HW3 — part 1

Cost funtion $J(w) = \frac{1}{2}\sum_{j=1}^{m}(w^T x^{(j)} - y^{(j)})^2$

$$= \frac{1}{2}(Xw - y)^T(Xw - y)$$

$$\therefore \nabla_w J(w) = \frac{\partial J(w)}{\partial w} = X^T(Xw - y)$$

$$\therefore H(J)(w) = \frac{\partial(\nabla_w J(w))}{\partial w} = X^T X$$

Newton's method:

$$w^* = w^{(0)} - H(J)(w^{(0)})^{-1} \nabla_w J(w^{(0)})$$

$$= w^{(0)} - (X^T X)^{-1} \cdot X^T(Xw^{(0)} - y)$$

$$= w^{(0)} - X^{-1}(X^T)^{-1}X^T X w^{(0)} + (X^T X)^{-1} X^T y$$

$$= w^{(0)} - X^{-1} X w^{(0)} + (X^T X)^{-1} X^T y$$

$$= w^{(0)} - w^{(0)} + (X^T X)^{-1} X^T y$$

$$= (X^T X)^{-1} X^T y$$

```
Training cost: [455.0, 390.0, 337.0, 295.0, 260.0, 232.0, 209.0, 191.0, 176.0, 164.0, 154.0, 146.0, 140.0, 135.0, 131.0, 127.0,
 124.0, 122.0, 120.0, 119.0, 118.0, 117.0, 116.0, 115.0, 115.0, 114.0, 114.0, 114.0, 114.0, 113.0, 113.0, 113.0, 113.0, 113.0, 1
13.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 113.0, 11
3.0, 113.0, 113.0]
```

```
Training cost: [0.22159473296466886, 0.22159852273429323, 0.22159344186891242, 0.22159722931307879, 0.2215921511297261З, 0.22159
593624956397, 0.221590860746982, 0.22159464354362018, 0.22158957072055202, 0.221593351195119, 0.22158828105030887, 0.22159205920
39323, 0.22158699173612439, 0.2215907675699316, 0.22158570277787129, 0.22158947629298897, 0.22158441417542191, 0.221588185372976
12, 0.22158312592864848, 0.22158689480976512]
```