

Math 189Z Homework 3: Reading Summaries

Daphne Poon

April 24 2020

1 Gene finding and the Hidden Markov models

This papers talks about how HMM is used to find genes in eukaryotes. There are two main tasks that use HMM: segmentation (identifying regions with different biochemical properties) and gene finding (differentiating between coding and non-coding regions of the genome).

To do segmentation, we run an iterative algorithm with a known observable sequence s (the sequence of genetic code), and guesses for the initial probabilities p , the transition matrix T and the emission matrix E . Using this information and something called the Viterbi algorithm, we can find h^* , the maximum likelihood hidden sequence. We then use this h^* and s to re-estimate p , T and E , and so on.

Gene finding (in eukaryotes) is a little more complicated, which is why it's not actually talked about in the paper.

2 Hidden Markov Model for Stock Trading

In this paper, several HMMs are tested to see which is the best for predicting stock performance. The optimal one, which is described in the paper, has the observable sequence at each timestep as a 4-tuple, and each timestep covered a month of time. The 4-tuple had information about the open, low, high or closing price of the S&P 500.

In this paper, the authors used the forward, backward and Baum-Welch algorithms to estimate the HMM parameters.

3 Final Project Reading

One of the sources we chose for our final project was 'Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand'. This paper was published mid-March, before social distancing became commonplace in Great Britain. It talks about the non-pharmaceutical interventions that would be effective to combat COVID-19. Two main strategies are identified: mitigation (slowing virus spread to reduce healthcare

demand) and suppression (attempting to stop spread altogether). This paper is focussed on making recommendations for the UK.

In the paper, they use a transmission model based on data from Wuhan, and they summarize that a multi-layering strategy (i.e. combining multiple interventions) is necessary to have significant impact on the transmission of the virus.