Cohesion Index Calculation

Johann Delcourt, Noam Miller, Iain Couzin, & Simon Garnier

March 14, 2017

Table of Contents

[Introduction 1](#_Toc477260745)

[Definition of 2](#_Toc477260746)

[Calculating 2](#_Toc477260747)

[Partitions 2](#_Toc477260748)

[Partitions when 3](#_Toc477260749)

[Partitions when 3](#_Toc477260750)

[as a function of and 4](#_Toc477260751)

[Calculating with R 5](#_Toc477260752)

[References 7](#_Toc477260753)

# Introduction

This note explains the steps to calculate a new index that estimates the cohesion of a group of individuals in a radial arm maze (note that this index can also be used for all types of mazes or arenas divided into distinct zones). This index is an alternative to the classical methods used to measure group cohesion based on topological and metric relationships between individuals (see [1]). These methods cannot be used reliably with a radial arm maze because individuals located in adjacent arms might be spatially close, yet incapable of interacting directly with each other.

The new cohesion index introduced here measures the ability of individuals to form cohesive groups in a radial arm maze. Its value is maximal (i.e. ) when all the animals are located in a single zone of the maze; it is minimal (i.e. ) when the animals are distributed uniformly between all the zones of the maze.

[Back to top](%22#%22)

# Definition of

Let be the number of individuals in each zone of the radial arm maze. For a maze with zones, the Euclidean distance between all can be calculated as:

Let be the value of for the least cohesive distribution possible of individuals between the zones of the maze (i.e. homogenous distribution of the individuals), and the total number of individuals in the maze. Therefore the cohesion index can be calculated as:

[Back to top](%22#%22)

# Calculating

## Partitions

depends on how individuals can be partitioned between zones in the maze.

In number theory and combinatorics, a partition of a positive integer , also called an integer partition, is a way of writing as a sum of positive integers [2–4]. Two sums that differ only in the order of their summands are considered to be the same partition. The number of distinct partitions of , without regard to the ordering of the terms, is defined as .

For example, 4 can be partitioned in five distinct ways;

Thus, the number of partitions . As partitions are not order-dependent, is the same partition as for instance. It is conventional to write the parts of a partition in descending order.

In R [5], the partitions package [6] can be used to obtain (with the function P) as well as all the possible partitions of (with the function parts). For instance, for :

if (!require(partitions)) {  
 install.packages("partitions")  
 library(partitions)  
}

## Loading required package: partitions

x <- 4  
P(x)

## [1] 5

parts(x)

##   
## [1,] 4 3 2 2 1  
## [2,] 0 1 2 1 1  
## [3,] 0 0 0 1 1  
## [4,] 0 0 0 0 1

(with each column corresponding to a different partition of ).

[Back to top](%22#%22)

## Partitions when

In the radial maze context, if the number of zones is superior or equal to the number of individuals , all the possible partitions of can be observed, and notably the more distributed partitions where each individual is alone in a zone). The number of partitions and the partitions themselves can therefore be obtained as described above.

[Back to top](%22#%22)

## Partitions when

If the number of zones is inferior to the number of individuals , then not all partitions can be observed. For instance, in the case of and , it is impossible to observe the partition . The total number of observable partitions in this case is therefore .

In R, the number of restricted partitions (i.e. the number of partitions that can be observed given and in our case) can be calculated using the function R in the partitions package. For instance, for and :

N <- 4  
Z <- 3  
  
R(Z, N, include.zero = TRUE)

## [1] 4

And the restricted partitions can be obtained with the function restrictedparts as follows:

restrictedparts(N, Z)

##   
## [1,] 4 3 2 2  
## [2,] 0 1 2 1  
## [3,] 0 0 0 1

(with each column corresponding to a different partition of in zones).

The Supplementary Table 1 below shows the number of restricted partitions for various values of and .

**Supplementary Table 1:** Partition number as a function of the number of individuals and the number of zones . In our experiment, the partition number is .

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Z=1 | Z=2 | Z=3 | Z=4 | Z=5 | Z=6 | Z=7 | Z=8 | Z=9 | Z=10 |
| **N=1** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **N=2** | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| **N=3** | 1 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| **N=4** | 1 | 3 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| **N=5** | 1 | 3 | 5 | 6 | 7 | 7 | 7 | 7 | 7 | 7 |
| **N=6** | 1 | 4 | 7 | 9 | 10 | 11 | 11 | 11 | 11 | 11 |
| **N=7** | 1 | 4 | 8 | 11 | 13 | 14 | 15 | 15 | 15 | 15 |
| **N=8** | 1 | 5 | 10 | 15 | 18 | 20 | 21 | 22 | 22 | 22 |
| **N=9** | 1 | 5 | 12 | 18 | 23 | 26 | 28 | 29 | 30 | 30 |
| **N=10** | 1 | 6 | 14 | 23 | 30 | 35 | 38 | 40 | 41 | 42 |

[Back to top](%22#%22)

## as a function of and

With R, can be easily calculated for a given and as follows (here and , as before):

N <- 4  
Z <- 3  
  
partitions <- restrictedparts(N, Z)  
Dc <- sqrt(apply(partitions ^ 2, 2, sum))  
min(Dc)

## [1] 2.44949

The Supplementary Table 2 below shows the values of for various values of and .

**Supplementary Table 2:** Minimal distance as a function of the number of individuals and the number of zones . When , . When , increases when decreases. For the particular case where there is just one zone, , in which case cannot be calculated. In our experiment, .

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Z=1 | Z=2 | Z=3 | Z=4 | Z=5 | Z=6 | Z=7 | Z=8 | Z=9 | Z=10 |
| **N=1** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **N=2** | 1 |  |  |  |  |  |  |  |  |  |
| **N=3** | 1 |  |  |  |  |  |  |  |  |  |
| **N=4** | 1 |  |  | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| **N=5** | 1 |  | 3 |  |  |  |  |  |  |  |
| **N=6** | 1 |  |  |  |  |  |  |  |  |  |
| **N=7** | 1 | 5 |  |  |  | 3 |  |  |  |  |
| **N=8** | 1 |  |  | 4 |  |  |  |  |  |  |
| **N=9** | 1 |  |  |  |  |  |  |  | 3 | 3 |
| **N=10** | 1 |  |  |  |  |  | 4 |  |  |  |

[Back to top](%22#%22)

# Calculating with R

For a same number of individuals, values can be different if the number of zones varies, but values are always between 0 (worst group cohesion possible) and 1 (best group cohesion possible). When , increases as decreases. In our experiment, with 7 zones and 10 fish, the partition number is 38, and not 42 as in or for example; as a consequence and varies between and .

For convenience, we provide with our study a R package - called projectRadial containing a function - called Ic - for computing for any observed partitions of fish in radial arm mazes with any arbitrary number of arms. Ic takes a single input, which should be a data frame containing the number of animals in each zone of the radial arm maze. The data frame should be organized as follows: each column must correspond to a zone of the radial arm maze and each row must correspond to a different observation.

An example of using Ic with actual data (20 fish in a radial maze with 6 arms and a central zone) is provided below:

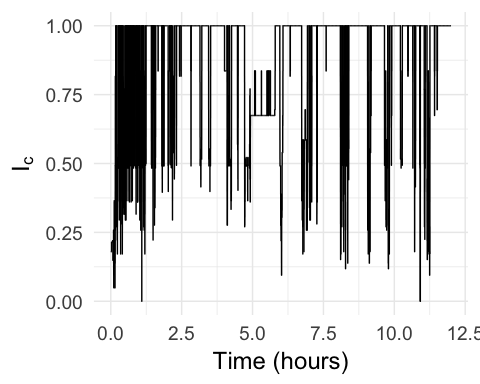
# Install and load the 'projecRadial' package  
if (!require(pacman)) {  
 install.packages("pacman")  
 library(partitions)  
}

## Loading required package: pacman

p\_load\_gh("sjmgarnier/projectRadial")  
  
  
# Calculate Ic with data provided in the 'projecRadial' package  
data\_path <- system.file("extdata", "sample\_data.csv", package = "projectRadial")  
dat <- read.csv(data\_path)  
  
dat$Ic <- Ic(dat[, 2:7])  
  
  
# Plot  
if (!require(ggplot2)) {  
 install.packages("ggplot2")  
 library(ggplot2)  
}

## Loading required package: ggplot2

ggplot(dat, aes(x = TimeInHours, y = Ic)) +  
 geom\_line() +   
 xlab("Time (hours)") +   
 ylab(bquote(I[c])) +  
 theme\_minimal(base\_size = 18)



[Back to top](%22#%22)

# References

1. Delcourt J, Poncin P. Shoals and schools: back to the heuristic definitions and quantitative references. Reviews in fish biology and fisheries. 2012;22: 595–619. doi:[10.1007/s11160-012-9260-z](https://doi.org/10.1007/s11160-012-9260-z)

2. Andrews GE. The Theory of Partitions [Internet]. Cambridge University Press; 1998. Available: <https://market.android.com/details?id=book-Sp7z9sK7RNkC>

3. Nathanson M. Partitions with parts in a finite set. Proceedings of the American Mathematical Society American Mathematical Society. ams.org; 2000;128: 1269–1273. Available: <http://www.ams.org/proc/2000-128-05/S0002-9939-00-05606-9/>

4. Andrews GE, Eriksson K. Integer Partitions [Internet]. Cambridge University Press; 2004. Available: <https://market.android.com/details?id=book-_5gsjNr0lWAC>

5. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available: <https://www.R-project.org/>

6. Hankin RKS. Additive integer partitions in r. Journal of Statistical Software, Code Snippets. 2006;16.