Computing coalescent intensities for an island model with varying population sizes and migration rates

Shyam Gopalakrishnan

February 21, 2012

1 Island model with K demes

Consider two lineages in an island model with K demes. Let N_i be the population size for the i-th deme and m_{ij} be the migration rate from deme i to deme j. Additionally, let the migration rates be symmetric, i.e., $m_{ij} = m_{ji}$.

2 Computing coalescent intensities

Let P_{ij}^t be the probability that two lines, one currently residing in deme i and the other in deme j, coalesce within t generations. For two lines sampled from the same deme i, this probability is written as P_{ii}^t.
Using a first step analysis, we can set up a regression to compute these coalescent probabilities.

$$\begin{aligned}
P_{ii}^t &= 1 \cdot P(C_i) + \sum_{k=1,k\neq i}^{K} P(i \to k) P_{ik}^{t-1} + \\
&\quad (1 - P(C_i) - \sum_{k=1,k\neq i}^{K} P(i \to k)) P_{ii}^{t-1} \qquad (1) \\
P_{ij}^t &= P(i \to j) P_{jj}^{t-1} + P(j \to i) P_{ii}^{t-1} + \sum_{k=1,k\neq i,j}^{K} (P(i \to k) P_{kj}^{t-1} + P(j \to k) P_{ik}^{t-1}) \\
&\quad + (1 - \sum_{k=1,k\neq i,j}^{K} (P(i \to k) + P(j \to k)) - P(i \to j) - P(j \to i)) P_{ij}^{t-1} \qquad (2)
\end{aligned}$$

Here, $i \to j$ signifies a migration from deme i to j and C_i signifies a coalescent event between the two lines when both reside in deme i. The migration and coalescent probabilities are given below.

$$
\begin{aligned}
P(C_i) &= \frac{1}{2N_i} &(3)\\
P(i \to j) &= n_i m_{ij} &(4)
\end{aligned}
$$

Noting that $n_i = 2$ when both lines are in the same deme and $n_i = n_j = 1$ when one line resides in demes i and j each, and substituting the above values into the recursion equations we get,

$$
\begin{aligned}
P_{ii}^t &= \frac{1}{2N_i} + 2\sum_{k=1,k\neq i}^{K} m_{ik}P_{ik}^{t-1} + (1 - \frac{1}{2N_i} - 2\sum_{k=1,k\neq i}^{K} m_{ik})P_{ii}^{t-1} &(5)\\
P_{ij}^t &= m_{ij}(P_{jj}^{t-1} + P_{ii}^{t-1}) + \sum_{k=1,k\neq i,j}^{K} (m_{ik}P_{kj}^{t-1} + m_{jk}P_{ik}^{t-1}) + \\
&\quad (1 - 2m_{ij} - \sum_{k=1,k\neq i,j}^{K} (m_{ik} + m_{jk}))P_{ij}^{t-1} &(6)
\end{aligned}
$$

The above probabilities can be computed, all at the same time, using a dynamic programming approach. The boundary conditions for these recursion equations are given at t=0, when $P_{ii}^0 = P_{ij}^0 = 0$.
We can modify these equations to pose this recursion as a matrix recursion.

$$
P^t = C + MP^{t-1} + (MP^{t-1})^T = C + MP^{t-1} + P^{t-1}M \tag{7}
$$

Here C is a diagonal matrix with $C_{ii} = 1/2N_i$, M is the *symmetric* migration matrix with diagonal elements $M_{ii} = 0.5 - \sum_{j=1,j\neq i}^{K} M_{ij}$ and P^{t-1} is the coalescent intensity matrix at $t-1$ generations. Adding a couple of steps to the previous equation, we get the two and three step recursion as given below.

$$
\begin{aligned}
P^t &= C + MC + M^2P^{t-2} + MP^{t-2}M + CM + MP^{t-2}M + P^{t-2}M^2\\
&= C + MC + CM + M^2P^{t-2} + 2MP^{t-2}M + P^{t-2}M^2 &(8)\\
P^t &= C + MC + CM + M^2C + M^3P^{t-3} + M^2P^{t-3}M + 2MCM + \\
&\quad 2M^2PM + 2MPM^2 + CM^2 + MP^{t-3}M^2 + P^{t-3}M^3\\
&= C + MC + CM + M^2C + 2MCM + CM^2 + \\
&\quad M^3P^{t-3} + 3M^2P^{t-3}M + 3MP^{t-3}M^2 + P^{t-3}M^3 &(9)
\end{aligned}
$$