Unsupervised Training

Given:

1. N – the number of states, e.g., 2, (s1 and s2)
2. V – the vocabulary, e.g., V={a,b}
3. O – observations, e.g., O=aaaaabbbbb
4. ~~State transitions, e.g., S=1121122222~~

Task: Estimate the following parameters

1. π_1, π_2
2. a_{11}, a_{12}, a_{22}, a_{21}
3. $b_1(a)$, $b_1(b)$, $b_2(a)$, $b_2(b)$

λ

How could this be possible?

Maximum Likelihood: $\lambda^* = \arg\max_\lambda p(O \mid \lambda)$