ESE 650, SPRING 2021

HOMEWORK 1

IAN MACDONALD [MACIAN@SEAS.UPENN.EDU],
COLLABORATORS: NONE

Solution 1 (Time spent: 9 hours)**.** -
Submitted on Gradescope

Solution 2 (Time spent: 8 hours). -

a)

Code submitted on Gradescope

Smoothing Probabilities (γ_k):

γ_1	0.55555556	0.44444444
γ_2	0.8	0.2
γ_3	0.8	0.2
γ_4	0.55555556	0.44444444
γ_5	0.16666667	0.83333333
γ_6	0.55555556	0.44444444
γ_7	0.16666667	0.83333333
γ_8	0.16666667	0.83333333
γ_9	0.16666667	0.83333333
γ_{10}	0.55555556	0.44444444
γ_{11}	0.16666667	0.83333333
γ_{12}	0.16666667	0.83333333
γ_{13}	0.16666667	0.83333333
γ_{14}	0.16666667	0.83333333
γ_{15}	0.16666667	0.83333333
γ_{16}	0.55555556	0.44444444
γ_{17}	0.55555556	0.44444444
γ_{18}	0.8	0.2
γ_{19}	0.8	0.2
γ_{20}	0.16666667	0.83333333

Forward Probabilities (α_k):

α_1	2.50000000e-01	2.00000000e-01
α_2	9.00000000e-02	2.25000000e-02
α_3	2.25000000e-02	5.62500000e-03
α_4	7.03125000e-03	5.62500000e-03
α_5	6.32812500e-04	3.16406250e-03
α_6	9.49218750e-04	7.59375000e-04
α_7	8.54296875e-05	4.27148438e-04

α_8	2.56289063e-05	1.28144531e-04
α_9	7.68867188e-06	3.84433594e-05
α_{10}	1.15330078e-05	9.22640625e-06
α_{11}	1.03797070e-06	5.18985352e-06
α_{12}	3.11391211e-07	1.55695605e-06
α_{13}	9.34173633e-08	4.67086816e-07
α_{14}	2.80252090e-08	1.40126045e-07
α_{15}	8.40756270e-09	4.20378135e-08
α_{16}	1.26113440e-08	1.00890752e-08
α_{17}	5.67510482e-09	4.54008386e-09
α_{18}	2.04303773e-09	5.10759434e-10
α_{19}	5.10759434e-10	1.27689858e-10
α_{20}	3.19224646e-11	1.59612323e-10

Backward Probabilities (β_k):

β_1	4.25632861e-10	4.25632861e-10
β_2	1.70253145e-09	1.70253145e-09
β_3	6.81012578e-09	6.81012578e-09
β_4	1.51336129e-08	1.51336129e-08
β_5	5.04453762e-08	5.04453762e-08
β_6	1.12100836e-07	1.12100836e-07
β_7	3.73669453e-07	3.73669453e-07
β_8	1.24556484e-06	1.24556484e-06
β_9	4.15188281e-06	4.15188281e-06
β_{10}	9.22640625e-06	9.22640625e-06
β_{11}	3.07546875e-05	3.07546875e-05
β_{12}	1.02515625e-04	1.02515625e-04
β_{13}	3.41718750e-04	3.41718750e-04
β_{14}	1.13906250e-03	1.13906250e-03
β_{15}	3.79687500e-03	3.79687500e-03
β_{16}	8.43750000e-03	8.43750000e-03
β_{17}	1.87500000e-02	1.87500000e-02
β_{18}	7.50000000e-02	7.50000000e-02
β_{19}	3.00000000e-01	3.00000000e-01
β_{20}	1.00000000e+00	1.00000000e+00

Point-wise most likely sequence of states:
'LA', 'LA', 'LA', 'LA', 'NY', 'LA', 'NY', 'NY', 'NY', 'LA', 'NY', 'NY', 'NY', 'NY', 'NY', 'LA', 'LA', 'LA', 'LA', 'NY'

b)
Proof:
$\xi_k(x,x') = P(x_k = x, x_{k+1} = x'|Y_1,...,Y_t,\lambda)$
$\xi_k(x,x') = P(x_{k+1} = x'|Y_1,...,Y_t,\lambda, x_k = x) * P(x_k = x|Y_1,...,Y_t,\lambda)$
$\xi_k(x,x') = P(x_k = x|Y_1,...,Y_t,\lambda) * P(x_{k+1} = x'|Y_1,...,Y_t,\lambda) * P(x_{k+1} = x|x_k = x)$
$\xi_k(x,x') = \gamma_k(x) * T_{x,x'} * \gamma_{k+1}(x')$
$\xi_k(x,x') = \frac{\alpha_k(x)\beta_k(x)}{\sum_x \alpha_t(x)} \frac{\alpha_{k+1}(x')\beta_{k+1}(x')}{\sum_x \alpha_t(x')} T_{x,x'}$
$\xi_k(x,x') = \eta\alpha_k(x)T_{x,x'}M_{x',y_{k+1}}\beta_{k+1}(x')$
c)
New Initial State (π'):

0.55555556	0.44444444

New Transition Matrix (T'):

0.47023206	0.52976794
0.3526061	0.6473939

New Observation Matrix (M'):

0.3902439	0.20325203	0.40650407
0.06779661	0.70621469	0.2259887

d)
Original Trajectory Probability: 3.7252903e-13
New Trajectory Probability: 7.67172593e-12

The new trajectory probability is higher than the original trajectory probability because our new hidden markov model λ' has been updated to better model the problem given the observations we have seen. Therefore, $P(Y_1, ..., Y_t)|\lambda < P(Y_1, ..., Y_t)|\lambda'$.

Solution 3 (Time spent: 3 hours)**.** -

a)

Using:

$\alpha_k(x) = P(Y_1, ..., Y_k, X_k = x)$

$\beta_k(x) = P(Y_{k+1}, ..., Y_t | X_k = x)$

$T_{ij} = P(X_{k+1} = x_j | X_k = x_i)$

$M_{ij} = P(Y_k = y_j | X_k = x_i)$

$\xi_k(x, x') = P(X_k = x, X_{k+1} = x' | Y_1, ..., Y_t)$

$\gamma_k(x) = P(X_k = x | Y_1, ..., Y_t)$

1)

$P(X_{k+1} = x_j | X_k = x_i, Y_1, ..., Y_t)$

$= \frac{P(X_{k+1}=x_j|X_k=x_i)P(X_{k+1}=x_j|Y_1,...,Y_t)}{P(X_k=x_i|Y_1,...,Y_t)}$

$= \frac{\gamma_{k+1}(x_j)T_{ij}}{\gamma_k(x_i)}$

$= \frac{\alpha_{k+1}(x_j)\beta_{k+1}(x_j)T_{ij}}{\alpha_k(x_i)\beta_k(x_i)}$

2)

$P(X_k = x_i | X_{k+1} = x_j, Y_1, ..., Y_t)$

$= \frac{X_k=x_i|X_{k+1}=x_j)P(X_k=x_i|Y_1,...,Y_t)}{P(X_{k+1}=x_j|Y_1,...,Y_t)}$

$= \frac{\gamma_k(x_i)T_{ij}}{\gamma_{k+1}(x_j)}$

$= \frac{\alpha_k(x_i)\beta_k(x_i)T_{ij}}{\alpha_{k+1}(x_j)\beta_{k+1}(x_j)}$

3)

$P(X_{k-1} = x_i, X_k = x_j, x_{k+1} = x_1 | Y_1, ..., Y_t)$

= $P(X_{k-1} = x_i, X_k = x_j | Y_1, ..., Y_t)P(X_k = x_j, x_{k+1} = x_1 | Y_1, ..., Y_t)$

= $\xi_{k-1}(x_i, x_j)\xi_k(x_j, x_k)$

= $\eta\alpha_{k-1}(x_i)T_{x_i,x_j}M_{x_i,y_k}\beta_k(x_j)\alpha_k(x_j)T_{x_j,x_k}M_{x_j}, y_{k+1}\beta_{k+1}(x_k)$

b)

In general, the solution of the decoding problem is not the same as the solution of the smoothing problem because the smoothing problem is solved independently for each time step, whereas the decoding problem models the joint probability of the states at all time steps simultaneously. This allows us to use all of the observations to model the most likely trajectory at once, rather than trying to find each state independently. These solutions would be the same if each state was independent of the previous state.

Solution 4 (Time spent: 2 hours)**.** -
Unbiased when $E[\hat{X}] = E[X]$
$E[\hat{X}] = E[X]$
$E[\hat{X}] = E[a_1 Y_1 + a_2 Y_2]$
$E[\hat{X}] = a_1 E[Y_1] + a_2 E[Y_2]$
$E[\hat{X}] = a_1 E[h_1 X + \epsilon_1] + a_2 E[h_2 X + \epsilon_2]$
$E[\hat{X}] = a_1(E[X] + E[\epsilon_1]) + 2a_2(E[X] + [\epsilon_2])$
$E[\hat{X}] = a_1 E[X] + 0 + 2a_2 E[X] + 0$
$E[\hat{X}] = (a_1 + 2a_2)E[X]$
$E[\hat{X}] = E[X]$ when $(a_1 + 2a_2) = 1$
$a_1 \sigma_2^2 = a_2 \sigma_1^2$
$a_1 = a_2 \frac{\sigma_1^2}{\sigma_2^2}$

Minimum value of estimation error = 0

b)
i) If $\sigma_2 > \sigma_1$ then we want to put more weight on a_1
ii) If $\sigma_2 = \sigma_1$ then it is even
iii) If $\sigma_1 > \sigma_2$ then we want to put more weight on a_2
This agrees with my intuition because we want to weigh more heavily the observations with less variance.