Q1 (a) $p(y|x) = \mathcal{N}(\mu = xb, \sigma^2 = e^{xa})$

$w = (a, b)$ & (a_1, b_1) weights

Well $D = \{\{(x_i, y_i)\}_{i=1}^n\}$

$p(y|x, w)$

Here $\mu = xb$ & $\sigma^2 = e^{xa}$.

$$\therefore w_{MLE} = \underset{w \in F}{\text{argmin}} - \sum_{i=1}^{n} \ln(p(y_i|x_i, w))$$

$$= \underset{w \in F}{\text{argmin}} - \sum_{i=1}^{n} \ln \frac{1}{\sqrt{2\pi e^{x_i^T a}}} \exp\left(-\frac{(y_i - x_i^T b)^2}{2e^{x_i^T a}}\right)$$

$$= \underset{w \in F}{\text{argmin}} - \sum_{n=1}^{n} \left(-\ln\left(\sqrt{2\pi e^{x_i^T a}}\right) - \frac{(y_i - x_i^T b)^2}{2e^{x_i^T a}}\right)$$

$$= \underset{w \in F}{\text{argmin}} \sum_{i=1} \ln\left(\sqrt{2\pi e^{x_i^T a}}\right) + \frac{(y_i - x_i^T b)^2}{2e^{x_i^T a}}$$

$$= \underset{w \in F}{\text{argmin}} \sum_{i=1}^{n} \ln(\sqrt{2\pi}) + \frac{1}{2} x_i^T a + \frac{(y_i - x_i^T b)^2}{2e^{x_i^T a}}$$

$$= \underset{w \in F}{\text{argmin}} \sum_{i=1}^{n} \left(\frac{1}{2} x_i^T a + \frac{(y_i - x_i^T b)^2}{2e^{x_i^T a}}\right)$$

$$= \underset{w \in F}{\text{argmin}} \sum_{i=1}^{n} \left(\frac{x_i^T a}{2} + \frac{1}{2}\frac{(y_i - x_i^T b)^2}{\cancel{2}e^{x_i^T a}}\right)$$

$$\frac{d\, c_t(w)}{da}$$

consider $c(w_t) = \frac{1}{2} x_i^T a + \frac{1}{2} \frac{(y_i - x_i^T b)^2}{e^{x_i^T a}}$

$$\frac{d\, c(w_t)}{da} = \frac{x_i^T}{2} + \frac{1}{2}(y_i - x_i^T b)^2 \cdot e^{-x_i^T a} \cdot -x_i^T$$

$$= \frac{x_i^T}{2}\left(1 - (y_i - x_i^T b)^2 \cdot e^{-x_i^T a}\right)$$

Similarly

$$\frac{d\, c(w_t)}{db} = 0 + \frac{1}{e^{x_i^T a}} \cdot \frac{\cancel{2}}{\cancel{2}} (x_i^T b - y_i)\, x_i^T$$

$$= \frac{1}{e^{x_i^T a}} (x_i^T b - y_i)\, x_i^T.$$

(b) Code

Distribution Regressor

(c) Mean - 0.234151

Standard error - 0.02816 } 20 epochs

~~Mean~~

~~Mean -~~ Mean - 0.231291

~~Standard error -~~ Stand error - 0.0237771 } 1 epoch

Mean Regressor

Mean - 0.197

Q2 (h)

	Linear	Polynomial
Average error	0.12324	0.074288
Standard error	0.0004832	0.00061126

(i)

Average error for Heuristic is the lowest but the standard error for Adagrad is the lowest with constant LR being the worst in both cases.