

Exercise: from sites to matrices

Sequences

Seq1 AAAAACCG
Seq2 AAAACCGG
Seq3 AATTGGGG
Seq4 ATTGTTGG

Alignment

Position	1	2	3	4	5	6	7	8
Seq1	A	A	A	A	A	C	C	G
Seq2	A	A	A	A	C	C	G	G
Seq3	A	A	T	T	G	G	G	G
Seq4	A	T	T	G	T	G	G	G

Count matrix

Position	1	2	3	4	5	6	7	8
A	4	3	2	2	1	0	0	0
C	0	0	0	0	1	2	1	0
G	0	0	0	1	1	2	3	4
T	0	1	2	1	1	0	0	0
Total	4	4	4	4	4	4	4	4

Frequency matrices

Pseudo-weight	0							
Position	1	2	3	4	5	6	7	8
A	1	0.75	0.5	0.5	0.25	0	0	0
C	0	0	0	0	0.25	0.5	0.25	0
G	0	0	0	0.25	0.25	0.5	0.75	1
T	0	0.25	0.5	0.25	0.25	0	0	0
Total	1	1	1	1	1	1	1	1

Pseudo-weight	2							
Position	1	2	3	4	5	6	7	8
A	0.75	0.583	0.417	0.417	0.25	0.083	0.083	0.083
C	0.083	0.083	0.083	0.083	0.25	0.417	0.25	0.083
G	0.083	0.083	0.083	0.25	0.25	0.417	0.583	0.75
T	0.083	0.25	0.417	0.25	0.25	0.083	0.083	0.083
Total	1	1	1	1	1	1	1	1

Pseudo-weight		10							
Position		1	2	3	4	5	6	7	8
A		0.464	0.393	0.321	0.321	0.250	0.179	0.179	0.179
C		0.179	0.179	0.179	0.179	0.250	0.321	0.250	0.179
G		0.179	0.179	0.179	0.250	0.250	0.321	0.393	0.464
T		0.179	0.250	0.321	0.250	0.250	0.179	0.179	0.179
Total		1	1	1	1	1	1	1	1

Pseudo-weight		100							
Position	1	2	3	4	5	6	7	8	
A	0.279	0.269	0.260	0.260	0.250	0.240	0.240	0.240	
C	0.240	0.240	0.240	0.240	0.250	0.260	0.250	0.240	
G	0.240	0.240	0.240	0.250	0.250	0.260	0.269	0.279	
T	0.240	0.250	0.260	0.250	0.250	0.240	0.240	0.240	
Total	1	1	1	1	1	1	1	1	

Weight matrices

Pseudo-weight	0	1.386						
Position	1	2	3	4	5	6	7	8
A	1.39	1.10	0.69	0.69	0.00	-Inf	-Inf	-Inf
C	-Inf	-Inf	-Inf	-Inf	0.00	0.69	0.00	-Inf
G	-Inf	-Inf	-Inf	0.00	0.00	0.69	1.10	1.39
T	-Inf	0.00	0.69	0.00	0.00	-Inf	-Inf	-Inf
Total	1.386	1.099	1.386	0.693	0	1.386	1.099	1.386

Pseudo-weight	1							
Position	1	2	3	4	5	6	7	8
A	1.10	0.85	0.51	0.51	0.00	-1.10	-1.10	-1.10
C	-1.10	-1.10	-1.10	-1.10	0.00	0.51	0.00	-1.10
G	-1.10	-1.10	-1.10	0.00	0.00	0.51	0.85	1.10
T	-1.10	0.00	0.51	0.00	0.00	-1.10	-1.10	-1.10
Total	-2.197	-1.35	-1.176	-0.588	0	-1.176	-1.35	-2.197

Residue	Prior
A	0.25
C	0.25
G	0.25
T	0.25

Pseudo-weight	0							
Position	1	2	3	4	5	6	7	8
A	1.20	0.92	0.51	0.51	-0.18	-Inf	-Inf	-Inf
C	-Inf	-Inf	-Inf	-Inf	0.22	0.92	0.22	-Inf
G	-Inf	-Inf	-Inf	0.22	0.22	0.92	1.32	1.61
T	-Inf	-0.18	0.51	-0.18	-0.18	-Inf	-Inf	-Inf
Total	1.204	0.734	1.022	0.552	0.082	1.833	1.545	1.609

Pseudo-weight	1							
Position	1	2	3	4	5	6	7	8
A	0.92	0.66	0.33	0.33	-0.18	-1.28	-1.28	-1.28
C	-0.88	-0.88	-0.88	-0.88	0.22	0.73	0.22	-0.88
G	-0.88	-0.88	-0.88	0.22	0.22	0.73	1.07	1.32
T	-1.28	-0.18	0.33	-0.18	-0.18	-1.28	-1.28	-1.28
Total	-2.116	-1.268	-1.094	-0.506	0.082	-1.094	-1.268	-2.116

Residue	Prior
A	0.3
C	0.2
G	0.2
T	0.3