Линейна регресия с множество променливи

Доц. д-р Ивайло Пенев

Кат. „Компютърни науки и технологии“

Понятия и означения

- Линейна регресия с множество променливи – многовариантна линейна регресия (multivariate linear regression)
- $x_j^{(i)}$ - стойност на променливата j в ред i от обучителните данни
- $x^{(i)}$ - вектор колона с всички променливи в ред i от обучителните данни
- m – брой обучителни данни (training examples)
- $n = \left|x^{(i)}\right|$ - брой променливи

Функция на хипотезата (Hypothesis function)

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \cdots + \theta_n x_n$$

- Пример:
  - θ_0 - базова цена на къща
  - θ_1 - цена на m^2
  - θ_2 - цена на етаж
  - …
  - x_1 - площ на къщата (m^2)
  - x_2 - брой етажи
  - …

Векторно представяне на функцията

- Векторизация на функцията на хипотезата за един обучителен пример (training example)

$$h_\theta(x) = \begin{bmatrix}\theta_0 & \theta_1 & \dots & \theta_n\end{bmatrix}\begin{bmatrix}x_0 \\ x_1 \\ \dots \\ x_n\end{bmatrix} = \theta^T x$$

- За удобство приемаме, че $x_0^{(i)} = 1$ за $(i\epsilon 1, \dots, m)$ – така векторите θ и x имат еднаква размерност

Матрично представяне на обучителните данни

- $X = \begin{bmatrix} x_0^{(1)} & x_1^{(1)} \\ x_0^{(2)} & x_1^{(2)} \\ x_0^{(3)} & x_1^{(3)} \end{bmatrix}$ - матрица с обучителни данни
- $\theta = \begin{bmatrix} \theta_0 \\ \theta_0 \end{bmatrix}$ - вектор с параметри (коефициенти) на функцията
- Функцията на хипотезата се изчислява като вектор mx1:

$$h_\theta(x) = X\theta$$

Функция на цената (Cost function)

- За вектор с параметри θ ф-та на цената е:

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

- Векторна форма на ф-та:

$$J(\theta) = \frac{1}{2m}(\mathrm{X}\theta - \vec{y})^T(\mathrm{X}\theta - \vec{y})$$

- $\vec{y}$ - вектор с всички стойности y

Градиентно спускане за множество променливи (Gradient Descent)

- За линейна регресия с множество променливи градиентното спускане има следният вид:

Повтаряй до достигане на минимум на ф-та на цената $J(\theta)$

{

$$\theta_0 \coloneqq \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right) x_0^{(i)}$$

$$\theta_1 \coloneqq \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right) x_1^{(i)}$$

$$\theta_2 \coloneqq \theta_2 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right) x_2^{(i)}$$

…

}

Обобщен вид на градиентното спускане

Повтаряй до достигане на минимум на ф-та на цената $J(\theta)$

{

$\theta_j \coloneqq \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)} \right) x_j^{(i)}$ за $j = 0,1,2, \dots, n$

}

Векторна форма на градиентното спускане

- $\theta = \theta - \alpha \nabla J(\theta)$, където

$$\nabla J(\theta) = \begin{bmatrix} \frac{\partial J(\theta)}{\partial \theta_0} \\ \frac{\partial J(\theta)}{\partial \theta_1} \\ \ldots \\ \frac{\partial J(\theta)}{\partial \theta_n} \end{bmatrix}$$ - вектор колона

Векторно представяне на градиента $\nabla J(\theta)$

- Компонент j на вектора $\nabla J(\theta)$ има следният вид:

$$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)x_j^{(i)} = \frac{1}{m}\sum_{i=1}^{m} x_j^{(i)}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)$$

- $x_j^{(i)} = \overrightarrow{x_j}$ - вектор с m елемента за променливата x_j
- $\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)$ - вектор с отклонения на предсказаната стойност $h_\theta\left(x^{(i)}\right)$ и коректната стойност $y^{(i)}$
- $\vec{y}$ - вектор с коректни стойности
- X - вектор с обучителни данни (training set)

$$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{1}{m}\overrightarrow{x_j}^T(X\theta - \vec{y})$$

$$\nabla J(\theta) = \frac{1}{m}X^T(X\theta - \vec{y})$$

Векторна форма на градиентното спускане

$$\theta \coloneqq \theta - \frac{\alpha}{m} X^T (X\theta - \vec{y})$$