Technologies

MÉTHODES SUPERVISÉES

$$Y \overset{?}{=} f(X_1, X_2, X_3, \dots\dots\dots\dots\dots, X_n)$$



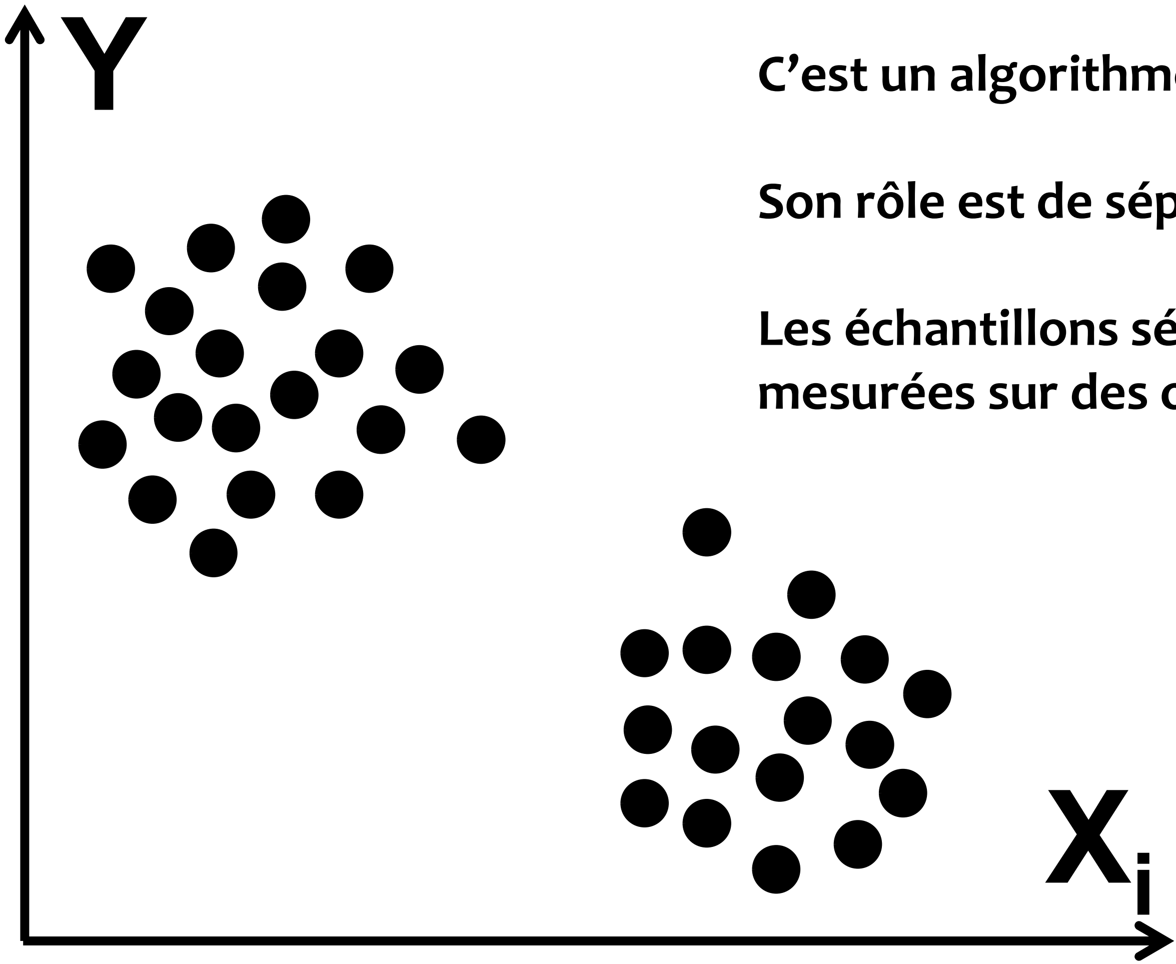
Arbre de décision	• Modèle à base de règles logiques • **ordre de séparation de données**
Régression Linéaire	• Modèle à base de coefficients estimateurs • **Modèle complet, avec sélection de variables, individus atypiques**
SVM	• **Classifieur Linéaire ou Classifieur Non Linéaire** • **Hyperplan séparateur, marge maximale**
Régression Logistique	• **Modèle à base de coefficients classifieurs** • **Fonction de lien Π, Fonction Gaussienne / LOGIT**
Analyse Linéaire Discriminante	• Modèle de Scoring
Réseaux de Neurones	• Modèle d’Apprentissage Artificielle
KNN	• Modèle de classement informé

EXEMPLE D'UTILISATION DU CLASSIFIEUR À VASTE MARGE

- Courrier électronique : filtrage des spams (spam / non spam)
- Classification message (urgent / non urgent)
- Recherche d'informations (correct / incorrect)
- Transformation de classifications multiples en classification binaire : 1 classe contre toutes les autres
- Classification émotions (positive / négative)

IDÉE DE BASE DU CLASSIFIEUR À VASTE MARGE

Qu'est-ce qu'un classifieur ?

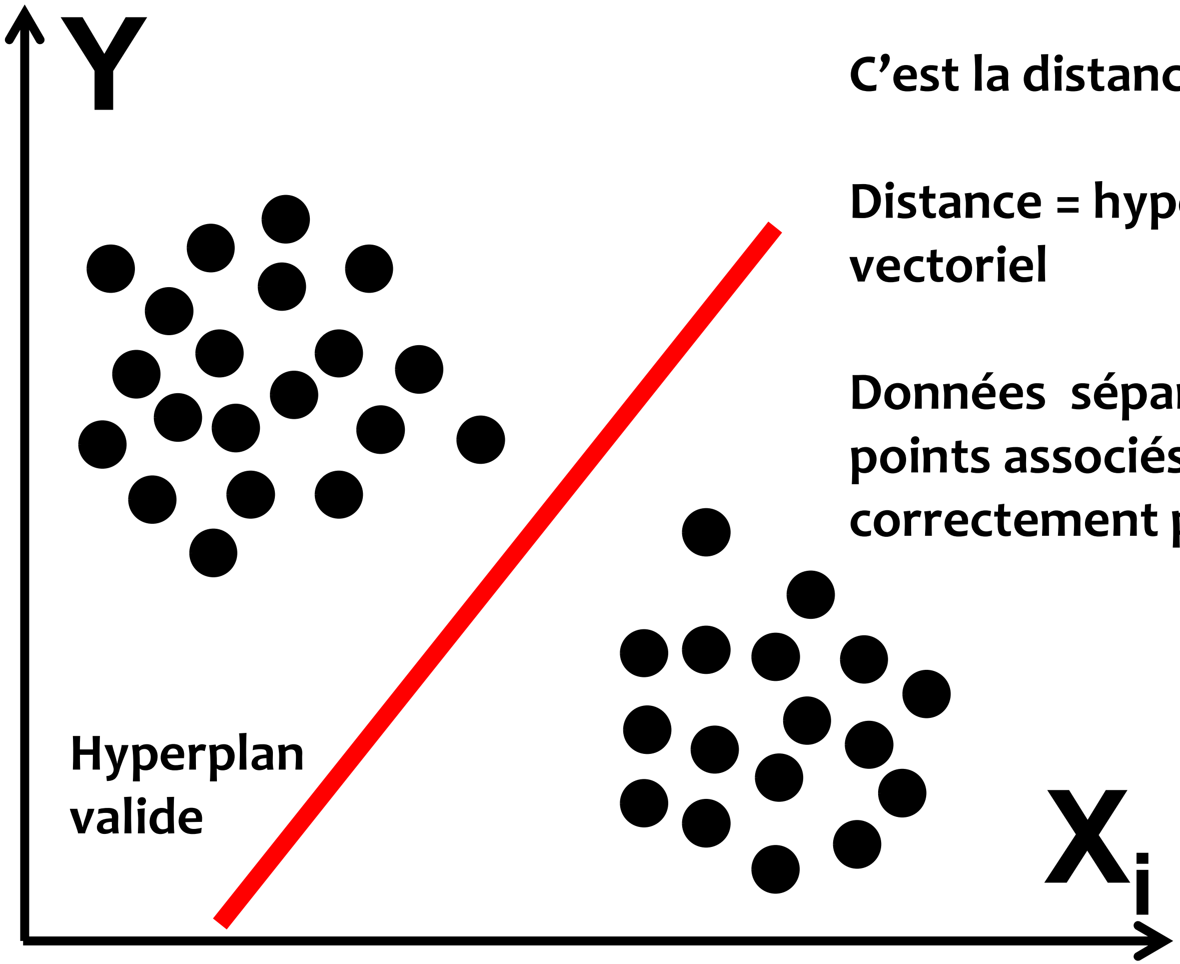


C'est un algorithmes de classement statistique.

Son rôle est de séparer des échantillons

Les échantillons séparés ont des propriétés similaires, mesurées sur des observations

Qu'est-ce qu'une marge ?

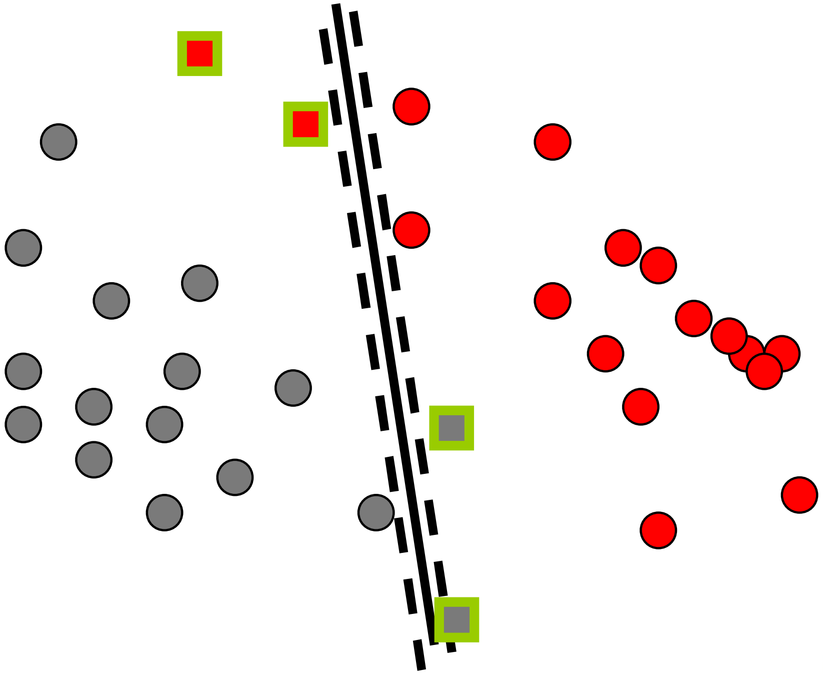


C'est la distance entre les échantillons séparés

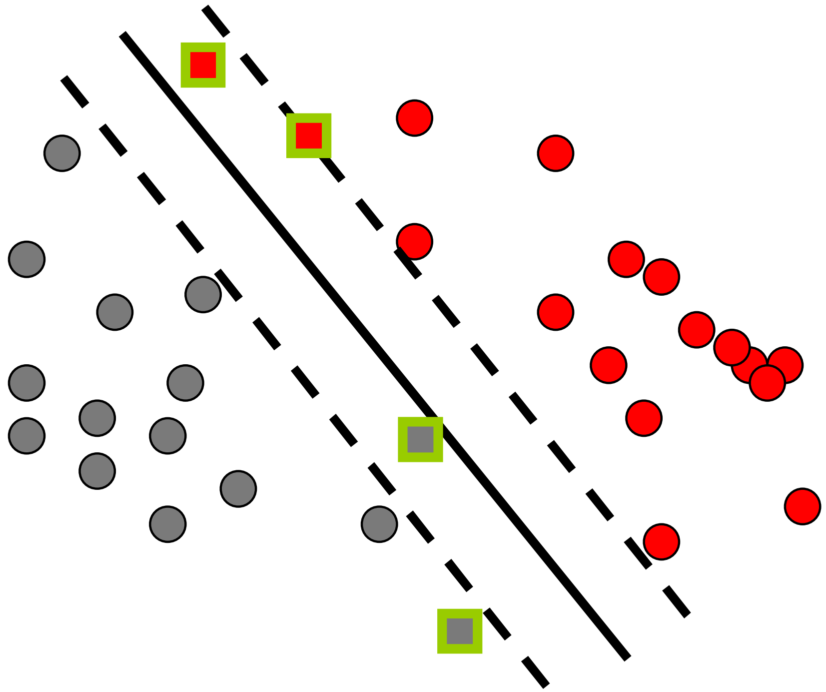
Distance = hyperplan : Plan de séparation - Espace vectoriel

Données séparables linéairement : si tous les points associés aux données peuvent être séparés correctement par une frontière linéaire

CHOIX D'UN CLASSIFIEUR À VASTE MARGE



MAUVAIS



BON

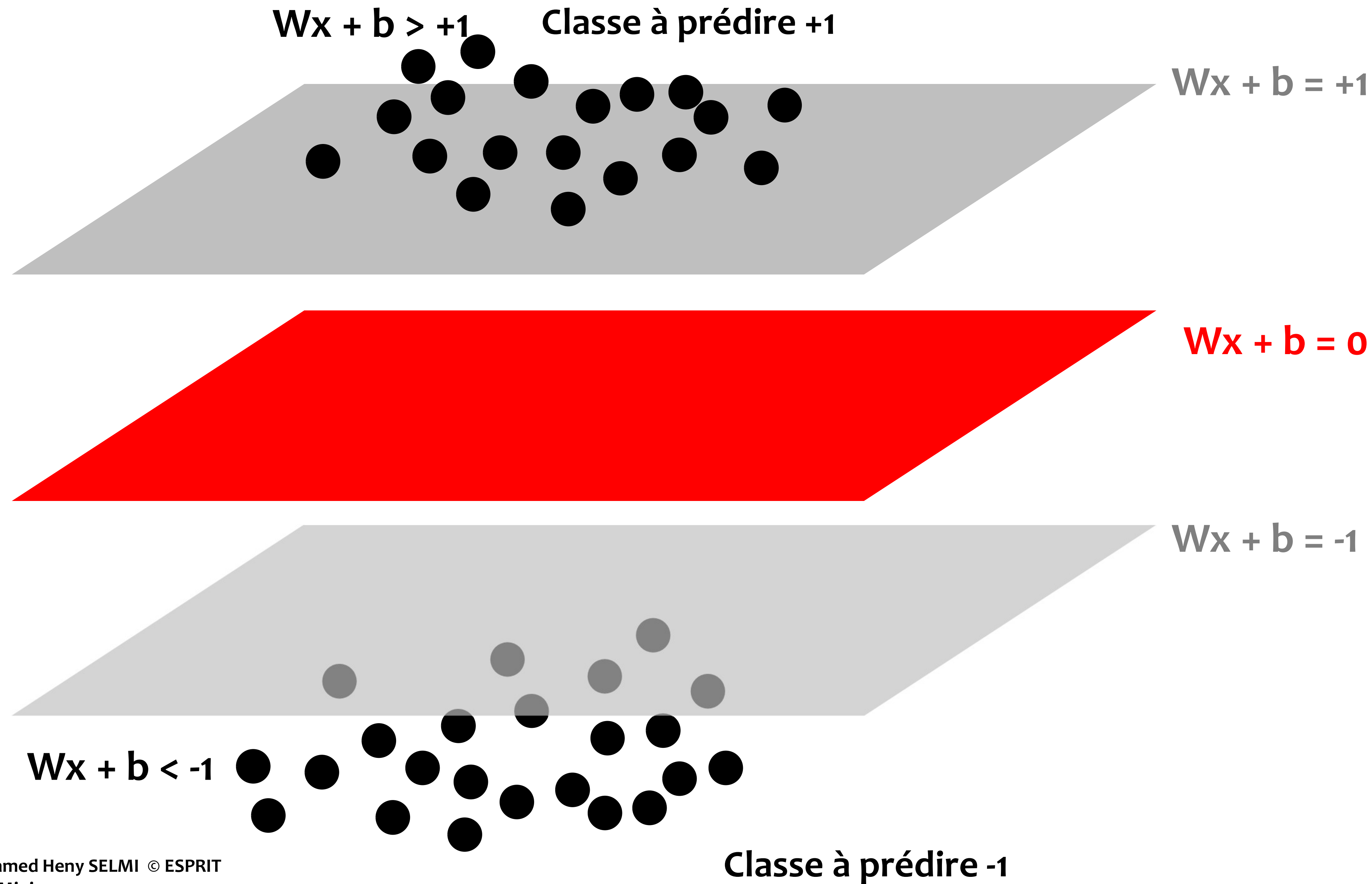
HYPERPLANS SÉPARATEURS

On cherche *h* sous forme d'une fonction linéaire : $h(x) = w.x + b$

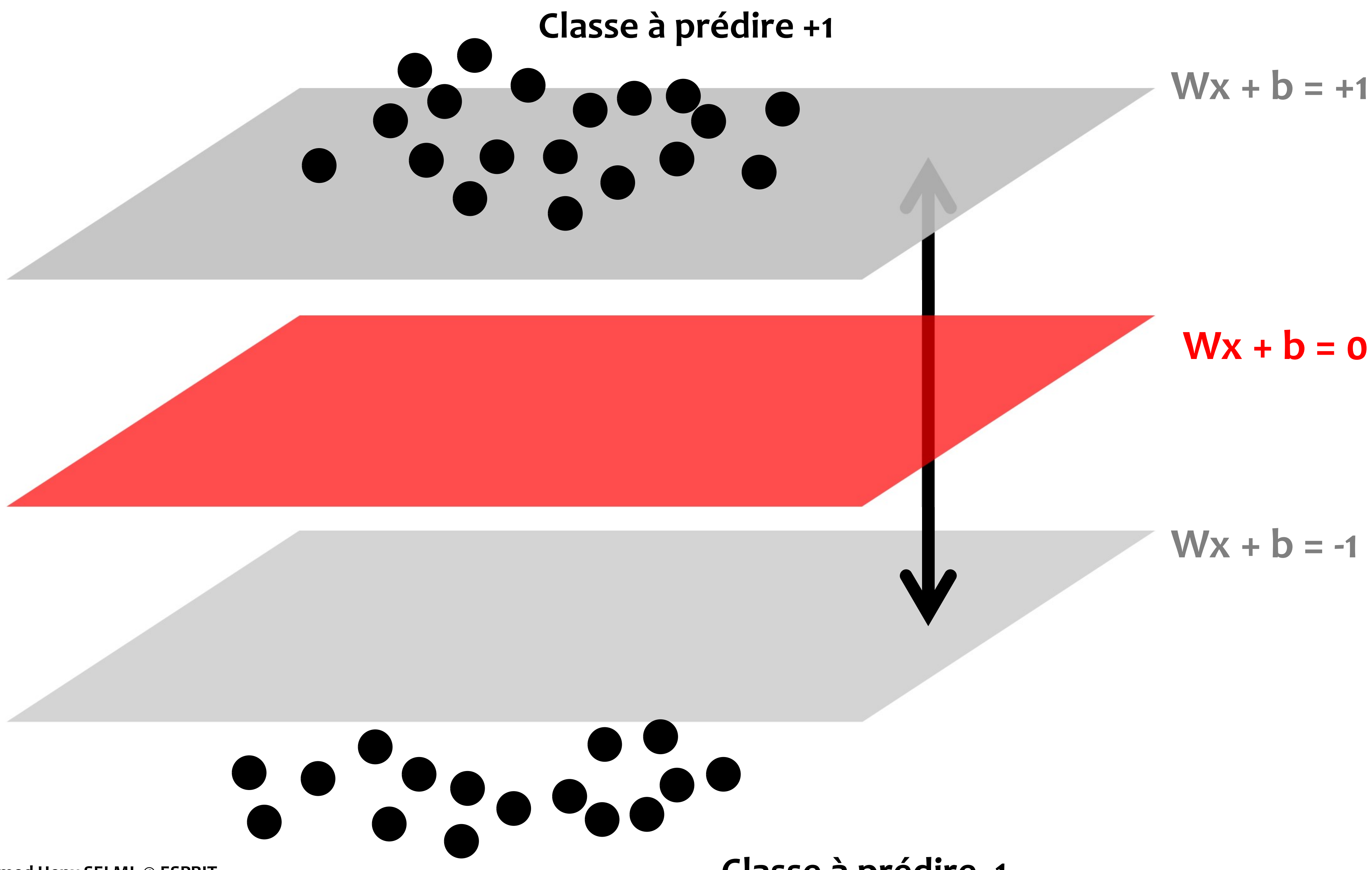
La *surface de séparation* est donc l'hyperplan $w.x + b = 0$

$w.x + b = 0$ *est valide si* x *est sur la marge*

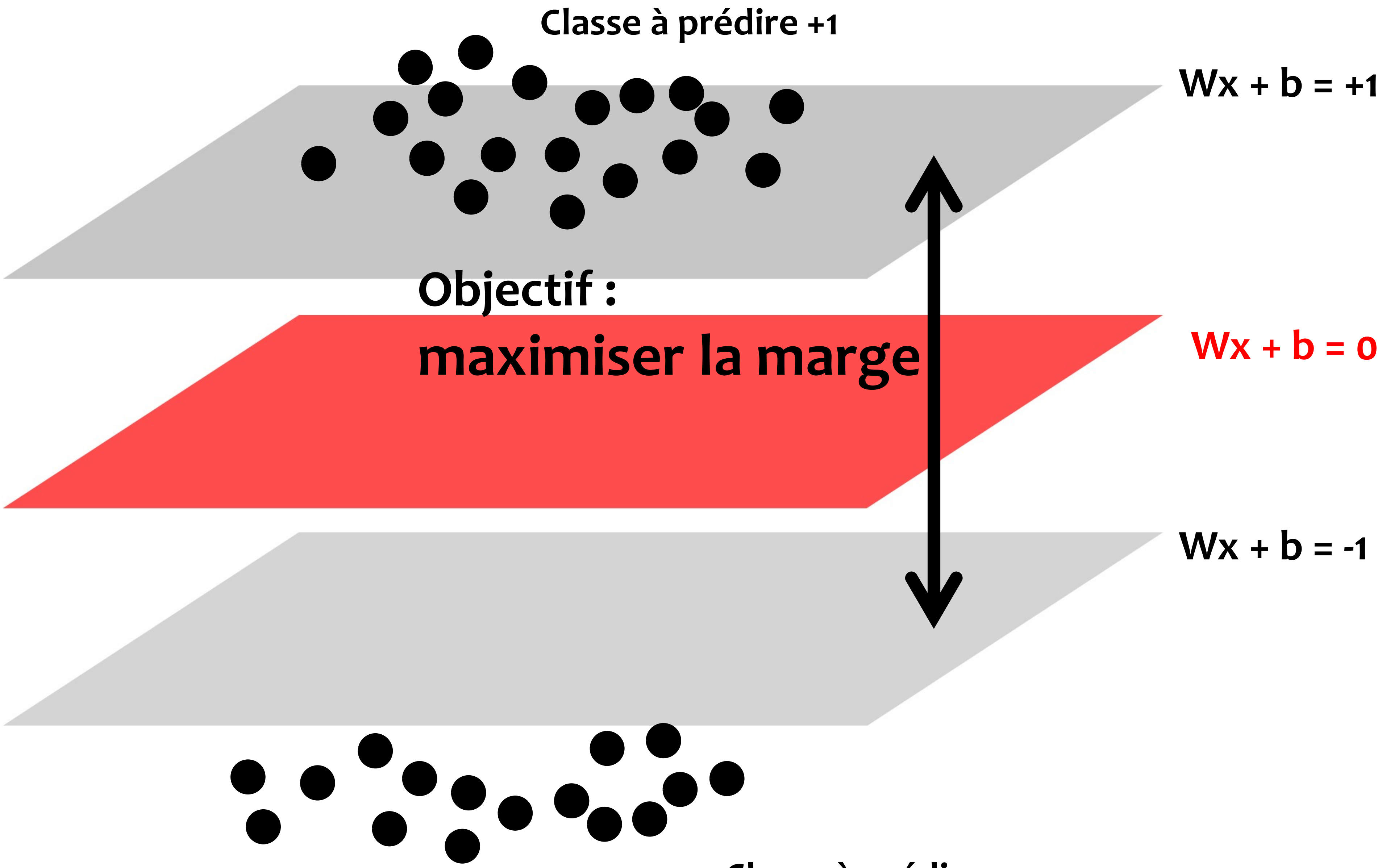
HYPERPLANS SÉPARATEURS



MARGE MAXIMALE ENTRE HYPERPLANS SÉPARATEURS



MARGE MAXIMALE ENTRE HYPERPLANS SÉPARATEURS



LES VECTEURS SUPPORTS MACHINES

Vecteurs supports (Support vector machine) :
Se sont les points les plus proches,
qui seuls sont utilisés pour la détermination de l'hyperplan

$$\overrightarrow{x_2} = \overrightarrow{x_1} + \vec{d}$$

$$\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{d}$$

$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\vec{d}$$

$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$

$$donc$$

$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$

$$wx_2 - wx_1 = 2$$

$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

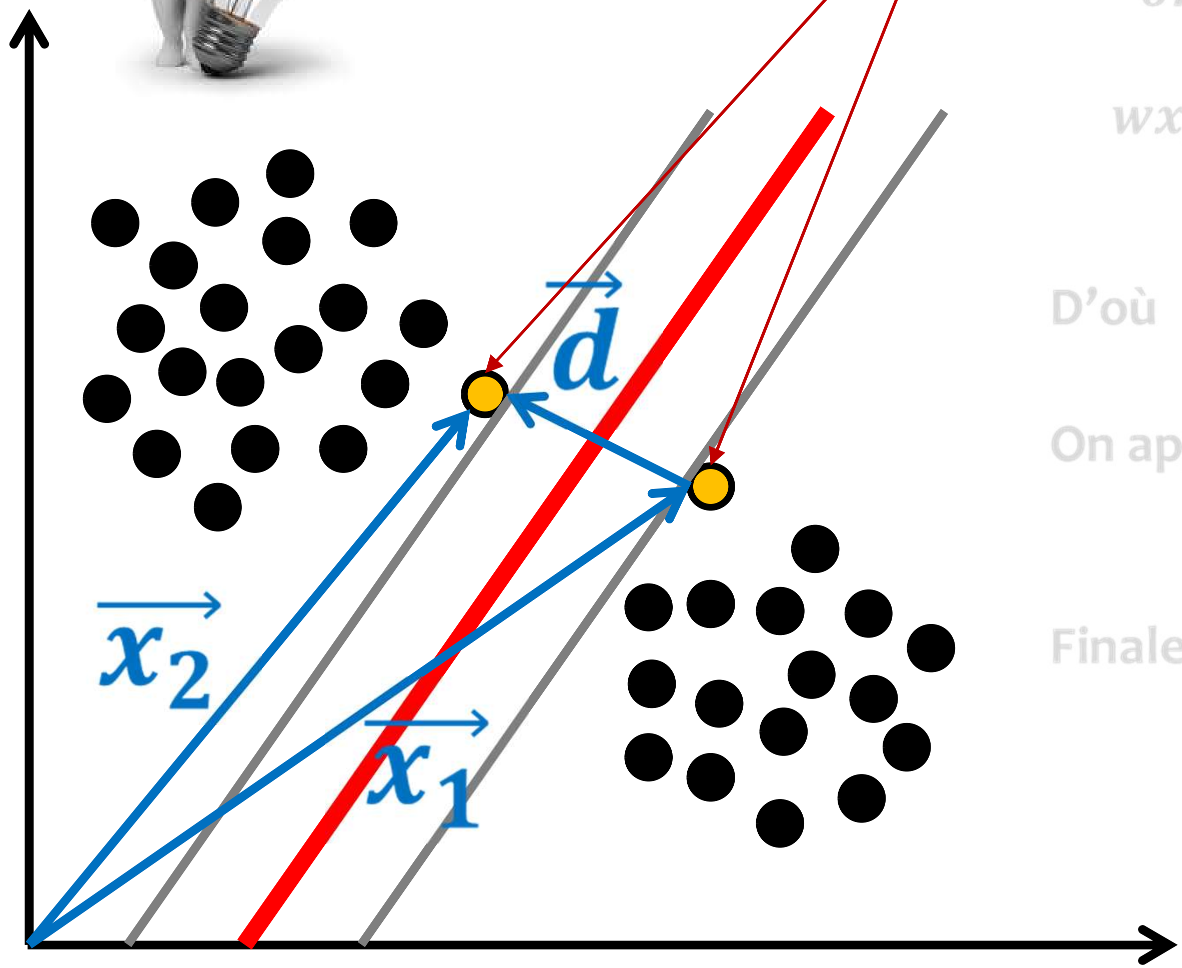
On applique la norme :

$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$

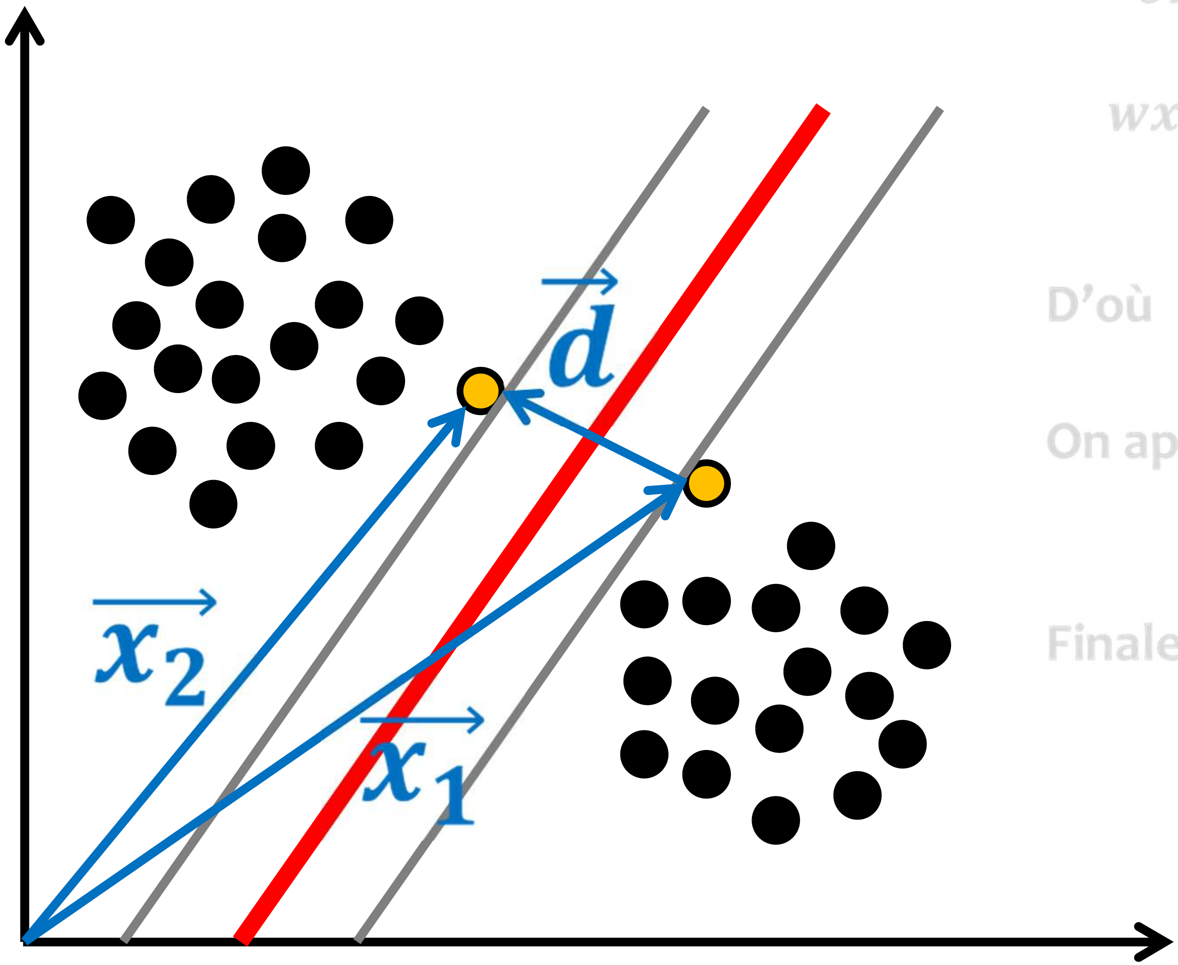
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$



CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\overrightarrow{x_2} = \overrightarrow{x_1} + \overrightarrow{d}$$

$$\overrightarrow{x_2} - \overrightarrow{x_1} = \overrightarrow{d}$$

$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\overrightarrow{d}$$

$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$

$$donc$$

$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$

$$wx_2 - wx_1 = 2$$

$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\overrightarrow{d}$$

On applique la norme :

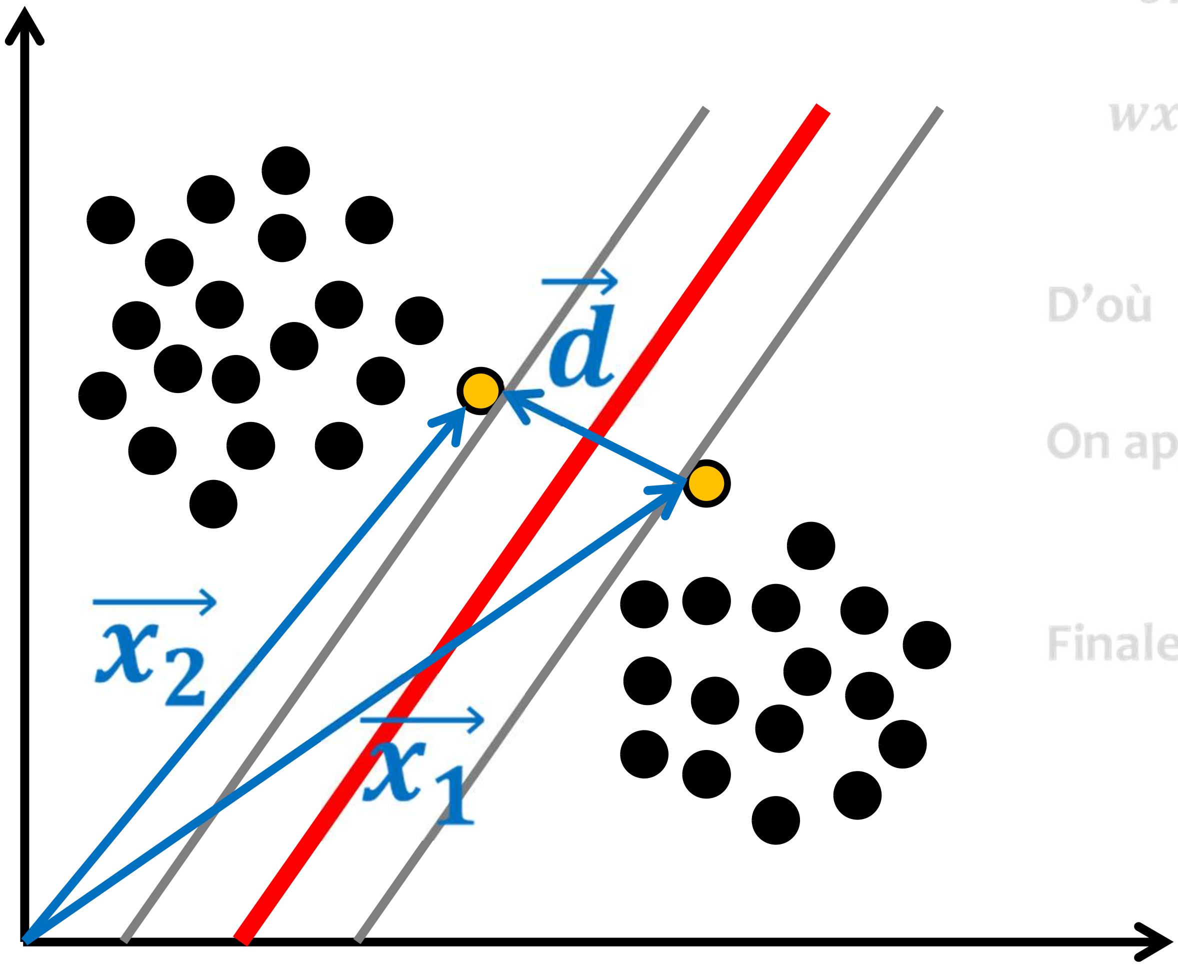
$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$

$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\overrightarrow{x_2} = \overrightarrow{x_1} + \overrightarrow{d}$$
$$\overrightarrow{x_2} - \overrightarrow{x_1} = \overrightarrow{d}$$
$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\overrightarrow{d}$$
$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$
$$donc$$
$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$
$$wx_2 - wx_1 = 2$$
$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\overrightarrow{d}$$

On applique la norme :

$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS

$$\overrightarrow{x_2} = \overrightarrow{x_1} + \vec{d}$$
$$\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{d}$$
$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\vec{d}$$

$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$
$$donc$$
$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$
$$wx_2 - wx_1 = 2$$
$$w(x_2 - x_1) = 2$$

D'où

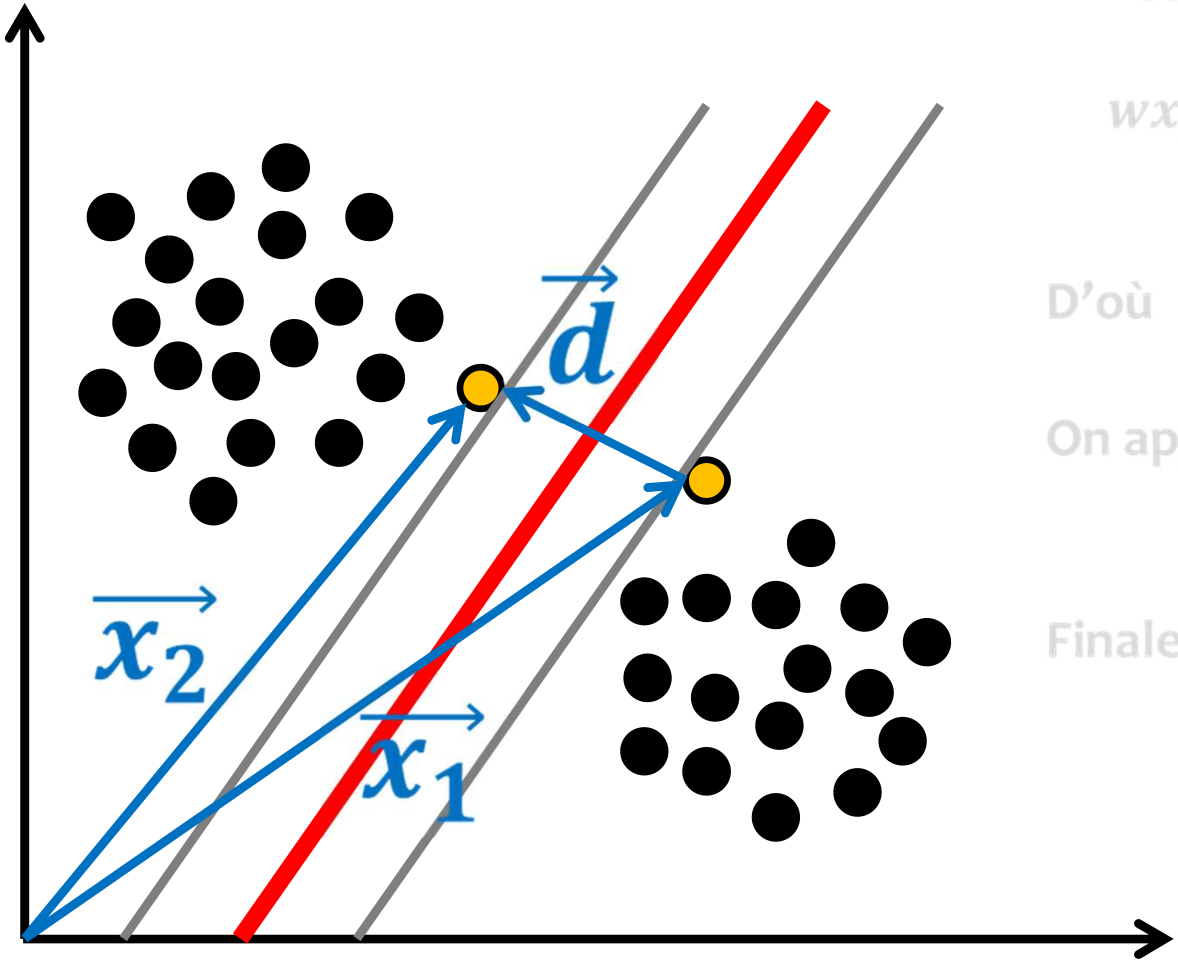
$$2 = w.\vec{d}$$

On applique la norme :

$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$



CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS

$$\overrightarrow{x_2} = \overrightarrow{x_1} + \vec{d}$$

$$\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{d}$$

$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\vec{d}$$

$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$

$$donc$$

$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$

$$wx_2 - wx_1 = 2$$

$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

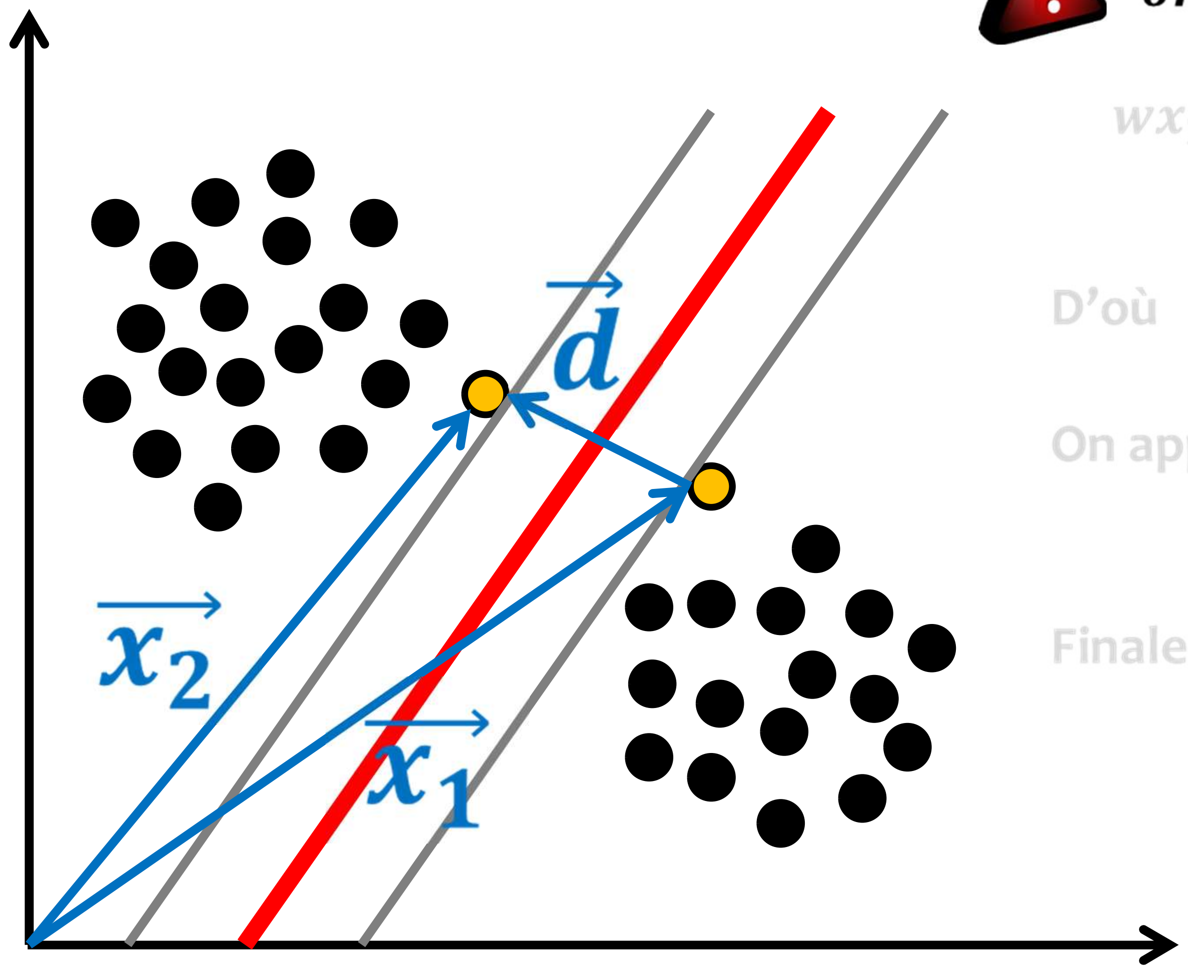
On applique la norme :

$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$

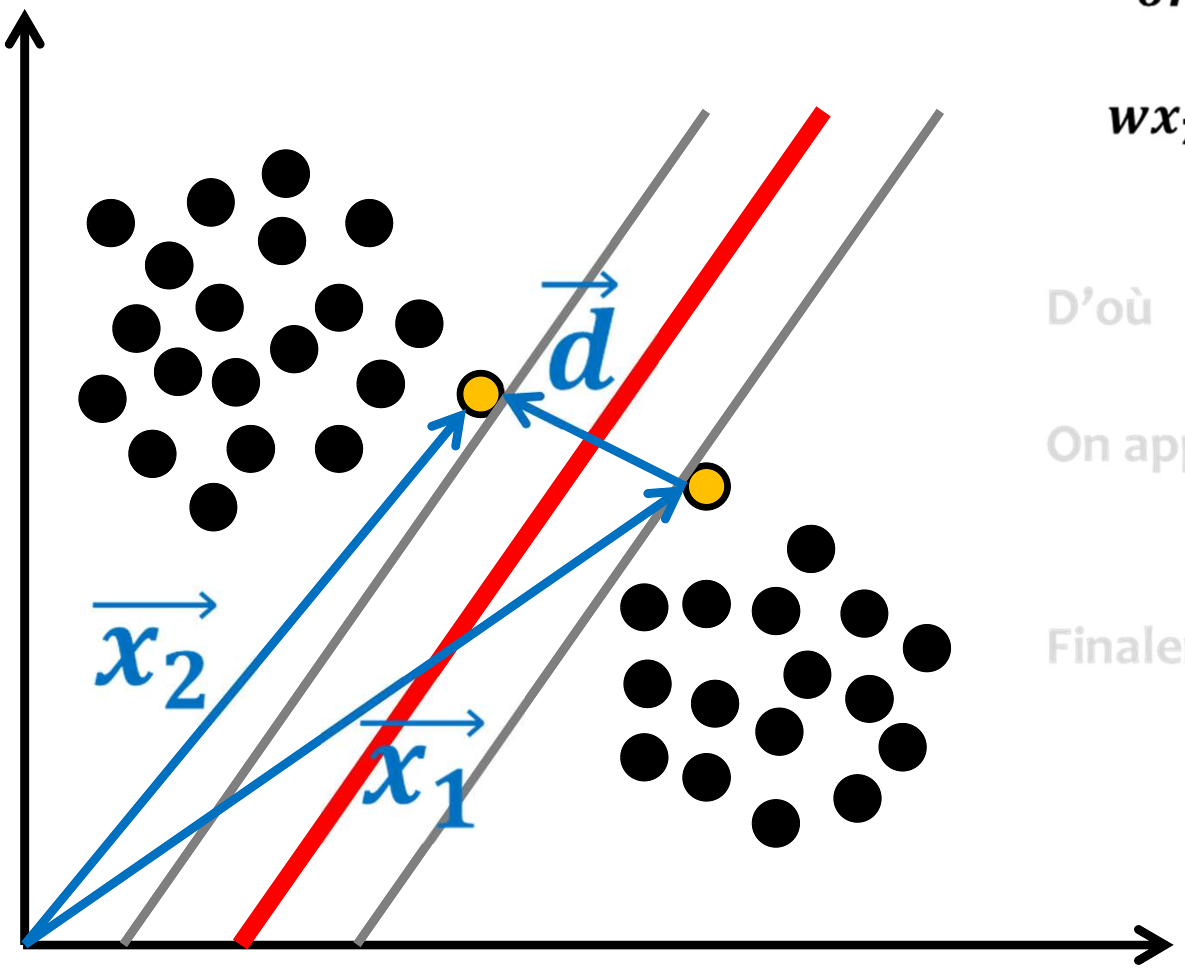
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$



CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\overrightarrow{x_2} = \overrightarrow{x_1} + \vec{d}$$

$$\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{d}$$

$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\vec{d}$$

$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$

$$donc$$

$$wx_2 + \cancel{b} - wx_1 - \cancel{b} = +1 - (-1) = 2$$

$$wx_2 - wx_1 = 2$$

$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

On applique la norme :

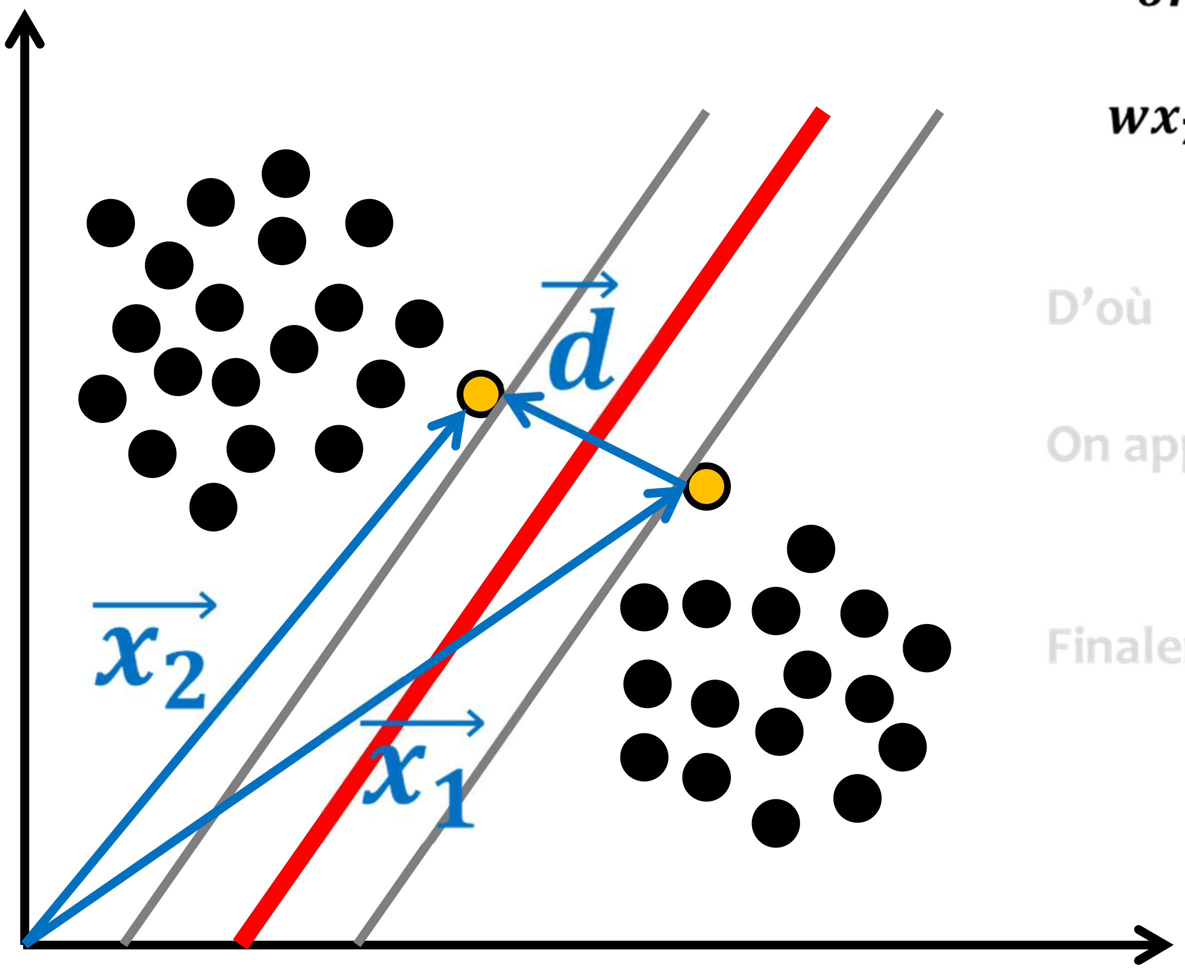
$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$

$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\overrightarrow{x_2} = \overrightarrow{x_1} + \vec{d}$$

$$\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{d}$$

$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\vec{d}$$

$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$

$$donc$$

$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$

$$wx_2 - wx_1 = 2$$

$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

On applique la norme :

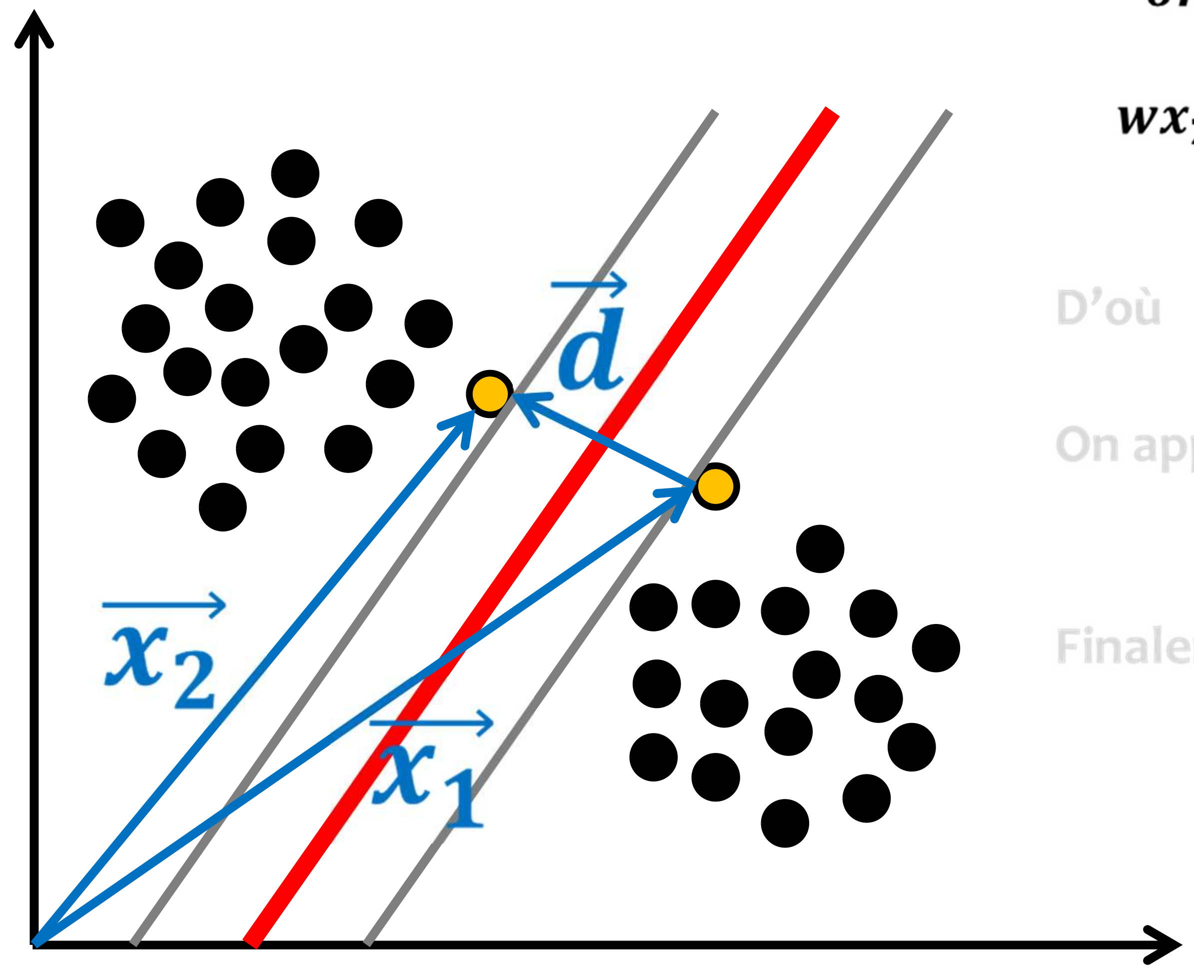
$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$

$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\vec{x_2} = \vec{x_1} + \vec{d}$$
$$\vec{x_2} - \vec{x_1} = \vec{d}$$
$$w.(\vec{x_2} - \vec{x_1}) = w.\vec{d}$$
$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$
$$donc$$
$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$
$$wx_2 - wx_1 = 2$$
$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

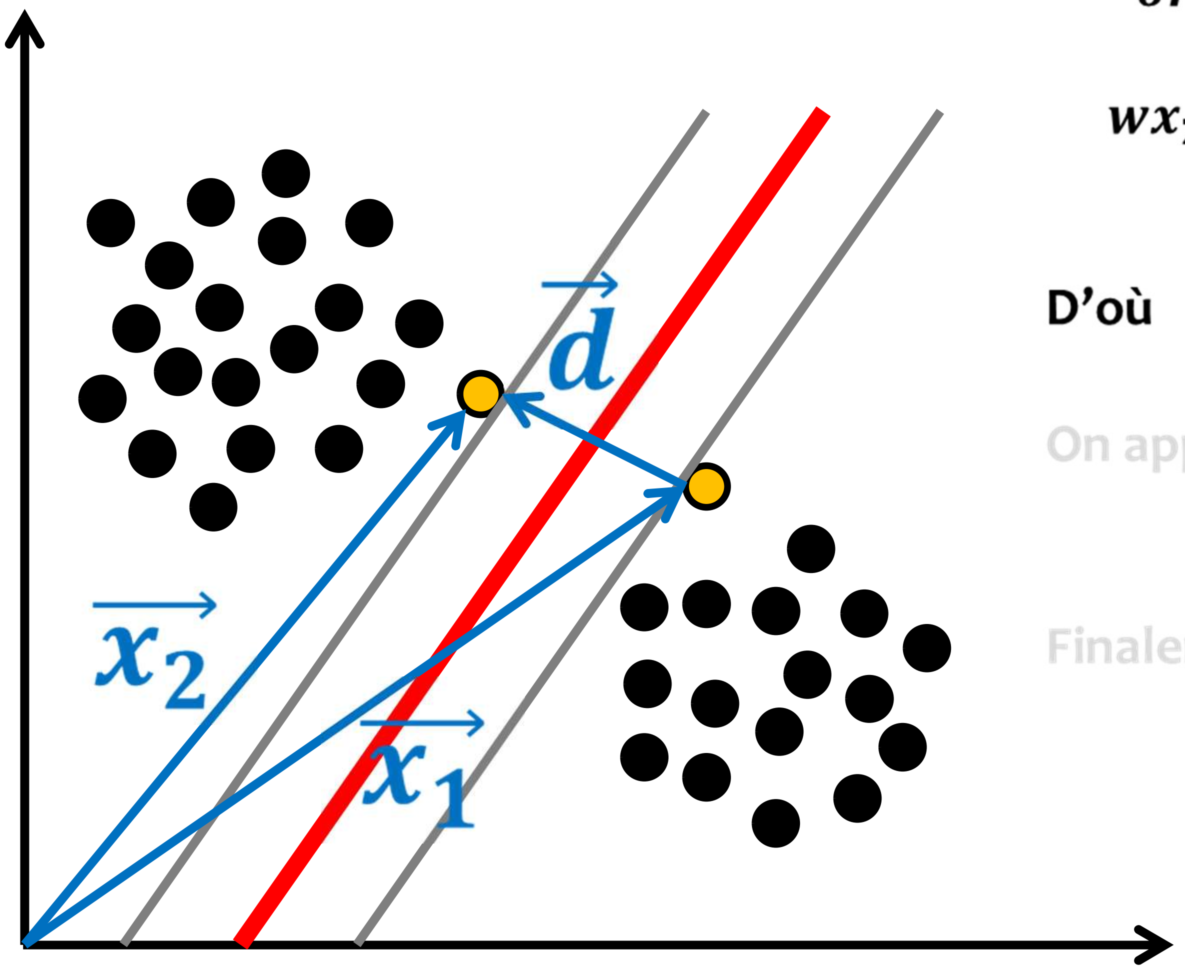
On applique la norme :

$$\|2\| = \|w\|.\|\vec{d}\|$$
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\overrightarrow{x_2} = \overrightarrow{x_1} + \vec{d}$$
$$\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{d}$$
$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\vec{d}$$
$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$
$$donc$$
$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$
$$wx_2 - wx_1 = 2$$
$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

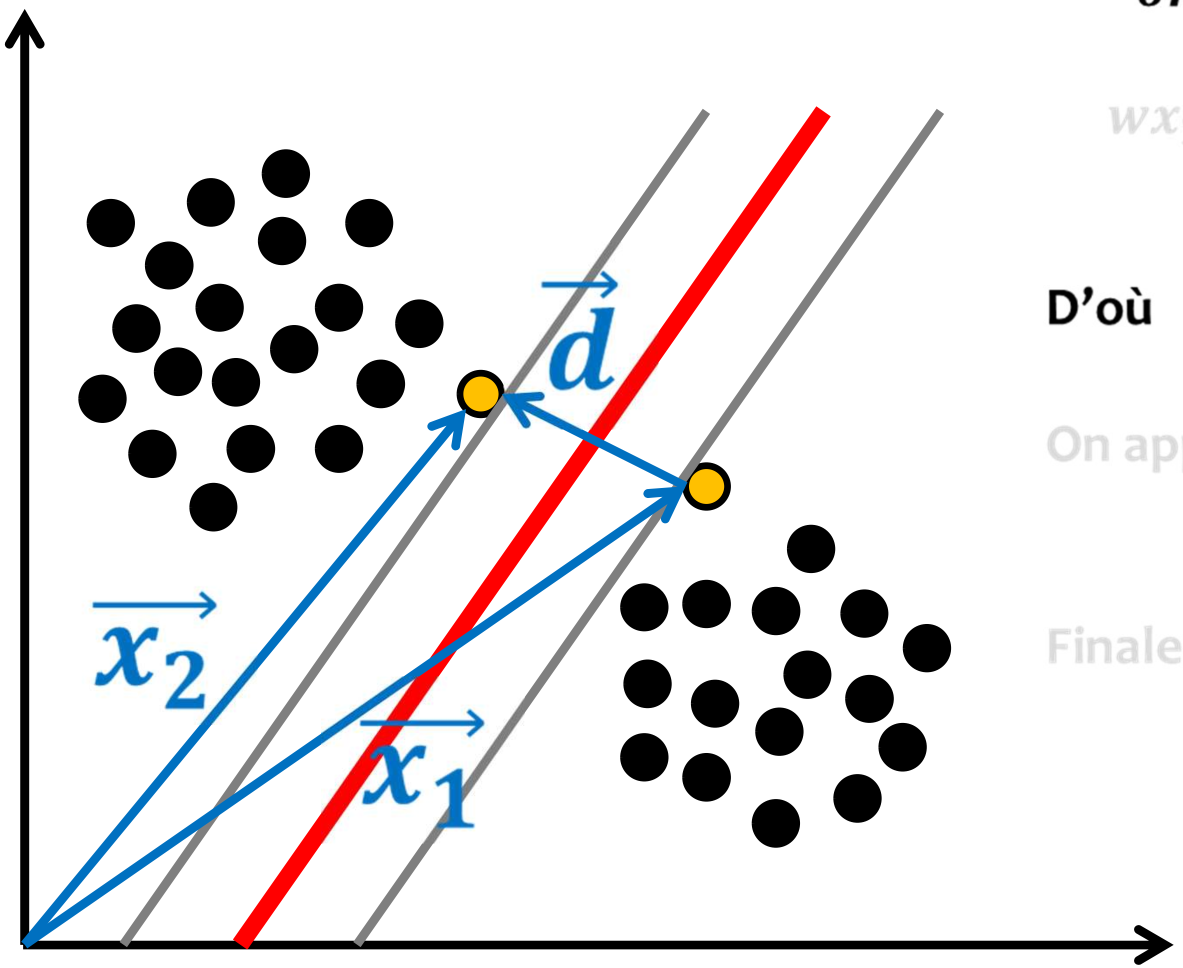
On applique la norme :

$$\|2\| = \|w\|.\|\vec{d}\|$$
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\overrightarrow{x_2} = \overrightarrow{x_1} + \vec{d}$$

$$\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{d}$$

$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\vec{d}$$

$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$

$$donc$$

$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$

$$wx_2 - wx_1 = 2$$

$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

On applique la norme :

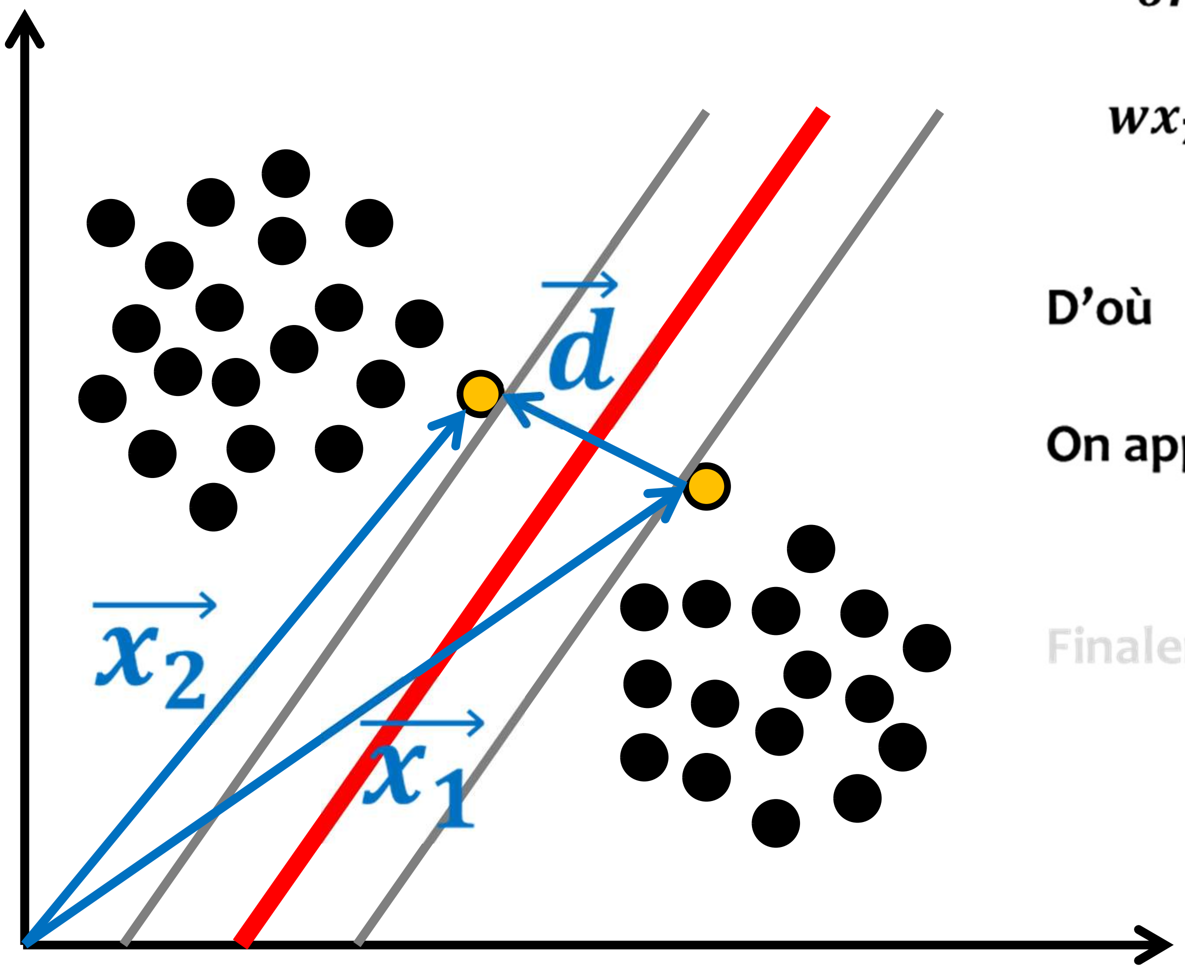
$$\|2\| = \|w\|.\|\overrightarrow{d}\|$$

$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\overrightarrow{x_2} = \overrightarrow{x_1} + \vec{d}$$
$$\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{d}$$
$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\vec{d}$$
$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$
$$donc$$
$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$
$$wx_2 - wx_1 = 2$$
$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

On applique la norme :

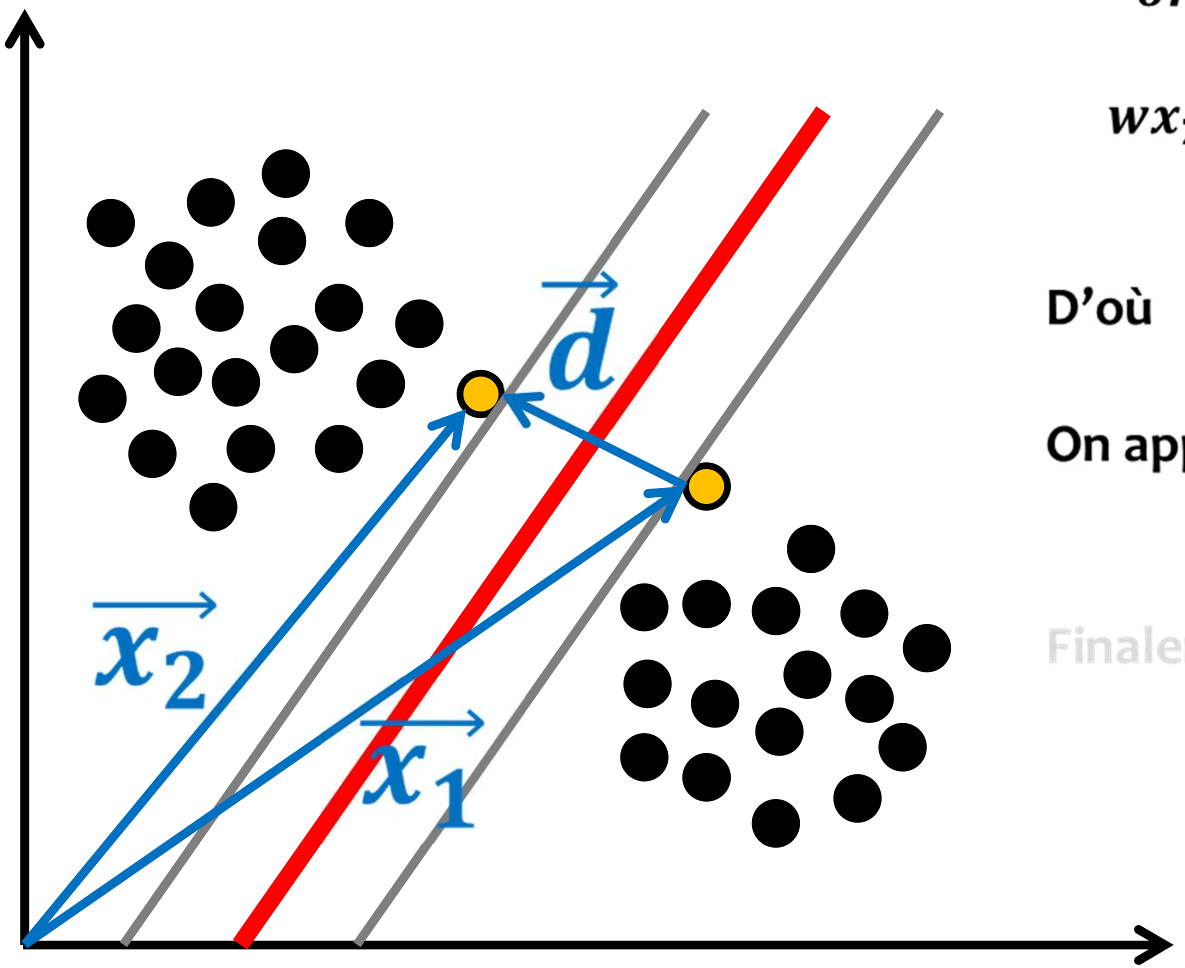
$$\|2\| = \|w\|.\|\vec{d}\|$$

$$2 = \|w\|.\,d$$

Finalement :

$$d = \frac{2}{\|w\|}$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\vec{x_2} = \vec{x_1} + \vec{d}$$
$$\vec{x_2} - \vec{x_1} = \vec{d}$$
$$w.(\vec{x_2} - \vec{x_1}) = w.\vec{d}$$
$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$
$$donc$$
$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$
$$wx_2 - wx_1 = 2$$
$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

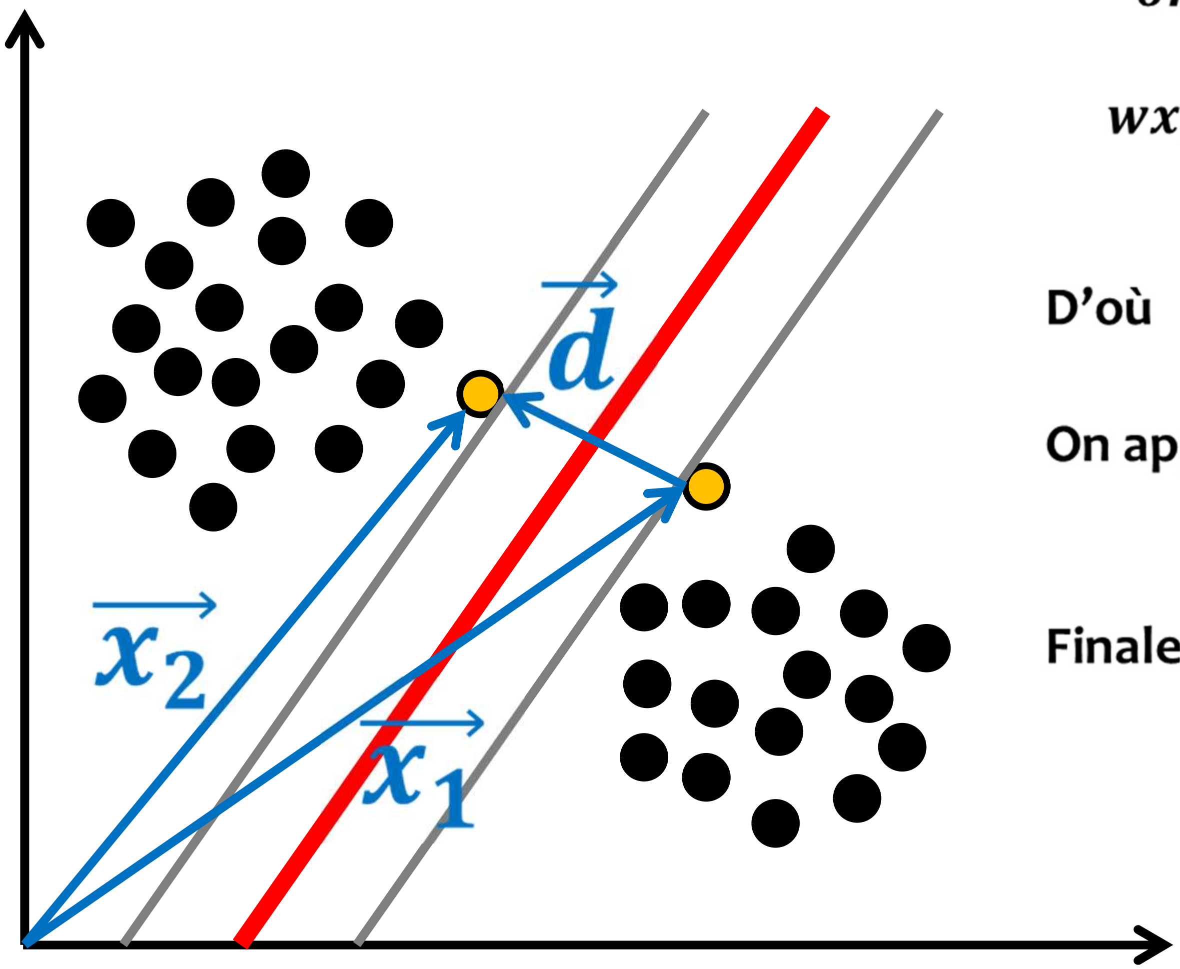
On applique la norme :

$$\|2\| = \|w\|.\,\overline{\|d\|}$$
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\overrightarrow{x_2} = \overrightarrow{x_1} + \vec{d}$$
$$\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{d}$$
$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\vec{d}$$
$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$
$$donc$$
$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$
$$wx_2 - wx_1 = 2$$
$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

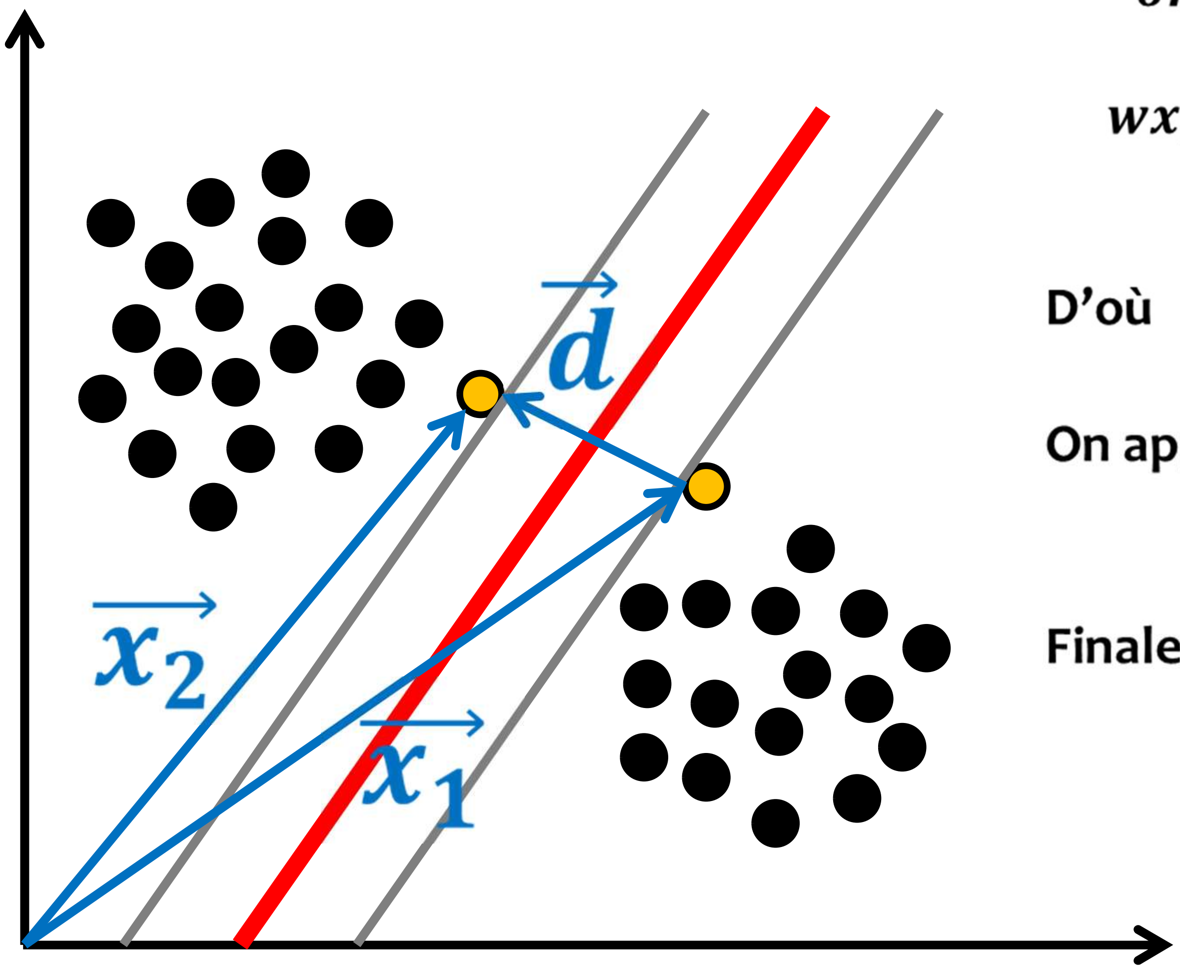
On applique la norme :

$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\overrightarrow{x_2} = \overrightarrow{x_1} + \vec{d}$$
$$\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{d}$$
$$w.(\overrightarrow{x_2} - \overrightarrow{x_1}) = w.\vec{d}$$
$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$
$$donc$$
$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$
$$wx_2 - wx_1 = 2$$
$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

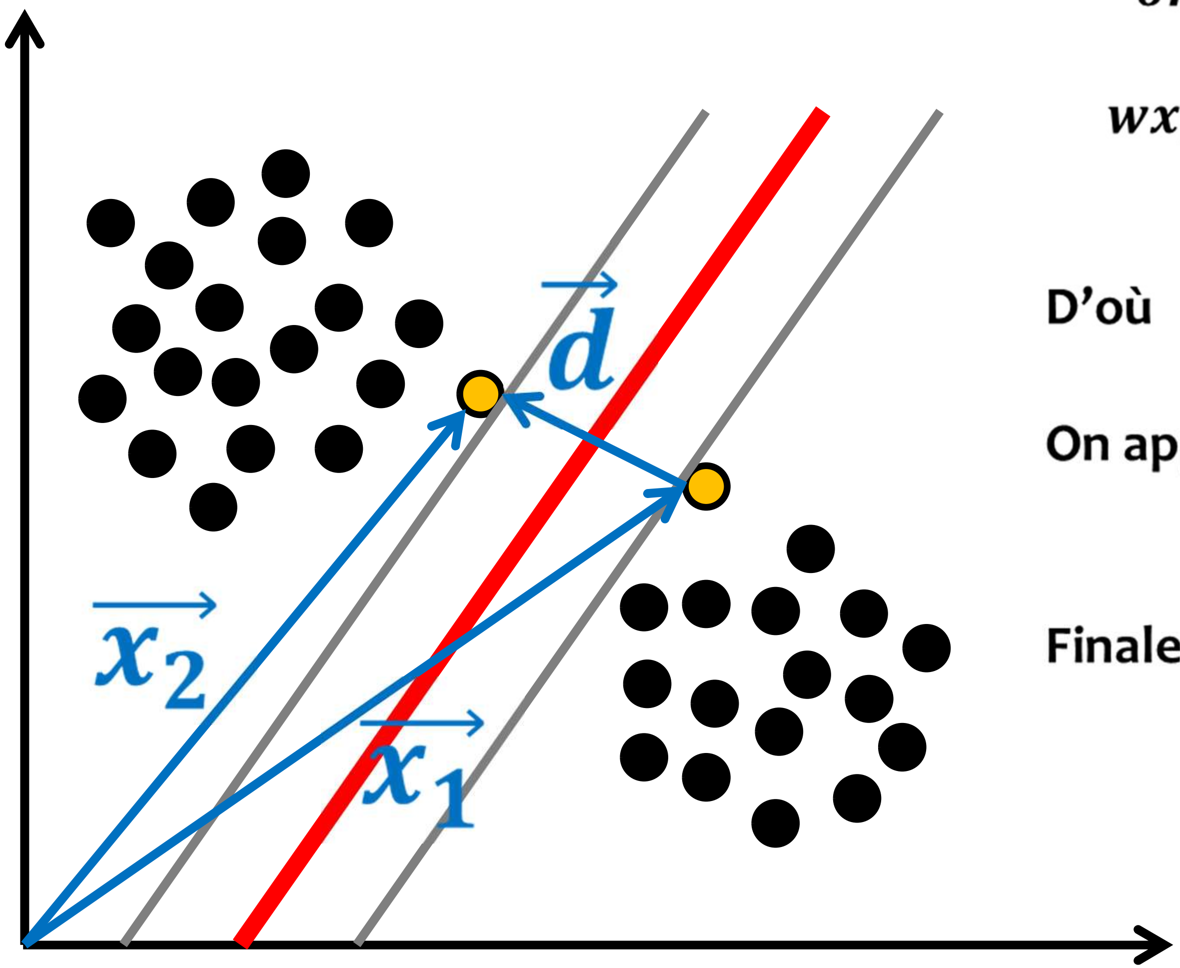
On applique la norme :

$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}\ \text{à}\ maximiser$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\vec{x_2} = \vec{x_1} + \vec{d}$$
$$\vec{x_2} - \vec{x_1} = \vec{d}$$
$$w.(\vec{x_2} - \vec{x_1}) = w.\vec{d}$$
$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$
$$donc$$
$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$
$$wx_2 - wx_1 = 2$$
$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

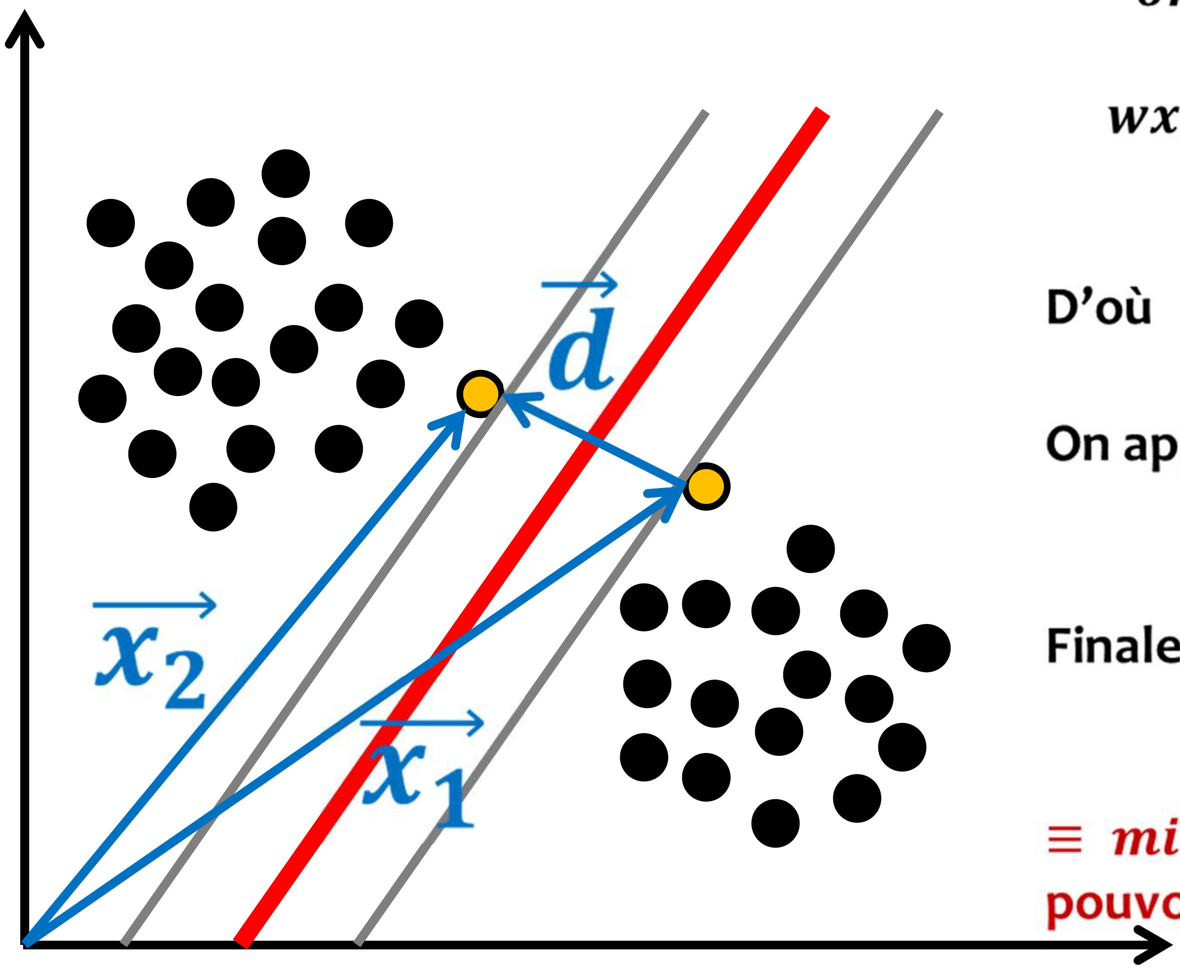
On applique la norme :

$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|} \text{ à } maximiser$$
$$\equiv\ minimiser\ \|w\|$$

CALCUL DE LA MARGE ENTRE HYPERPLANS SÉPARATEURS



$$\vec{x_2} = \vec{x_1} + \vec{d}$$
$$\vec{x_2} - \vec{x_1} = \vec{d}$$
$$w.(\vec{x_2} - \vec{x_1}) = w.\vec{d}$$
$$or\ wx_2 + b = +1\ et\ wx_1 + b = -1$$
$$donc$$
$$wx_2 + b - wx_1 - b = +1 - (-1) = 2$$
$$wx_2 - wx_1 = 2$$
$$w(x_2 - x_1) = 2$$

D'où

$$2 = w.\vec{d}$$

On applique la norme :

$$\|2\| = \|w\|.\overrightarrow{\|d\|}$$
$$2 = \|w\|.\ d$$

Finalement :

$$d = \frac{2}{\|w\|}\ \text{à}\ maximiser$$

$\equiv$ *minimiser* $\|w\|$ tout en préservant le pouvoir de classification

OPTIMISATION DE LA MARGE

Afin d'optimiser la marge, il faut réaliser les deux objectifs suivants :

1

Classer correctement les individus

$$\forall\, y_i = +1, wx_i + b \geq +1$$
$$\forall\, y_i = -1, wx_i + b \leq -1$$

$$y_i\,(wx_i + b) \geq +1\ \forall i$$

2

Maximiser la marge

$$d = \frac{2}{\|w\|} \text{ à } maximiser$$

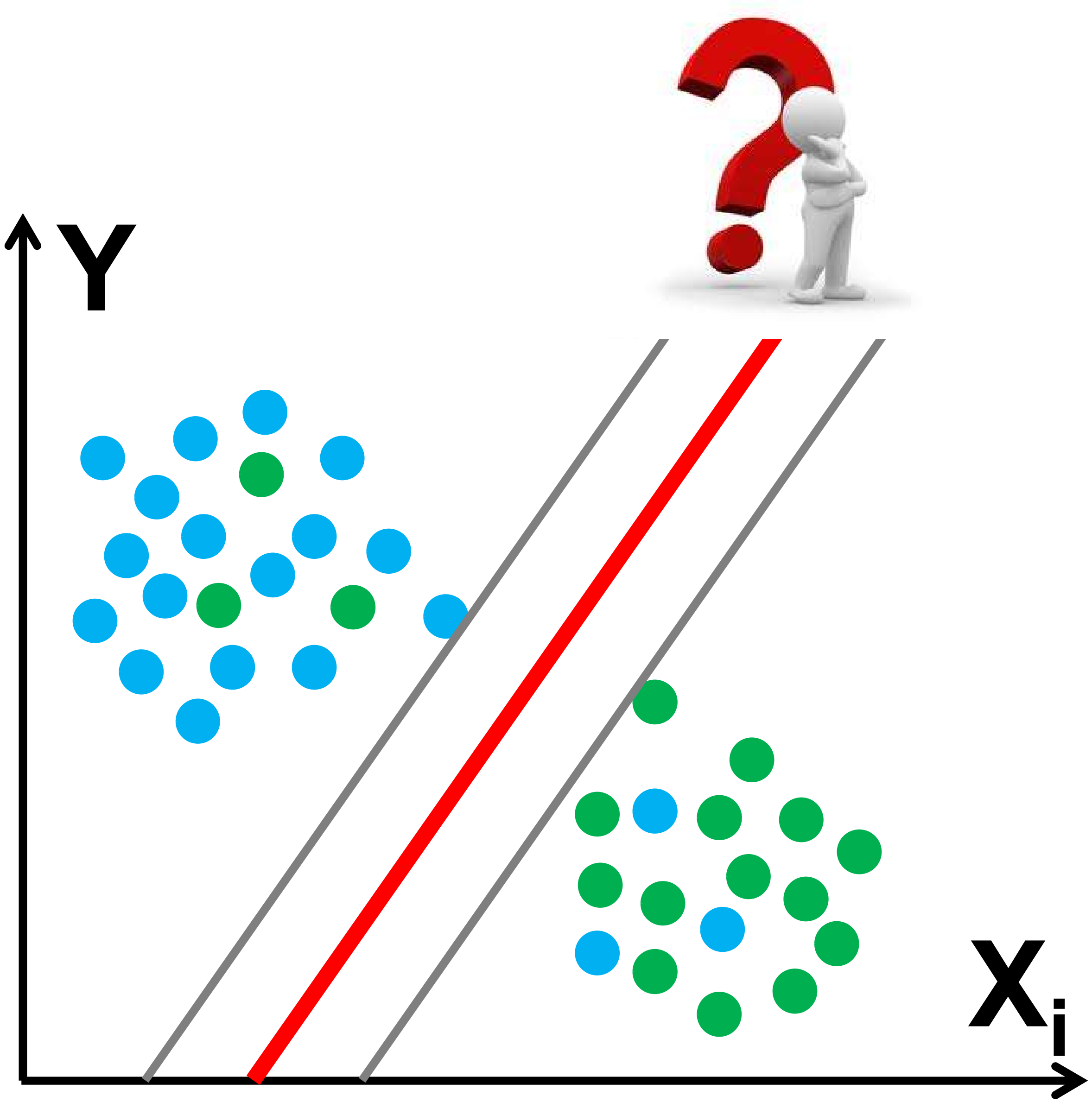
$$minimiser\ \|w\| \equiv minimiser\ \frac{\|w\|^2}{2}$$

$$\|w\|^2\ pour\ \text{é}viter\ la\ racine$$

$$\frac{1}{2}\ reste\ utile\ en\ cas\ de\ d\text{é}rivation$$

$$\frac{1}{2}\|w\|^2 = \frac{1}{2}\left(\sqrt{\sum_i {w_i}^2}\right)^2 = \frac{1}{2}\sum_i {w_i}^2$$

PROBLÉMATIQUE DES INDIVIDUS MAL PLACÉS

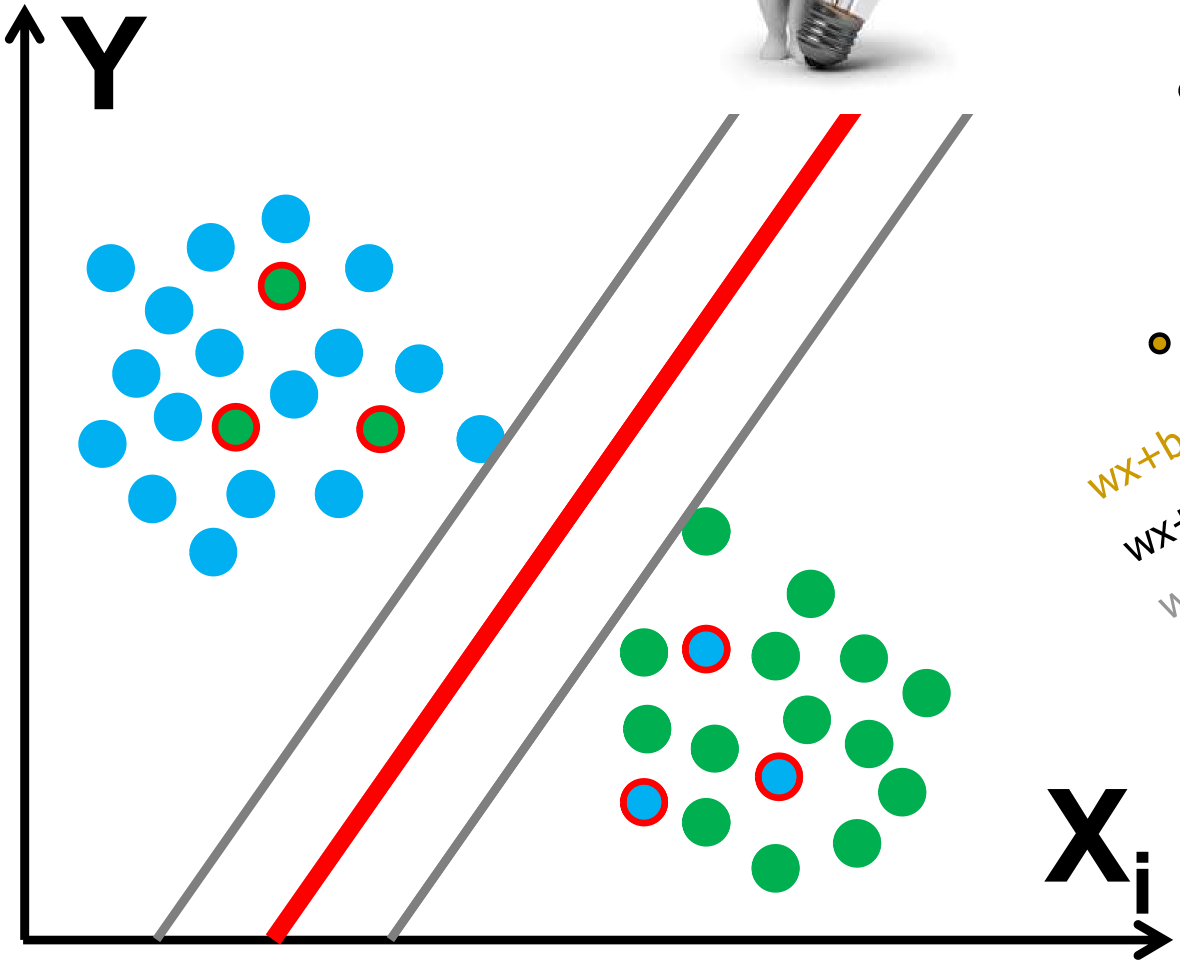


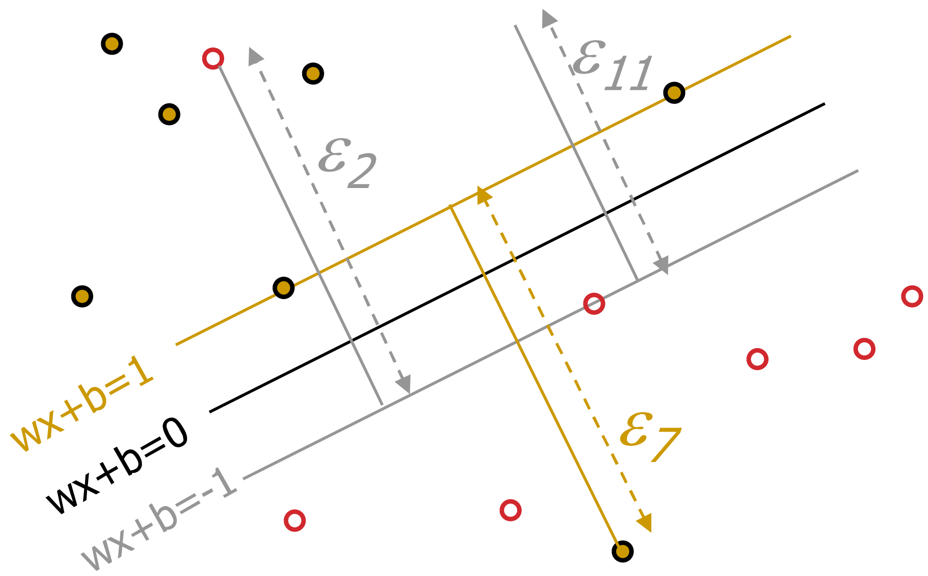
PROBLÉMATIQUE DES INDIVIDUS MAL PLACÉS

$$minimiser \frac{1}{2} \|w\|^2 + C.\sum_{1}^{N} \varepsilon_i$$

$$(taux\ d'erreur\ de\ classification)$$

$$N : nombre\ d'individus\ mal\ classés$$





OPTIMISATION DE LA MARGE

Afin d'optimiser la marge, il faut réaliser les deux objectifs suivants :

1

Classer correctement les individus

$$\forall\, y_i = +1, wx_i + b \geq +1$$
$$\forall\, y_i = -1, wx_i + b \leq -1$$

$$y_i\,(wx_i + b) \geq +1 - \varepsilon_i \quad \forall i, \varepsilon_i$$

2

Maximiser la marge

$$d = \frac{2}{\|w\|} \text{ à } maximiser$$

$$minimiser\ \|w\| \equiv minimiser\ \frac{\|w\|^2}{2}$$

$$\|w\|^2\ pour\ \text{é}viter\ la\ racine$$

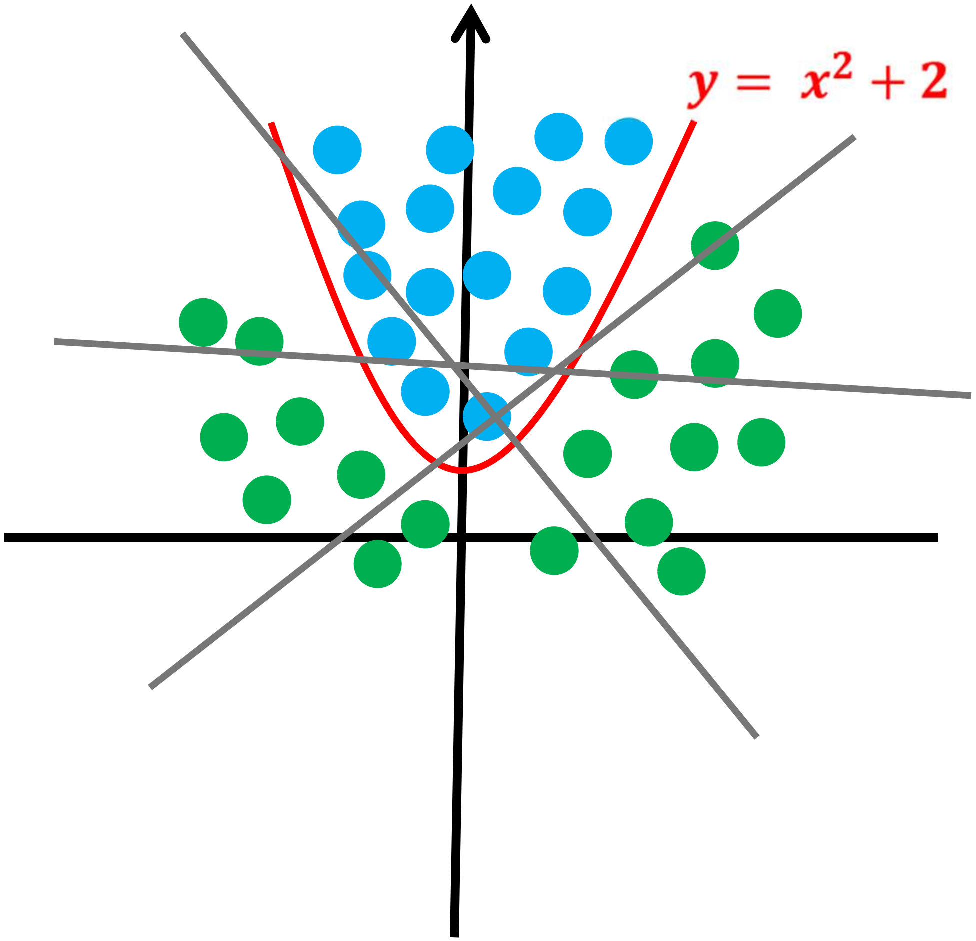
$$\frac{1}{2}\ reste\ utile\ en\ cas\ de\ d\text{é}rivation$$

$$\frac{1}{2}\|w\|^2 + C.\sum_1^N \varepsilon_i$$

C paramètre de control du sur apprentissage

PROBLÈME LINÉAIREMENT NON SÉPARABLE

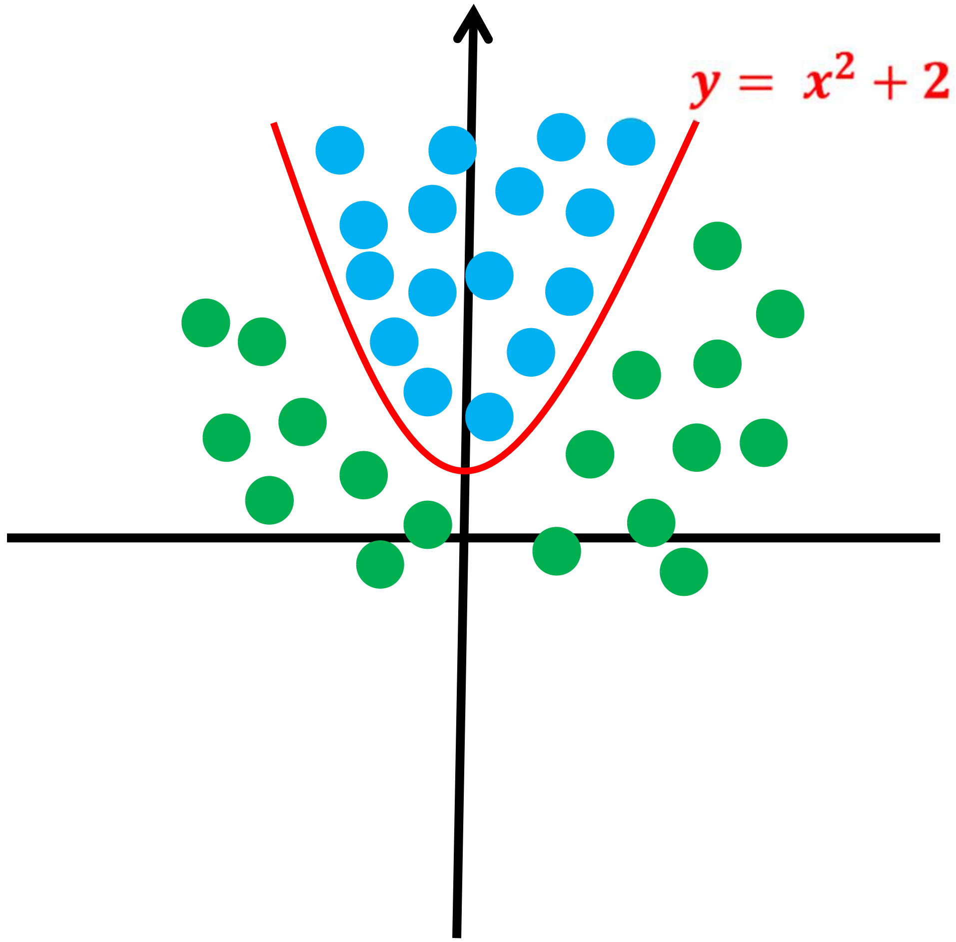
Exemple



la séparation linéaire est impossible :
Recours à d'autres type de classifieurs

PROBLÈME LINÉAIREMENT NON SÉPARABLE

Exemple

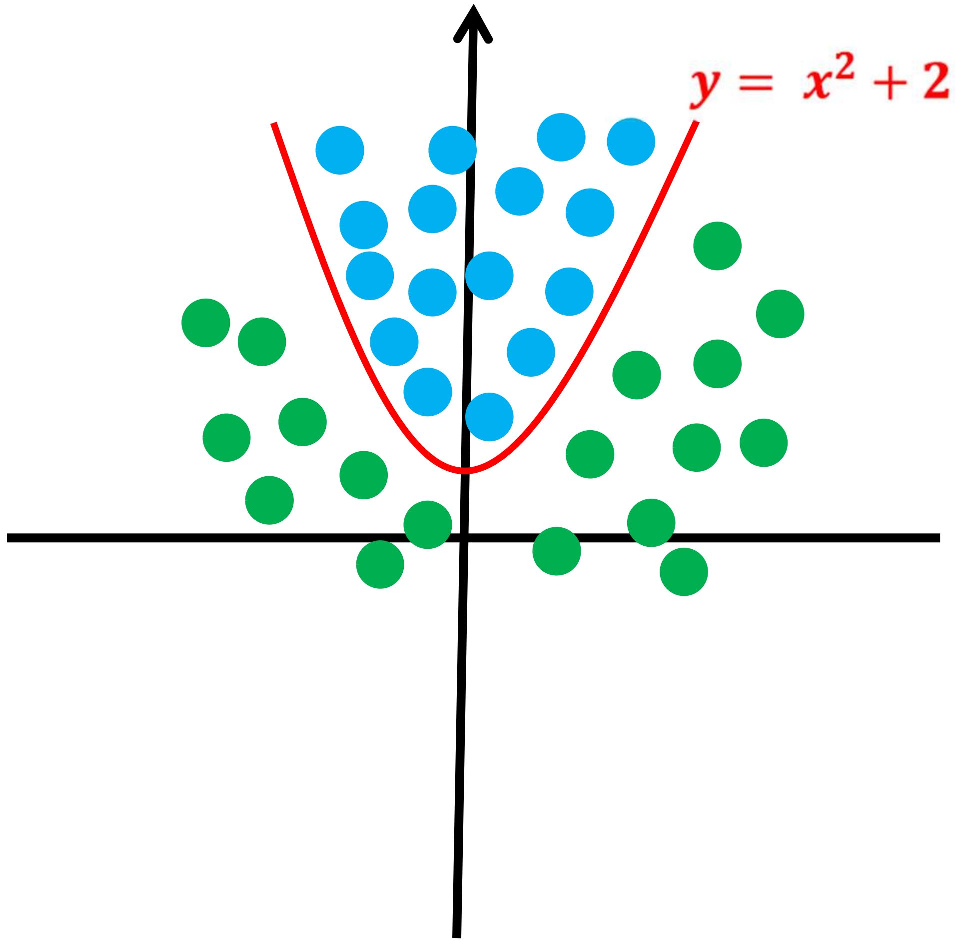


la séparation linéaire est impossible :
Recours à d'autres type de classifieurs

Classifieur non linéaire

PROBLÈME LINÉAIREMENT NON SÉPARABLE

Exemple



la séparation linéaire est impossible :
Recours à d'autres type de classifieurs

Classifieur non linéaire

Fonction Noyau

EXEMPLES DE FONCTION NOYAU

- Linéaire : $K(\mathbf{x_i},\mathbf{x_j}) = \mathbf{x_i}^T\mathbf{x_j}$

- Polynomiale : $K(\mathbf{x_i},\mathbf{x_j}) = (1+ \mathbf{x_i}^T\mathbf{x_j})^p$

- Gaussienne :

$$K(\mathbf{x}_\mathbf{i}, \mathbf{x}_\mathbf{j}) = \exp\left(-\frac{\left\|\mathbf{x}_\mathbf{i} - \mathbf{x}_\mathbf{j}\right\|^2}{2\sigma^2}\right)$$