57. for multiple linear regression,

$$\left[\hat{Y}_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} \cdots \beta_p X_{p,i} + \varepsilon_i\right]$$

now, here ε_i has gaussian distribution with 0 mean and σ^2, unknown parameter, so joint density for ε_i is

$$L(\beta, \sigma^2 | Y, X) = \frac{1}{(\sqrt{2\pi}\,\sigma)^n} \exp\left(-\sum \frac{\varepsilon_i^2}{2\sigma^2}\right)$$

Also, $\left[\varepsilon_i = Y_i - (\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} \cdots \beta_p X_{p,i})\right]$

$$\ln(L(\beta, \sigma^2 | Y, X)) = -\frac{n}{2}\ln 2\pi - n\ln\sigma - \frac{1}{2\sigma^2}\sum \varepsilon_i^2$$

So, maximizing the $\ln(L(\beta, \sigma^2 | Y, X))$ is same as minimizing

$$SS(\beta) = \sum \varepsilon_i^2$$

$$\left[SS(\beta) = \sum \left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{j,i} - Y_i\right)^2\right]$$

hence, MLE for the matrix of the coefficients is same as that obtained via solving the normal equation.