

1월 . ISL 3.5

5. Consider the fitted values that result from performing linear regression without an intercept. In this setting, the i th fitted value takes the form

$$\hat{y}_i = x_i \hat{\beta},$$

where

$$\hat{\beta} = \left(\sum_{i=1}^n x_i y_i \right) / \left(\sum_{i'=1}^n x_{i'}^2 \right). \quad (3.38)$$

Show that we can write

$$\hat{y}_i = \sum_{i'=1}^n a_{i'} y_{i'}.$$

What is $a_{i'}$?

$$\begin{aligned} \hat{y}_i &= x_i \hat{\beta} \\ &= x_i \frac{\sum_{i'=1}^n x_{i'} y_{i'}}{\sum_{i'=1}^n x_{i'}^2} \\ &= x_i \cdot \frac{\sum_{i'=1}^n x_{i'} y_{i'}}{\sum_{j=1}^n x_j^2} \\ &= \sum_{i'=1}^n \frac{x_i \cdot x_{i'} \cdot y_{i'}}{\sum_{j=1}^n x_j^2} \quad \leftarrow a_{i'} = \frac{x_i \cdot x_{i'}}{\sum_{j=1}^n x_j^2} \\ &= \sum_{i'=1}^n a_{i'} y_{i'} \end{aligned}$$

ESL 3.4

Ex. 3.4 Show how the vector of least squares coefficients can be obtained from a single pass of the Gram-Schmidt procedure (Algorithm 3.1). Represent your solution in terms of the QR decomposition of \mathbf{X} .

$$\begin{aligned} \mathbf{X} &= \mathbf{Q}\mathbf{R} \quad \text{plug in!} \\ \hat{\beta} &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \\ &= ((\mathbf{Q}\mathbf{R})^T \mathbf{Q}\mathbf{R})^{-1} (\mathbf{Q}\mathbf{R})^T \mathbf{y} \end{aligned}$$

$$\begin{aligned} &\star \underline{\mathbf{Q}^T \mathbf{Q} = \mathbf{I}} \\ &(\because \mathbf{Q} \text{ is orthonormal basis vectors}) \\ &\quad \text{or (주어진 행렬이므로)} \end{aligned}$$

$$= (R^T \underline{Q^T Q} R)^{-1} R^T Q^T y$$

$$= R^{-1} (R^T)^{-1} R^T Q^T y$$

$$= R^{-1} Q^T y \quad \square$$

```
In [1]: # Data Import
import ssl
import pandas as pd

ssl._create_default_https_context = ssl._create_unverified_context #Github에서 데이터를 바로 불러오도록 하는 세팅입니다. 해당 코드 무시하고 데이터 받아서 쓰셔도 됩니다!

data = pd.read_csv('https://github.com/YonseiESC/ESC-21SUMMER/blob/main/week1/HW/week1_data.csv?raw=True')
y = data['mpg']
x = data.drop(['mpg'],axis=1)
```

```
In [2]: idx = x['horsepower']!= '?'
x = x[idx]
y = y[idx]
```

```
In [3]: x = x.astype({'horsepower': 'float'})
```

```
In [4]: import numpy as np
# numpy 모듈만을 이용해주세요.
def YourOwnRegression(x,y):
    x = np.array(x)
    y = np.array(y)
    XTX = np.dot(np.transpose(x), x)
    betahat = np.dot(np.dot(np.linalg.inv(XTX), np.transpose(x)), y)
    yhat = np.dot(x, betahat)
    return betahat, yhat
# 결과물 반환
bhat, yhat = YourOwnRegression(x,y)
print("beta hat: ", bhat,'\n', "yhat: ", yhat)
```

```
beta hat: [-0.5226089    0.01022108 -0.020873   -0.00639456 -0.05202195  0.61025869]
yhat: [15.93081361 14.45720405 16.11263762 15.93670419 16.10071197 10.51021782
 10.27543153 10.53128391  9.67525181 13.49634208 15.59946068 15.17857818
 14.95056695 18.23756934 23.85149163 20.70116218 21.04691637 22.47738545
 25.4075601  27.85849004 21.93938914 23.54965655 23.61026333 24.5700506
 22.02475307  7.48992449  8.73758029  8.68094775  6.39448743 26.01781879
 25.50668622 25.43458909 22.95714525 17.50355598 18.56684981 18.98997897
 18.64504732 11.74185816 10.43958016 12.2762232 12.39844199  7.02172716
  8.71466792  6.09084578 20.89073634 24.7795066 18.89340516 20.0843144
 25.76559234 26.24104628 26.30753902 26.59195154 28.28090927 29.28279009
 28.2609542  27.13912286 25.6470868 26.69569542 26.07663641 24.9880403
 26.2074288 11.93647037 11.52899822 12.73330182 13.0723569 15.65493241
  9.60277136 10.60920749 10.79899231 10.95329347 25.45996983 14.19610698
 13.24891771 11.63170439 13.0781754 21.23759877 24.50545484 21.19494316
 26.4550875 25.15275892 25.40418994 24.27095036 26.56185652 26.71587043
 13.39863077 16.2651911 14.74101165 13.99404444 15.68483539  8.35588404
 12.1559116 12.08192325 12.63929331  9.52710571  8.09046672 15.38896808
 20.7055219 19.98343495 22.03288824 21.95534612 22.05191343 28.9278806
  8.64356655  8.94866353 10.06261966 10.76884679 23.08246175 26.05144779
 26.01877494 25.52924044 27.53361915 26.19737345 24.2284252 26.29152354
 14.1399229 11.80551478 29.17763168 27.47281611 24.48514487 22.211482
 18.18001246 23.66132966 21.80938422 16.18633836 21.37003802 22.92450763
 20.27483103 29.01987066 26.1144079 29.60478409 25.79483636 17.42925127
 18.20305761 18.16965645 13.93463354 10.62233937 11.89477047 10.6524927
 12.92654367 27.29002111 29.11881818 26.97024946 31.35712023 28.84708125
 28.08036911 28.1370547 27.59838633 25.57416431 26.13565935 28.85620844
 21.24635801 20.04621563 20.64756809 22.47771079 11.68048944 13.0191654
 12.19083021 11.60515866 16.62277687 17.11534294 18.13462735 17.75872575
 22.48538772 20.67124626 21.09405268 28.38999264 25.6178386 23.45163703
 25.94471506 25.06438254 28.04340891 25.7120626 22.52118358 30.02318129
 21.7091604 24.86417853 23.27558145 23.28386212 24.73315284 31.11419428
 27.02526368 28.70906066 26.57792197 28.30139996 28.71233562 14.76463858
 14.84864669 16.74288177 14.99473679 21.98169626 21.41233074 23.42150596
 23.09392064 29.90480163 29.05099549 30.6230356 31.72965517 19.18911918
 20.22456108 19.35231654 22.55428679 31.49394493 30.08735698 29.08175554
 27.10777245 22.52930717 16.43688261 21.62943493 23.04180738 17.16039851
 13.38138443 16.19537407 17.04478509 17.57825889 30.44250197 29.8139795
 31.85863268 28.62881666 30.99577532 17.43872127 16.22016388 15.84658308
 14.94255744 20.751491 21.24312777 19.93366366 20.83473714 15.57784903
 15.61161917 14.64046595 14.81182982 31.10306759 26.25259415 28.89497299
 26.0337752 30.58653253 30.07282718 30.94607365 29.52523606 24.56685474
 26.54846444 25.84824622 31.61605985 32.87446415 31.51916094 30.71307435
 32.82411068 21.45671946 19.17701507 20.11026212 21.03033425 23.4041659
 24.95271522 26.90952532 21.84884557 23.53592664 22.19836586 24.23446653
 20.34711601 22.00581861 20.92208997 20.65519199 22.5275121 16.94488046
 30.45340282 27.78764395 29.22915666 30.00329119 28.04418604 26.49072908
 26.10686219 28.54732333 25.25241894 22.73501516 25.65655504 20.72705133
 31.43713179 30.57609185 23.48396909 25.27825721 26.3338669 23.74246252
 22.79209671 19.07725247 19.90688066 18.75951203 19.18011155 15.71621531
 17.98573432 20.31674699 18.70734857 32.50985671 32.22227707 32.45712629
 27.83310428 22.24218693 19.15384865 24.39039603 21.78419026 30.97738025
 31.20867101 31.71621295 31.22474446 27.87898482 27.26090106 26.50623701
 28.75362625 31.67091587 32.82504694 31.90800579 32.48874873 28.41207605
 27.03039926 26.14656464 23.51339355 31.28036395 28.07931906 29.22595587
 29.68156125 30.99424682 31.82881074 27.47919606 31.75096112 32.18671672
 30.48403252 26.14671839 24.9076443 33.71404606 31.67048481 33.75190878
 25.44572433 29.75593179 29.36847327 30.7931412 30.22783049 29.42774108
 29.5118248 27.68686195 31.04492087 34.83572405 34.04069724 34.82406496
 32.80271279 33.2140812 32.94723852 33.3257177 32.15965292 33.06580622
 30.68952944 32.1263293 31.98760285 30.97773269 28.97782474 29.22077581
 25.39617051 24.96456589 26.39147031 25.8288513 23.70096728 21.82763916
 26.04576044 23.80012019 29.58125213 29.40946443 31.00734511 30.01258259
 30.59867295 29.19007052 28.28783476 34.02222172 33.56574566 33.94865969
 33.35582651 33.00291086 32.77391248 32.62463215 32.35858312 34.13676256
 34.1003472 33.88249939 26.77440244 27.64530275 30.34069094 28.04577073
 29.6541942 31.74589579 27.85165521 28.93428241 32.95667079 32.29831861
 29.77500853 29.05306756]
```

```
In [ ]:
```