1. $L(\beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta)\right\}.$

Suppose $Y = X\beta + e \quad \vec{e} \sim N(\vec{0}, \sigma^2 I).$

$l = \log L(\beta, \sigma^2).$

$$\frac{\partial l}{\partial \beta} = \frac{1}{2\sigma^2}(2X'Y - 2X'X\beta) = 0$$

$$\frac{\partial l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}(Y - X\beta)'(Y - X\beta) = 0.$$

$\left(\frac{\partial^2 l}{\partial \beta^2}\right)$

$$\begin{cases} \frac{\partial l}{\partial \beta} = 0 \\ \frac{\partial l}{\partial \sigma^2} = 0 \end{cases} \Rightarrow (\hat{\beta}, \hat{\sigma}^2).$$

$$X'Y = X'X\hat{\beta} \Rightarrow \hat{\beta} = (X'X)^{-1}X'Y. = \text{OLS}$$

$(\hat{\sigma}^2_{MLE} \neq \hat{\sigma}^2_{OLS}).$

$\hat{\beta}_{MLE} = \hat{\beta}_{OLS}$ depends on Normal assumption.

$\max_{\beta} \exp\left\{-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta)\right\}.$ pdf of error term.

$\sigma^2 > 0.$

$\arg\min_x f(x) = x$ s.t. $f(x)$ minimized.

$$\arg\max_{\beta} \exp\left\{-\frac{1}{2\sigma^2}(Y - X\beta')(Y - X\beta)\right\} = \arg\min_{\beta} (Y - X\beta)'(Y - X\beta).$$

OLS.

$Y = X\beta + e$. $Ee_i = 0$. $Var(e_i) = \sigma^2$. $Cov(e_i, e_j) = 0$. $i \neq j$.

$Cov(\hat{Y}, \hat{e})$. $Cov(Y, \hat{e})$. $Cov(\hat{Y}, e)$. $Cov(Y, e)$.

$\hat{Y} = X\hat{\beta} = X(X'X)^{-1}X'Y. = HY$. $\hat{e} = Y - \hat{Y}$.

hat matrix H: projection operator

$\hat{Y} = HY \in C(X)$. H properties:

① idempotent : $H^2 = H$.

② symmetric : $H^T = H$.

③ perpendicular projection operator:

1° $v \in C(X) \Rightarrow Hv = v$.

2° $v \perp C(X) \Rightarrow Hv = 0$

$Cov(\hat{Y}, \hat{e}) = Cov(HY, Y - \hat{Y}) = \underline{H\,Cov(Y,Y) - H\,Cov(Y, HY)}$.

$Cov(Y, Y) = Var(Y) = Var(X\beta + e)$

$= Var(e) = \sigma^2 I$.

$= H \cdot \sigma^2 I - H \cdot \sigma^2 I \cdot H^T = \sigma^2(H - H) = 0$.

$H = X(X'X)^{-1}X'$.

$\hat{Y}, \hat{e}$ are uncorrelated.

$Cov(Y, \hat{e}) = Cov(Y, Y - \hat{Y}) = \sigma^2 I - \sigma^2 I \cdot H = \sigma^2(I - H)$.

$Cov(Y, e) = Cov(X\beta + e, e) = Cov(e, e) = \sigma^2 I$.

Assignment

Model	EV's	df	RSS
$y=\beta_0+e \rightarrow$	Null (const.)	46	$\boxed{1101.377}$ = SYY
	X_1	45	✓
	X_2	45	✓
	X_3	45	X
$y=\beta_0+\beta_1X_1+\beta_2X_2+e$	X_1, X_2	44	871.01
	X_2, X_3	44	$\boxed{}$
	$X_1\ X_3$	44	35.637
	X_1, X_2, X_3	43	② 34.874

$n = 44+2+1 = 47$.

✓ $H_0: Y=\beta_0+e$ v.s. $H_1: Y=\beta_0+\beta_2X_2+\beta_3X_3$.

	df	SS	MS	F	p-value
Reg	2				
Residuals	44				
Tot	46	1101.377			

✓ $H_0: Y=\beta_0+\beta_1X_1+\beta_3X_3+e$ v.s. $H_1: Y=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3$.

	df	SS	MS	F	p-val.
Reg	1	0.763			
Residuals	43	$\boxed{34.874}$			
Tot	44	$\boxed{35.637}$			

$SS_{reg} = \boxed{RSS_1} - RSS_2$.

$RSS_1 = RSS_2 + SS_{reg}$.

$$\rho^2 = R^2 = \frac{SS_{reg}}{SYY} = \frac{SYY - RSS}{SYY}$$

$SYY = \Sigma(y_i - \bar{y})^2$.

$\hat{\rho}(y, x_1) = 0.3273. \Rightarrow R^2 = \rho^2 = 0.3273^2 \approx 0.1071$

$$\frac{SYY - RSS}{SYY} = 0.1071.$$

$$\Rightarrow RSS = (1 - 0.1071) \cdot SYY \approx 983. \cdots$$

$\vec{Y} \in \mathbb{R}^n$. $E\vec{Y} = \vec{\mu}$. $Var(\vec{Y}) = \Sigma$.

$f(\vec{Y}) = Y^T A Y$. A is symmetric

$$\begin{aligned}
E(f(Y)) &= E(Y^T A Y) = E((Y-\mu+\mu)^T A (Y-\mu+\mu)). \\
&= E\{(Y-\mu)^T A (Y-\mu) + (Y-\mu)^T A\mu + \mu^T A (Y-\mu) + \mu^T A \mu\} \\
&= E\{(Y-\mu)^T A (Y-\mu)\} + E(Y-\mu)^T A\mu + E\mu^T A (Y-\mu) + E\mu^T A \mu. \\
&= E\{(Y-\mu)^T A (Y-\mu)\} + \underbrace{E(Y-\mu)^T}_{0} \cdot A\mu + \mu^T A \underbrace{E(Y-\mu)}_{0} + \mu^T A \mu. \\
&= E\{\underbrace{(Y-\mu)^T A (Y-\mu)}_{①}\} + \underbrace{\mu^T A \mu}_{②}.
\end{aligned}$$

$$\begin{aligned}
(Y-\mu)^T A (Y-\mu) &= tr((Y-\mu)^T A (Y-\mu)) \qquad tr(AB) = tr(BA). \\
&= tr(A(Y-\mu)(Y-\mu)^T).
\end{aligned}$$

$$\begin{aligned}
E(tr(A(Y-\mu)(Y-\mu)^T)) &= tr(E(A(Y-\mu)(Y-\mu)^T)) \\
&= tr(A\, E((Y-\mu)(Y-\mu)^T)) \\
&= tr(A\Sigma). ①
\end{aligned}$$

$$E(f(\vec{Y})) = E(Y^T A Y) = \mu^T A \mu + tr(A\Sigma).$$