Hidden Markov models (continued)

When dealing with probabilities that are extremely small but are non-zero, limitations of computers are apt to result in the treatment by the computers of these probabilities as being exactly zero rather than as being slightly greater than 0.

This kind of problem is known as an **underflow** error.

Underflow errors can cause havoc in maximum likelihood and Bayesian analyses.

As mentioned in your text, underflow errors can be avoided by working in logarithms, when possible.

Consider the main recursion in the forward algorithm

$$\Pr\left(x^{i+1}, y_{i+1}\right) = \sum_{y_i} \Pr\left(x^i, y_i\right) p_{x_{i+1} y_{i+1}} \rho_{y_i y_{i+1}}$$

Let $y_{i'}$ be the value of y_i that satisfies

$$\max_{y_i} \Pr\left(x^i, y_i\right)$$

and let

$$m_i = \Pr\left(x^i, y_{i'}\right)$$

We get

$$\begin{aligned}
&\Pr\left(x^{i+1}, y_{i+1}\right) = \sum_{y_i} \Pr\left(x^i, y_i\right) p_{x_{i+1} y_{i+1}} \rho_{y_i y_{i+1}} \\
&= e^{log(m_i) - log(m_i)} \sum_{y_i} \Pr\left(x^i, y_i\right) p_{x_{i+1} y_{i+1}} \rho_{y_i y_{i+1}} \\
&= e^{log(m_i) - log(m_i)} \sum_{y_i} exp(log(\Pr\left(x^i, y_i\right) p_{x_{i+1} y_{i+1}} \rho_{y_i y_{i+1}})) \\
&= e^{log(m_i)} \sum_{y_i} exp(-log(m_i) + log(\Pr\left(x^i, y_i\right) p_{x_{i+1} y_{i+1}} \rho_{y_i y_{i+1}})) \\
&= e^{log(m_i)} \sum_{y_i} exp(-log(m_i) + log(\Pr\left(x^i, y_i\right)) + log(p_{x_{i+1} y_{i+1}} \rho_{y_i y_{i+1}}))
\end{aligned}$$

So,

$$\begin{aligned} log(\Pr\left(x^{i+1}, y_{i+1}\right)) &= log(m_i) + log(\sum_{y_i} exp(-log(m_i) + \\ &\quad log(\Pr\left(x^i, y_i\right)) + log(p_{x_{i+1}y_{i+1}}\rho_{y_i y_{i+1}}))) \end{aligned}$$

This approach for avoiding underflow errors generally works quite well but it can be computationally expensive because calculating logarithms is computationally demanding.

Other possible approaches for avoiding underflow errors are less computationally demanding but are harder somewhat more complicated to implement.