を出力するアルゴリズムである。ここに、$\hat{\theta}(X \cdot x^{t-1})$ は $X \cdot x^{t-1}$ からの θ の最尤推定量である。

データ列 $x_1 x_2, \ldots$ が与えられるとき、$\mathcal{P}_k$ を用いる逐次的正規化最尤予測アルゴリズムを構成せよ。

問題 3) s 個の異なる確率モデル $P(X_t : M_1), \ldots, P(X_t; M_s)$ があり $(t = 1, 2, \ldots)$, それらは $0 < \alpha < 1$ を未知パラメータであるとして、

変化点検知
(モデル変化)

$$P(M_j | M_i) = \begin{cases} \frac{\alpha}{s-1} & (j \neq i) \\ 1 - \alpha & (j = i) \end{cases}$$

の確率で遷移し合うとする。

「素直に適用しましょう」

今、データ列 $x^n = x_1, \ldots, x_n$ が与えらたとき、モデルの変化点が 1 つだけであると分かっているとして、MDL 原理の考え方に従って、この変化点を求めるアルゴリズムを示せ。

問題 4) ある定義域 X 上の k_i 次元のパラメトリックな確率モデル $P(X; \theta, k_i)$ が s 個用意されているとする $(i = 1, \ldots, s)$, $k_i \neq k_j (i \neq j)$. データが $x_1, x_2, \ldots$ と逐次的に与えられる下で、各モデルではベイズ予測を行ない、これらを統合して集合的予測することを考える。

ここで、**ベイズ予測**とは、各時刻 t について、各 i に対して、

$$P_{\mathrm{Bayes}}(X | x^{t-1} : k_i) = \int P(X; \theta, k_i) P(\theta | x^{t-1}) d\theta$$

なる予測分布を出力するアルゴリズムである。$P(\theta | x^{t-1})$ は事前分布 $p(\theta)$ を与えた上で計算できる、θ の x^{t-1} からの事後確率である。

また、**集合的予測**とは、各時刻 t に対し、s 個の予測分布のそれぞれに重み $v_{i,t}$ を与えて線形に予測分布を出力するアルゴリズムである。すなわち、時刻 t における出力予測分布を $\hat{P}_t(X | x^{t-1})$ とおくと、

aggregating algorithm
各Expert : Bayes予測
Expertの統合 : Bayes
⇓
予測損失はどうなる?

$$\hat{P}_t(X | x^{t-1}) = \sum_{i=1}^{s} v_{i,t} P_{\mathrm{Bayes}}(X | x^{t-1}; k_i).$$

ただし、重みは以下の規則で逐次的に更新する。

$$\begin{aligned} v_{i,t} &= \frac{w_{i,t}}{W_t}, \quad W_t = \sum_{i=1}^{s} w_{i,t} \\ w_{i,t+1} &= w_{i,t} P_{\mathrm{Bayes}}(x_t | x^{t-1}; k_i) \\ t &\leftarrow t + 1 \end{aligned}$$