Capítulo 1

Anexo

1.1. Descripción del algoritmo MCMC

1.1.1. Modelo

$$
\begin{aligned}
y_i|f(x_i), z_i, \sigma_k^* &\sim w_p^{AL}(\varepsilon_i|\sigma_{z_i}),\\
f(X)|M(X), \lambda &\sim \mathcal{GP}(f(X)|M(X), K(X, X|\lambda)),\\
\lambda &\sim GI(c_\lambda, d_\lambda),\\
z_i|\pi &\sim Mult(\pi),\\
\pi|\alpha &\sim GEM(\alpha),\\
\sigma_k^*|c_{DP}, d_{DP} &\sim GI(\sigma_k|c_{DP}, d_{DP}),
\end{aligned}
$$

donde $k(x_i, x_j|\lambda) = \lambda \times exp\{-\|x_i - x_j\|_2\}$ y $\varepsilon_i = y_i - f(x_i)$.

1.1.2. Actualización del error

Recordando que los centros de masa y los pesos del Proceso de Dirichlet son independientes, pueden ser actualizados de forma independiente, con el inconveniente de que hay un número infinito de parámetros que actualizar. Para resolverlo, se utilizará el algoritmo de truncamiento propuesto por KALLI, adaptado para el modelo propuesto en esta tesis, de la manera siguiente.

Sea $\xi_1, \xi_2, \xi_3, ...$ una secuencia positiva, generalmente elegida determinista y decreciente. Sea N una variable aleatoria con soporte en los enteros positivos, una variable auxiliar incorporada al modelo.

Actualización de los centros de masa

Para cada $k \in \{1, 2, ..., N\}$,

$$\sigma_k | \{\varepsilon_i, z_i | z_i = k\}, c, d \sim GI(\bar{c}_{DP}, \bar{d}_{DP}),$$
$$\bar{c}_{DP} = c_{DP} + |\{i | z_i = k\}|,$$
$$\bar{d}_{DP} = d_{DP} + p \left[\sum_{\{i | z_i = k,\ \varepsilon_i \geq 0\}} \varepsilon_i \right] + (1 - p) \left[\sum_{\{i |\ z_i = k,\ \varepsilon_i < 0\}} -\varepsilon_i \right].$$

Actualización de los pesos

Sea $\hat{\pi}_k = \beta_k \prod_{j=1}^{k-1} (1 - \beta_j)$, de modo que para cada $k \in \{1, 2, ..., N\}$,

$$\beta_k | \{z_i\}, a, b \sim Beta(\bar{a}, \bar{b}),$$
$$\bar{a} = 1 + |\{i | z_i = k\}|,$$
$$\bar{b} = \alpha + |\{i | z_i > k\}|.$$

Entonces, se calcula

$$\pi_k = \frac{\bar{\pi_k}}{\sum_{j=1}^{N} \bar{\pi_j}}$$

Actualización de las clases y variables de truncamiento

Para cada $i \in \{1, ..., m\}$, se obtiene

$$u_i \sim U(0, \xi_{z_i}),$$

valor que se utiliza para actualizar la probabilidad de pertenencia a cada clase de la siguiente forma. Para cada $k \in \{1, 2, ..., N\}$,

$$P(z_i = k|\varepsilon_i, \pi_k, \sigma_k) \propto \mathbb{1}(\xi_k > u_i) \cdot \frac{\pi_k}{\xi_k} \cdot w_p^{AL}(\varepsilon_i|\sigma_k).$$

Posteriormente se actualiza

$$N = \text{máx}\{N_i|N_i = \text{máx}\{j|\xi_j > u_i\}, i \in \{1, ..., m\}\}.$$

1.1.3. Actualización de la tendencia

Se define la variable aleatoria auxiliar:

$$b_i \sim \begin{cases} \frac{p}{\sigma_i} & prob = P(\varepsilon_i \geq 0) = 1 - p \\ -\frac{1-p}{\sigma_i} & prob = P(\varepsilon_i < 0) = p \end{cases},$$

de forma que $b = [b_1, ..., b_m]^T$.

Actualización de f(x)

Es posible calcular que

$$f(x)|Y, X, M(X), b, \lambda \sim TruncNormal(\bar{M}(X, b), K(X, X|\lambda), \rho, \eta),$$
$$\bar{M}(X, b) = M(X) + K(X, X|\lambda)b,$$
$$\rho_i = \begin{cases} -\infty & \text{si } b_i > 0 \\ y_i & \text{si } b_i < 0 \end{cases},$$
$$\eta_i = \begin{cases} y_i & \text{si } b_i > 0 \\ \infty & \text{si } b_i < 0 \end{cases},$$

donde ρ es el vector de límites inferiores y η es el vector de límites superiores.

Actualización del parámetro de escala

Por otro lado, se puede obtener que

$$P(\lambda|X, M(X), f(X), b, c_\lambda, d_\lambda) \propto \lambda^{-\bar{c}_\lambda - 1} \cdot exp\left\{-\frac{\bar{d}_\lambda}{\lambda}\right\} \cdot exp\left\{-\bar{B}\lambda\right\},$$

$$\bar{c}_\lambda = c_\lambda + \frac{p}{2},$$

$$\bar{d}_\lambda = d_\lambda + \bar{F},$$

$$\bar{F} = \frac{1}{2}(f(X) - M(X))^T[K(X, X|\lambda = 1)^{-1}](f(X) - M(X)),$$

$$\bar{B} = \frac{1}{2}b^T[K(X, X|\lambda = 1)]b.$$