RLEMMC EM Derivation

Orhan Sönmez

August 1, 2017

Likelihood and Log Likelihood

$$
\begin{aligned}
\pi^* &= \arg\max_\pi \mathcal{L}(\pi) \\
&= \arg\max_\pi P(r=1;\pi) \\
&= \arg\max_\pi \sum_T \sum_{x_{0:T}} \sum_{a_{0:T}} P(r=1, x_{0:T}, a_{0:T}, T;\pi) \\
&= \arg\max_\pi \log \sum_T \sum_{x_{0:T}} \sum_{a_{0:T}} P(r=1, x_{0:T}, a_{0:T}, T;\pi)
\end{aligned}
$$

Jensen's Inequality and Lower Bound

$$
\begin{aligned}
&\log \sum_T \sum_{x_{0:T}} \sum_{a_{0:T}} P(r=1, x_{0:T}, a_{0:T}, T;\pi) \\
&\quad = \log \sum_T \sum_{x_{0:T}} \sum_{a_{0:T}} P(r=1, x_{0:T}, a_{0:T}, T;\pi) \frac{q(x_{0:T}, a_{0:T}, T)}{q(x_{0:T}, a_{0:T}, T)} \\
&\quad = \log \left\langle \frac{P(r=1, x_{0:T}, a_{0:T}, T;\pi)}{q(x_{0:T}, a_{0:T}, T)} \right\rangle_{q(x_{0:T}, a_{0:T}, T)} \\
&\quad \geq \left\langle \log \frac{P(r=1, x_{0:T}, a_{0:T}, T;\pi)}{q(x_{0:T}, a_{0:T}, T)} \right\rangle_{q(x_{0:T}, a_{0:T}, T)} && \text{(Jensen's Inequality)} \\
&\quad = B(\pi) && \text{(Lower Bound)}
\end{aligned}
$$

Maximum Lower Bound

$$\begin{aligned}
B^*(\pi) &= \langle \log P(r=1, x_{0:T}, a_{0:T}, T; \pi) \rangle_{P(x_{0:T}, a_{0:T}, T | r=1; \pi')} \\
&\quad - \langle \log P(x_{0:T}, a_{0:T}, T | r=1; \pi') \rangle_{P(x_{0:T}, a_{0:T}, T | r=1; \pi')} \qquad \text{(Does not depend on } \pi\text{)} \\
&=^+ \langle \log P(r=1, x_{0:T}, a_{0:T}, T; \pi) \rangle_{P(x_{0:T}, a_{0:T}, T | r=1; \pi'))}
\end{aligned}$$

where $q^*(x_{0:T}, a_{0:T}, T) = P(x_{0:T}, a_{0:T}, T | r=1; \pi')$. (Derivative wrt $q = 0$)

Update Rule

$$\pi^{(k)} \leftarrow \arg\max_{\pi^{(k)}} \left\langle \log P(r=1, x_{0:T}, a_{0:T}, T; \pi^{(k)}) \right\rangle_{P(x_{0:T}, a_{0:T}, T | r=1; \pi^{(k-1)})}$$