Homework 5 Jeffrey Kesley

Question One

a) The fisher linear discriminant method projects a set of sample points of d dimensions, onto a projected line that has the most seperation in d-1 dimensions.

This is done through finding some w, which will project our points and satisfy the ~~linear~~ criterion function for its maximum. Using the projected means and within class scatter, as well as the between class scatter, we see the ration between these projected values gives us

$J(w) = \frac{w^t S_B w}{w^t S_W w}$, where w is the solution from our $w = S_W^{-1}(m_1 - m_2)$ to maximize $J(\cdot)$.

This w is used for our projected means, points, and is only relevant for its direction it tells us, not magnitude.

b) First we find our means, $m1 = \begin{pmatrix} 1.8\bar{3}\bar{3} \\ 2.00 \end{pmatrix}$

$m2 = \begin{pmatrix} 3.8\bar{3}\bar{3} \\ 3.00 \end{pmatrix}$

Next we find our scatters, ~~which is the sum of~~

$S_i = \sum_{x \in D_i} (x - m_i)(x - m_i)^t$, Here $S_1 = \begin{bmatrix} 4.833 & -1 \\ -1.00 & 8 \end{bmatrix}$

$S_2 \begin{bmatrix} 8.8\bar{3}\bar{3} & 3 \\ 3.00 & 12 \end{bmatrix}$

$S_W = S_1 + S_2 = \begin{bmatrix} 13.66 & 2 \\ 2.00 & 20 \end{bmatrix}$

Then we solve for w

$$w = S_w^{-1}(m_1 - m_2)$$

$$w = \begin{pmatrix} -.141089 \\ -.03589 \end{pmatrix}$$

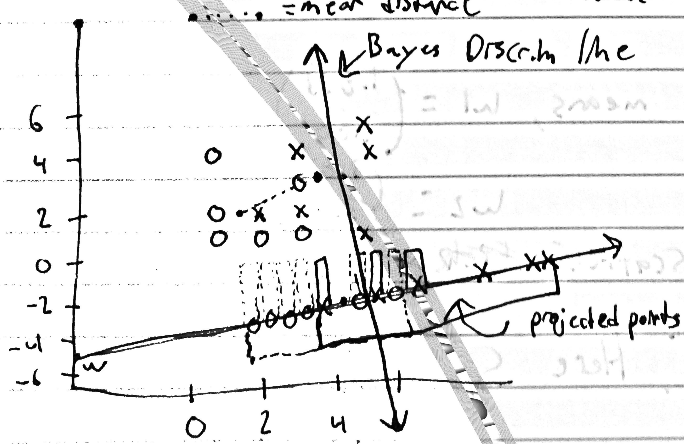
Finally for our S_B, between class scatter

$$S_B = (m_1 - m_2)(m_1 - m_2)^t = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$$

$$\text{our } J(w) = \frac{\begin{pmatrix} -.141089 \\ -.03589 \end{pmatrix}^t \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \begin{pmatrix} -.141089 \\ -.03589 \end{pmatrix}}{\begin{pmatrix} -.141089 \\ -.03589 \end{pmatrix}^t \begin{bmatrix} 13.66 & 2.00 \\ 2.00 & 20.00 \end{bmatrix} \begin{pmatrix} -.141089 \\ -.03589 \end{pmatrix}} = .3180693$$

i) Our projection line simply uses .3180693 as a coef for each X

...... = mean distance   added an intercept value for clarity. −4 for each y value



We assume uniform distribution within our max/min for all points

$Priors_1 = Priors_2$

If we increase our w magnitude by 10 when we project our points, as well as move them down −4 to match our w change.

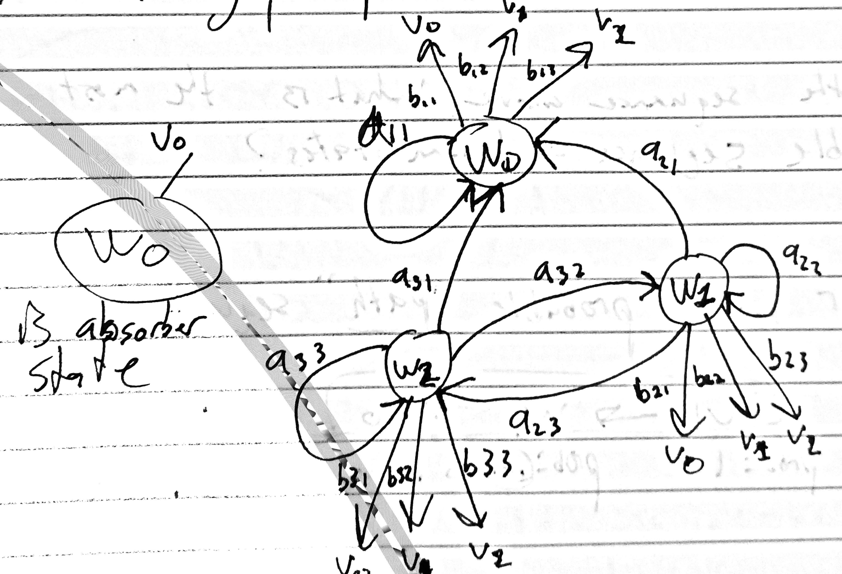
Our $y = w^t x$ projected points are passed to our criterion $J(w)$ to give us our one dimensional mapping on the projection line.

Question 2

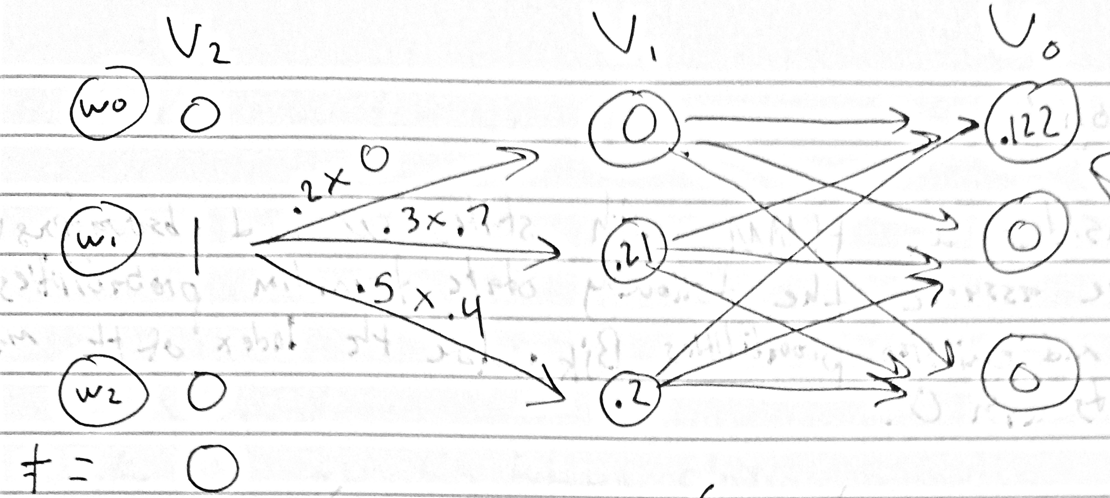
Consider an HMM with states ω and observations V. We assume the following state transition probabilities A_{ij} and emission probabilities B_{jk} where the index of the matrices starts with 0.

$$A_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ .2 & .3 & .5 \\ .4 & .5 & .1 \end{pmatrix} \qquad B_{jk} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & .7 & .3 \\ 0 & .4 & .6 \end{pmatrix}$$

a) Show a graph representation of this hidden Markov Model.



b) suppose $t=0$, $= w_1$, what is the prob it generates the particular sequence $V = \{v_2, v_1, v_0\}$



$1 \times (0 \times 1 + .21 \times .2 + .2 \times .4) =$

the prob of this $\boxed{\text{sequence } V = \{v_2, v_1, v_0\}} = .122$

() Given the sequence above, what is the most probable sequence of hidden states?

Our most "probable "path" sequence of hidden states

is $\boxed{w_1 \rightarrow w_1 \rightarrow w_0}$ as

prob = 1 prob = .21 vs .2 w_2 and .122 vs 0

We used forward algorithm...