1

$$\min_{\vec{v}:||\vec{v}||=1} J(\vec{v}) = \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} ||\vec{x}_i - (\vec{v}^T \vec{x}_t)\vec{v}||^2 \right)$$

In the above notation $\min_{\vec{v}:||\vec{v}||=1} J(\vec{v})$ means that we minimize $J(\vec{v})$ *such that* the length of $\vec{v}$ is 1. Continuing our derivation:

$$\begin{aligned}
\min_{\vec{v}:||\vec{v}||=1} J(\vec{v}) &= \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} ||\vec{x}_i - (\vec{v}^T \vec{x}_i)\vec{v}||^2 \right) \\
&= \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} [\vec{x}_i - (\vec{v}^T \vec{x}_i)\vec{v}]^T [\vec{x}_i - (\vec{v}^T \vec{x}_i)\vec{v}] \right) \\
&= \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} [\vec{x}_i^T - (\vec{v}^T \vec{x}_i)\vec{v}^T][\vec{x}_i - (\vec{v}^T \vec{x}_i)\vec{v}] \right) \\
&= \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} \vec{x}_i^T \vec{x}_i - (\vec{v}^T \vec{x}_i)\vec{v}^T \vec{x}_i - \vec{x}_i^T (\vec{v}^T \vec{x}_i)\vec{v} + (\vec{v}^T \vec{x}_i)\vec{v}^T (\vec{v}^T \vec{x}_i)\vec{v} \right) \\
&= \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} \vec{x}_i^T \vec{x}_i - \vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v} - \vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v} + (\vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v})\vec{v}^T \vec{v} \right)
\end{aligned}$$

In the last line we have used the fact that $(\vec{v}^T \vec{x}_i)\vec{v}^T \vec{x}_i = \vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v}$, $\vec{x}_i^T (\vec{v}^T \vec{x}_i)\vec{v} = \vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v}$ and $(\vec{v}^T \vec{x}_i)\vec{v}^T (\vec{v}^T \vec{x}_i)\vec{v} = (\vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v})\vec{v}^T \vec{v}$

We can then write:

$$\begin{aligned}
\min_{\vec{v}:||\vec{v}||=1} J(\vec{v}) &= \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} \vec{x}_i^T \vec{x}_i - \vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v} - \vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v} + (\vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v})\vec{v}^T \vec{v} \right) \\
&= \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} \vec{x}_i^T \vec{x}_i + \vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v}(-2 + \vec{v}^T \vec{v}) \right) \\
&= \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} \vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v}(-2 + \vec{v}^T \vec{v}) \right) \\
&= \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} \vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v}(-2 + 1) \right) \\
&= \min_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} -\vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v} \right) \\
&= \max_{\vec{v}:||\vec{v}||=1} \left(\sum_{i=1}^{n} \vec{v}^T \vec{x}_i \vec{x}_i^T \vec{v} \right) \\
&= \max_{\vec{v}:||\vec{v}||=1} \left(\vec{v}^T X X^T \vec{v} \right)
\end{aligned}$$

2 for conveniencewe set

S1=A S2=B

Forward:

$$\alpha_1^A = 0.6 * 0.7 = 0.42$$

$$\alpha_1^B = 0.4 * 0.8 = 0.32$$

$$\alpha_2^A = 0.3 * (0.42 * 0.9 + 0.32 * 0.2) = 0.132$$

$$\alpha_2^B = 0.2 * (0.42 * 0.1 + 0.8 * 0.32) = 0.06$$

$$\alpha_3^A = 0.7 * (0.132 * 0.9 + 0.06 * 0.2) = 0.09$$

$$\alpha_3^B = 0.8 * (0.132 * 0.1 + 0.8 * 0.06) = 0.048$$

P=0.09+0.048=0.138

Backward:

$$\beta_3^A = 1$$

$$\beta_3^B = 1$$

$$\beta_2^A = 0.9 * 0.7 * 1 + 0.1 * 0.8 * 1 = 0.71$$

$$\beta_2^B = 0.2 * 0.7 * 1 + 0.8 * 0.8 * 1 = 0.78$$

$$\beta_1^A = 0.9 * 0.71 * 0.3 + 0.1 * 0.78 * 0.2 = 0.2$$

$$\beta_1^B = 0.2 * 0.71 * 0.3 + 0.8 * 0.2 * 0.78 = 0.1678$$

P=0.32*0.1678+0.42*0.2=0.138

Forward and backward agree each other

Forward-backward

$$\alpha_1^A \beta_1^A = 0.42 * 0.2 = 0.084$$

$$\alpha_1^B \beta_1^B = 0.32 * 0.1678 = 0.054$$

$$\alpha_2^A \beta_2^A = 0.132 * 0.71 = 0.09$$

$$\alpha_2^B \beta_2^B = 0.06 * 0.78 = 0.046$$

$$\alpha_3^A \beta_3^A = 0.09$$

$$\alpha_3^B \beta_3^B = 0.048$$

So it should be A AA

Viterbi

$$V_1^A = 0.7 * 0.6 = 0.42$$

$$V_1^B = 0.4 * 0.8 = 0.32$$

$$V_2^A = 0.9 * 0.42 * 0.3 = 0.11$$

$$V_2^B = 0.1 * 0.42 * 0.2 = 0.008$$

$$V_3^A = 0.9 * 0.11 * 0.7 = 0.07$$

$$V_3^B = 0.1 * 0.11 * 0.8 = 0.008$$

So it should be A AA

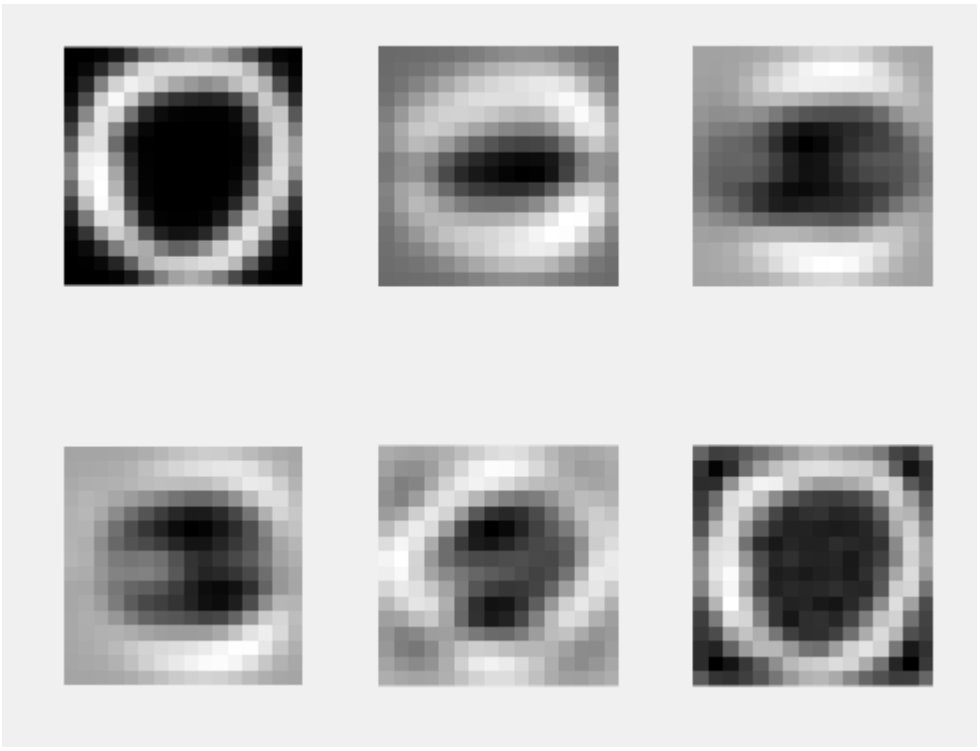
3 2)

3) K=original ,1,3,5,15,100

#250

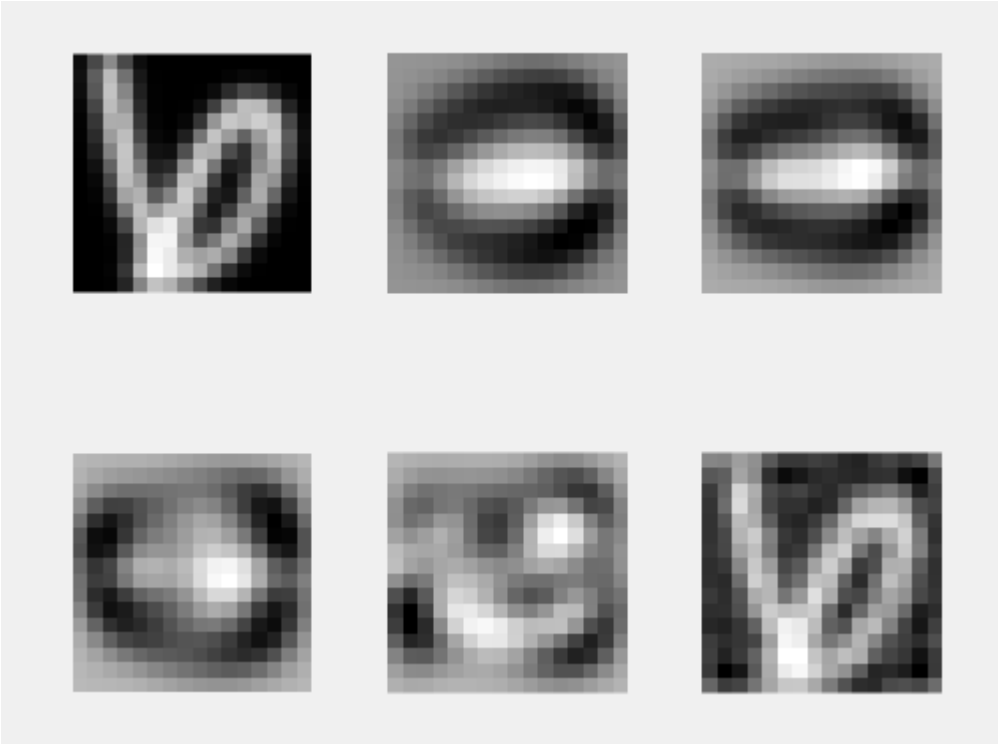
#300

450



#500

#3000



4)

the time for training data k=1,3,5,15,100 are:

0.1404, 0.2652,0.4680,0.8892,6.0528

The accuracy are:

0.6014,0.8134,0.9204,0.9666,0.9686

the time fortest data k=1,3,5,15,100 are:

0.0312, 0.0624,0.1872,0.2028,1.0296

The accuracy are:

0.5623,0.7514,0.8527,0.9223,0.9445

We can see that k increase or data amount increase, we need more time to do the computation.