Implementation of the EM algorithm for the moving transmitter DPD algorithm

January 24, 2011

The EM algorithm is an efficient iterative procedure to compute the Maximum Likelihood estimate in the presence of missing or hidden data. Each iteration of the EM algorithm consists of two processes: The E-step and the M-step. In the expectation step, or E-step, the missing data are estimated given the observed data and current estimate of the model parameters.

This is achieved by using the conditional expectation. In the M-step, the likelihood function is maximized under the assumption that the missing data variables are known. Convergence to a local maximum is assured since the algorithm is guaranteed to increase the likelihood at each iteration.

In the derivation of the algorithm, $\mathbf{x}$ is defined as the random vector that results from a parametrized family, defined by the parameters vector θ. We wish to find θ that maximizes $P(\mathbf{x}|\theta)$.

We define the likelihood function $L(\theta)$ as follows:

$$L(\theta) \triangleq \ln P(\mathbf{x}|\theta) \tag{1}$$

We denote a hidden random vector $\mathbf{z}$, of hidden, unknown or latent variables that simplify the maximization of the likelihood function.

We start with at initial value for the parameters vector θ. In each step of the algorithm we improve the likelihood of the estimation by choosing θ_{n+1} as follows:

$$
\begin{aligned}
\theta_{n+1} &= \tag{2}\\
&= \operatorname{argmax}_\theta \{ \mathrm{E}_{\mathbf{z}|\mathbf{x},\theta_\mathbf{n}} \{ \ln P(\mathbf{x}, \mathbf{z}|\theta) \} \}\\
&= \operatorname{argmax}_\theta \left\{ \sum_{\mathbf{z}} P(\mathbf{z}|\mathbf{x}, \theta_\mathbf{n}) \ln P(\mathbf{x}, \mathbf{z}|\theta) \right\}
\end{aligned}
$$

The two steps of the EM algorithm appear in the equation above. At the E step, the conditional expectation $\mathrm{E}_{\mathbf{z}|\mathbf{x},\theta_\mathbf{n}} \{ \ln P(\mathbf{x}, \mathbf{z}|\theta) \}$ is evaluated, and then at the M step, this expression is maximized with respect to θ.

A possible implementation of the EM algorithm for our problem is achieved by defining the velocity of the transmitter as the missing variable. Thus, we define:

$$\mathbf{x} = \mathbf{r} \tag{3}$$
$$\mathbf{z} = \vec{v} \tag{4}$$
$$\theta = \vec{p} \tag{5}$$

Where $\mathbf{r}$ represents the received signal, $\vec{v}$ represents the velocity of the transmitter, and $\vec{p}$ represents the position of the transmitter.

Notice that in the suggested model, we treat the position as the unknown parameter, while we treat the velocity as the missing random variable. Thus, in order to implement the EM algorithm, the probabilistic characteristics of the

velocity as a random variable should be defined.

We can then see that:

$$P(\mathbf{x}, \mathbf{z}|\theta) = P(\mathbf{r}|\vec{v}, \vec{p})P(\vec{v}|\vec{p}) \tag{6}$$

Since we consider the velocity and the position of the transmitter to be independent, we can write:

$$P(\mathbf{x}, \mathbf{z}|\theta) = P(\mathbf{r}|\vec{v}, \vec{p})P(\vec{v}) \tag{7}$$

Using Bayes's rule, we can also see that:

$$P(\mathbf{z}|\mathbf{x}, \theta_{\mathbf{n}}) = \frac{P(\mathbf{x}|\mathbf{z}, \theta_{\mathbf{n}})P(\mathbf{z}|\theta_{\mathbf{n}})}{\sum_{\mathbf{z}'} P(\mathbf{x}|\mathbf{z}', \theta_{\mathbf{n}})P(\mathbf{z}'|\theta_{\mathbf{n}})} \tag{8}$$

So that:

$$P(\mathbf{z}|\mathbf{x}, \theta_{\mathbf{n}}) = \frac{P(\mathbf{r}|\vec{v}, \vec{p}_n)P(\vec{v})}{\sum_{\vec{v}'} P(\mathbf{r}|\vec{v}', \vec{p}_n)P(\vec{v}')} \tag{9}$$

We can now write:

$$E_n(\vec{p}) \triangleq \mathrm{E}_{\vec{v}|\mathbf{r}, \vec{p}_n}\{\ln P(\mathbf{r}, \vec{v}|\vec{p})\} = \sum_{\vec{v}}\{\frac{P(\mathbf{r}|\vec{v}, \vec{p}_n)P(\vec{v})}{\sum_{\vec{v}'}\{P(\mathbf{r}|\vec{v}', \vec{p}_n)P(\vec{v}')\}} \ln\{P(\mathbf{r}|\vec{v}, \vec{p})P(\vec{v})\}\} \tag{10}$$

If we define:

$$\tilde{E}_n(\vec{p}) \triangleq \sum_{\vec{v}}\{\frac{P(\mathbf{r}|\vec{v}, \vec{p}_n)P(\vec{v})}{\sum_{\vec{v}'}\{P(\mathbf{r}|\vec{v}', \vec{p}_n)P(\vec{v}')\}} \ln\{P(\mathbf{r}|\vec{v}, \vec{p})\}\} \tag{11}$$

We can see that $\tilde{E}_n$ and E_n have the same maximum with respect to $\vec{p}$, since we eliminated an expression that does not depend on $\vec{p}$.

The expressions of the form $P(\mathbf{r}|\vec{v}, \vec{p})$ are well known to us from the development of the DPD algorithm, and are the expressions for the likelihood function. $P(\vec{v})$ is the assumed velocity probability density function, which can be defined as a constant probability for a defined range of possible velocities, or in a more complex manner, given there is some a-priori information about the velocity.

And so, for the M step:

$$\vec{p}_{n+1} = \mathrm{argmax}_{\vec{p}}\tilde{E}_n(\vec{p}) \tag{12}$$

Convergence of the method is also guaranteed even if the maximum of the function is not found at the M-step, but rather a position $\vec{p}_{n+1}$ is found, for which:

$$\tilde{E}_n(\vec{p}_{n+1}) > \tilde{E}_n(\vec{p}_n) \tag{13}$$