③ a) $\min_{\beta} : \|X\beta - y\|_2^2 = \sum (x_i\beta - y_i)^2$

$$\|X\beta - y\|_2^2 = (X\beta - y)^T (X\beta - y) = (\beta^T X^T - y^T)(X\beta - y)$$
$$= \beta^T X^T X\beta - 2\beta^T X^T y + y^T y$$

$$\nabla(\|X\beta - y\|_2^2) = 2X^T X\beta - 2X^T y = 0$$
$$\Rightarrow X^T X\beta = X^T y$$

$$\Rightarrow \boxed{\hat{\beta}_{LS} = (X^T X)^{-1} X^T y}$$

b) $\min_{\beta} : \|X\beta - y\|_2^2 + \lambda\|\beta\|_2^2 = \sum_i (x_i\beta - y_i)^2 + \lambda \sum_i \beta_i^2$

$$= \beta^T X^T X\beta - 2\beta^T X^T y + y^T y + \lambda\beta^T I\beta$$
$$\beta^T (X^T X + \lambda I)\beta - 2\beta^T X^T y + y^T y$$

$$\nabla(\|X\beta - y\|_2^2 + \lambda\|\beta\|_2^2) = 2(X^T X + \lambda I)\beta - 2X^T y = 0$$
$$\Rightarrow (X^T X + \lambda I)\beta = X^T y$$

$$\Rightarrow \boxed{\hat{\beta}_{RR} = (X^T X + \lambda I)^{-1} X^T y}$$