4 Duality

Duality can often be interpreted as a one-to-one correspondence between two 'objects'. Each bijection H between two sets A and B is a duality between A and B. We also say A and B are dual w.r.t. H. There exist more involved dualities. The uniqueness theorem for Fourier transforms is for example a duality between the probability measures and their Fourier transforms. We will now learn about a duality for Markov chains. The following Definition is taken from Liggett's book 'Interacting Particle Systems'.

4.1 Def. (Duality)
Two Markov chains $(X_n)_{n\in\mathbb{N}_0}$ with state space S_1 and $(Y_n)_{n\in\mathbb{N}_0}$ with state space S_2 are called dual w.r.t. a bounded function $H : S_1 \times S_2 \to \mathbb{R}$, if

$$\mathrm{E}(H(X_n, y)\,|\,X_0 = x) \;=\; \mathrm{E}(H(x, Y_n)\,|\,Y_0 = y) \qquad \forall\, n \in \mathbb{N}_0, x \in S_1, y \in S_2.$$

Rem.
For discrete time HMCs with the same finite state space S one can view $H : S^2 \to \mathbb{R}$ as a matrix with entries $h_{ij} := H(i, j)$. Duality then means

$$\sum_{j\in S} \pi_{ij}^{(n)} h_{jk} \;=\; \mathrm{E}(H(X_n, k)\,|\,X_0 = i) \;=\; \mathrm{E}(H(i, Y_n)\; Y_0 = k) \;=\; \sum_{j\in S} h_{ij} p_{kj}^{(n)}.$$

Thus, $\Pi^n H = H(P^n)^\top$ $\forall\, n \in \mathbb{N}_0$, where $\Pi = (\pi_{ij})_{i,j\in S}$ and $P = (p_{ij})_{i,j\in S}$ are the corresponding transition matrices. If the matrix H is non-singular, then P can be computed from Π and vice versa, and H is a bijection.

4.2 Theorem (Duality for Cannings Models, Sampling Duality)
There exists a non-singular, left lower, triangular matrix $H = (h_{ij})_{i,j\in\{0,\ldots,N\}}$ such that, for every Cannings model, $\Pi H = HP^\top$. The matrix H is, for example, given by

$$h_{ij} \;:=\; \frac{\binom{i}{j}}{\binom{N}{j}}, \qquad i, j \in \{0, \ldots, N\}.$$

Rem.
The matrix H can be interpreted probabilistically as follows. Sample randomly (without replacement) j balls from a box containing i black and $N-i$ white balls. Then, h_{ij} is the probability that all j sampled balls are black (hypergeometric distribution). For a numerical calculation of h_{ij} the following formula is useful.

$$h_{ij} \;=\; \prod_{k=0}^{j-1} \frac{i-k}{N-k}\,.$$

Cannings (1974) and Gladstien (1976, 1977, 1978) verified Theorem 4.2 essentially by direct computations of ΠH and $HP^\top$ using the formulas for the transition probabilities π_{ij} and p_{ij} of the forward and the backward process respectively. The following probabilistic proof (M., 1999) does not use these formulas and instead exploits the so-called 'random assignment condition' of Cannings models.

Proof.
Clearly, H is left lower triangular with $\det H = \prod_{i=0}^N h_{ii} = \prod_{i=0}^N 1/\binom{N}{i} \neq 0$. Fix $i,k \in \{0,\ldots,N\}$ and $n,r \in \mathbb{N}_0$ with $n \leq r$. Let E denote the event that in gen. n there exist exactly k ancestors, i.e. $R_{r-n}^{(r)} = k$, and that all these k individuals are a descendant of the individuals $1,\ldots,i$ of gen. 0. The probability of E can be expressed in two ways.

$$E \;=\; \bigcup_{j=0}^N \{R_{r-n}^{(r)} = k \text{ and } X_n^{(i)} = j, \text{and each of these } k \text{ ancestors is one of the } j \text{ descendants}\}$$

and, hence, (under the assumptions of the model)

$$P(E) \;=\; \sum_{j=0}^N P(R_{r-n}^{(r)} = k)\, P(X_n^{(i)} = j) \frac{\binom{j}{k}}{\binom{N}{k}} \;=\; P(R_{r-n}^{(r)} = k) \sum_{j=0}^N \pi_{ij}^{(n)}\, h_{jk}.$$

On the other hand,

$$E \;=\; \bigcup_{j=0}^N \{R_{r-n}^{(r)} = k \text{ and } R_r^{(r)} = j, \text{and each of these } j \text{ ancestors is one of the individuals } 1,\ldots,i\}$$

and therefore

$$P(E) \;=\; \sum_{j=0}^N \frac{\binom{i}{j}}{\binom{N}{j}} P(R_r^{(r)} = j, R_{r-n}^{(r)} = k) \;=\; \sum_{j=0}^N h_{ij}\, P(R_r^{(r)} = j, R_{r-n}^{(r)} = k).$$

The resulting equality divided[3] by $P(R_{r-n}^{(r)} = k)$ yields $\sum_{j=0}^N \pi_{ij}^{(n)} h_{jk} = \sum_{j=0}^N h_{ij} p_{kj}^{(n)}$. Thus, $\Pi^n H = H(P^n)^\top$. □

Rem.
Note that H does not depend on the particular Cannings model. The inverse H^{-1} is a left lower triangular matrix with integer entries

$$(H^{-1})_{ij} \;=\; (-1)^{i-j}\binom{i}{j}\binom{N}{i}.$$

[3]One can achieve $P(R_{r-n}^{(r)} = k) \neq 0$ (even $P(R_{r-n}^{(r)} = k) = 1$) by considering the case $r := n$ and the ancestral process starting at the state k, so $P(R_0^{(n)} = k) = 1$.

4.3 Corollary (Eigenvalue Theorem, Cannings, 1974)
The transition matrix Π *of the forward process has the eigenvalues* $\lambda_i = \mathrm{E}(\nu_1 \cdots \nu_i)$, $i \in \{0, \ldots, N\}$.

Cannings' original proof from 1974 uses an expansion argument leading to a triangular structure. The following proof uses duality.

Proof.
H non-singular. $\Rightarrow$ Π has the same eigenvalues as $P^\top$. P is a left lower triangular matrix. $\Rightarrow$ The eigenvalues are $\lambda_i = p_{ii} = \mathrm{E}(\nu_1 \cdots \nu_i)$ by Lemma 3.3.2. □

4.4 Example
By Example 3.3.3, for the WFM, $\lambda_i = p_{ii} = S(i,i)(N)_i N^{-i} = (N)_i N^{-i}$, $i \in \{0, \ldots, N\}$.

Rem.

1. Let λ be an eigenvalue of $P^\top$ (and hence also of Π) and x a corresponding right eigenvector. $\Rightarrow$ $\Pi H x = H P^\top x = H \lambda x = \lambda H x$. $\Rightarrow$ Hx is a right eigenvector of Π to the eigenvalue λ. The map $x \mapsto Hx$ is hence an isomorphism between the two eigenspaces to λ of $P^\top$ and Π respectively. In particular, these eigenspaces have the same dimension. Another interesting implication of the duality is, that Π is diagonalisable if and only if $P^\top$ (and hence also P) is diagonalisable.

2. There exists not only one matrix H satisfying $\Pi H = H P^\top$. Of some interest is hence the subspace

$$U \;:=\; \{H \,|\, \Pi H = H P^\top\} \;=\; \{H \,|\, \Pi^n H = H (P^n)^\top \; \forall\, n \in \mathbb{N}\}.$$

   Typical questions: Basis and dimension of U? Which $H \in U$ are non-singular?

   It can be shown (proof not provided here) that $\dim U = N+3$, if the non-unit eigenvalues of Π (or $P^\top$) are all real and pairwise distinct. This is for example the case for the MM and WFM.

3. Other matrices H are known for particular Cannings models. For example, for the WFM, $\Pi H = H P^\top$ for $h_{ij} := (i/N)^j$ or for $h_{ij} := (1 - i/N)^j$.

   The entries of these matrices can be interpreted via urn models of the type 'sampling with replacement'.

   The algebraic relation $\Pi^n H = H(P^n)^\top$, that is, $\sum_{j=0}^N \pi_{ij}^{(n)} h_{jk} = \sum_{j=0}^N h_{ij} p_{kj}^{(n)}$, can be rewritten probabilistically in terms of expectations. For instance, if $h_{ij} = (i/N)^j$ (WFM), then

$$\mathrm{E}\left(\left(\frac{X_n^{(i)}}{N}\right)^k\right) \;=\; \mathrm{E}\left(\left(\frac{i}{N}\right)^{R_n} \,\middle|\, R_0 = k\right)$$

or short,

$$\mathrm{E}^i((\tfrac{X_n}{N})^k) \;=\; \mathrm{E}^k((\tfrac{i}{N})^{R_n}).$$

The left-hand side is the k-th moment of the frequency of descendants in gen. n. The right-hand side is the pgf of R_n, conditional on $R_0 = k$, evaluated at i/N. Relations of this form are also known for the K-allele model, where not only two types (descendant or no descendant), but $K \in \mathbb{N}$ types are distinguished.