Exercise: Distances
Sep 26th, 2018
Deadline: Oct 1st, 2018

Table 1 shows counters for different species of plants in different sites (s1...s30). Calculate the distances among the pairs of sites (e.g., (s1,s2), etc) using the distance metrics:

- Euclidean
- Minkowski
- Mahalonobis
- Bray-Curtis index
- Chi-square

SITE	SPECIES COUNTS				
NO.	a	b	c	d	e
s1	0	2	9	14	2
s2	26	4	13	11	0
s3	0	10	9	8	0
s4	0	0	15	3	0
s5	13	5	3	10	7
s6	31	21	13	16	5
s7	9	6	0	11	2
s8	2	0	0	0	1
s9	17	7	10	14	6
s10	0	5	26	9	0
s11	0	8	8	6	7
s12	14	11	13	15	0
s13	0	0	19	0	6
s14	13	0	0	9	0
s15	4	0	10	12	0
s16	42	20	0	3	6
s17	4	0	0	0	0
s18	21	15	33	20	0
s19	2	5	12	16	3
s20	0	10	14	9	0
s21	8	0	0	4	6
s22	35	10	0	9	17
s23	6	7	1	17	10
s24	18	12	20	7	0
s25	32	26	0	23	0
s26	32	21	0	10	2
s27	24	17	0	25	6
s28	16	3	12	20	2
s29	11	0	7	8	0
s30	24	37	5	18	1

Table 1: species counters
(Source: http://www.econ.upf.edu/~michael/stanford/Stanford_Week1.pdf)

Rank the pairs of instances according to the measured distances. Plot the similarities for each distance used.

What can you conclude?

Is it possible to choose the best distance metric for this particular set of data?

Is it possible to compare distances among species instead of comparing regions? What would be the meaning? Repeat this exercise for the species distances.