

Exercise: from sites to matrices

Sequences

Seq1 AAAAACCG  
Seq2 AAAACCGG  
Seq3 AATTGGGG  
Seq4 ATTGTGGG

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 1

| Position | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    |
|----------|------|------|------|------|------|------|------|------|
| A        | 0.85 | 0.65 | 0.45 | 0.45 | 0.25 | 0.05 | 0.05 | 0.05 |
| C        | 0.05 | 0.05 | 0.05 | 0.05 | 0.25 | 0.45 | 0.25 | 0.05 |
| G        | 0.05 | 0.05 | 0.05 | 0.25 | 0.25 | 0.45 | 0.65 | 0.85 |
| T        | 0.05 | 0.25 | 0.45 | 0.25 | 0.25 | 0.05 | 0.05 | 0.05 |
| 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

| Position | 1     | 2     | 3     | 4     | 5    | 6     | 7     | 8     |
|----------|-------|-------|-------|-------|------|-------|-------|-------|
| A        | 0.60  | 0.48  | 0.30  | 0.30  | 0.00 | -Inf  | -Inf  | -Inf  |
| C        | -Inf  | -Inf  | -Inf  | -Inf  | 0.00 | 0.30  | 0.00  | -Inf  |
| G        | -Inf  | -Inf  | -Inf  | 0.00  | 0.00 | 0.30  | 0.48  | 0.60  |
| T        | -Inf  | 0.00  | 0.30  | 0.00  | 0.00 | -Inf  | -Inf  | -Inf  |
| Total    | 0.602 | 0.477 | 0.602 | 0.301 | 0    | 0.602 | 0.477 | 0.602 |

Pseudo-weight 1

| Position | 1      | 2      | 3      | 4      | 5    | 6      | 7      | 8      |
|----------|--------|--------|--------|--------|------|--------|--------|--------|
| A        | 0.53   | 0.41   | 0.26   | 0.26   | 0.00 | -0.70  | -0.70  | -0.70  |
| C        | -0.70  | -0.70  | -0.70  | -0.70  | 0.00 | 0.26   | 0.00   | -0.70  |
| G        | -0.70  | -0.70  | -0.70  | 0.00   | 0.00 | 0.26   | 0.41   | 0.53   |
| T        | -0.70  | 0.00   | 0.26   | 0.00   | 0.00 | -0.70  | -0.70  | -0.70  |
| Total    | -1.565 | -0.983 | -0.887 | -0.444 | 0    | -0.887 | -0.983 | -1.565 |

| 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        | 0.52  | 0.40  | 0.22  | 0.22  | -0.08 | -Inf  | -Inf  | -Inf  |
| C        | -Inf  | -Inf  | -Inf  | -Inf  | 0.10  | 0.40  | 0.10  | -Inf  |
| G        | -Inf  | -Inf  | -Inf  | 0.10  | 0.10  | 0.40  | 0.57  | 0.70  |
| T        | -Inf  | -0.08 | 0.22  | -0.08 | -0.08 | -Inf  | -Inf  | -Inf  |
| Total    | 0.523 | 0.319 | 0.444 | 0.24  | 0.035 | 0.796 | 0.671 | 0.699 |

Pseudo-weight 1

| Position | 1     | 2      | 3      | 4      | 5     | 6      | 7      | 8     |
|----------|-------|--------|--------|--------|-------|--------|--------|-------|
| A        | 0.45  | 0.34   | 0.18   | 0.18   | -0.08 | -0.78  | -0.78  | -0.78 |
| C        | -0.60 | -0.60  | -0.60  | -0.60  | 0.10  | 0.35   | 0.10   | -0.60 |
| G        | -0.60 | -0.60  | -0.60  | 0.10   | 0.10  | 0.35   | 0.51   | 0.63  |
| T        | -0.78 | -0.08  | 0.18   | -0.08  | -0.08 | -0.78  | -0.78  | -0.78 |
| Total    | -1.53 | -0.948 | -0.852 | -0.408 | 0.035 | -0.852 | -0.948 | -1.53 |

| Residue | Prior |
|---------|-------|
| A       | 0.3   |
| C       | 0.2   |
| G       | 0.2   |
| T       | 0.3   |