1 Dynamics model

The dynamics model is an autoregressive model of order p:

$$\mathbf{x}_t = \mathbf{b} + \sum_{k=1}^{p} \mathbf{A}_{p+1-k}\mathbf{x}_{t-k} + \epsilon$$

where ϵ is zero-mean Gaussian noise with isotropic covariance $\mathbf{\Sigma}_x = \sigma_x^2\mathbf{I}$. Maximum likelihood estimation for this model given training data consisting of length-$p+1$ windows of data is a matter of solving the linear least squares problem

$$\min_{\mathbf{A}_{1:p},\mathbf{b}} \sum_{i=1}^{N} \left\| \mathbf{x}_{p+1}^{(i)} - \mathbf{b} - \sum_{k=1}^{p} \mathbf{A}_k \mathbf{x}_k^{(i)} \right\|^2$$

and then setting σ^2 to $1/N$ times the minimal value of this objective.

2 Observation model

The observation model is a two-component mixture, consisting of a Gaussian distribution centered at the state and a uniform distribution to account for outliers.

Let $\{\mathbf{z}_i\}_{i=1}^N$ be the observed (filtered) points on the screen. Each $\mathbf{z}_i$ has an associated latent indicator variable y_i that signifies which mixture component it belongs to:

$$y_i = 1[\mathbf{x}_i \text{ is not noise}]$$

The prior probability of these variables is

$$p(y_i = 1) = \pi$$

If $\mathbf{z}_i$ is not noise, it is normally distributed around the true position $\mathbf{x}$ with covariance $\mathbf{\Sigma}_z = \sigma_z^2\mathbf{I}$. Otherwise, it is uniformly distributed throughout the screen with density ρ:

$$p(\mathbf{z}_i|\mathbf{x}, y_i = 1) = \mathcal{N}(\mathbf{x}, \mathbf{\Sigma}_z)$$
$$p(\mathbf{z}_i|\mathbf{x}, y_i = 0) = \rho$$

3 Inference

EM algorithm for MAP estimation

Complete-data likelihood

$$p(\mathbf{z}_{1:N}, y_{1:N}|\mathbf{x})p(\mathbf{x}) = p(\mathbf{x}) \prod_i p(\mathbf{z}_i, y_i|\mathbf{x})$$
$$p(\mathbf{z}_i, y_i|\mathbf{x}) = (\pi\mathcal{N}(\mathbf{z}_i; \mathbf{x}, \mathbf{\Sigma}_z))^{y_i} \left((1-\pi)\rho\right)^{1-y_i}$$

Each iteration of the EM algorithm proceeds by maximizing the expected complete-data log likelihood with the latent variables drawn from their conditional distribution given the observations and settings of the parameters from the previous iteration.

$$\begin{aligned}
Q(\mathbf{x}, \mathbf{x}^{(k-1)}) &= \mathbb{E}\left[\log p(\mathbf{z}_{1:N}, y_{1:N}|\mathbf{x})p(\mathbf{x})\right] \\
&= \log p(\mathbf{x}) + \sum_i \mathbb{E}\left[\log p(\mathbf{z}_i, y_i|\mathbf{x})\right] \\
&= \log p(\mathbf{x}) + \sum_i r_i \log \mathcal{N}(\mathbf{z}_i; \mathbf{x}, \mathbf{\Sigma}_z) + \text{const}
\end{aligned}$$

where

$$r_i = \frac{\mathcal{N}(\mathbf{z}_i; \mathbf{x}^{(k-1)}, \mathbf{\Sigma}_z)\pi}{\mathcal{N}(\mathbf{z}_i; \mathbf{x}^{(k-1)}, \mathbf{\Sigma}_z)\pi + \rho(1-\pi)}$$

In the E-step, we compute these r_i values, and in the M-step, we compute $\mathbf{x}^{(k)}$ by maximizing $Q(\mathbf{x}, \mathbf{x}^{(k-1)})$:

$$\begin{aligned}
\mathbf{x}^{(k)} &\leftarrow \arg\max_{\mathbf{x}} Q(\mathbf{x}, \mathbf{x}^{(k-1)}) \\
&= \arg\min_{\mathbf{x}} \, (\mathbf{x} - \mathbf{x}_0)^T \mathbf{\Sigma}_x^{-1} (\mathbf{x} - \mathbf{x}_0) + \sum_i r_i (\mathbf{z}_i - \mathbf{x})^T \mathbf{\Sigma}_z^{-1} (\mathbf{z}_i - \mathbf{x}) \\
&= \left(\mathbf{\Sigma}_x^{-1} + \mathbf{\Sigma}_z^{-1} \sum_i r_i\right)^{-1} \left(\mathbf{\Sigma}_x^{-1}\mathbf{x}_0 + \mathbf{\Sigma}_z^{-1} \sum_i r_i \mathbf{z}_i\right) \\
&= \frac{\sigma_z^2 \mathbf{x}_0 + \sigma_x^2 \sum_i r_i \mathbf{z}_i}{\sigma_z^2 + \sigma_x^2 \sum_i r_i}
\end{aligned}$$