

Ejercicio 6

a) Calcular las matrices de distancias y/o similitudes correspondientes a las medidas de asociacion SM, Jaccard, Rao, Nei, Cuerda, Prevosti y Rogers, en caso de utilizacion de un marcador codominante (dos alelos alternativos, A_1 asociado con A y A_2 asociado con a)

SIMILARIDADES

Table 1: SM

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.00	0.70	0.55	0.50	0.70	0.65	0.80	0.65	0.70	0.80	0.55	0.75	0.65	0.50	0.65
2	0.70	0.00	0.45	0.50	0.80	0.75	0.80	0.75	0.60	0.70	0.65	0.65	0.65	0.50	0.65
3	0.55	0.45	0.00	0.65	0.55	0.70	0.55	0.60	0.75	0.55	0.70	0.60	0.50	0.65	0.60
4	0.50	0.50	0.65	0.00	0.50	0.55	0.50	0.55	0.50	0.60	0.55	0.65	0.45	0.60	0.55
5	0.70	0.80	0.55	0.50	0.00	0.75	0.70	0.75	0.70	0.60	0.65	0.75	0.75	0.50	0.55
6	0.65	0.75	0.70	0.55	0.75	0.00	0.75	0.80	0.85	0.65	0.70	0.70	0.70	0.55	0.70

Table 2: JACCARD

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.00	0.62	0.50	0.44	0.62	0.61	0.75	0.61	0.65	0.73	0.53	0.69	0.59	0.44	0.59
2	0.62	0.00	0.42	0.44	0.73	0.71	0.75	0.71	0.56	0.62	0.61	0.59	0.59	0.44	0.59
3	0.50	0.42	0.00	0.59	0.50	0.67	0.53	0.58	0.71	0.50	0.67	0.56	0.47	0.59	0.56
4	0.44	0.44	0.59	0.00	0.44	0.53	0.47	0.53	0.47	0.53	0.53	0.59	0.42	0.53	0.50
5	0.62	0.73	0.50	0.44	0.00	0.71	0.65	0.71	0.65	0.53	0.61	0.69	0.69	0.44	0.50
6	0.61	0.71	0.67	0.53	0.71	0.00	0.72	0.78	0.82	0.61	0.68	0.67	0.67	0.53	0.67

Table 3: RAO

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.00	0.50	0.45	0.4	0.50	0.55	0.60	0.55	0.55	0.55	0.50	0.55	0.50	0.40	0.50
2	0.50	0.00	0.40	0.4	0.55	0.60	0.60	0.60	0.50	0.50	0.55	0.50	0.50	0.40	0.50
3	0.45	0.40	0.00	0.5	0.45	0.60	0.50	0.55	0.60	0.45	0.60	0.50	0.45	0.50	0.50
4	0.40	0.40	0.50	0.0	0.40	0.50	0.45	0.50	0.45	0.45	0.50	0.50	0.40	0.45	0.45
5	0.50	0.55	0.45	0.4	0.00	0.60	0.55	0.60	0.55	0.45	0.55	0.55	0.55	0.40	0.45
6	0.55	0.60	0.60	0.5	0.60	0.00	0.65	0.70	0.70	0.55	0.65	0.60	0.60	0.50	0.60

DISTANCIAS

Table 4: NEI

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
01	0.00	0.24	0.43	0.55	0.24	0.22	0.12	0.37	0.18	0.17	0.56	0.21	0.28	0.55	0.35
02	0.24	0.00	0.52	0.55	0.11	0.15	0.18	0.15	0.32	0.31	0.29	0.28	0.21	0.46	0.35
03	0.43	0.52	0.00	0.35	0.35	0.27	0.38	0.35	0.15	0.35	0.19	0.41	0.59	0.35	0.41
04	0.55	0.55	0.35	0.00	0.55	0.56	0.49	0.37	0.49	0.38	0.46	0.35	0.61	0.31	0.52
05	0.24	0.11	0.35	0.55	0.00	0.15	0.25	0.15	0.25	0.46	0.37	0.14	0.14	0.55	0.35
06	0.22	0.15	0.27	0.56	0.15	0.00	0.16	0.13	0.09	0.22	0.21	0.27	0.19	0.37	0.27

Table 5: CUERDA

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
01	0.00	0.44	0.56	0.62	0.44	0.43	0.33	0.51	0.40	0.38	0.59	0.41	0.47	0.62	0.51
02	0.44	0.00	0.61	0.62	0.33	0.36	0.38	0.36	0.50	0.48	0.47	0.47	0.43	0.58	0.51
03	0.56	0.61	0.00	0.51	0.53	0.44	0.53	0.50	0.36	0.53	0.40	0.54	0.62	0.51	0.54
04	0.62	0.62	0.51	0.00	0.62	0.59	0.58	0.53	0.58	0.54	0.56	0.51	0.64	0.50	0.59
05	0.44	0.33	0.53	0.62	0.00	0.36	0.44	0.36	0.44	0.57	0.51	0.36	0.36	0.62	0.53
06	0.43	0.36	0.44	0.59	0.36	0.00	0.36	0.33	0.28	0.43	0.40	0.44	0.40	0.53	0.44

Table 6: PREVOSTI

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
01	0.00	0.27	0.41	0.45	0.27	0.32	0.18	0.32	0.27	0.18	0.41	0.23	0.32	0.45	0.32
02	0.27	0.00	0.50	0.45	0.18	0.23	0.18	0.23	0.36	0.27	0.32	0.32	0.32	0.45	0.32
03	0.41	0.50	0.00	0.32	0.41	0.27	0.41	0.36	0.23	0.41	0.27	0.36	0.45	0.32	0.36
04	0.45	0.45	0.32	0.00	0.45	0.41	0.45	0.41	0.45	0.36	0.41	0.32	0.50	0.36	0.41
05	0.27	0.18	0.41	0.45	0.00	0.23	0.27	0.23	0.27	0.36	0.32	0.23	0.23	0.45	0.41
06	0.32	0.23	0.27	0.41	0.23	0.00	0.23	0.18	0.14	0.32	0.27	0.27	0.27	0.41	0.27

Table 7: ROGERS

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
01	0.00	0.27	0.41	0.45	0.27	0.32	0.18	0.32	0.27	0.18	0.41	0.23	0.32	0.45	0.32
02	0.27	0.00	0.50	0.45	0.18	0.23	0.18	0.23	0.36	0.27	0.32	0.32	0.32	0.45	0.32
03	0.41	0.50	0.00	0.32	0.41	0.27	0.41	0.36	0.23	0.41	0.27	0.36	0.45	0.32	0.36
04	0.45	0.45	0.32	0.00	0.45	0.41	0.45	0.41	0.45	0.36	0.41	0.32	0.50	0.36	0.41
05	0.27	0.18	0.41	0.45	0.00	0.23	0.27	0.23	0.27	0.36	0.32	0.23	0.23	0.45	0.41
06	0.32	0.23	0.27	0.41	0.23	0.00	0.23	0.18	0.14	0.32	0.27	0.27	0.27	0.41	0.27

b Analizar la relacion entre las distintas medidas de distancias/similaridad recurriendo a correlacion entre matrices. Comentar al respecto y sacar conclusiones

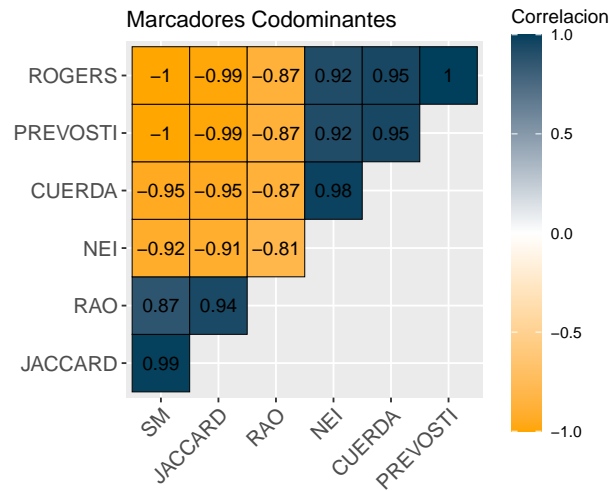


Figure 1: Correlacion entre las medidas de asociacion

Para poder evaluar y comparar las distintas medidas de distancias y similaridad se obtiene una matriz de

correlaciones, la cual nos va a permitir comprender de manera mas sencilla cuales son las medidas que tienen un comportamiento analogo o semejante y cuales son las que mas difieren. Dentro de las medidas de similitudes calculadas podemos decir que las tres son bastante similar, ya que presentan correlaciones mayores a 0.87, pero las medidas Jaccard y Simple Matching son las que mas se asemejan, siendo practicamente iguales. Considerando ahora solo las medidas de distancias, tambien podemos observar que todas son muy parecidas, siendo la de Rogers analoga con la de Prevosti.

Por otra parte, cabe destacar que la medida Simple Matching coincide con las medidas de distancias de Rogers y Prevosti. Luego, dada la complejidad en el calculo de las ultimas dos es preferible trabajar con Simple Matching, ya que su formula es mas sencilla y se obtendra el mismo resultado.

Otro aspecto importante a destacar es que la medida de Rao es la que mas difiere del resto, tanto para las medidas de similitud como para las de distancias.

c) Repetir el ejercicio suponiendo la utilización de un marcador dominante.

SIMILARIDADES

Table 8: SM

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.0	0.7	0.4	0.5	0.7	0.6	0.7	0.6	0.6	0.7	0.5	0.8	0.6	0.5	0.6
2	0.7	0.0	0.3	0.6	0.8	0.7	0.8	0.9	0.5	0.8	0.8	0.7	0.5	0.6	0.7
3	0.4	0.3	0.0	0.5	0.3	0.6	0.5	0.4	0.6	0.3	0.5	0.4	0.4	0.7	0.6
4	0.5	0.6	0.5	0.0	0.4	0.5	0.6	0.5	0.3	0.6	0.4	0.7	0.5	0.6	0.7
5	0.7	0.8	0.3	0.4	0.0	0.7	0.6	0.7	0.5	0.6	0.6	0.7	0.7	0.4	0.5
6	0.6	0.7	0.6	0.5	0.7	0.0	0.9	0.8	0.8	0.7	0.7	0.6	0.8	0.7	0.8

Table 9: JACCARD

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.00	0.57	0.25	0.37	0.50	0.50	0.62	0.50	0.43	0.57	0.44	0.67	0.50	0.44	0.50
2	0.57	0.00	0.22	0.50	0.67	0.62	0.75	0.86	0.37	0.71	0.75	0.57	0.44	0.56	0.62
3	0.25	0.22	0.00	0.37	0.12	0.50	0.44	0.33	0.43	0.22	0.44	0.25	0.33	0.62	0.50
4	0.37	0.50	0.37	0.00	0.25	0.44	0.56	0.44	0.22	0.50	0.40	0.57	0.44	0.56	0.62
5	0.50	0.67	0.12	0.25	0.00	0.57	0.50	0.57	0.29	0.43	0.50	0.50	0.57	0.33	0.37
6	0.50	0.62	0.50	0.44	0.57	0.00	0.88	0.75	0.71	0.62	0.67	0.50	0.75	0.67	0.75

Table 10: RAO

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.0	0.4	0.2	0.3	0.3	0.4	0.5	0.4	0.3	0.4	0.4	0.4	0.4	0.4	0.4
2	0.4	0.0	0.2	0.4	0.4	0.5	0.6	0.6	0.3	0.5	0.6	0.4	0.4	0.5	0.5
3	0.2	0.2	0.0	0.3	0.1	0.4	0.4	0.3	0.3	0.2	0.4	0.2	0.3	0.5	0.4
4	0.3	0.4	0.3	0.0	0.2	0.4	0.5	0.4	0.2	0.4	0.4	0.4	0.4	0.5	0.5
5	0.3	0.4	0.1	0.2	0.0	0.4	0.4	0.4	0.2	0.3	0.4	0.3	0.4	0.3	0.3
6	0.4	0.5	0.4	0.4	0.4	0.0	0.7	0.6	0.5	0.5	0.6	0.4	0.6	0.6	0.6

DISTANCIAS

Table 11: NEI

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.00	0.36	0.92	0.69	0.36	0.51	0.36	0.51	0.51	0.36	0.69	0.22	0.51	0.69	0.51
2	0.36	0.00	1.20	0.51	0.22	0.36	0.22	0.11	0.69	0.22	0.22	0.36	0.69	0.51	0.36
3	0.92	1.20	0.00	0.69	1.20	0.51	0.69	0.92	0.51	1.20	0.69	0.92	0.92	0.36	0.51
4	0.69	0.51	0.69	0.00	0.92	0.69	0.51	0.69	1.20	0.51	0.92	0.36	0.69	0.51	0.36
5	0.36	0.22	1.20	0.92	0.00	0.36	0.51	0.36	0.69	0.51	0.51	0.36	0.36	0.92	0.69
6	0.51	0.36	0.51	0.69	0.36	0.00	0.11	0.22	0.22	0.36	0.36	0.51	0.22	0.36	0.22

Table 12: CUERDA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.00	0.55	0.77	0.71	0.55	0.63	0.55	0.63	0.63	0.55	0.71	0.45	0.63	0.71	0.63
2	0.55	0.00	0.84	0.63	0.45	0.55	0.45	0.32	0.71	0.45	0.45	0.55	0.71	0.63	0.55
3	0.77	0.84	0.00	0.71	0.84	0.63	0.71	0.77	0.63	0.84	0.71	0.77	0.77	0.55	0.63
4	0.71	0.63	0.71	0.00	0.77	0.71	0.63	0.71	0.84	0.63	0.77	0.55	0.71	0.63	0.55
5	0.55	0.45	0.84	0.77	0.00	0.55	0.63	0.55	0.71	0.63	0.63	0.55	0.55	0.77	0.71
6	0.63	0.55	0.63	0.71	0.55	0.00	0.32	0.45	0.45	0.55	0.55	0.63	0.45	0.55	0.45

Table 13: PREVOSTI

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.0	0.3	0.6	0.5	0.3	0.4	0.3	0.4	0.4	0.3	0.5	0.2	0.4	0.5	0.4
2	0.3	0.0	0.7	0.4	0.2	0.3	0.2	0.1	0.5	0.2	0.2	0.3	0.5	0.4	0.3
3	0.6	0.7	0.0	0.5	0.7	0.4	0.5	0.6	0.4	0.7	0.5	0.6	0.6	0.3	0.4
4	0.5	0.4	0.5	0.0	0.6	0.5	0.4	0.5	0.7	0.4	0.6	0.3	0.5	0.4	0.3
5	0.3	0.2	0.7	0.6	0.0	0.3	0.4	0.3	0.5	0.4	0.4	0.3	0.3	0.6	0.5
6	0.4	0.3	0.4	0.5	0.3	0.0	0.1	0.2	0.2	0.3	0.3	0.4	0.2	0.3	0.2

Table 14: ROGERS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.0	0.3	0.6	0.5	0.3	0.4	0.3	0.4	0.4	0.3	0.5	0.2	0.4	0.5	0.4
2	0.3	0.0	0.7	0.4	0.2	0.3	0.2	0.1	0.5	0.2	0.2	0.3	0.5	0.4	0.3
3	0.6	0.7	0.0	0.5	0.7	0.4	0.5	0.6	0.4	0.7	0.5	0.6	0.6	0.3	0.4
4	0.5	0.4	0.5	0.0	0.6	0.5	0.4	0.5	0.7	0.4	0.6	0.3	0.5	0.4	0.3
5	0.3	0.2	0.7	0.6	0.0	0.3	0.4	0.3	0.5	0.4	0.4	0.3	0.3	0.6	0.5
6	0.4	0.3	0.4	0.5	0.3	0.0	0.1	0.2	0.2	0.3	0.3	0.4	0.2	0.3	0.2

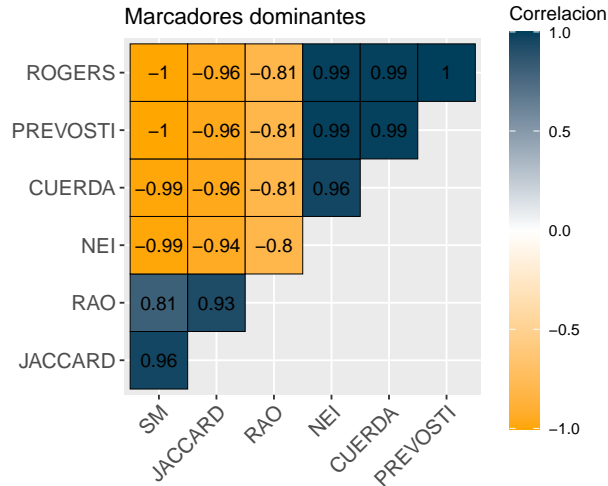


Figure 2: Correlacion entre las medidas de asociacion

Al utilizar marcadores moleculares dominantes podemos observar algunos cambios en la matriz de correlacion con respecto a la calculada con marcadores dominantes.

Las medidas de similaridad presentan mas discrepancias entre ellas, si bien siguen presentando altas correlaciones, observamos que las mismas son menores en comparacion a la primera codificacion utilizada.

Por el contrario, si nos centramos en las medidas de distancias observamos que las correlaciones son aun mas fuertes, asemejandose aun mas Nei con Rogers y Prevosti.

Nuevamente, es notable que la medida de similaridad de Rao es la que mas difiere del resto de las medidas, tanto de distancias como de similaridad.

d) Analizar para cada una de las distancias/similaridades estudiadas la correlación entre los resultados obtenidos para marcadores codominantes y dominantes.

Table 15: Correlacion

	Codominante vs Dominante
SM	0.71
JACCARD	0.62
RAO	0.50
NEI	0.62
CUERDA	0.65
PREVOSTI	0.71
ROGERS	0.71

Al realizar la correlacion de las medidas obtenidas a traves de las dos codificaciones propuestas, podemos observar que el comportamiento de las distintas medidas varia de acuerdo a la codificacion que se este utilizando. La medida de similaridad Simple Matching y las de distancias Prevosti y Rogers son las que presentan correlacion mas alta, mientras que la medida de similaridad de Rao es la que presenta la correlacion mas baja.