Extension of ensemble Gaussian mixture filter for vector observations

January 11, 2013

1 Ensemble Gaussian Mixture Filter derivation

Consider the discrete time state evolution given by

$$X_{n+1} = X_n + V_n \tag{1}$$

where $X_n \in \mathbb{R}^N$, $V_n \sim N(0, R_X)$, $R_X \in \mathbb{R}^{N\times N}$. At time 0, the state is assumed to be distributed according to a L component Gaussian mixture as follows -

$$\pi(X_0) = \sum_{l=1}^{L} \alpha_{l,0}\pi_l(X_0) \tag{2}$$

The observations are available at each discrete time step n and are given by

$$Y_n = h(X_n) + W_n \tag{3}$$

where $Y_n, W_n \in \mathbb{R}^M$, $W_n \sim N(0, R_Y)$ and $R_Y \in \mathbb{R}^{M\times M}$. When the observation map $h()$ is linear i.e. $h(X_n) = HX_n$, $H \in \mathbb{R}^{M\times N}$ and the state evolution and observations (prior and the likelihood) are of the above form, the posterior distribution at every time step n will also be a Gaussian mixture as follows -

$$\pi(X_n) = \sum_{l=1}^{L} \alpha_{l,n}\pi_{l,n}(X_n) \tag{4}$$

where $\pi_{l,n}(X)$ is the normal distribution $N(\mu_{l,n}, P_{l,n})$.

A continuous time formulation of the data assimilation step (Bayes update) gives the following equations (Reich, 2011) in artificial time $s \in [0, 1]$

$$\frac{dX}{ds} = g(X, s) \tag{5}$$

$$\nabla_X \cdot (\pi g) = \pi(S - \mathbb{E}_\pi[S]) \tag{6}$$

where $S(X, Y_{obs}) = \frac{1}{2}(h(X) - Y_{obs})^T R_Y^{-1}(h(X) - Y_{obs})$ is the negative log-likelihood function. The random vectors $X(s = 0)$ and $X(s = 1)$ are distributed according to the available prior and the required posterior distribution. Solving for $g(X, s)$ from the partial differential equation (PDE) in 6 provides a way to numerically solve for X(s) using the ordinary differential equation (ODE) in 5. The solution to equation 6 is not unique and we make appropriate choices/assumptions to solve for $g()$. For example, we can impose the additional constraint that g is also the minimizer of the kinetic energy defined as

$$\mathcal{T}(v) = \frac{1}{2}\int v^T M v d\pi \tag{7}$$

where $M \in \mathbb{R}^{N\times N}$ is a positive definite matrix. Under such constraint it can be shown that (Villani, 2003) $g = M^{-1}\nabla_X\psi$ where the potential $\psi(X, s)$ is the solution of the elliptic PDE,

$$\nabla_X \cdot (\pi M^{-1}\nabla_X\psi) = \pi(S - \mathbb{E}_\pi[S]) \tag{8}$$

In the simple case of a single component Gaussian mixture ($L = 1$) where the prior is distributed as $\mathcal{N}(\mu, P)$ and $M^{-1} = P(s)$ it can be shown (Beregman and Reich, 2010a,b)

$$\frac{dX}{ds} = -\frac{1}{2}P(s)H^T R_Y(HX(s) + H\mu(s) - 2Y_{obs}) \tag{9}$$

In case of a general L component Gaussian mixture we decompose the vector field g into two components (Reich, 2011)

$$\frac{dX}{ds} = g(X, s) = u_A(X, s) + u_B(X, s) \tag{10}$$

and define

$$u_A(X) = \sum_{l=1}^{L} \frac{\alpha_l \pi_l(X)}{\pi(X)} P_l \nabla_X \psi_{A,l}(X) \tag{11}$$

$$u_B(X) = \sum_{l=1}^{L} \frac{\alpha_l \pi_l(X)}{\pi(X)} P_l \nabla_X \psi_{B,l}(X) \tag{12}$$

where we have dropped the s for simplicity. Substituting in equation 6 we get

$$\nabla_X \cdot \left(\sum_{l=1}^{L} \alpha_l \pi_l(X) P_l \nabla_X \psi_{A,l}(X) + \sum_{l=1}^{L} \alpha_l \pi_l(X) P_l \nabla_X \psi_{B,l}(X) \right) = \sum_{l=1}^{L} \alpha_l \pi_l(X)(S(X) - \mathbb{E}_\pi[S]) \tag{13}$$

$$= \sum_{l=1}^{L} \alpha_l \pi_l(X)(S(X) - \mathbb{E}_{\pi_l}[S]) + \sum_{l=1}^{L} \alpha_l \pi_l(X)(\mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S]) \tag{14}$$

Thus we have,

$$\nabla_X \cdot \{\pi_l(X) P_l \nabla_X \psi_{A,l}(X)\} = \pi_l(X)(S(X) - \mathbb{E}_{\pi_l}[S]) \quad l = 1, 2...L \tag{15}$$

and

$$\nabla_X \cdot \{\pi_l(X) P_l \nabla_X \psi_{B,l}(X)\} = \pi_l(X)(\mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S]) \quad l = 1, 2...L \tag{16}$$

The equation 15 is similar to the case of single component Gaussian and hence from 9 and 11 we have the solution,

$$u_A(X, s) = -\frac{1}{2} \sum_{l=1}^{L} \frac{\alpha_l(s) \pi_l(X, s)}{\pi(X, s)} P_l(s) H^T R_Y^{-1} [HX(s) + H\mu_l(s) - 2Y_{obs}] \tag{17}$$

The equation 16 needs to be solved to have the complete solution.

1.1 Scalar observation

In case of scalar observation $Y \in \mathbb{R}$, the PDE can be reduced to a ODE and explicitly solved (Reich, 2011). To derive this we assume the potential $\psi_{B,l}$ to be of the following form

$$\psi_{B,l}(X) = \hat{\psi}_{B,l}(HX - H\mu_l) = \hat{\psi}_{B,l}(Y - Y_l) \tag{18}$$

where $Y = HX$ and $Y_l = H\mu_l$. Hence we have $\nabla_X \psi_{B,l}(X) = H^T \frac{d\hat{\psi}_{B,l}}{dY}(Y - Y_l)$. Thus the PDE in equation 16 simplifies to the following ODE,

$$-(Y - Y_l)\frac{d\hat{\psi}_{B,l}}{dY}(Y - Y_l) + HP_l H^T \frac{d^2 \hat{\psi}_{B,l}}{dY^2}(Y - Y_l) = \mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S] \tag{19}$$

Under the initial condition $\frac{d\hat{\psi}_{B,l}}{dY}(Y - Y_l)|_{Y=Y_l} = 0$ we can solve the above differential equation to obtain

$$\frac{d\hat{\psi}_{B,l}}{dY}(Y - Y_l) = \frac{1}{2} \frac{\mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S]}{HP_l H^T} \frac{\mathrm{Erf}((Y - Y_l)/\sqrt{2\sigma_l^2})}{\pi_l(Y)} \tag{20}$$

with the PDF

$$\pi_l(Y) = \frac{1}{\sqrt{2\pi\sigma_l^2}} \exp\left(-\frac{(Y-Y_l)^2}{2\sigma_l^2}\right) \tag{21}$$

and $\sigma_l^2 = HP_lH^T$. The standard error function is given by

$$\mathrm{Erf}(Y) = \frac{2}{\sqrt{\pi}} \int_0^Y e^{-s^2} ds, \quad Y > 0 \tag{22}$$

$$\mathrm{Erf}(Y) = -\mathrm{Erf}(-Y), \quad Y < 0 \tag{23}$$

$$\tag{24}$$

Thus the expression for $u_B(X, s)$ can be written as

$$u_B(X, s) = \frac{1}{2} \sum_{l=1}^{L} \frac{\alpha_l(s)\pi_l(X, s)}{\pi(X, s)} P_l(s) H^T \frac{\mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S]}{HP_lH^T} \frac{\mathrm{Erf}((Y - Y_l)/\sqrt{2\sigma_l^2})}{\pi_l(Y)} \tag{25}$$

1.2 Vector observation

For the general case of vector observations the above technique does not reduce the PDE to ODE. Equation 16 can be expanded as

$$\{\nabla_X \pi_l(X)\}\cdot P_l \nabla_X \psi_{B,l}(X) + \pi_l(X) \nabla_X \cdot \{P_l \nabla_X \psi_{B,l}(X)\} = \pi_l(X)(\mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S]) \tag{26}$$

$$-\pi_l(X)(X - \mu_l)^T P_l^{-1} P_l \nabla_X \psi_{B,l}(X) + \pi_l(X) \nabla_X \cdot \{P_l \nabla_X \psi_{B,l}(X)\} = \pi_l(X)(\mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S]) \tag{27}$$

$$-(X - \mu_l)^T \nabla_X \psi_{B,l}(X) + \nabla_X \cdot \{P_l \nabla_X \psi_{B,l}(X)\} = \mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S] \tag{28}$$

This is a second order PDE with all the cross terms present in general. To simplify the problem we employ the following transformation of variables. Let $X = Q_l Z$ where $P_l = Q_l Q_l^T$ is the Cholesky decomposition and it always exists since P_l is a positive definite matrix. Matrix Q_l is a lower triangular matrix with positive diagonal entries and is invertible. Hence we have the relation $Z = Q_l^{-1} X$ and we define $\gamma_l = Q_l^{-1} \mu_l$. Let the transformation induce the function $\phi_l(Z) = \psi_{B,l}(X)|_{X=Q_l Z}$ in the variable Z and we have the following relations

$$Q_l^T \nabla_X \psi_{B,l}(X) = \nabla_Z \phi_l(Z) \tag{29}$$

$$\nabla_X \cdot \{P_l \nabla_X \psi_{B,l}(X)\} = \nabla_Z^2 \phi_l(Z) \tag{30}$$

Thus equation 28 becomes

$$-(Z - \gamma_l)^T \nabla_Z \phi_l(Z) + \nabla_Z^2 \phi_l(Z) = \mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S] \tag{31}$$

If vector $Z = (z_1, z_2..z_N)$, then the above equation can be written as

$$-\sum_{i=1}^{N} (z_i - \gamma_{l,i}) \frac{\partial \phi_l}{\partial z_i} + \sum_{i=1}^{N} \frac{\partial^2 \phi_l}{\partial z_i^2} = \mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S] \tag{32}$$

$$-\sum_{i=1}^{N} \{(z_i - \gamma_{l,i}) \frac{\partial \phi_l}{\partial z_i} + \frac{\partial^2 \phi_l}{\partial z_i^2}\} = \mathbb{E}_{\pi_l}[S] - \mathbb{E}_\pi[S] = \sum_{i=1}^{N} C_i \tag{33}$$

$$-(z_i - \gamma_{l,i}) \frac{\partial \phi_l}{\partial z_i} + \frac{\partial^2 \phi_l}{\partial z_i^2} = C_i \quad i = 1, 2...N \tag{34}$$

If a solution $\phi_l(Z)$ satisfies the system of PDE's given in 34, it also satisfies the PDE in equation 33. It can be seen that the individual equations in 34 are similar to equation 19 obtained in the scalar observation case. If we assume that $\frac{\partial \phi_l}{\partial z_i} = f_i(z_i)$ i.e. the individual partial derivatives of ϕ with respect to z_i are functions of only the variable z_i then we can solve the individual equations in the above system independently. With the initial condition set to $f_i(z_i = 0) = 0$, we obtain the solution

$$f_i(z_i) = \frac{\partial \phi_l}{\partial z_i} = \frac{1}{2}\frac{C_i}{\pi_s(z_i - \gamma_{l,i})}\left(\mathrm{Erf}\left[\frac{(\mathrm{z_i} - \gamma_{\mathrm{l,i}})}{\sqrt{2}}\right] + \mathrm{Erf}\left[\frac{\gamma_{\mathrm{l,i}}}{\sqrt{2}}\right]\right) \qquad i = 1, 2...N \tag{35}$$

where π_s is the standard normal density function. Note that the solution depends on the choice of the initial condition. A different choice, for Ex. $f_i(z_i = \gamma_{l,i}) = 0$ gives the solution

$$f_i(z_i) = \frac{\partial \phi_l}{\partial z_i} = \frac{1}{2}\frac{C_i}{\pi_s(z_i - \gamma_{l,i})}\left(\mathrm{Erf}\left[\frac{(\mathrm{z_i} - \gamma_{\mathrm{l,i}})}{\sqrt{2}}\right]\right) \qquad i = 1, 2...N \tag{36}$$

The final solution is completed from equation 12 by observing that $\nabla_X \psi_{B,l}(X) = (Q_l^T)^{-1} \nabla_Z \phi_l(Z)|_{Z=Q_l^{-1}X}$.