1 The Expectation-Maximization Algorithm

Likelihood: $L(\theta; X, Z) = p(X, Z \,|\, \theta)$, where θ is the random vector of model parameters, X is the (non-random) vector of observed data, and Z is the random vector of unobservable variables.

The Expectation-Maximization (EM) Algorithm is an algorithm that solves the optimization (maximization) problem for a marginal likelihood (or probability):

$$L(\theta; X) \;=\; p(X \,|\, \theta) \;=\; \int p(X, Z \,|\, \theta)\,\mathrm{d}Z$$

More specifically, the EM Algorithm *attempts* to compute:

$$\widehat{\theta} \;:=\; \underset{\theta}{\operatorname{argmax}}\,\{\, L(\theta\,;\,X)\,\} \;=\; \underset{\theta}{\operatorname{argmax}}\,\{\, p(X \,|\, \theta)\,\} \;=\; \underset{\theta}{\operatorname{argmax}}\left\{\,\int p(X, Z \,|\, \theta)\,\mathrm{d}Z\,\right\}$$

In practice, the EM Algorithm produces estimates of a local maximum $\widehat{\theta}$ of $L(\theta; X)$.

The Expectation-Maximization (EM) Algorithm

Choose (arbitrarily) an initial value θ_0 for θ. Choose (arbitrarily) a termination threshold $\tau > 0$. Generate the sequence $\{\,\theta_t\,\}$, for $t = 1, 2, 3, \ldots$, by iterating through the following two-step procedure:

1. **Expectation Step:** Compute the following expectation value (as a function of θ):

$$Q(\theta \,|\, \theta_t) \;:=\; E_{Z|X,\theta_t}\{\,\log L(\theta\,;\,X, Z)\,\} \;=\; \int [\log L(\theta\,;\,X, Z)]\, p(Z \,|\, X, \theta_t)\mathrm{d}Z \tag{1.1}$$

2. **Maximization Step:** Solve the following optimization (maximization) problem to obtain θ_{t+1}:

$$\theta_{t+1} \;:=\; \underset{\theta}{\operatorname{argmax}}\,\{\, Q(\theta \,|\, \theta_t)\,\} \tag{1.2}$$

Terminate the EM Algorithm when

$$\left|\,\frac{\theta_{t+1} - \theta_t}{\theta_t}\,\right| \;\leq\; \tau \tag{1.3}$$

Theorem 1.1 *The sequence θ_1, θ_2, θ_3, … produced by the EM Algorithm satisfies the following:*

$$\log p(X \,|\, \theta_{t+1}) \;\geq\; \log p(X \,|\, \theta_t), \quad \text{for each } \; t = 1, 2, 3, \ldots$$

PROOF First, observe that:

$$\begin{aligned}
\log p(X \,|\, \theta) \;&=\; \log\left(\frac{p(X,\theta)}{p(\theta)}\right) \;=\; \log\left(\frac{p(X,Z,\theta)}{p(\theta)}\,\frac{p(X,\theta)}{p(X,Z,\theta)}\right) \\
&=\; \log\left(p(X, Z \,|\, \theta)\right) - \log\left(p(Z \,|\, X, \theta)\right)
\end{aligned}$$

Taking expectation on both sides with respect to $p(Z\,|\,X,\theta_t)\,\mathrm{d}Z$ yields:

$$\begin{aligned}
E_{Z|X,\theta_t}\{\log p(X\,|\,\theta)\} &= E_{Z|X,\theta_t}\{\log(p(X,Z\,|\,\theta))\} - E_{Z|X,\theta_t}\{\log(p(Z\,|\,X,\theta))\} \\
\int\{\log p(X\,|\,\theta)\}\,p(Z\,|\,X,\theta_t)\,\mathrm{d}Z &= \int\{\log(p(X,Z\,|\,\theta))\}\,p(Z\,|\,X,\theta_t)\,\mathrm{d}Z - \int\{\log(p(Z\,|\,X,\theta))\}\,p(Z\,|\,X,\theta_t)\,\mathrm{d}Z \\
\log p(X\,|\,\theta) &= Q(\theta\,|\,\theta_t) + H(\theta\,|\,\theta_t)
\end{aligned}$$

where $H(\theta\,|\,\theta_t)$ is defined as follows:

$$H(\theta\,|\,\theta_t) \;:=\; -\int\{\log(p(Z\,|\,X,\theta))\}\,p(Z\,|\,X,\theta_t)\,\mathrm{d}Z$$

CLAIM 1: $H(\theta\,|\,\theta_t) \geq H(\theta_t\,|\,\theta_t)$, for any θ.

Indeed,

This completes the proof of **CLAIM 1**.

Now, the following equation

$$\log p(X\,|\,\theta) \;=\; Q(\theta\,|\,\theta_t) + H(\theta\,|\,\theta_t) \tag{1.4}$$

holds for any value of θ; in particular, it holds for θ_t:

$$\log p(X\,|\,\theta_t) \;=\; Q(\theta_t\,|\,\theta_t) + H(\theta_t\,|\,\theta_t) \tag{1.5}$$

Subtracting Equation (1.5) from Equation (1.4) yields:

$$\log p(X\,|\,\theta) - \log p(X\,|\,\theta_t) \;=\; (\,Q(\theta\,|\,\theta_t) - Q(\theta_t\,|\,\theta_t)\,) + (\,H(\theta\,|\,\theta_t) - H(\theta_t\,|\,\theta_t)\,) \tag{1.6}$$

Thus, **CLAIM 1** implies:

$$\log p(X\,|\,\theta) - \log p(X\,|\,\theta_t) \;\geq\; Q(\theta\,|\,\theta_t) - Q(\theta_t\,|\,\theta_t) \tag{1.7}$$

Since, by definition, $\theta_{t+1} := \underset{\theta}{\operatorname{argmax}}\,\{\,Q(\theta\,|\,\theta_t)\,\}$, we therefore have:

$$\log p(X\,|\,\theta_{t+1}) - \log p(X\,|\,\theta_t) \;\geq\; Q(\theta_{t+1}\,|\,\theta_t) - Q(\theta_t\,|\,\theta_t) \;\geq\; 0 \tag{1.8}$$

This proves the Theorem. $\square$

References