

Chapter 5

Time dependent probabilities

5.1 Transition rates

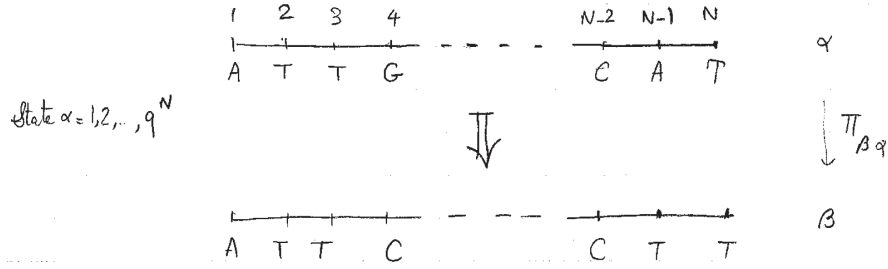
5.1.1 Evolving sequence

As organisms reproduce, the underlying genetic information is passed on to subsequent generations. The copying of the genetic content is not perfect, and leads to a diverse and evolving population of organisms after many generations. The changes are stochastic, and are thus appropriately described by evolving probability distributions. After motivating such evolving probabilities in the contexts of DNA and populations, we introduce the mathematical tools for treating them.

Let us consider the evolution of probabilities in the context of the simplified model of N independently distributed sites. We model mutations by assuming that at subsequent time-steps (generations) each site may change its state (independent of the other sites), say from α to β with a *transition probability* $\pi_{\beta\alpha}$. The $m \times m$ such elements form the *transition probability matrix* π . (Without the assumption that the sites evolve independently, we would have constructed a much larger ($m^N \times m^N$) matrix $\mathbf{\Pi}$. With the assumption of independence, this larger matrix is a direct product of transition matrices for individual sites, i.e. $\mathbf{\Pi} = \pi_1 \otimes \pi_2 \otimes \cdots \otimes \pi_N$, with π_i a $m \times m$ matrix acting on site i .) With the transition probability matrix, we can track the evolution of the probabilities as

$$p_\alpha(\tau + 1) = \sum_{\beta=1}^m \pi_{\alpha\beta} p_\beta(\tau), \quad \text{or in matrix form} \quad \vec{p}(\tau + 1) = \pi \vec{p}(\tau) = \pi^\tau \vec{p}(1), \quad (5.1.1)$$

where the last identity is obtained by recursion, assuming that the transition probability matrix remains the same for all generations, i.e. does not change with time.



Probabilities must be normalized to unity, and thus the transition probabilities are constrained by

$$\sum_{\alpha} \pi_{\alpha\beta} = 1, \quad \text{or} \quad \pi_{\beta\beta} = 1 - \sum_{\alpha \neq \beta} \pi_{\alpha\beta}. \quad (5.1.2)$$

The last expression formalizes the statement that the probability to stay in the same state is the complement of the probabilities to make a change. Using this result, we can rewrite Eq. (5.1.1) as

$$p_{\alpha}(\tau + 1) = p_{\alpha}(\tau) + \sum_{\beta \neq \alpha} [\pi_{\alpha\beta} p_{\beta}(\tau) - \pi_{\beta\alpha} p_{\alpha}(\tau)]. \quad (5.1.3)$$

5.1.2 Steady state

Because of the conservation of probability in Eqs. (5.1.2) and (5.1.13), the transition probability matrix π , and by extension the rate matrix \mathbf{R} have a left-eigenvector $\overleftarrow{v}^* = (1, 1, \dots, 1)$ with eigenvalues of unity and zero respectively, i.e.

$$\overleftarrow{v}^* \pi = \overleftarrow{v}^* \quad , \quad \text{and} \quad \overleftarrow{v}^* \mathbf{R} = 0. \quad (5.1.4)$$

For each eigenvalue of a matrix there is both a left eigenvector and a right eigenvector. The matrices π and \mathbf{R} thus must also have a right-eigenvector \overrightarrow{p}^* such that

$$\pi \overrightarrow{p}^* = \overrightarrow{p}^* \quad , \quad \text{and} \quad \mathbf{R} \overrightarrow{p}^* = 0. \quad (5.1.5)$$

The elements of the vector \overrightarrow{p}^* represent the *steady state probabilities* for the process. These probabilities no longer change with time. In many cases, the Perron–Frobenius theorem ensures that the vector \overrightarrow{p}^* is unique, with positive elements as appropriate to probabilities.¹ Since the matrix π is not symmetric, the remaining eigenvalues need not be real, but must occur in complex conjugate pairs. These remaining eigenvalues of the transition matrix have magnitude less than unity, and determine how an initial vector of probabilities approaches the steady state.

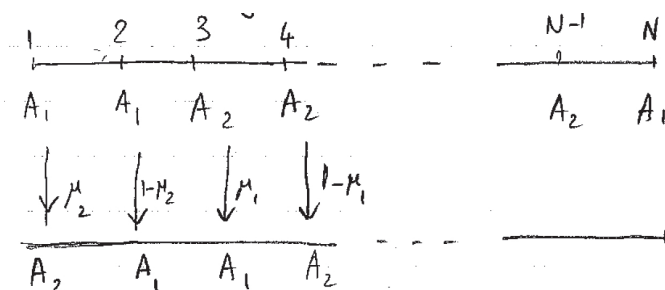
¹A condition for the validity of the above is that it any state should be accessible (through a set of transitions) from any other states. Consequences of violation of this condition will be explored later in connection with genetic drift and absorbing states.

From Eq. (5.1.3) we observe that a particular steady state \vec{p}^* can be achieved through transition probabilities that satisfy the so-called condition of *detailed balance*,

$$\pi_{\alpha\beta}p_{\beta}^* = \pi_{\beta\alpha}p_{\alpha}^*. \quad (5.1.6)$$

This condition holds for systems in thermal equilibrium, where the steady-state probabilities are constrained to satisfy the Boltzmann distribution ($p_{\alpha}^* \propto \exp(-\beta E_{\alpha})$). Equation (5.1.6) requires equal probability fluxes between any pair of states. More general steady states can be formed with the probability flux circulating along triplets and larger sets. Indeed, such circulation of flux can be used as indicator of non-equilibrium steady states.

5.1.3 Evolving binary sequence



As a simple example, consider a *binary* sequence (i.e. $m = 2$) with independent states A_1 or A_2 at each site.² Let us assume that the state A_1 can “mutate” to A_2 at a rate μ_2 , while state A_2 may change to A_1 with a rate μ_1 . The probabilities $p_1(t)$ and $p_2(t)$ now evolve in time as

$$\frac{d}{dt} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} -\mu_2 & \mu_1 \\ \mu_2 & -\mu_1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}. \quad (5.1.7)$$

The above 2×2 transition rate matrix has the following two eigenvectors

$$\begin{pmatrix} -\mu_2 & \mu_1 \\ \mu_2 & -\mu_1 \end{pmatrix} \begin{pmatrix} \frac{\mu_1}{\mu_1 + \mu_2} \\ \frac{\mu_2}{\mu_1 + \mu_2} \end{pmatrix} = 0, \quad \text{and} \quad \begin{pmatrix} -\mu_2 & \mu_1 \\ \mu_2 & -\mu_1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -(\mu_1 + \mu_2) \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (5.1.8)$$

As anticipated, there is an eigenvector \vec{p}^* with eigenvalue of zero; the elements of this vector have been normalized to add to unity, as required for probabilities. We have not normalized the second eigenvector, whose eigenvalue $-(\mu_1 + \mu_2)$ determines the rate of approach to steady state.

²Clearly with the assumption of independence we are really treating independent sites, and the insistence on a sequence may appear frivolous. The advantage of this perspective, however, will become apparent in the next section.

To demonstrate evolution of probabilities with time, let us start with a sequence that is purely A_1 , i.e. with $p_1 = 1$ and $p_2 = 0$ at $t = 0$. The formal solution to the linear differential equation (5.1.7) is

$$\begin{pmatrix} p_1(t) \\ p_2(t) \end{pmatrix} = \exp \left[t \begin{pmatrix} -\mu_2 & \mu_1 \\ \mu_2 & -\mu_1 \end{pmatrix} \right] \begin{pmatrix} p_1(0) \\ p_2(0) \end{pmatrix}. \quad (5.1.9)$$

Decomposing the initial state as a sum over the eigenvectors, and noting the action of the rate matrix on each eigenvector from Eq. (5.1.8), we find

$$\begin{aligned} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} &= \exp \left[t \begin{pmatrix} -\mu_2 & \mu_1 \\ \mu_2 & -\mu_1 \end{pmatrix} \right] \left[\begin{pmatrix} \frac{\mu_1}{\mu_1 + \mu_2} \\ \frac{\mu_2}{\mu_1 + \mu_2} \end{pmatrix} + \frac{\mu_2}{\mu_1 + \mu_2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right] \\ &= \begin{pmatrix} \frac{\mu_1}{\mu_1 + \mu_2} + e^{-(\mu_1 + \mu_2)t} \frac{\mu_2}{\mu_1 + \mu_2} \\ \frac{\mu_2}{\mu_1 + \mu_2} - e^{-(\mu_1 + \mu_2)t} \frac{\mu_2}{\mu_1 + \mu_2} \end{pmatrix}. \end{aligned} \quad (5.1.10)$$

At long times the probabilities to find state A_1 or A_2 are in the ratios μ_1 to μ_2 as dictated by the steady state eigenvector. The rate at which the probabilities converge to this steady state is determined by the second eigenvalue $-(\mu_1 + \mu_2)$.

5.1.4 The Master equation

In many circumstances of interest the probabilities change slowly and continuously over time, in which case we introduce a time interval Δt between subsequent generations, and write

$$\frac{p_\alpha(\tau + 1) - p_\alpha(\tau)}{\Delta t} = \sum_{\beta \neq \alpha} \left[\frac{\pi_{\alpha\beta}}{\Delta t} p_\beta(\tau) - \frac{\pi_{\beta\alpha}}{\Delta t} p_\alpha(\tau) \right]. \quad (5.1.11)$$

In the limit of small Δt , $[p_\alpha(\tau + 1) - p_\alpha(\tau)]/\Delta t \approx dp_\alpha/dt$, while

$$\frac{\pi_{\alpha\beta}}{\Delta t} = R_{\alpha\beta} + \mathcal{O}(\Delta t) \quad \text{for } \alpha \neq \beta, \quad (5.1.12)$$

are the off-diagonal elements of the matrix \mathbf{R} of *transition probability rates*. The diagonal elements of the matrix describe the depletion rate of a particular state, and by conservation of probability must satisfy, as in Eq. (5.1.2),

$$\sum_{\alpha} R_{\alpha\beta} = 0, \quad \text{or} \quad R_{\beta\beta} = - \sum_{\alpha \neq \beta} R_{\alpha\beta}. \quad (5.1.13)$$

We thus arrive at

$$\frac{dp_\alpha(t)}{dt} = \sum_{\beta \neq \alpha} (R_{\alpha\beta} p_\beta(t) - R_{\beta\alpha} p_\alpha(t)) \quad , \quad (5.1.14)$$

which is known as the *Master equation*.

5.1.5 Mutating Population

The previous example of a binary sequence of length N can be recast and interpreted in terms of the evolution of a (haploid) population of fixed size as follows. Let us assume that A_1 and A_2 denote two forms of a particular allele. In each generation any individual is replaced by an offspring that mostly retains its progenitor's allele, but may mutate to the other form at some rate. In this model the total population size is fixed to N , while the sub-populations N_1 and N_2 may vary. A particular state of the population is thus described by $N_1 = n$ and $N_2 = N - n$, and since $n = 0, 1, \dots, N$ there are $N + 1$ possible states. At a particular time, the system may be in any one of these states with probability $p(n, t)$, and we would like to follow the evolution of these probabilities.

After an individual replication event (A_1 to A_1 at rate $-\mu_2$, A_1 to A_2 at rate μ_2 , A_2 to A_1 at rate μ_1 , or A_2 to A_2 at rate $-\mu_1$), the number N either stays the same, or changes by unity. Thus the transition rate matrix only has non-zero terms along or adjoining to the diagonal. For example

$$R_{n,n+1} = \mu_2(n + 1), \quad \text{and} \quad R_{n,n-1} = \mu_1(N - n + 1), \quad (5.1.15)$$

where the former indicates that a population of $n + 1$ A_1 s can decrease by unity if any one of them mutates to A_2 , while a population with $n - 1$ A_1 s increases by unity if any of A_2 s mutates to A_1 . The diagonal terms are obtained from the normalization condition in Eq. (5.1.13) resulting in the Master equation

$$\frac{dp(n, t)}{dt} = \mu_2(n + 1)p(n + 1) + \mu_1(N - n + 1)p(n - 1) - \mu_2np(n) - \mu_1(N - n)p(n), \quad (5.1.16)$$

for $0 < n < N$, and with boundary terms

$$\frac{dp(0, t)}{dt} = \mu_2p(1) - \mu_1Np(0), \quad \text{and} \quad \frac{dp(N, t)}{dt} = \mu_1p(N - 1) - \mu_2Np(N). \quad (5.1.17)$$

The above equation looks complicated, but in fact admits a simple solution as justified in the different context described next.