Notes

Valentin Dreismann - 1864900

July 12, 2019

1 Derivation of Log-Likelihood

1.1 Forward Algorithm

Let $u \in \mathbb{R}^n$ be the initial distribution (which *can* be the *stationary distribution* δ), Γ the *transition probability matrix* and $P(x)$ ....

Then the forward algorithm is defined as follows:

$$\alpha_0 := u$$
$$\alpha_i := \alpha_{i-1}\,\Gamma^t\,P(x) \quad \text{for}\, i \geq 1$$

The latter can be expressed as follows:

$$\alpha_i(k) = \sum_{j=1}^{m} \alpha_{i-1}(j)\,\Gamma^{d_i}(j,k)\,P_{(k,k)}(x_i)$$

We leverage the following identity:

$$log_b\left(\sum_{i=0}^{N} a_i\right) = log_b(a_0) + log_b\left(1 + \sum_{i=1}^{N} b^{log_b(a_i) - log_b(a_0)}\right) \tag{1}$$

In particular, we note that the expression on the right-hand side can be evaluated by *exclusively* relying on the logarithm of a_i.

Let $\beta_{j,k,i} := \Gamma^{d_i}(j,k)P_{(k,k)}(x_i)$. Note that neither $\Gamma^{d_i}(j,k)$ nor $P_{(k,k)}(x_i)$ depend on the position within the MC; $\alpha_{i-1}(j)$ does, however.

Then, we have:

$$\alpha_i(k) = \sum_{j=1}^{m} \underbrace{\alpha_{i-1}(j)\,\beta_{j,k,i}}_{a_{j,k}} \tag{2}$$

By combining equations (2) and (1) we obtain

$$
\begin{aligned}
log_b\left(\sum_{j=0}^{m} a_j\right) &= log_b\left(\alpha_{i-1}(0)\,\beta_{0,k,i}\right) \,+\, log_b\left(1+\sum_{i=1}^{m} b^{log_b(\alpha_{i-1}(j)\,\beta_{j,k,i})-log_b(\alpha_{i-1}(0)\,\beta_{0,k,i})}\right)\\
&= log_b(\alpha_{i-1}(0)) + log_b(\beta_{0,k,i}) + log_b\left(1+\sum_{i=1}^{m} b^{log_b(\alpha_{i-1}(j))+log_b(\beta_{j,k,i})-log_b(\alpha_{i-1}(0))-log_b(\beta_{0,k,i})}\right)
\end{aligned}
\tag{3}
$$

Further setting

$$
\tilde{\alpha_i}(k) := log\left(\alpha_i(k)\right) = log_b\left(\sum_{j=0}^{m} a_{j,k}\right)
$$

we obtain

$$
\tilde{\alpha}_i(k) = \tilde{\alpha}_{i-1}(0) + log\left(\beta_{0,k,i}\right) + log\left(1+\sum_{j=1}^{m} b^{\tilde{\alpha}_{i-1}(j)+log_b(\beta_{j,k,i})-\tilde{\alpha}_{i-1}(0)-log_b(\beta_{0,k,i})}\right)
$$

w.l.o.g. the maximum element is the first (0).