Problem 1

For OLS,

$$\begin{aligned}
\underset{\beta}{argmin}\sum_{i=1}^{n}(y_i - f(x_i))^2 &= \underset{\beta}{argmin}(y - X\beta)^T(y - X\beta) \\
\nabla_\beta(y - X\beta)^T(y - X\beta) &= \nabla_\beta(y^Ty + \beta^TX^TX\beta - \beta^TX^Ty - y^TX\beta) \\
&= 2(X^TX\beta - X^Ty)
\end{aligned}$$

Let $\nabla_\beta(y - X\beta)^T(y - X\beta) = 0$, we get $\hat{\beta} = (X^TX)^{-1}X^Ty$.

Problem 2

Since $y = X\beta + \epsilon$,

$$\begin{aligned}
Cov(\hat{\beta}) &= Cov[(X^TX)^{-1}X^Ty] \\
&= Cov[\beta + (X^TX)^{-1}X^T\epsilon] \\
&= (X^TX)^{-1}X^TCov(\epsilon)X(X^TX)^{-1} \\
&= (X^TX)^{-1}\sigma^2
\end{aligned}$$

Problem 3

$$SS_{tot} = \sum_{i=1}^{n}(y_i - \bar{y})^2 \quad = y^Ty - n\bar{y}^2 \tag{1}$$

Since $E(y - \hat{y}) = E(y - X\beta) = E(\epsilon) = 0, \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \hat{y}_i$.

$$\begin{aligned}
SS_{reg} &= \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \\
&= \hat{y}^T\hat{y} + n\bar{y}^2 - 2\bar{y}\sum_{i=1}^{n}\hat{y}_i \\
&= \hat{y}^T\hat{y} - n\bar{y}^2
\end{aligned} \tag{2}$$

Notice that $\hat{y}^T(y - \hat{y}) = \beta^TX^T[y - X(X^TX)^{-1}X^Ty] = 0$.

$$\begin{aligned}
SS_{res} &= \sum_{i=1}^{n}(y_i - \hat{y})^2 \\
&= y^Ty - \hat{y}^T\hat{y} + \hat{y}^T(y - \hat{y}) \\
&= y^Ty - \hat{y}^T\hat{y}
\end{aligned} \tag{3}$$

Form (1), (2) and (3), we get $SS_{tot} = SS_{reg} + SSres$.

Problem 4

(1). Assume that ϵ and $f(x^*) - \hat{f}(x^*)$ are independent.

$$\begin{aligned}
E[Y - \hat{f}(x^*)]^2 &= E[f(x^*) - \hat{f}(x^*) + \epsilon]^2 \\
&= E[f(x^*) - \hat{f}(x^*)]^2 + E(\epsilon^2) + 2E(\epsilon)E[f(x^*) - \hat{f}(x^*)] \\
&= E[f(x^*) - \hat{f}(x^*)]^2 + E(\epsilon^2)
\end{aligned}$$

Since $E(\epsilon^2)$ is constant, minimizing $E[Y - \hat{f}(x^*)]^2$ is equivalent to minimizing $E[f(x^*) - \hat{f}(x^*)]^2$. We take the expectation with regard to r.m. x^*.

(2).

$$\begin{aligned}
E[f(x^*) - \hat{f}(x^*)]^2 &= E[f(x^*) - E(\hat{f}(x^*)) + E(\hat{f}(x^*)) - \hat{f}(x^*)]^2 \\
&= (f(x^*) - E\hat{f}(x^*))^2 + E[E(\hat{f}(x^*)) - \hat{f}(x^*)]^2 \\
&\quad + 2[f(x^*) - E(\hat{f}(x^*))]E[E(\hat{f}(x^*)) - \hat{f}(x^*)] \\
&= (f(x^*) - E\hat{f}(x^*))^2 + Var[\hat{f}(x^*)]^2
\end{aligned}$$