1. a) As $\sum_{i=1}^{n} y_i = n\bar{y}$, $\sum_{i=1}^{n} x_i = n\bar{x}$

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n}(y_i-\hat{y}_i) = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\left(\bar{y} + r\sigma_y \frac{x_i-\bar{x}}{\sigma x}\right) = n\bar{y} - n\cdot\bar{y} - r\frac{\sigma y}{\sigma x}\, n\bar{x} - n\bar{x} = 0$$

b) As $\sum_{i=1}^{n}(y_i-\hat{y}) = 0$ $\quad \sum_{i=1}^{n} y_i = \sum_{i=1}^{n}\hat{y}$ $\quad \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y} = \frac{1}{n}\sum_{i=1}^{n}\hat{y} = \bar{\hat{y}}$

c) As $\hat{y} = \bar{y} + \frac{r\sigma x}{\sigma y}(x-\bar{x})$

when $x = \bar{x}$, $\hat{y} = \bar{y} + r\frac{\sigma y}{\sigma x}(\bar{x}-\bar{x}) = \bar{y}$, $(\bar{x},\bar{y})$ is on the regression line

2. a) For the OLS estimator $\hat{\theta}_0, \hat{\theta}_1$ minimize the sum of residuals,

$$\sum e_i^2 = [e_1\, e_2 \cdots e_n]\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = e^T e = (Y-\hat{Y})^T(Y-\hat{Y})$$

This is like a quardratic funtion.

$\hat{Y} = \hat{\theta}_0 1 + \hat{\theta}_1 \vec{x}$. to find the min. take derivative with respect to $\hat{\theta}_1$

$$\frac{d}{d\hat{\theta}_1}(Y-\hat{Y})^T(Y-\hat{Y}) = -2\vec{x}(Y-\hat{Y})$$

then set it to zero and solve for $\hat{\theta}_1$

So that we have $Y-\hat{Y}=0$, $\sum e_i^2 = 0$

b) $\hat{\theta}_0 = \bar{y} - \hat{\theta}_1 x$, $\hat{\theta}_1 = r\frac{\sigma y}{\sigma x}$

$e = y - \bar{y} + r\frac{\sigma y}{\sigma x}(\bar{x}-x)$ $\qquad \vec{x}^T e = \sum_{i=1}^{n}\left(y_i - \bar{y} + r\frac{\sigma y}{\sigma x}(\bar{x}-x_i)\right)x_i$

$$= \sum x_i y_i - x_i\bar{y} + r\frac{\sigma y}{\sigma x}(x_i\bar{x} - x_i^2)$$

$nr\,\sigma x\,\sigma y = \sum x_i y_i - x_i\bar{y} - y_i\bar{x} + \bar{x}\bar{y}$

$\vec{x}^T e = nr\sigma_x\sigma y + r\frac{\sigma y}{\sigma x}\sum(\bar{x}^2 - x_i^2)$

$= n\sigma x^2 + \sum\bar{x}^2 - \sum x_i^2$

$\sigma x^2 = \frac{1}{n}\sum x_i^2 - \frac{1}{n}\sum\bar{x}^2$ $\quad$ As $Var(x) = E(x^2) - E(x)^2$

$\vec{x}^T e = 0$ $\qquad$ so $\vec{x}$ and e are orthogonal

c) Because $\vec{x}^T e = 0$ and $\hat{Y}$ lays on span(x).

$\hat{Y}$ and e are orthogonal

3.

$$\frac{d}{d\gamma} R(\gamma) = \frac{d}{d\gamma} \frac{1}{n}\sum_{i=1}^{n} (y_i - \gamma x_i)^2 = \frac{1}{n}\sum_{i=1}^{n} [-2x_i(y_i - \gamma x_i)]$$

to reach the minimum value.

$$\frac{d}{d\gamma} R(\gamma) = 0 = \sum_{i=1}^{n} (x_i^2 \gamma - x_i y_i) = 0$$

$$\gamma = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

4. a) False. $\hat{y} = \frac{\sum x_i y_i}{\sum x_i^2} x$ eg. two sample points $(3,4)$, $(2,2)$

$\hat{\gamma} = \frac{16}{13}$ $\hat{y} = \frac{16}{13} x$ gives $\hat{y}_1 = \frac{48}{13}$ $\hat{y}_2 = \frac{32}{13}$ $e \neq 0$

b) True. Because the geometric relation still holds

c) True $\hat{\gamma}$ is still on span(x), e and $\vec{v}$ are orthogonal.

d) False. same example in a) $\bar{x} = 2.5$ $\bar{y} = 3$

$\hat{y} = \frac{16}{13}\bar{x} \neq \bar{y}$

5. a) quadratic γ

b) $g_i(\gamma) = \frac{1}{n}(y_i - \gamma x_i)^2$

$$\frac{d}{d\gamma} g_i(\gamma) = \frac{-2x_i}{n}(y_i - \gamma x_i) = \frac{2}{n}(x_i^2 \gamma - x_i y_i)$$

$$\frac{d}{d\gamma^2} g_i(\gamma) = \frac{2}{n} x_i^2 \geq 0$$ so g_i is a convex function

c) Because $\frac{d}{dx^2} g(x)$ is non-negative, the $\frac{d}{dx} g(x)$ will always be ascending

when $\frac{d}{dx} g(x) < 0$, $g(x)$ is decreasing, $\frac{d}{dx} g(x) > 0$, $g(x)$ is increasing.

then when $\frac{d}{dx} g(x) = 0$, $g(x)$ will reach the minimum values

d) i. $g(cx_1 + (1-c)x_2) \le cg(x_1) + (1-c)g(x_2)$

$h(cx_1 + (1-c)x_2) \le ch(x_1) + (1-c)h(x_2)$

so that $g(x) + h(x)$ also holds

$f(x)$ is a convex function

ii. Because the inequality always holds, all the sum function will give the same result

e) Becausen $\frac{1}{n}(y_i - \gamma x_i)^2$ is convex

$MSE(\gamma) = \frac{1}{n}\sum(y_i - \gamma x_i)^2$ is convex

so $\frac{d}{d\gamma}MSE(\gamma)$ will give the minimum point