

Exercice de programmation 1 : Régression linéaire

Dans cet exercice, vous allez mettre en œuvre la régression linéaire et voir comment elle fonctionne sur des données.

Régression linéaire avec une seule variable

Dans cette partie de l'exercice, vous allez mettre en œuvre la régression linéaire avec une variable pour prédire les bénéfices d'un camion de nourriture. Supposons que vous êtes le CEO d'une franchise de restaurant et que vous envisagez d'ouvrir un nouveau point de vente dans différentes villes. La chaîne possède déjà des camions dans plusieurs villes et vous disposez de données sur les bénéfices et la population de ces villes.

Vous souhaitez utiliser ces données pour vous aider à choisir la ville dans laquelle vous allez vous développer. à la ville suivante.

Le fichier **ex1data1.txt** contient l'ensemble de données pour notre problème de régression linéaire. La première colonne correspond à la population d'une ville et la seconde à le bénéfice d'un food truck dans cette ville. Une valeur négative pour le bénéfice indique une perte.

Avant de commencer toute tâche, il est souvent utile de comprendre les données en les visualisant. Pour cet ensemble de données, vous pouvez utiliser un diagramme de dispersion pour visualiser les données. données, puisqu'il n'y a que deux propriétés à représenter (le bénéfice et la population). (De nombreux problèmes que vous rencontrerez dans la vie courante sont multidimensionnels et ne peuvent pas être représentés sur un graphique à deux dimensions).

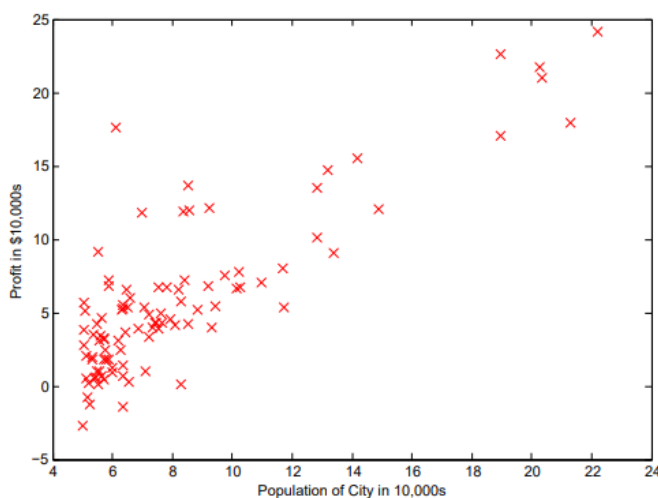


Figure 1: Scatter plot of training data

Dans cette partie, vous allez ajuster les paramètres de régression linéaire θ à notre jeu de données en utilisant la descente de gradient.

L'objectif de la régression linéaire est de minimiser la fonction de coût

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

où l'hypothèse $h_{\theta}(x)$ est donnée par le modèle linéaire

$$h_{\theta}(x) = \theta^T x = \theta_0 + \theta_1 x_1$$

Rappelons que les paramètres de votre modèle sont les valeurs θ_j . Ce sont les valeurs que vous allez ajuster pour minimiser le coût $J(\theta)$. Une façon de le faire est d'utiliser l'algorithme de descente de gradient par lots. Dans l'algorithme de descente de gradient par lots, chaque itération effectue la mise à jour

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \quad (\text{simultaneously update } \theta_j \text{ for all } j).$$

A chaque étape de la descente de gradient, vos paramètres θ_j se rapprochent des valeurs optimales qui permettront d'obtenir le coût le plus bas $J(\theta)$.

Nous initialisons les paramètres initiaux à 0 et le taux d'apprentissage alpha à 0,01.

Lorsque vous effectuez une descente de gradient pour apprendre à minimiser la fonction de coût $J(\theta)$, il est utile de surveiller la convergence en calculant le coût. Dans cette section, vous allez implémenter une fonction pour calculer $J(\theta)$ afin de pouvoir vérifier la convergence de votre implémentation de la descente par gradient.

Vos valeurs finales pour θ seront également utilisées pour faire des prédictions sur les bénéfices dans des zones de 35 000 et 70 000 habitants.

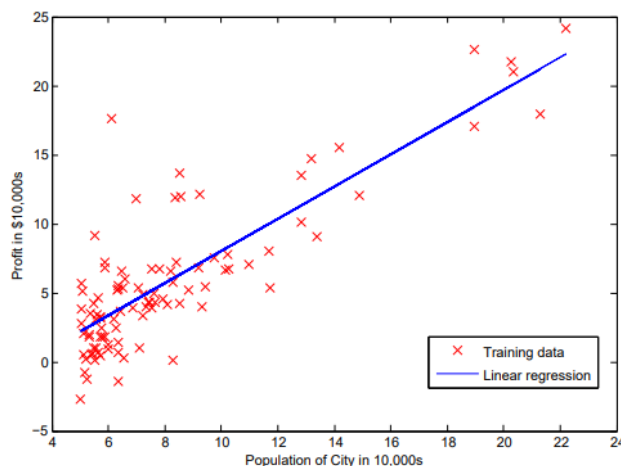


Figure 2: Training data with linear regression fit

Linear regression with multiple variables

Dans cette partie, vous allez mettre en œuvre la régression linéaire avec des variables multiples pour prédire les prix des maisons. Supposons que vous vendiez votre maison et que vous vouliez savoir quel serait le bon prix du marché. Une façon de le faire est d'abord de collecter des informations sur les maisons vendues récemment et de faire un modèle des prix des maisons.

Le fichier **ex1data2.txt** contient un ensemble d'apprentissage des prix des logements à Portland, dans l'Oregon. La première colonne est la taille de la maison (en pieds carrés), la deuxième colonne est le nombre de chambres et la troisième colonne est le prix de la maison.

En regardant les valeurs, on constate que la taille des maisons est environ 1000 fois supérieure au nombre de chambres. Lorsque les caractéristiques diffèrent d'un ordre de grandeur, le fait d'effectuer d'abord une mise à l'échelle des caractéristiques peut permettre à la descente de gradient de converger beaucoup plus rapidement.

Votre tâche est de

- Soustraire la valeur moyenne de chaque caractéristique de l'ensemble de données.
- Après avoir soustrait la moyenne, mettez à l'échelle (divisez) les valeurs des caractéristiques par leurs "écarts types" respectifs.

Précédemment, vous avez implémenté la descente de gradient sur un problème de régression univariée. La seule différence maintenant est qu'il y a une caractéristique de plus dans la matrice X . La fonction d'hypothèse et la règle de mise à jour de la descente de gradient par lots restent inchangées.