Additional information about the paper **Koichiro Yamauchi, "Optimal Incremental Learning under Covariate Shift", Memetic Computing 2009**.

This text describes the parameter values for the Expectation and Maximization (EM) algorithm used in the paper. Since these parameters are for the traditional EM algorithm, their detailed explanations were omitted in the paper.

I think this information is useful to re-execute the method in your computer. If you want to know the whole algorithm, please refer to the paper.

dataset	σ_0	θ	$M^{'}_{MAX}$	σ_{min}
mackey-glass	0.001	0.5	10	0.001
cpu-performance	0.01	0.01	10	0.01
auto-mpg	0.001	0.3	10	0.01
ConcreteData	0.01	0.001	8	0.01
servo	0.01	0.001	10	0.01
housing	1	0.001	8	0.01
ForestFires	0.1	0.001	10	0.1
hearta1	0.1	0.001	10	0.1

Table 1: Datasets used and corresponding parameters.

- σ_0 denotes the initial values for the diagonal elements of variance-covariance matrix's for the mixture of Gaussian.
- θ is a threshold for the stopping condition of the EM algorithm. Therefore, if the total amount of the change in model-parameters is less than θ, the EM algorithm was terminated.
- $M^{'}_{MAX}$ is the maximum number of models in the mixture of Gaussian. Therefore, the system searched an appropriate size of the mixture of Gaussian in the interval $[1, M^{'}_{MAX}]$.
- σ_{min} is the minimum value for the diagonal elements of variance-covariance matrix's. This lower limit value is need to avoid the overflow during the calculation. Note that the variance-covariance matrix was calculated as the diagonal matrix for simplicity in the experiment.

The algorithm 1 in the next page gives more detailed information about them.

Algorithm 1 Constructing mixture of Gaussian (EM algorithm)

Require: buffer B, initial variance σ_0^2, maximum number of models M'_{Max}, θ for stop condition
  for $m = 1$ to M'_{Max} **do**
    {Initialize centers and variances}
    for $b = 1$ to m **do**
      $\boldsymbol{u}_b \leftarrow \boldsymbol{x}_b$ $\quad \{\boldsymbol{u}_b$ is the center of the b-th model, $\boldsymbol{x}_b \in B\}$
      $\sigma_b \leftarrow \sigma_0$
      $\pi_b \leftarrow 1$
      $b \leftarrow b + 1$
    end for
    {Expectation and Maximization (EM) algorithm}
    ChangeInParameters $\leftarrow$ DoubleMAX
    $n \leftarrow 1$
    while ChangeInParameters $> \theta$ **do**
      $\boldsymbol{\theta}_{prev} \leftarrow \boldsymbol{\theta}^{(n-1)}(m)$ $\{\boldsymbol{\theta}^{(n)}(m)$ is the parameter vector of the mixture of Gaussian having m models at the n-th step.$\}$
      execute E-step
      execute M-step ($\boldsymbol{\theta}^{(n)}(m)$ is obtained)
      ChangeInParameters $\leftarrow \|\boldsymbol{\theta}_{prev} - \boldsymbol{\theta}^{(n)}(m)\|$
      $n \leftarrow n + 1$
    end while
    Calculate $AIC(m)$. $\{AIC(m)$ is the AIC value of the mixture of Gaussians having m models.$\}$
    if $m > 1$ and $AIC(m) \geq AIC(m-1)$ **then**
      RETURN $\boldsymbol{\theta}(m-1)$ (return previous parameter vector).
      TERMINATE
    end if
    $m \leftarrow m + 1$
  end for

Koichiro Yamauchi
Chubu University
1200, Matsumoto-cho, Kasugai-shi, Aichi
487-8501, JAPAN