Properties of the hat matrix H (Exercise A)

If X is full rank, i.e. if $rk(X) = p$, then the following decomposition holds true

$$X = QR,$$

with $Q \in \mathbb{R}^{m \times p}$ with $rk(Q) = p$ and s.t. $Q^T Q = I_p$.

H is symmetric

Indeed, we have that

$$H^T = \left[X(X^TX)^{-1}X^T\right]^T = (X^T)^T \underbrace{\left[(X^TX)^{-1}\right]^T}_{\text{symmetric}} X^T = X(X^TX)^{-1}X^T = H$$

$$\Longrightarrow \boxed{H = H^T}$$

The trace of H is p

$tr(H) = tr(QQ^T) = tr(Q^TQ) = tr(I_p) = p.$ (Proof based on QR)

as shown in the slides

$tr(H) = tr(X(X^TX)^{-1}X^T) = tr(X^TX \cdot (X^TX)^{-1}) = tr(I_p) = p.$ (Alternative proof)

H is idempotent

$H^2 = H \cdot H = Q \underbrace{Q^T Q}_{= I_p} Q^T = QQ^T = H \Longrightarrow \boxed{H = H^2}$ (Proof based on QR)

$H^2 = H \cdot H = X(X^TX)^{-1}\underbrace{X^TX(X^TX)^{-1}}_{= I_p}X^T = X(X^TX)^{-1}X^T = H$ (Alternative proof)

The rank of H is p. (Theory on eigenvalue is required)

Suppose λ is an eigenvalue of the symmetric matrix $H \in \mathbb{R}^{m \times m}$. Then, it means there is a vector $x \in \mathbb{R}^m$ such that $x \neq 0_m$ and

$$Hx = \lambda x.$$

We proved that $H^2 = H$, therefore

$$\lambda x = Hx = H \cdot \underbrace{Hx}_{\lambda x} = H \lambda x = \lambda \underbrace{Hx}_{\lambda x} = \lambda^2 x.$$

In particular, we have shown that $\lambda x = \lambda^2 x$. This equality is true iff

$$\lambda x = \lambda^2 x \iff \lambda^2 x - \lambda x = 0 \iff \lambda(\lambda - 1) x = 0.$$

From here, we conclude that the eigenvalues are necessarily $\lambda \in \{0, 1\}$, i.e. either equal to 0 or to 1. Incidentally, this shows that H is positive semi-definite.

Now, recall that $rk(H) =$ "# number of strictly positive eigenvalues" and that $tr(H) = \sum_{i=1}^{m} \lambda_i = p$.

eigenvalues of H

Hence, there are p eigenvalues equal to 1 and $m-p$ eigenvalues equal to 0. In particular, $rk(H) = p$.

The sums of the rows (and the columns) of H is 1, if there is the intercept

The matrix H is <u>projecting</u> the values of y into those spanned by X. Hence, if the intercept is included in the model, we have

$$H 1_m = 1_m.$$