$E(\hat{\varepsilon}^T\hat{\varepsilon}) = (n-p)\sigma^2$

$y = X\beta + \varepsilon$, $E(\varepsilon) = 0$ $Var(\varepsilon) = \sigma^2 I$

$HX = X(X^TX)^{-1}X^TX = X$

$\therefore (I-H)X = 0$

$\hat{y} = X\hat{\beta} = X(X^TX)^{-1}X^Ty = Hy$

$\hat{\varepsilon} = y - \hat{y} = y - Hy = (I-H)y = My = M\varepsilon$

$\hat{\varepsilon}^T\hat{\varepsilon} = \varepsilon^T M^T M \varepsilon = \varepsilon^T M \varepsilon$

- $H^T = X(X^TX)^{-1}X^T = H$

$HH = X(X^TX)^{-1}X^TX(X^TX)^{-1}X^T = X(X^TX)^{-1}X^T = H$

$M^T = M$

$MM = M$

$E(\hat{\varepsilon}^T\hat{\varepsilon}) = E(\varepsilon^T M \varepsilon) = E[tr(\varepsilon^T M \varepsilon)]$

$\underset{①}{=} E[tr(M\varepsilon\varepsilon^T)]$

$= tr[E(M\varepsilon\varepsilon^T)]$

$= tr[M E(\varepsilon\varepsilon^T)]$

$\underset{②}{=} tr[\sigma^2 M] \underset{③}{=} \sigma^2 tr(M)$

$\underset{④}{=} \sigma^2 \cdot (n-p)$

① $tr(AB) = tr(BA)$

② $E(\varepsilon\varepsilon^T) = \sigma^2 I$

③ $tr(aA) = a\,tr(A)$

a: scalar

④ $tr(A+B) = tr(A) + tr(B)$

$tr(I_{n\times n}) = n$

$tr(H) = rank(X) = p$