An Application of Markov Chain Monte Carlo to Community Ecology

Sophia & Cathy

The ecologists' question: How species distribute themselves across different islands?

Co-occurrence matrix: helps ecologists understand how species distribute themselves across different islands. This distribution can reveal patterns of colonization, extinction, and competition among species.

	Islands																	
Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	total
A	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17
B	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	14
C	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	14
D	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	13
E	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	12
F	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	11
G	1	1	1	1	1	0	1	1	0	0	0	0	0	1	0	1	1	10
H	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	0	0	10
I	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	10
J	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	6
K	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	2
L	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
M	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
total	11	10	10	10	10	9	9	9	8	8	7	4	4	4	3	3	3	122

Figure 1. The co-occurrence matrix of the finches on the Galapagos islands.

A binary matrix

1 ⇒ presence of species

0 ⇒ absence of species

Row ⇒ a different species of finch

Column ⇒ a different island in the Galápagos.

Checkerboards

Species	Islands 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	total
A	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17
B	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	14
C	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	14
D	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	13
E	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	12
F	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	11
G	1	1	1	1	1	0	1	1	0	0	0	0	0	1	0	1	1	10
H	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	0	0	10
I	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	10
J	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	6
K	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	2
L	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
M	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
total	11	10	10	10	10	9	9	9	8	8	7	4	4	4	3	3	3	122

Figure 1. The co-occurrence matrix of the finches on the Galapagos islands.

Checkerboards:

A pair of species that never occur together on an island. ⇒ A pair of rows with product zero.

Checkerboard Unit: 2 x 2 matrix of the form

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad or \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Ex: species J and K, and islands 4 and 14.

Larger total number of checkerboard units ⇒ Higher degree of competition

Ex: 10 checkerboards, 333 checkerboard units, out of $\binom{13}{2}\binom{17}{2} = 10608$

The ecologists' question: how likely is it that a random 13 × 17 matrix of 0s and 1s will have at least ten checkerboards?

Intuition: list all the matrices, then find the proportion of having >10 checkerboards.

Unavailable: there are more than 6.71×10^{16}

Solution: generate random matrices uniformly ⇒ assign equal probabilities to all the matrices, then utilize ***the law of large numbers*** to estimate the probability.

$$\hat{p} = \frac{\#\text{ random matrices with requisite level of competition}}{\text{total }\#\text{ random matrices generated}}$$

Markov Chain Monte Carlo

Goal: generate random samples from a probability distribution where direct sampling is difficult.

Markov Chain Property: reach a desired distribution as the number of steps increases.

In this paper:

- Use MCMC to generate random matrices with fixed row and column totals.
- Help scientists figure out if the distribution of species happens by chance or is influenced by specific ecological rules and interactions.

Markov Chain Monte Carlo Simulation

Example 5.1 Bob's daily lunch choices at the cafeteria are described by a Markov chain with transition matrix

$$P = \begin{array}{c} \\ \text{Yogurt} \\ \text{Salad} \\ \text{Hamburger} \\ \text{Pizza} \end{array} \begin{array}{cccc} \text{Yogurt} & \text{Salad} & \text{Hamburger} & \text{Pizza} \end{array} \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 0 & 1/4 & 1/2 \\ 1/4 & 0 & 1/4 & 1/2 \end{pmatrix}.$$

Yogurt costs \$3.00, hamburgers cost \$7.00, and salad and pizza cost \$4.00 each. Over the long term, how much, on average, does Bob spend for lunch?

Solution Let

$$r(x) = \begin{cases} 3, & \text{if } x = \text{yogurt}, \\ 4, & \text{if } x = \text{salad or pizza}, \\ 7, & \text{if } x = \text{hamburger}. \end{cases}$$

The lunch chain is ergodic with stationary distribution

Yogurt	Salad	Hamburger	Pizza
7/65	2/13	18/65	6/13

1. Constructs an ergodic Markov chain whose limiting distribution is π.
2. Run the chain long enough for the chain to converge to its limiting distribution.
3. **Strong Law of Large Number**: limiting properties of ergodic Markov chains have some similarities to i.i.d. Sequences. Let's see.

Strong Law of Large Number

If X_1, X_2, ... is an i.i.d. sequence with common mean $\mu < \infty$, then the strong law of large numbers says that, with probability 1,

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \mu.$$

Equivalently, let Y be a random variable with the same distribution as the Y_i and assume that r is a bounded, real-valued function. Then, $r(Y_1)$, $r(Y_2)$, ... is also an i.i.d. sequence with finite mean, and, with probability 1,

$$\lim_{n\to\infty} \frac{r(Y_1) + \cdots + r(Y_n)}{n} = E(r(Y)).$$

Strong Law of Large Number for Markov Chains

Assume that $X_0, X_1, \ldots$ is an ergodic Markov chain with stationary distribution π. Let r be a bounded, real-valued function. Let X be a random variable with distribution π. Then, with probability 1,

$$\lim_{n\to\infty} \frac{r(X_1) + \cdots + r(X_n)}{n} = E(r(X)),$$

where $E(r(X)) = \sum_j r(j)\pi_j$.

Back to our example.

Markov Chain Monte Carlo Simulation

Solution Let

$$r(x) = \begin{cases} 3, & \text{if } x = \text{yogurt}, \\ 4, & \text{if } x = \text{salad or pizza}, \\ 7, & \text{if } x = \text{hamburger}. \end{cases}$$

The lunch chain is ergodic with stationary distribution

Yogurt	Salad	Hamburger	Pizza
7/65	2/13	18/65	6/13

With probability 1, Bob's average lunch cost converges to

$$\sum_x r(x)\pi_x = 3\left(\frac{7}{65}\right) + 4\left(\frac{2}{13} + \frac{6}{13}\right) + 7\left(\frac{18}{65}\right) = \$4.72 \text{ per day.}$$

1. Constructs an ergodic Markov chain whose limiting distribution is π.
2. Run the chain long enough for the chain to converge to its limiting distribution.
3. **Strong Law of Large Number**:

$$E(r(X)) = \sum_j r(j)\pi_j.$$

Problem

What algorithm should we use to generate random matrices?

Ecologist used the 2 x 2 swap

ALGORITHM 22 (2 × 2 swap)

1. **Random:** Choose two rows uniformly at random without replacement; choose two columns uniformly at random without replacement.
2. **Swap:** Is the resulting 2 × 2 submatrix swappable?
   a. No: Return to step 1. (Do not count this as a swap.)
   b. Yes: Make the swap.

2 x 2 swap

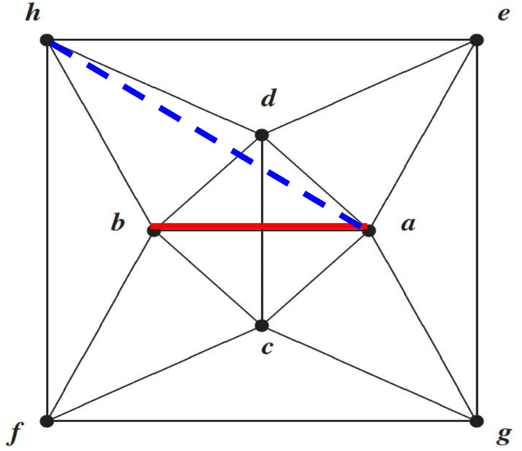
Suppose we have matrices A(r,c) where r = (3,2,1) and c = (2,2,1,1)

$$a = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \; b = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \; c = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \; d = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

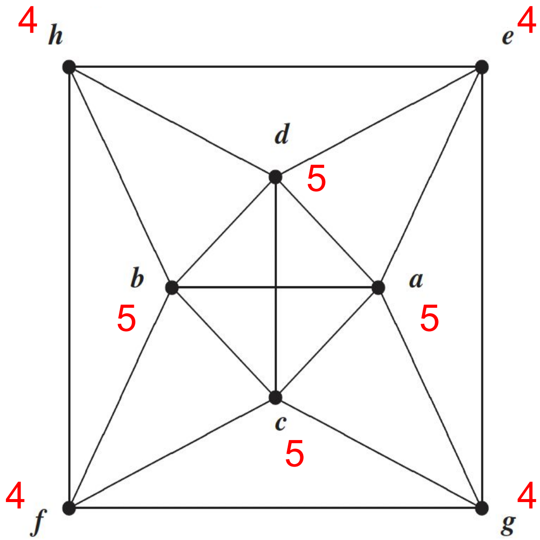
$$e = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \; f = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \; g = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \; h = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$a = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \; b = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$a = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \; h = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$



2 x 2 swap



$$\pi_i = d_i / \sum d_i$$

$$P = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{array} \begin{array}{c} \begin{array}{cccccccc} a & b & c & d & e & f & g & h \end{array} \\ \begin{pmatrix} 0 & 1/5 & 1/5 & 1/5 & 1/5 & 0 & 1/5 & 0 \\ 1/5 & 0 & 1/5 & 1/5 & 0 & 1/5 & 0 & 1/5 \\ 1/5 & 1/5 & 0 & 1/5 & 0 & 1/5 & 1/5 & 0 \\ 1/5 & 1/5 & 1/5 & 0 & 1/5 & 0 & 0 & 1/5 \\ 1/4 & 0 & 0 & 1/4 & 0 & 0 & 1/4 & 1/4 \\ 0 & 1/4 & 1/4 & 0 & 0 & 0 & 1/4 & 1/4 \\ 1/4 & 0 & 1/4 & 0 & 1/4 & 1/4 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 1/4 & 1/4 & 0 & 0 \end{pmatrix} \end{array}$$

$$\pi = (\frac{5}{36}, \frac{5}{36}, \frac{5}{36}, \frac{5}{36}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9})$$

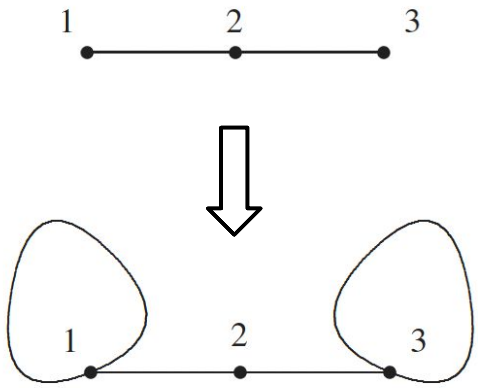
π is not uniform $\Rightarrow$ matrices generated by the 2 x 2 swap algorithm has ***bias***

How to solve this problem?

2 x 2 swap hold

<u>ALGORITHM 22H</u>

1. **Random.** Choose two rows uniformly at random without replacement; choose two columns uniformly at random without replacement.
2. **Swap or hold.** Is the resulting submatrix swappable?
    a. No. Stay in place, but count this as a step of the walk.
    b. Yes. Make the swap.



$$\boldsymbol{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix} \end{array} \quad \Rightarrow \quad \boldsymbol{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \end{array}$$

$$\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$$

Metropolized 2 x 2 swap

ALGORITHM 22M (Metropolized 2×2 swap)

1. **Random:** Choose two rows uniformly at random without replacement; choose two columns uniformly at random without replacement.
2. **Proposed:** Is the resulting 2×2 submatrix swappable?
   a. No: Return to step 1. (Do not count this as a step.)
   b. Yes: Take the 1s complement as a *proposed* swap.
3. **Metropolis:** Let $\boldsymbol{a}_j$ be the matrix obtained from $\boldsymbol{a}_i$ by making the swap in 2b. Let d_j and d_i be their respective vertex degrees, and let $\alpha = \min\{1, d_i/d_j\}$. Generate a uniform random number u in [0, 1].
   a. If $u \leq \alpha$, *accept*: move from $\boldsymbol{a}_i$ to $\boldsymbol{a}_j$.
   b. If $u > \alpha$, *reject*: remain at $\boldsymbol{a}_i$ (self-loop).

$d_i < d_j$: $\alpha = d_i/d_j$, $u > \alpha$, remain at a_i

$d_i > d_j$: $\alpha = 1$, $u \leq \alpha = 1$, move from a_i to a_j

This algorithm is faster than the previous 2 x 2 swap hold algorithm.

Conclusion

The ecologists' question: how likely is it that a random 13 × 17 matrix of 0s and 1s will have at least ten checkerboards?

MCMC

The chance that a random matrix has at least ten checkerboards is only ***0.03***. The distribution of finch species among the Galapagos ***shows evidence of competition***.