Biostat 140.668 Problem set 2

Hidden Markov models

1. Consider a single intercross individual derived from two inbred mouse strains. Consider M ordered markers, and let $G_i \in \{AA, AB, BB\}$ denote the (phase-unknown) genotype of the individual at marker i. Show that the G_i form a Markov chain.

2. Fill in the details in the Viterbi algorithm to calculate

$$\hat{g} = \arg \max_{g_1,\ldots,g_M} \Pr(G_1 = g_1, \ldots, G_M = g_M \mid \boldsymbol{O})$$