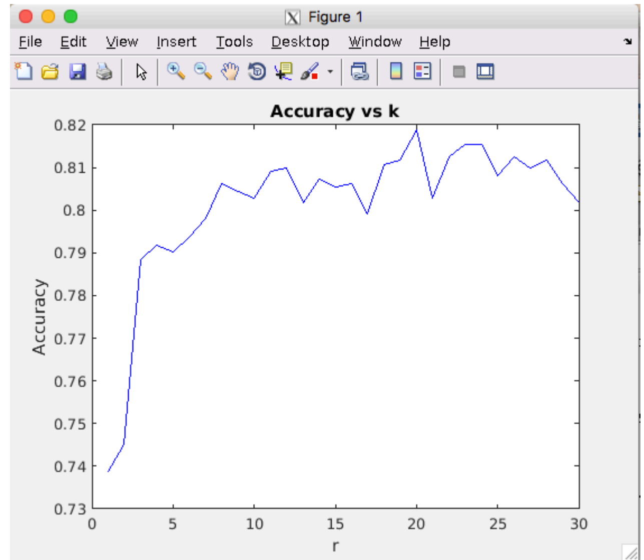
1.)

Accuracies for k = 1 – 30.
Highest average accuracy is achieved when k = 20.

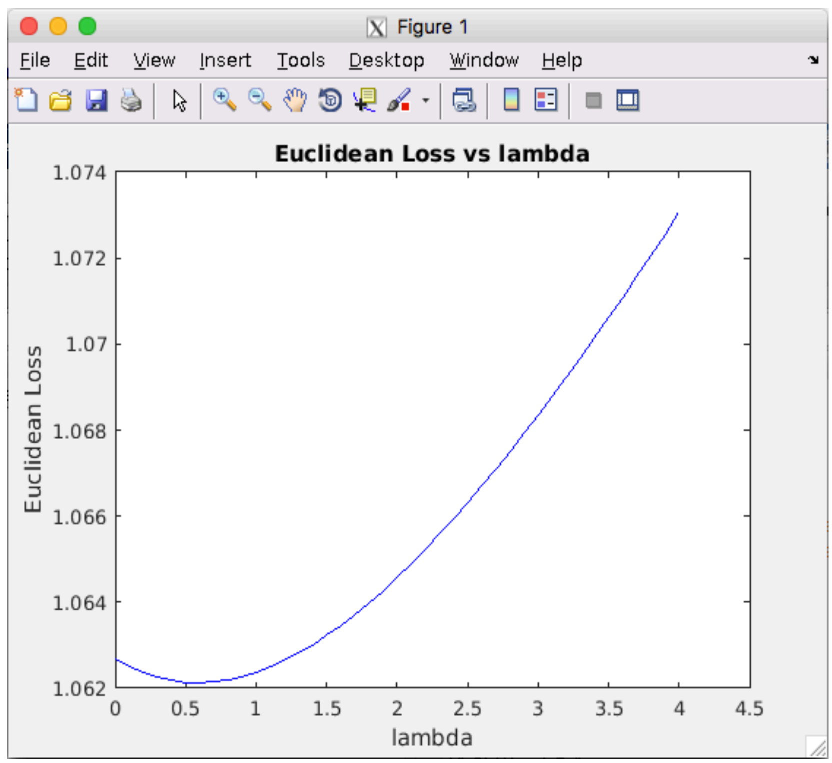
0.7387	0.7450	0.7883	0.7919	0.7901	0.7937
0.7982	0.8063	0.8045	0.8027	0.8090	0.8099
0.8018	0.8072	0.8054	0.8063	0.7991	0.8108
0.8117	**0.8189**	0.8027	0.8126	0.8153	0.8153
0.8081	0.8126	0.8099	0.8117	0.8063	0.8018



2.)

Losses for lambda = 0.0,0.1,0.2, ..., 3.9,4.0
Lowest average loss is achieved when lambda = 0.6.

```
lambda 0.00000000        avgLoss 1.06269953
lambda 0.10000000        avgLoss 1.06251424
lambda 0.20000000        avgLoss 1.06236716
lambda 0.30000000        avgLoss 1.06225651
lambda 0.40000000        avgLoss 1.06218060
lambda 0.50000000        avgLoss 1.06213780
lambda 0.60000000        avgLoss 1.06212658
lambda 0.70000000        avgLoss 1.06214546
lambda 0.80000000        avgLoss 1.06219305
lambda 0.90000000        avgLoss 1.06226800
lambda 1.00000000        avgLoss 1.06236906
lambda 1.10000000        avgLoss 1.06249500
lambda 1.20000000        avgLoss 1.06264467
lambda 1.30000000        avgLoss 1.06281695
lambda 1.40000000        avgLoss 1.06301079
lambda 1.50000000        avgLoss 1.06322518
lambda 1.60000000        avgLoss 1.06345917
lambda 1.70000000        avgLoss 1.06371182
lambda 1.80000000        avgLoss 1.06398225
lambda 1.90000000        avgLoss 1.06426964
lambda 2.00000000        avgLoss 1.06457317
lambda 2.10000000        avgLoss 1.06489207
lambda 2.20000000        avgLoss 1.06522562
lambda 2.30000000        avgLoss 1.06557310
lambda 2.40000000        avgLoss 1.06593386
lambda 2.50000000        avgLoss 1.06630723
lambda 2.60000000        avgLoss 1.06669262
lambda 2.70000000        avgLoss 1.06708942
lambda 2.80000000        avgLoss 1.06749708
lambda 2.90000000        avgLoss 1.06791505
lambda 3.00000000        avgLoss 1.06834283
lambda 3.10000000        avgLoss 1.06877991
lambda 3.20000000        avgLoss 1.06922583
lambda 3.30000000        avgLoss 1.06968012
lambda 3.40000000        avgLoss 1.07014236
lambda 3.50000000        avgLoss 1.07061212
lambda 3.60000000        avgLoss 1.07108901
lambda 3.70000000        avgLoss 1.07157265
lambda 3.80000000        avgLoss 1.07206267
lambda 3.90000000        avgLoss 1.07255871
lambda 4.00000000        avgLoss 1.07306045
```



③ To minimize $L(w,b)$, set $\frac{\partial L(w,b)}{\partial w} = \frac{\partial L(w,b)}{\partial b} = 0$

$$\begin{aligned}
\frac{\partial L(w,b)}{\partial w} &= \frac{\partial}{\partial w}\left(\sum_{n=1}^{m} r_n (y_n - w x_n + b)^2\right) \\
&= \sum_{n=1}^{m} r_n \frac{\partial}{\partial w}(y_n - w x_n + b)^2 \\
&= \sum_{n=1}^{m} r_n 2 (y_n - w x_n + b)(-x_n) \\
&= -\sum_{n=1}^{m} 2 r_n x_n (y_n - w x_n + b) \\
&= 0 \qquad \text{(I)}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial L(w,b)}{\partial b} &= \frac{\partial}{\partial b}\left(\sum_{n=1}^{m} r_n (y_n - w x_n + b)^2\right) \\
&= \sum_{n=1}^{m} r_n \frac{\partial}{\partial b}(y_n - w x_n + b)^2 \\
&= \sum_{n=1}^{m} r_n 2 (y_n - w x_n + b) \\
&= \sum_{n=1}^{m} 2 r_n (y_n - w x_n + b) \\
&= 0 \qquad \text{(II)}
\end{aligned}$$

$$\text{(I)} \Rightarrow \sum_{n=1}^{m} (r_n x_n y_n - w r_n x_n^2 + b r_n x_n) = \sum_{n=1}^{m} r_n x_n y_n - w \sum_{n=1}^{m} r_n x_n^2 + b \sum_{n=1}^{m} r_n x_n = 0 \qquad \text{(III)}$$

$$\text{(II)} \Rightarrow \sum_{n=1}^{m} (r_n y_n - w r_n x_n + r_n b) = \sum_{n=1}^{m} r_n y_n - w \sum_{n=1}^{m} r_n x_n + b \sum_{n=1}^{m} r_n = 0 \qquad \text{(IV)}$$

$$\begin{aligned}
(\Sigma r_n)\text{(III)} - (\Sigma r_n x_n)\text{(IV)} &= \Sigma r_n (\Sigma r_n x_n y_n - w \Sigma r_n x_n^2 + b \Sigma r_n x_n) \\
&\quad - \Sigma r_n x_n (\Sigma r_n y_n - w \Sigma r_n x_n + b \Sigma r_n) \\
&= \Sigma r_n \Sigma r_n x_n y_n - w \Sigma r_n \Sigma r_n x_n^2 - \Sigma r_n x_n \Sigma r_n y_n + w (\Sigma r_n x_n)^2 \\
&= w\left[(\Sigma r_n x_n)^2 - \Sigma r_n \Sigma r_n x_n^2\right] - \left[\Sigma r_n x_n \Sigma r_n y_n - \Sigma r_n \Sigma r_n x_n y_n\right] \\
&= 0
\end{aligned}$$

$$\Rightarrow \boxed{w = \frac{\Sigma r_n x_n \Sigma r_n y_n - \Sigma r_n \Sigma r_n x_n y_n}{(\Sigma r_n x_n)^2 - \Sigma r_n \Sigma r_n x_n^2}}$$

$$\begin{aligned}
(\Sigma r_n x_n)\text{(III)} - (\Sigma r_n x_n^2)\text{(IV)} &= \Sigma r_n x_n (\Sigma r_n x_n y_n - w \Sigma r_n x_n^2 + b \Sigma r_n x_n) \\
&\quad - \Sigma r_n x_n^2 (\Sigma r_n y_n - w \Sigma r_n x_n + b \Sigma r_n) \\
&= b(\Sigma r_n x_n \Sigma r_n x_n - \Sigma r_n \Sigma r_n x_n^2) - (\Sigma r_n x_n^2 \Sigma r_n y_n - \Sigma r_n x_n \Sigma r_n x_n y_n) \\
&= 0
\end{aligned}$$

$$\Rightarrow \boxed{b = \frac{\Sigma r_n x_n^2 \Sigma r_n y_n - \Sigma r_n x_n \Sigma r_n x_n y_n}{(\Sigma r_n x_n)^2 - \Sigma r_n \Sigma r_n x_n^2}}$$

$* \quad \Sigma \equiv \sum_{n=1}^{m} \quad *$

(b) Say $y = f(\bar{x}_n) + \varepsilon$ with $\varepsilon = N(0, \theta_n^2)$

Then $Pr(y \mid \bar{X}, w, \theta) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}\,\theta_n} e^{-(y_n - w x_n + b)^2 / 2\theta_n^2}$

$$
\begin{aligned}
w^* &= \arg\max_{w,b} \left(\prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}\,\theta_n} e^{-(y_n - w x_n + b)^2 / 2\theta_n^2} \right) \\
&= \arg\max_{w,b} \left(\prod_{n=1}^{N} \frac{1}{\theta_n} e^{-(y_n - w x_n + b)^2 / 2\theta_n^2} \right) \\
&= \arg\max_{w,b} \left(\prod_{n=1}^{N} \frac{1}{\theta_n} \prod_{n=1}^{N} e^{-(y_n - w x_n + b)^2 / 2\theta_n^2} \right) \\
&= \arg\max_{w,b} \prod_{n=1}^{N} e^{-(y_n - w x_n + b)^2 / 2\theta_n^2} \\
&= \arg\max_{w,b} \sum_{n=1}^{N} -(y_n - w x_n + b)^2 / 2\theta_n^2 \\
&= \arg\max_{w,b} \sum_{n=1}^{N} -(y_n - w x_n + b)^2 / \theta_n^2 \\
&= \arg\min_{w,b} \sum_{n=1}^{N} r_n (y_n - w x_n + b)^2 \quad \text{with} \quad r_n = \frac{1}{\cancel{2}\theta_n^2}
\end{aligned}
$$

$\Rightarrow$ The objective is equivalent.

Variance of measure n $= \theta_n^2 = \frac{1}{\cancel{2} r_n}$