

TP 7 – Régression linéaire univariée

Exercice 1 :

Utiliser la matrice de donnée data1. On souhaite prédire la valeur de « Y » à partir de « X » en utilisant une régression linéaire du type « $Y = aX + b$ ». Pour ce jeu de donnée : « $X=0$ » implique « $Y=0$ », on sait donc que « $b=0$ ». Il ne reste que la valeur de « a » à trouver.

1. Représenter les données sur un graphique à 2 dimensions.
2. Utiliser la fonction « `lm()` » pour faire une régression linéaire. Regarder les valeurs obtenues pour « a » et « b ». Représenter cette régression sur le graphique. Calculer la fonction de cout.
3. Calculer la fonction de cout pour « a » variant de -2 à +2. Représenter sur un graphique la fonction de cout en fonction de « a ». Identifier visuellement la valeur optimale de « a ».
4. Faire une descente du gradient pour trouver la valeur optimale de « a ». Calculer et afficher sur un graphique la fonction de cout à chaque itération. Tester différentes valeurs de pas dans la descente du gradient.

Exercice 2 :

Utiliser la matrice de donnée data1. On souhaite prédire la valeur de « Y » à partir de « X » en utilisant une régression linéaire du type « $Y = aX+b$ ». Pour ce jeu de donnée on n'a aucune information sur a et b, il faut donc trouver les valeurs de ces deux paramètres.

1. Représenter les données sur un graphique à 3 dimensions.
2. Utiliser la fonction `lm()` pour faire une régression linéaire. Regarder les valeurs obtenues pour « a » et « b ». Calculer la fonction de cout.
3. Calculer la fonction de cout pour « a » variant de -2 à +2 et « b » de -5 à +5. Représenter sur un graphique la fonction de cout en fonction de « a » et « b ». Identifier visuellement les valeurs optimales de « a » et « b ».
4. Faire une descente du gradient pour trouver les valeurs optimales de « a » et « b ». Calculer et afficher sur un graphique la fonction de cout à chaque itération. Tester différentes valeurs de pas dans la descente du gradient.