

1. θ 는 편향 θ_0 와 특성 별 가중치 $\theta_1, \dots, \theta_n$ 으로 구성된 파라미터 벡터이고, x 는 $x_0(=1), x_1, \dots, x_n$ 으로 구성된 샘플의 특성 벡터라고 할 때, 다중선형회귀모형 가설은 $h_{\theta}(x) = \theta^T x$ 로 표현된다. $(x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)}, y^{(i)}) = (x^{(i)}, y^{(i)}), i = 1, \dots, m$ 의 훈련자료가 주어졌을 때, 이를 아래와 같은 행렬표현식으로 정의하자.

$$X = \begin{pmatrix} 1 & x_1^{(1)} & \dots & x_n^{(1)} \\ 1 & x_1^{(2)} & \dots & x_n^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^{(m)} & \dots & x_n^{(m)} \end{pmatrix}, y = \begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{pmatrix}$$

주어진 훈련자료로 최소제곱법을 이용하여 $h_{\theta}(x)$ 에서의 파라미터 벡터 θ 추정하는 행렬 표현식이 $\hat{\theta} = (X^T X)^{-1} X^T y$ 가 됨을 보여라. (단, 최소제곱법은 $Q(\theta) = (y - X\theta)^T (y - X\theta)$ 가 최소가 되도록 하는 θ 를 찾는 문제임)

2. 다음의 코드를 실행하여, 다중선형회귀모형 훈련과 평가를 위한 데이터셋 Xtrain, ytrain, Xtest, ytest를 생성하고 물음에 답하여라.

```
1 import numpy as np
2 np.random.seed(123)
3 Xtrain = 2 * np.random.rand(100, 3)
4 ytrain = 6 + Xtrain @ np.array([[3], [2], [5]]) + np.random.randn(100, 1)
5 Xtest = 2 * np.random.rand(20, 3)
6 ytest = 6 + Xtest @ np.array([[3], [2], [5]]) + np.random.randn(20, 1)
```

- (1) 훈련자료(Xtrain, ytrain)를 입력하고 반복횟수 iters를 입력하면, 배치 경사하강법(학습율 = 0.01을 이용할 것)을 적용하여 다중선형회귀모형을 훈련한 뒤, 편향과 가중치 파라미터 추정치를 아래와 같은 형태로 출력해주는 함수 gradient_descent_steps를 만들어라. 이 함수를 이용하여 주어진 훈련자료를 적합한 뒤 파라미터 추정치를 아래와 같이 w_pred로 저장하여라.

```
1 w_pred = gradient_descent_steps( Xtrain, ytrain, iters=5000 )
2 print( w_pred )
```

[[5.78346749]
[2.98606662]
[2.27738749]
[4.88477677]]

- (2) 평가자료(Xtest, ytest)를 이용하여, 위 (1)에서 계산된 파라미터 추정치 w_pred로 정의된 다중선형회귀모델의 성능을 MSE로 평가하여라.