FAA partie 1

Matthieu Caron

March 19, 2016

0.1 TP 1 : Calcul de performance

Voici les différents résultats obtenus avec les différentes mesures de performance.

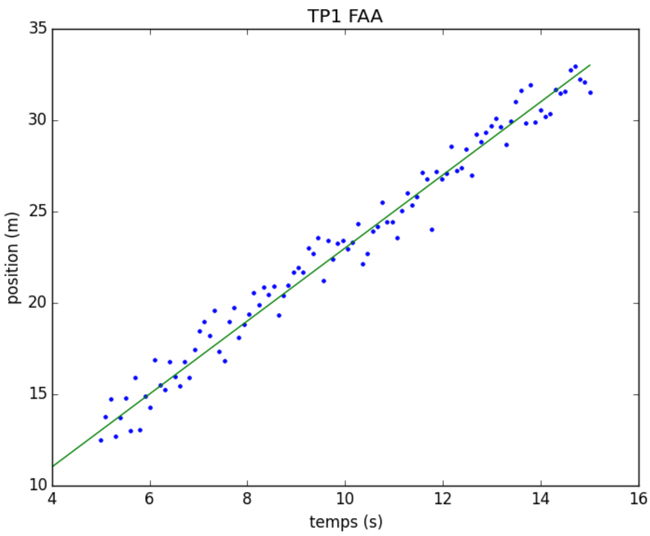
$$J_{abs} = 0.73987984094$$

$$J_{l1} = 0.0896787983772$$

$$J_{l2} = 0.804228687838$$

$$J_{l\infty} = 2.51624302238$$

Figure 1: Comparaison entre $2 * x + 3$ et les points générés

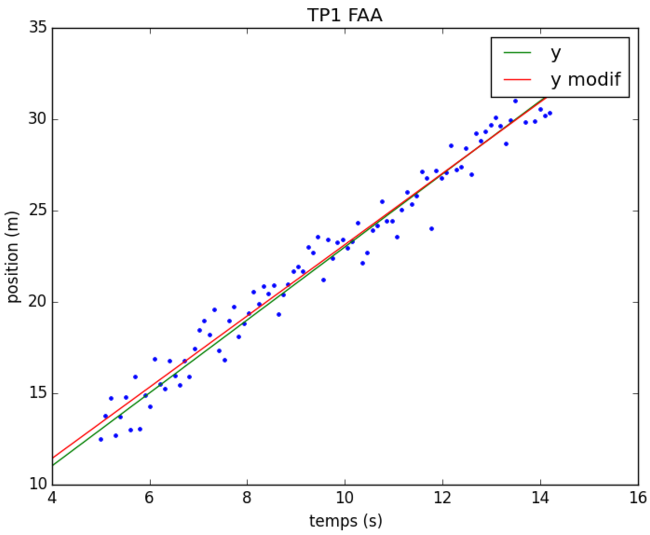


0.2 TP 2 : Moindres Carrés

Les valeurs qui ont permis de générer les points sont 2 et 3 mais il existe un meilleur vecteur teta pour aproximer les points obtenus. Comme la fonction est une fonction linéaire on peut l'approximer avec les moindres carrés. Notre fonction $2 * x + 3$ devient maintenant :

$$1.95293789 * x + 3.59623499$$

Figure 2: Nouvelle approximation



Voici les résultats des mesures de performance après les moindres carrés.

$$J_{abs} = 0.727356264922$$

$$J_{l1} = 0.0877279862436$$

$$J_{l2} = 0.769619957035$$

$$J_{l\infty} = 2.55866627298$$

Et enfin les différences avec les résultats du tp1.

$$diff(J_{abs}) = 0.0125235760181$$

$$diff(J_{l1}) = 0.00195081213361$$

$$diff(J_{l2}) = 0.0346087308024$$

$$diff(J_{l\infty}) = 0.0424232506015$$

0.3 TP 3 : Descente de gradient

J'ai implémenté la descente de gradient globale qui évalue donc tout le jeu de donné avant d'apprendre et j'ai aussi implémenté la descente de gradient stochastique qui évalue une donné au hasard et apprend tout de suite après. Comme on peut l'observer sur les figures, la descente stochastique est bruité. J'ai aussi fait varier le pas d'apprentissage alpha de la forme $\alpha = \frac{1}{i*100+t}$ en fonction de i et voici le résultat.

Figure 3: Descente de gradient globale

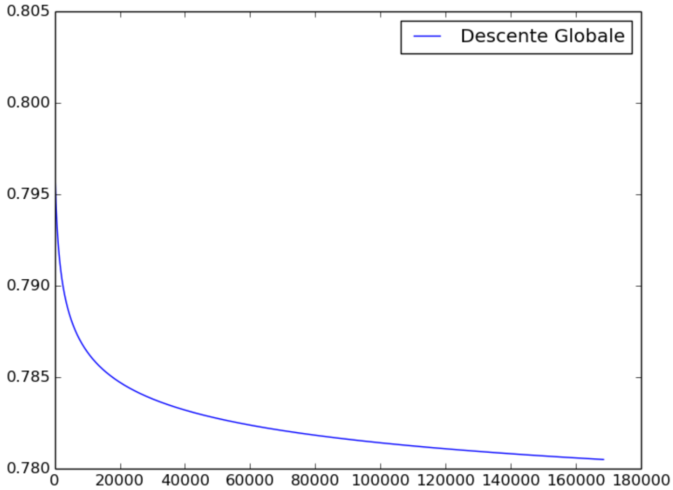


Figure 4: Descente de gradient stochastique

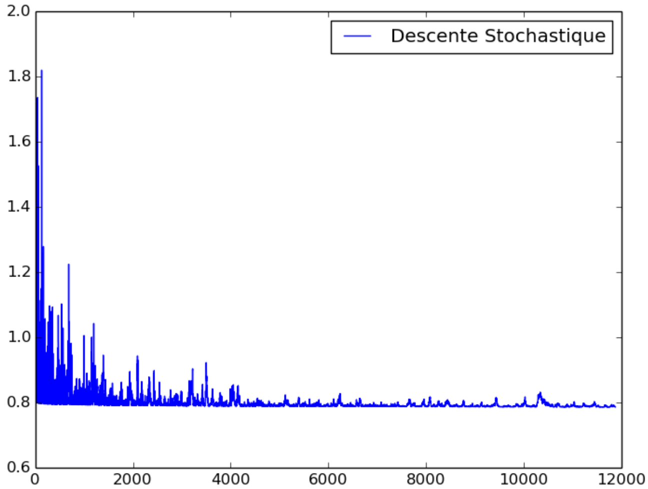
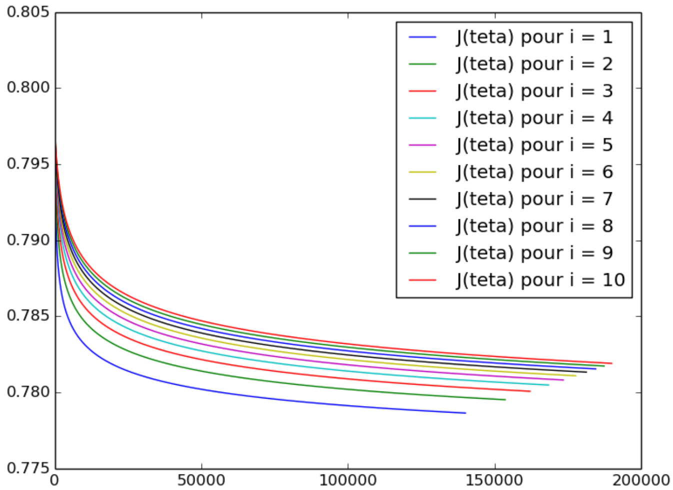


Figure 5: Differents départs pour alpha



0.4 TP4 : Généralisation et sur apprentissage

Avec la méthode des moindres carré il est possible d'approximer des ensembles de points comme on peut l'observer si dessous, cependant en machine learning il existe le phénomène de sur apprentissage c est a dire qu'on perd toute notion de prédiction, qu'une fois sorti du set d'apprentissage on ne déduit pas la bonne valeur. Pour vérifier que l'on a pas sur-appris on fait de la cross-validation, en général on garde 10 pour cent des données pour vérifier si on arrive a prédire les bonnes valeurs après apprentissage.

Figure 6: matrix0

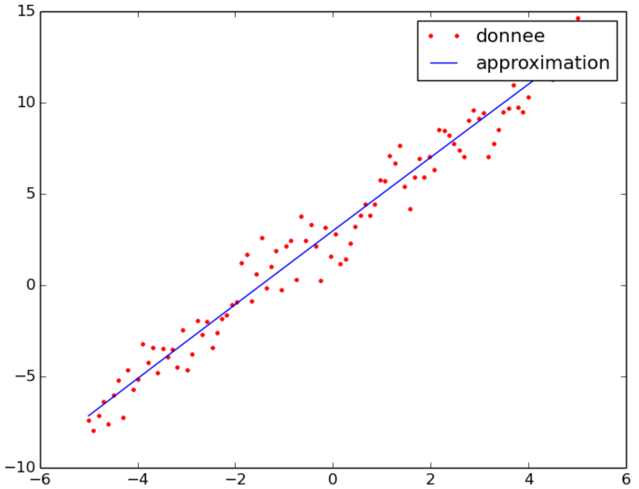


Figure 7: matrix1

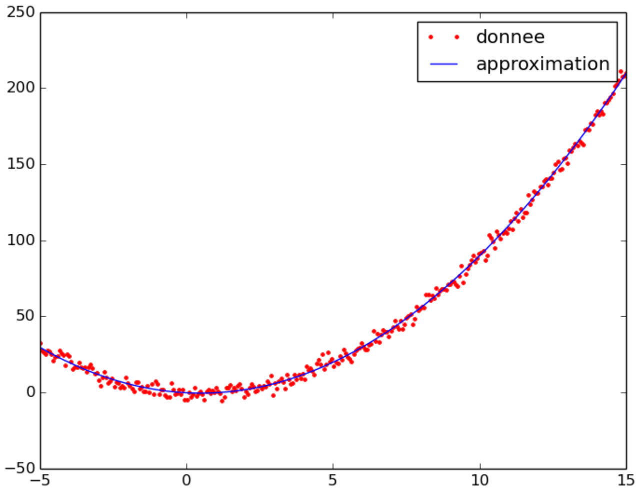


Figure 8: matrix2