5.1 Transition rates

1. *Point mutations in DNA:* Since the four nucleotides in DNA have different chemical compositions and energetics, they could mutate at different rates. We shall explore whether, without natural selection at work, such preferential mutation may lead to different compositions of nucleotides.

   (a) Consider a simple model in which all *transitions* (i.e. mutations between purines A and G, or between pyramidines T and C) occur with probability q, while *transversions* (i.e. any mutation from a purine to a pyrimidine or vice versa) occur with probability p, in each generation. Write down the 4×4 (Markov) transition matrix, Π_1, that relates the frequencies of nucleotides ($p_{\rm A}$, $p_{\rm G}$, $p_{\rm T}$, $p_{\rm C}$) from one generation to the next. (Make sure that the normalization condition $p_{\rm A} + p_{\rm G} + p_{\rm T} + p_{\rm C} = 1$ is preserved.)

   (b) Find the eigenvalues of the transition matrix Π_1. (**Hint:** You should be able to simply guess the eigenvectors by considering the symmetries of the matrix.) item Find the matrix $\Pi_t = \Pi_1^t$, describing the evolution of probabilities after t generations.

   (c) Show that in steady state (after many duplications), all nucleotides occur with the same frequency. Estimate the number of generations (as a function of p and q) needed to reach such a steady state.

   (d) You should be able to convince yourself that for any model in which mutation rates between pairs of bases are the same in the forward and backward directions, all nucleotides are equally likely in the steady state. However, in the human genome the nucleotides C and G occur less often than A and T. This is partly due to methylation of successive CG pairs which makes them more susceptible to mutations. To mimic this asymmetry, consider an unrealistic model in which transversions from A to C and T to G occur with probability p_+, while the reverse transversions (from C to A or G to T) occur at a lower probability of p_-. (The other transversions occur at rate p, and transitions at rate q as before.) Write the modified transfer matrix corresponding to this model, and obtain the resulting frequencies of nucleotides in steady state.

*****

2. *Activation/deactivation reaction:* Many molecules in biology can be made active or inactive through the addition of a phosphate group. The enzyme that adds the phosphate group is usually termed a kinase, while a phosphatase removes this group. Let us consider a case where a finite number N of such molecules within a cell can be exchanged between the two forms at rates a and b, i.e.

$$\mathrm{A} \rightleftharpoons_b^a \mathrm{B} \quad ,$$

   where we have folded the probabilities to encounter the enzymes in the reaction rates.

(a) Write down the Master equation that governs the evolution of the probabilities $p(N_A = n, N_B = N - n, t)$.

(b) Assuming that initially all molecules are in state A, i.e. $p(n, t = 0) = \delta_{n,N}$, find $p(n, t)$ at all times. You may find it easier to guess the solution, but should then check that it satisfies the equations obtained before.

*****