Question: *Let X be the Markov chain on $\{0,1\}$ with $\mathbb{P}(X_1 = 1|X_0 = 0) = a$ and $\mathbb{P}(X_1 = 0|X_0 = 1) = b$. Find the stationary distribution, and an explicit expression for the total variation distance between the distribution of X_n and the stationary distribution as a function of n and the starting point.*

Solution: The characterstic equation for the transition matrix is $(1-a-\lambda)(1-b-\lambda)-ab = 0$ and hence the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 1-(a+b)$. Suppose for now that $a+b \neq 0$ and $a+b \neq 2$. Then we must have $0 < |1-(a+b)| < 1$ since all eigenvalues have magnitude less than 1.

The eigenvectors are $(1,1)$ and $(a,-b)$. Thus

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} = \begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1-(a+b) \end{pmatrix} \begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix}^{-1}$$

Noting that

$$\begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix}^{-1} = \begin{pmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{1}{a+b} & -\frac{1}{a+b} \end{pmatrix}$$

We get

$$P^n = \begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix} \begin{pmatrix} 1^n & 0 \\ 0 & [1-(a+b)]^n \end{pmatrix} \begin{pmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{1}{a+b} & -\frac{1}{a+b} \end{pmatrix}$$

After multiplying the matrices out, we find that

$$P^n = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}^n = \frac{1}{a+b} \begin{pmatrix} b + a[1-(a+b)]^n & a - a[1-(a+b)]^n \\ b - b[1-(a+b)]^n & a + b[1-(a+b)]^n \end{pmatrix}$$

Thus

$$\lim_{n\to\infty} P^n = \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix}$$

and hence $\pi = (\frac{b}{a+b}, \frac{a}{a+b})$. We can now find the total variation distance.

$$d_{TV}(P_0^n, \pi) = \sum_{k=0,1} |p_{0k}(n) - \pi_k| = 2\frac{a}{a+b}|1-(a+b)|^n$$

$$d_{TV}(P_1^n, \pi) = \sum_{k=0,1} |p_{1k}(n) - \pi_k| = 2\frac{b}{a+b}|1-(a+b)|^n$$

If $a+b = 0$, then since $a, b \geq 0$ we have $a = b = 0$, and hence the chain is completely stationary and thus any initial distribution is a stationary distribution. If $a+b = 2$, then since $a, b \leq 1$ we have $a = b = 1$ and hence the chain is aperiodic, with period 2, so it does not converge to a stationary distribution.

Question: *Take a graph G and a pallet of k colors, and call any assignment of colors to the nodes of the graph such that no two adjacent nodes have the same color a coloring of the graph. Describe an MCMC algorithm that will sample (approximately) uniformly chosen colorings of the graph, as long as k is large enough, and discuss possible problems that may occur for smaller k.*

Solution: One possible algorithm is to choose some initial assignment of colors that satisfies the requirement that no two adjacent nodes have the same color. Such a coloring may not be possible. For example, if G is a planar graph, i.e. if no two edges cross each other, then if k is $1, 2$, or 3 then it may not be possible to find such an assignment of colors, by the *map coloring theorem* (see e.g. wikipedia). The algorithm proceeds in the following way. At each stage, select a node uniformly at random, and a uniformly random color from the k colors. Propose changing the chosen node to the chosen color. If this results in another assignment of colors such that no two adjacent nodes have the same color, accept the proposition, otherwise, reject the proposition and do not change any colors.

This is suitable because our proposition probabilities $h(x, y)$ are clearly symmetric (the probabillity that a node of color a is chosen and made into color b is the same as the probability that the same node, when it has color b, is chosen and turned into color a). Also, since we are sampling from the uniform distribution, $\pi(x) = \frac{1}{|S|}$ for every $x \in S$, where S is the set of assignment of colors such that no two adjacent nodes have the same color, i.e. the set of 'colorings'. Thus our acceptance probablity, if $x, y \in S$, is givn by

$$a(x, y) = \min\left\{1, \frac{\pi(y)h(y, x)}{\pi(x)h(x, y)}\right\} = \min\{1, 1\} = 1$$

which is exactly the scheme described above.

This will work as long as the chain is irreducible, i.e. as long as it is possible to get to any coloring using this system, which it may not always be if k is too small. For example, if a node is surrouned by $k - 1$ other nodes, each of a different color, then we may not change the color of that node, or in fact any of its neighbours, and we are stuck! One possible resolution to such problems is to try changing the colors of multiple nodes at once. In fact, it can be shown that the chain is always irreducible under the changing one at a time system if we have $k \geq d + 1$, where d is the maximum degree of any vertex of G.

Another important thing is that the node to be changed is chosen randomly - otherwise there are examples where the chain stops being irreducible.

Note finally that in order to avoid bias towards states with more "close by" colorings, we must not only propose colorings, but also unnaceptable arrangements of colors as well, and then stay where we are if we propose an unnaceptable arrangement.