MATH 448 — Lecture 13

§4.4 (continued): Powers of a Markov matrix and the Kimura model

Terms

- A *Jukes–Cantor matrix* is a matrix $M(\alpha)$ whose diagonal entries are all $1-\alpha$ and whose other entries are all $\alpha/3$. Bases are in the order A, G, C, T. Put $\mu := 1 - \frac{4}{3}\alpha$.
- A *Kimura 2-parameter matrix* is a matrix $M(\beta,\gamma)$ in which each transition (A↔G, C↔T) has probability β, each of the two transversions from a base has probability γ, and the diagonal entries are $1-\beta-2\gamma$.

Rules

- The entry in row i, column j of M^t is $P(S_t = i \mid S_0 = j)$, and $\mathbf{p}_t = M^t\mathbf{p}_0$.
- Let I be the identity matrix and Q the matrix with every entry $\frac{1}{4}$. Then $M(\alpha) = \mu I + (1-\mu)Q$: the diagonal entries of the right side are $\mu + \frac{1}{4}(1-\mu) = 1-\alpha$ and the others are $\frac{1}{4}(1-\mu) = \alpha/3$. Since $Q^2 = Q$ and $Q\mathbf{v} = \frac{1}{4}(v_1+v_2+v_3+v_4)(1,1,1,1)^{\mathsf{T}}$, $M(\alpha)$ has eigenvalue 1 with eigenvector $(1,1,1,1)$ and eigenvalue μ with eigenvectors all $\mathbf{v}$ whose entries sum to 0, such as $(1,-1,0,0)$, $(1,0,-1,0)$, $(1,0,0,-1)$.
- $M(\alpha_1)M(\alpha_2) = \mu_1\mu_2 I + (1-\mu_1\mu_2)Q = M(\alpha_3)$, where $\mu_3 = \mu_1\mu_2$, by expanding the product and using $Q^2 = Q$.
- Repeating, $M(\alpha)^t = \mu^t I + (1-\mu^t)Q = M(\alpha_t)$ with $\mu_t = \mu^t$, that is, $\alpha_t = \frac{3}{4}(1-\mu^t)$. So $M(\alpha)^t$ is a Jukes–Cantor matrix with a different parameter. Its diagonal entries are $\frac{1}{4} + \frac{3}{4}\mu^t$ and its other entries are $\frac{1}{4} - \frac{1}{4}\mu^t$.

Problem 13.1. Let $M = M(.03)$, with .97 on the diagonal and .01 in every other position. (a) Show that $(1,-1,0,0)$ is an eigenvector of M and find its eigenvalue. (b) Find the $(1,1)$ entry of M^2 directly, and then from the formula $\frac{1}{4} + \frac{3}{4}(1-\frac{4}{3}\alpha)^2$.

Solution. (a) $M(1,-1,0,0)^{\mathsf{T}} = (.97-.01,\ .01-.97,\ .01-.01,\ .01-.01) = .96\,(1,-1,0,0)$, so the eigenvalue is $.96 = 1-\frac{4}{3}(.03)$. (b) The $(1,1)$ entry of M^2 is the first row times the first column: $.97^2 + 3(.01)^2 = .9409 + .0003 = .9412$. The formula gives $\frac{1}{4} + \frac{3}{4}(.96)^2 = .25 + .6912 = .9412$.

Now do Exercise 13.1 from Homework 8.

Problem 13.2. Let $\alpha = .003$ in the Jukes–Cantor model. A site has base A at time 0. (a) Find the probability that it has base A at time 100, and the probability that it has base T at time 100. (b) What happens to the entries of M^t as t grows, and what does that say about a site?

Solution. (a) Here $1 - \frac{4}{3}\alpha = .996$ and $.996^{100} \approx .6698$. The probability of A is $\frac{1}{4} + \frac{3}{4}(.6698) \approx .752$, and the probability of T is $\frac{1}{4} - \frac{1}{4}(.6698) \approx .083$. (Check: $.752 + 3(.083) \approx 1$.) (b) Since $|1 - \frac{4}{3}\alpha| < 1$, $(1-\frac{4}{3}\alpha)^t \to 0$, so every entry of M^t approaches $\frac{1}{4}$. Every column becomes the equilibrium $(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4})$: after enough time, the base at a site does not depend on the base it started with.

Now do Exercise 13.2 from Homework 8.

Problem 13.3. Two Jukes–Cantor processes act in succession, with $\alpha_1 = .02$ and $\alpha_2 = .03$. (a) Find α_3 so that $M(\alpha_1)M(\alpha_2) = M(\alpha_3)$. (b) Check $1-\alpha_3$ against the $(1,1)$ entry of the product, computed directly.

Solution. (a) $1 - \frac{4}{3}\alpha_3 = (1-\frac{4}{3}(.02))(1-\frac{4}{3}(.03)) = (.97333)(.96) = .9344$, so $\alpha_3 = \frac{3}{4}(1-.9344) = .0492$. This is $\alpha_1 + \alpha_2 - \frac{4}{3}\alpha_1\alpha_2 = .05 - .0008$. (b) The $(1,1)$ entry is $(1-\alpha_1)(1-\alpha_2) + 3\cdot\frac{\alpha_1}{3}\cdot\frac{\alpha_2}{3} = (.98)(.97) + \frac{(.02)(.03)}{3} = .9506 + .0002 = .9508$, and $1-\alpha_3 = 1 - .0492 = .9508$.

Problem 13.4. Take the Kimura 2-parameter matrix with $\beta = .06$ and $\gamma = .01$:

$$M = \begin{pmatrix} .92 & .06 & .01 & .01 \\ .06 & .92 & .01 & .01 \\ .01 & .01 & .92 & .06 \\ .01 & .01 & .06 & .92 \end{pmatrix}.$$

(a) Check that each column sums to 1 and that $\mathbf{q} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ is an equilibrium. (b) Show that $(1, 1, -1, -1)$ is an eigenvector and find its eigenvalue. (c) Show that $(1, -1, 0, 0)$ is an eigenvector and find its eigenvalue. (d) For which β and γ is M a Jukes–Cantor matrix, and what is α then?

Solution. (a) .92+.06+.01+.01 = 1. Each row sums to .92+.06+.01+.01 = 1, so $M\mathbf{q} = \mathbf{q}$. (b) $M(1, 1, -1, -1)^{\mathsf{T}} = (.92 + .06 - .01 - .01,\ \dots) = (.96,\ .96,\ -.96,\ -.96)$, so the eigenvalue is $.96 = 1 - 4\gamma$. (c) $M(1, -1, 0, 0)^{\mathsf{T}} = (.92 - .06,\ .06 - .92,\ 0,\ 0) = .86\,(1, -1, 0, 0)$, so the eigenvalue is $.86 = 1 - 2\beta - 2\gamma$. (d) When $\beta = \gamma$, every off-diagonal entry is β and the diagonal entries are $1 - 3\beta$, so $M = M(\alpha)$ with $\alpha = 3\beta$.

Now do Exercise 13.3 from Homework 8.