

Séparateur à Vaste Marge : une méthode à noyaux pour la discrimination

Gaëlle Loosli - LISC

L'apprentissage automatique offre un grand panel de méthodes pour la discrimination binaire. Nous nous intéressons ici au plus populaire des algorithmes d'apprentissage à noyaux : le SVM (Support Vector Machine ou encore Séparateur à Vaste Marge).

Dans un premier temps nous justifierons cette popularité et préciserons à quel type de problème le SVM est le plus adapté.

Dans un second temps, nous expliquerons plus précisément ce que sont les noyaux et comment fonctionnent les SVM.

Enfin, nous discuterons des limitations (en particulier celles liées aux problèmes de dimension des données) et des dérivations issues des SVMs (comment traiter des données en ligne, partiellement étiquetées, etc.).

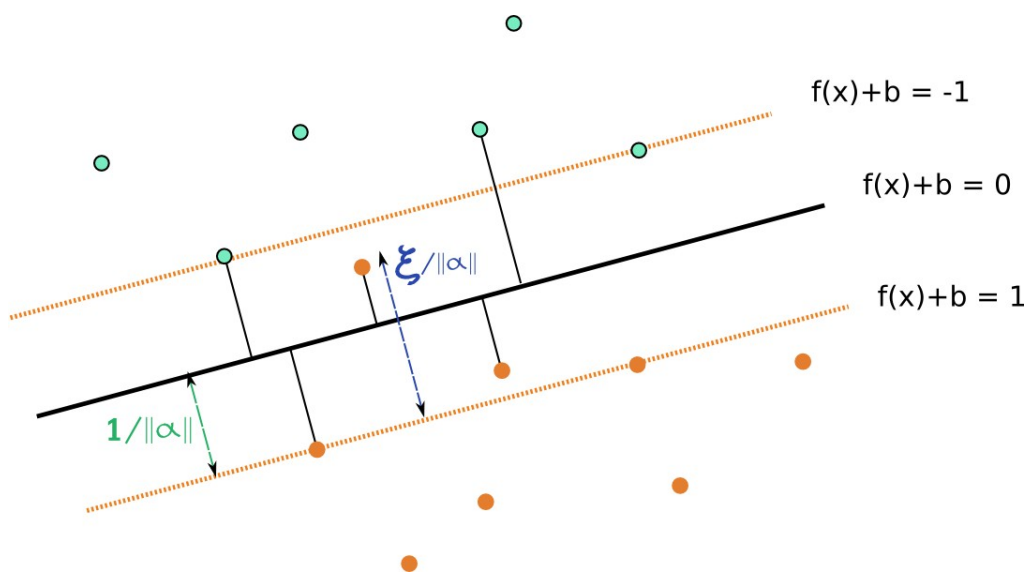


Illustration 1: Cette figure illustre les principales caractéristiques des SVMs. La ligne noire est la frontière de décision. Les lignes orangées en pointillé représentent les marges. Les vecteurs supports sont les points qui se trouvent sur les marges (ou à l'intérieur). Les points situés à l'extérieur de ces marges ne participent pas à la définition de la frontière de décision : c'est la propriété de parcimonie des SVMs.

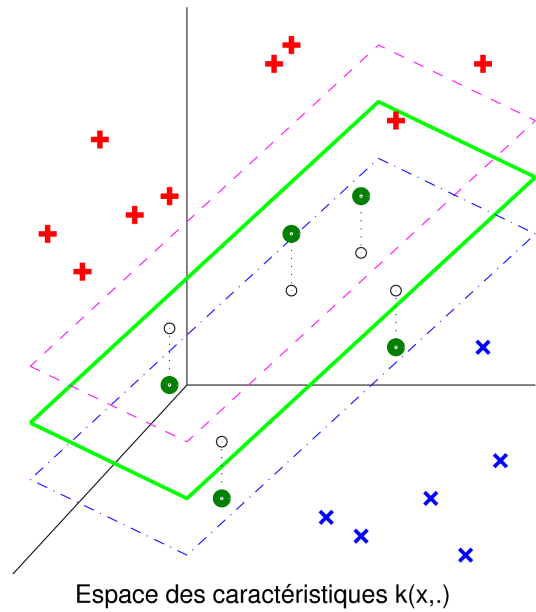
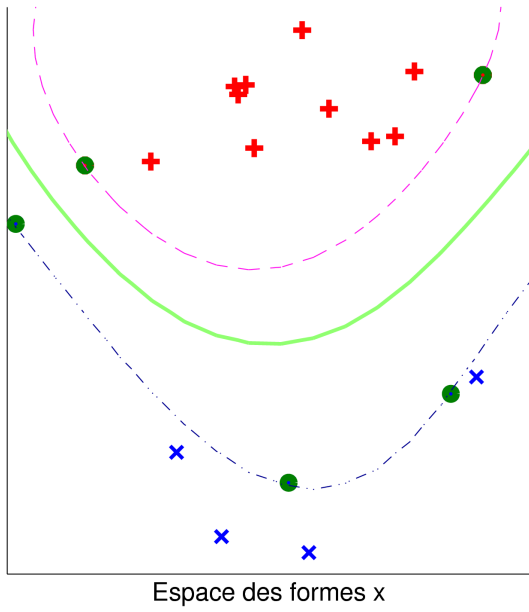


Illustration 2: Effet de l'utilisation des noyaux. A gauche sont représentés les points dans leur espace de départ et la frontière associée. A droite sont représentés les mêmes points dans l'espace induit par l'utilisation du noyau. Dans cet espace, les données sont linéairement séparables. Il est donc possible d'exprimer la solution comme une combinaison linéaire.