1 EM with stochastic sampling

1.1 The 1D problem

The governing equation for the 1D problem can be represented as:

$$\dot{x} = f(x; \beta) \tag{1}$$

Considering a parameteric model for the system, $f(x)$ can be represented as an additive expression of a class of basis functions:

$$f(x) = \sum_{i=1}^{N} \beta_i \phi_i(x) \tag{2}$$

The data is given in the form of a time series, $\mathbf{x}$ at discrete time points. For simplicity, let us assume the observations are collected at equispaced times, jh for $0 \leq j \leq J$. Thus the observed data can be represented as $\mathbf{x} = x_0, x_1, \cdots, x_J$.

We start by introducing data at intermediate time points. In between 2 observed data points, x_i and x_{i+1}, we introduce F sampled values, $z_{i1}, z_{i2}, \cdots, z_{iF}$. The sampled values are created using a diffusion bridge. The ith diffusion bridge sample is depends on the observed data, x and parameter θ of function $f(x; \theta)$.

$$z^{(i)} \sim z \,|\, x, \beta^{(k)} \tag{3}$$

The observed data and sampled data can be combined together, where $M = J + F * J$

$$x_1, z_{1,1}, \cdots, z_{1,F}, x_2, z_{2,1}, \cdots, z_{2,F}, x_3, \cdots, x_J \to y_1^{(i)}, y_2^{(i)}, \cdots, y_M^{(i)} \tag{4}$$

The Expectation-Maximization algorithm consists of two steps, computing the expected log likelihood function, Q and maximizing this function with respect to the parameters, β_i. Using the complete data, y means the expected log likelihood is the summation of the complete log likelihood over all y_i.

$$Q = \sum_{y^{(i)}} \text{complete log likelihood}$$

$$Q = \sum_i \sum_j \frac{(y_{j+1}^{(i)} - y_j^{(i)} - f(y_j^{(i)})h)^2}{2\sigma^2 h}$$

For the M step of the Expectation-Maximization algorithm, the complete log likelihood, Q is minimized with respect to the paramteres β

$$\min_\beta \frac{1}{2\sigma^2 h} \sum_i \sum_j (y_{j+1}^{(i)} - y_j^{(i)} - h \sum_{k=1}^{M} \beta_k \phi_k(y_j^{(i)}))^2$$

Differentiating with respect to each component of parameter, β_ℓ and setting the value to zero gives the maximum value, thus providing a direct way to evaluate the M step

$$\frac{\partial(\cdot)}{\partial\beta_\ell} = \frac{1}{2\sigma^2 h}\sum_i\sum_j(y_{j+1}^{(i)} - y_j^{(i)} - h\sum_{k=1}^{M}\beta_k\phi_k(y_j^{(i)})) * h\phi_\ell(y_j^{(i)}) = 0$$

Rearranging the terms gives

$$\sum_i\sum_j(\frac{y_{j+1}^{(i)} - y_j^{(i)}}{h})\phi_\ell(y_j^{(i)}) = \sum_k \beta_k\phi_k(y_j^{(i)})\phi_\ell(y_j^{(i)})$$

This system can be written as $r = A * b$ matrix solution. The resulting value of ϕ is used in the next iteration of the EM algorithm as the iterate is chosen as:

$$\beta^{(k+1)} = \arg\max_\beta Q(\beta, \beta^{(k)})$$