Question: How can you modify this for $K=2,3,\ldots$ time steps?

S_0: $\vec{p}_0$

M

S_1 $\vec{p}_1$

S_2 $\vec{p}_2$

S_3 p_3

$\vdots$

S_K

$\Delta t = 1$ time step

$\vec{p}_1 = \vec{p}_0 M$

$\vec{p}_2 = \vec{p}_1 M = \vec{p}_0 M^2$

$\vec{p}_3 = \vec{p}_0 M^3$

$\vec{p}_K = \vec{p}_0 M^K$

S_0

S_K

The formulation thus far is for is for "DISCRETE TIME"

or modelling evolution from tail to tip of an edge.

Alternatively, there is the CONTINUOUS TIME formulation

$S_0 = S(t=0)$

$S_K = S(t=t_K)$

In this setup, we introduce a rate matrix Q

rates = derivatives!

$$Q = \begin{pmatrix} q_{RR} & q_{RY} \\ q_{YR} & q_{YY} \end{pmatrix}$$

the off-diagonal entries are non-negative $q_{YR}, q_{RY} \geq 0$ and

row sums equal 0

q_{RY} = rate at which Rs are converted to Ys ≥ 0

q_{YR} = " " " Ys are converted to Rs ≥ 0

units

$\dfrac{\text{substitution per site}}{\text{unit time}}$

$q_{RR} \leq 0$ rate at which leaving R state

$q_{YY} \leq 0$

Importantly, $q_{RR} + q_{RY} = 0$ rates balance

added to Y

$\approx$ lost to R class

The root distribution is now a function of time

$$\vec{p}_0 = \vec{p}(t=0) = \left(p_R(0), p_Y(0)\right) \quad \text{and at time } t$$

$$\vec{p}_t = \left(p_R(t), \quad p_Y(t)\right)$$

distribution of purines, pyrimidines at time $t \geq 0$.

The distribution of states satisfies the following system of differential equations:

$$(*) \quad \begin{aligned} \frac{d}{dt} p_R(t) &= p_R(t)\, q_{RR} + p_Y(t)\, q_{YR} \\ \frac{d}{dt} p_Y(t) &= \underbrace{p_R(t)\, q_{RY}} + \underbrace{p_Y(t)\, q_{YY}} \end{aligned}$$

$$\begin{pmatrix}\text{current freq.}\\ \text{of } R\end{pmatrix}\begin{pmatrix}\text{rate of}\\ \text{conversion } R \to Y\end{pmatrix} + \begin{pmatrix}\text{current freq}\\ \text{of } Y\end{pmatrix}\begin{pmatrix}\text{rate of} & \text{loss}\\ & Y\end{pmatrix}$$

In matrix form, the right hand side is:

$$\vec{p}(t)\,Q$$

Check:

$$\left(p_R(t) \quad p_Y(t)\right)\begin{pmatrix} q_{RR} & q_{RY} \\ q_{YR} & q_{YY}\end{pmatrix}$$

Thus, $(*)$ is the differential equation

$$\vec{p}\,'(t) = p(t)\,Q \qquad \text{with} \qquad \vec{p}(0) = \left(p_R(0), \; p_Y(0)\right)$$

Using matrix exponentials, the solution is

$$\vec{p}(t) = \vec{p}(0)\, e^{Qt}$$

with $$e^{Qt} = 1 + (Qt) + \frac{(Qt)^2}{2!} + \frac{(Qt)^3}{3!} + \frac{(Qt)^4}{4} + \cdots$$

$$\vec{p}(t) = \vec{p}_0 e^{Qt} = \vec{p}_0 M(t)$$

distribution of states at time t

initial state substitution

product

Markov model at time t $\quad M(t) = e^{Qt}$

Demo: $\quad t=0 \quad \vec{p}_0 = (.7, .3) \qquad Q = \begin{pmatrix} -.1 & .1 \\ .2 & -.2 \end{pmatrix}$

$t \quad M(t) = e^{Qt}$

Note:

- e^{Qt} is a Markov matrix
- How should the diagonal entries of $M(1)$ and $M(2)$ compare?

  $t=1 \qquad t=2$

- MATH: computing matrix exponentials:

  a) Q is diagonalizable $\qquad Q = S \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_3 \end{pmatrix} S^{-1}$

  b) $M = e^{Qt} = I + (Qt) + \dfrac{(Qt)^2}{2!} + \dfrac{(Qt)^3}{3!} + \ldots$

$$= I + S\Lambda t S^{-1} + S\frac{(\Lambda t)^2}{2!}S^{-1} + S\frac{(\Lambda t)^3}{3!}S^{-1} + \ldots$$

$$\Lambda t = \begin{pmatrix} \lambda_1 t & & 0 \\ & \lambda_2 t & \\ 0 & & \lambda_3 t \end{pmatrix}$$

$$= S\left[I + \Lambda t + \frac{(\Lambda t)^2}{2!} + \frac{(\Lambda t)^3}{3!} + \frac{(\Lambda t)^4}{4} + \ldots\right] S^{-1}$$

$$= S e^{\Lambda t} S^{-1} = S \begin{pmatrix} 1 & & & 0 \\ & e^{-\lambda_1 t} & & \\ & & e^{-\lambda_2 t} & \\ 0 & & & e^{-\lambda_3 t} \end{pmatrix} S^{-1}$$

i.e. diagonalize Q, exponentiate diagonal entries of Λ, ...

Summary: Continuous-time formulation

Initial base distribution $\vec{p}_0 = (p_R(0), p_Y(0))$ $t=0$.

At any time $t>0$, the Markov transition matrix is given by

$M(t) = e^{Qt}$ for a fixed rate matrix Q.