

Consider a hidden Markov model with 2 possible hidden states (non-coding = 1 and coding = 2) and 4 possible observations (A, C, G, T). Assuming its parameters as the following

- Initial hidden state probabilities $\pi = [0.7, 0.3]$
- Transition probabilities $A = \begin{bmatrix} 0.96 & 0.04 \\ 0.15 & 0.85 \end{bmatrix}$
- Emission probabilities $B = \begin{bmatrix} 0.4 & 0.1 & 0.1 & 0.4 \\ 0.2 & 0.3 & 0.3 & 0.2 \end{bmatrix}$

- (a) Calculate $P(\text{TCGA}|\lambda)$
- (b) How many possible explanations (hidden state sequences) for ACATCGTCGGT?
- (c) Calculate the best hidden state sequence for GCGCATTAAATTAGGCGTCGTAGTTCCTT.
- (d) How to deal with the underflow problem, especially when applied a hidden Markov model to a very long sequence?

Submission: (1) Your codes, (2) your answers to the above three questions in a written report.

<https://powcoder.com>

Assignment Project Exam Help
Assignment Project Exam Help
Add WeChat powcoder

<https://powcoder.com>

Add WeChat powcoder