YUIMAにおけるモデル選択

江口 翔一

大阪大学 数理・データ科学教育研究センター

YUIMAチュートリアル 確率微分方程式のデータサイエンス入門

2018. 12. 2.

概要

1. 情報量規準
2. 関数 “IC”
3. シミュレーション・実習

はじめに

- データ $\mathrm{x}_n = (x_1, x_2, \ldots, x_n)$
- 適当な候補モデルを想定し，候補モデルの中からデータ x_n を生成したモデルを表すために**最も適したモデル**を見つけたい

真のモデル: $g_n(x)$

$\Longleftrightarrow$

最適なモデル?

候補モデル: $\mathcal{M}_1, \cdots, \mathcal{M}_K$

$$\mathcal{M}_i = \{f_{i,n}(x, \theta_i) | \theta_i \in \Theta_i \subset \mathbb{R}^{p_i}\}$$

$\Rightarrow$ モデルの推定: $f_{i,n}(x, \hat{\theta}_{i,n})$ ($\hat{\theta}_{i,n}$: 最尤推定量)

$\Rightarrow$ モデル $f_{i,n}(x, \hat{\theta}_{i,n})$ のよさを評価する基準が必要: **情報量規準**

赤池情報量規準 AIC

$$\mathrm{x}_n \xrightarrow{\text{モデルの推定}} f_{i,n}(x, \hat{\theta}_{i,n})$$

- **赤池情報量規準 (AIC): モデルの評価を予測の観点から捉える**

  $\Rightarrow$ **将来真のモデルから獲得したデータ z の従う真のモデル $g_n(z)$ を $f_{i,n}(z, \hat{\theta}_{i,n})$ で予測したときの平均的なよさ (悪さ) を評価**

- **Kullback-Leibler 情報量 (K-L 情報量): 分布の近さを測る尺度**

$$KL(g_n; f_{i,n}) = \mathbb{E}_Z \left(\log \frac{g_n(Z)}{f_{i,n}(Z, \hat{\theta}_{i,n})} \right)$$
$$= \mathbb{E}_Z\{\log g_n(Z)\} - \mathbb{E}_Z\{\log f_{i,n}(Z, \hat{\theta}_{i,n})\}$$

  - **性質:** 「$KL(g_n; f_{i,n}) \geq 0$」，「$KL(g_n; f_{i,n}) = 0 \Leftrightarrow g_n = f_{i,n}$」

$\Rightarrow$ $\mathbb{E}_Z\{\log f_{i,n}(Z, \hat{\theta}_{i,n})\}$ **の有効な推定量の導出**

$$n\mathbb{E}_Z\{\log f_{i,n}(Z,\hat{\theta}_{i,n})\} \xleftrightarrow{\text{推定値}} \mathbb{H}_{i,n}(\hat{\theta}_{i,n}) = \sum_{j=1}^{n} \log f_{i,n}(x_j,\hat{\theta}_{i,n})$$

$\Rightarrow$ **各候補モデルの $\mathbb{H}_{i,n}(\hat{\theta}_{i,n})$ の比較で十分? $\cdots$ NO**

$\uparrow$

$\mathbb{H}_{i,n}(\hat{\theta}_{i,n})$ は $n\mathbb{E}_Z\{\log f_{i,n}(Z,\hat{\theta}_{i,n})\}$ の推定量としてバイアスを持つ

$\Rightarrow$ **バイアスの補正が必要:**

$$\mathbb{E}_{\mathbf{X}_n}\left[\mathbb{H}_{i,n}(\hat{\theta}_{i,n}) - n\mathbb{E}_Z\left\{\log f_{i,n}(Z,\hat{\theta}_{i,n})\right\}\right] \approx p_i$$

$\Rightarrow$ **AIC の導出:**

$$\mathrm{AIC}_n^{(i)} = -2\mathbb{H}_{i,n}(\hat{\theta}_{i,n}) + 2p_i, \quad i = 1,\ldots,K$$

- **各候補モデルごとに AIC を計算し, 最小の値をとるモデルを選択**

Bayes 情報量規準 BIC

- Bayes 情報量規準 (BIC): モデルの評価をモデルの事後確率から捉える
- データ x_n に対する候補モデル $\mathcal{M}_i$ の事後確率: $i = 1, \ldots, K$

$$\mathbb{P}(\mathcal{M}_i|\mathrm{x}_n) = \frac{h_{i,n}(\mathrm{x}_n)\mathfrak{p}_i}{\sum_{i=1}^{K} h_{i,n}(\mathrm{x}_n)\mathfrak{p}_i}$$

$\cdots$ データ x_n が得られたとき, それがモデル $\mathcal{M}_i$ から生起するものである確率

$\Rightarrow$ 確率が最大となるモデルを選択

- $h_{i,n}(\mathrm{x}_n)$: データ x_n に対するモデル $\mathcal{M}_i$ の周辺尤度

$$h_{i,n}(\mathrm{x}_n) = \int_{\Theta_i} \exp\left\{\mathbb{H}_{i,n}(\theta_i)\right\} \pi_{i,n}(\theta_i) d\theta_i$$

- $\pi_{i,n}(\theta_i)$: パラメータ θ_i の事前分布, $\mathfrak{p}_i$: モデル $\mathcal{M}_i$ の生起確率

事後確率: $i = 1, \dots, K$

$$\mathbb{P}(\mathcal{M}_i|\mathrm{x}_n) = \frac{h_{i,n}(\mathrm{x}_n)\mathfrak{p}_i}{\sum_{i=1}^{K} h_{i,n}(\mathrm{x}_n)\mathfrak{p}_i}$$

- 事後確率の最大化 $\Rightarrow$ 分子の最大化 (分母は全てのモデルで共通)

$$\Downarrow \ (\mathfrak{p}_1 = \cdots = \mathfrak{p}_K)$$

モデルの周辺尤度 $h_{i,n}$ の最大化

$$\Downarrow$$

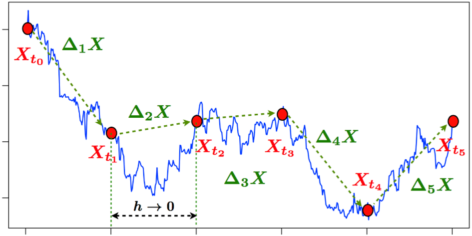
対数周辺尤度 $\log h_{i,n}$ を使いやすい形で表現

$\Rightarrow$ BIC の導出:

$$\mathrm{BIC}_n^{(i)} = -2\mathbb{H}_{i,n}(\hat{\theta}_{i,n}) + p_i \log n, \quad i = 1, \dots, K$$

- 各候補モデルごとに BIC を計算し, 最小の値をとるモデルを選択

モデル設定: 確率微分方程式モデル



- データ: $\mathbf{X}_n = (X_{t_j})_{j=0}^n$
  - $t_j = jh_n$
  - $T_n = nh_n \to \infty,\ nh_n^2 \to 0$

- 候補モデル $\mathcal{M}_{k,\ell}$ $(k \le K, \ell \le L)$:

$$dX_t = a_k(X_t, \alpha_k)dt + b_\ell(X_t, \beta_\ell)dw_t, \quad t \in [0, T_n]$$

  - 候補ドリフト項: $a_1(X_t, \alpha_1), \ldots, a_K(X_t, \alpha_K)$ $\cdots$ 選択の対象
  - 候補拡散項: $b_1(X_t, \beta_1), \ldots, b_L(X_t, \beta_L)$ $\cdots$ 選択の対象
  - パラメータ: $\theta_{k,\ell} = (\alpha_k, \beta_\ell)$ $\cdots$ 推定の対象
  - w: 標準 Wiener 過程

確率微分方程式モデルにおける情報量規準: 同時

$$dX_t = a(X_t, \alpha)dt + b(X_t, \beta)dw_t$$

- $\mathcal{L}(X_{t_j} | X_{t_{j-1}}) \approx N\big(X_{t_{j-1}} + h_n a_{j-1}(\alpha), h_n S_{j-1}(\beta)\big)$

$$\mathbb{H}_n(\theta) = \sum_{j=1}^{n} \log \phi\Big(X_{t_j}; X_{t_{j-1}} + h_n a_{j-1}(\alpha), h_n S_{j-1}(\beta)\Big)$$

  - $f_{j-1}(\cdot) = f(X_{t_{j-1}}, \cdot)$, $S_{j-1}(\beta) = b_{j-1} b_{j-1}^{\top}(\beta)$
  - $\theta = (\alpha, \beta) \in \mathbb{R}^{p_\alpha} \times \mathbb{R}^{p_\beta}$, $\hat{\theta}_n \in \mathrm{argmax}_\theta \, \mathbb{H}_n(\theta)$

CIC (Uchida(2010))：AIC 型情報量規準

$$\mathrm{CIC}_n = -2\mathbb{H}_n(\hat{\theta}_n) + 2(p_\alpha + p_\beta)$$

---

quasi-BIC and BIC (E and Masuda(2018))：BIC 型情報量規準

$$\mathrm{QBIC}_n = -2\mathbb{H}_n(\hat{\theta}_n) + \log\left|-\partial_\alpha^2 \mathbb{H}_n(\hat{\theta}_n)\right| + \log\left|-\partial_\beta^2 \mathbb{H}_n(\hat{\theta}_n)\right|$$

$$\mathrm{BIC}_n = -2\mathbb{H}_n(\hat{\theta}_n) + p_\alpha \log T_n + p_\beta \log n$$

確率微分方程式モデルにおける情報量規準: 二段階

候補モデル $\mathcal{M}_{k,\ell}$ $(k \le K, \ell \le L)$:

$$dX_t = a_k(X_t, \alpha_k)dt + b_\ell(X_t, \beta_\ell)dw_t$$

(i) α を含まない (補助) 対数尤度関数 $\mathbb{H}_{1,n}^{(\ell)}(\beta_\ell)$ を用いて拡散項を選択:

$$\mathrm{QBIC}_n^{(k)} = -2\mathbb{H}_{1,n}^{(k)}(\hat{\beta}_{\ell,n}) + \log\left| -\partial_{\beta_\ell}^2 \mathbb{H}_{1,n}^{(\ell)}(\hat{\beta}_{\ell,n})\right|$$

$$\{\ell_n^*\} = \operatorname*{argmin}_{1\le\ell\le L} \mathrm{QBIC}_n^{(\ell)}$$

(ii) ドリフト項を選択: $\underline{dX_t = a_k(X_t, \alpha_k)dt + b_{\ell_n^*}(X_t, \hat{\beta}_{\ell_n^*,n})dw_t}$

$$\mathrm{QBIC}_n^{(k|\ell_n^*)} = -2\mathbb{H}_n^{(k,\ell_n^*)}(\hat{\alpha}_{k,n}, \hat{\beta}_{\ell_n^*,n} big) + \log\left| -\partial_{\beta_\ell}^2 \mathbb{H}_n^{(k_n^*,\ell)}(\hat{\alpha}_{k,n}, \hat{\beta}_{\ell_n^*,n})\right|$$

$$\{k_n^*\} = \operatorname*{argmin}_{1\le k\le K} \mathrm{QBIC}_n^{(k|\ell_n^*)}$$

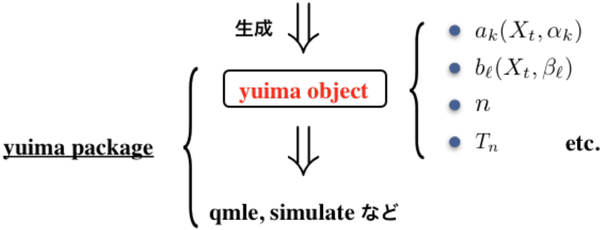
※ BIC についても同様の二段階選択が可能

1. 情報量規準
2. 関数 “IC”
3. シミュレーション・実習

yuima object

候補モデル： $\mathcal{M}_{k,\ell} = \{a_k(\cdot,\cdot), b_\ell(\cdot,\cdot)\},\quad k = 1,\dots,K,\quad \ell = 1,\dots,L$

$$dX_t = a_k(X_t,\alpha_k)dt + b_\ell(X_t,\beta_\ell)dw_t,\quad t\in[0,T_n]$$



- 候補モデルごとに yuima object を生成 (生成例は次のスライド)
  - モデルの各項 (拡散項 a やドリフト項 b), データ数 n などの情報を持つ
- yuima object はパラメータの推定やデータ生成ための関数に利用可能

生成例

$$dX_t = \alpha_1 X_t dt + \sqrt{\beta_1} dw_t, \quad t \in [0, T_n]$$

- $n = 1000$：データ数, $T_n = n^{1/3}$

```
> ymodel <- setModel(drift = "alpha1*x", diffusion = "sqrt(beta1)")
 警告メッセージ:
 yuima.warn("Solution variable (lhs) not specified. Trying to use state variables.") で:

YUIMA: Solution variable (lhs) not specified. Trying to use state variables.

> ysamp <- setSampling(Terminal = 1000^(1/3), n = 1000)
 警告メッセージ:
 yuima.warn("'delta' (re)defined.") で:
YUIMA: 'delta' (re)defined.

> yuima <- setYuima(model = ymodel, sampling = ysamp)
```

関数 "IC"

```
IC <- function(yuima, data, start, lower, upper,
                                joint = FALSE, rcpp = FALSE)
```

- **yuima: モデルの設定 (yuima object)**
- **data: データセット** $(X_{t_0}, X_{t_1}, \ldots, X_{t_n})$
- **start, lower, upper: パラメータの推定のための名前付きリスト**

```
> init <- list(alpha1 = -1, beta1 = 0.5)
> lower <- list(alpha1 = -3, beta1 = 0.001)
> upper <- list(alpha1 = -0.001, beta1 = 3)
```

- **joint: パラメータの推定を同時に行うか，または二段階で行うか**
- **rcpp: C++ code を使用するか否か**

返り値

- パラメータの推定量 ($par)
- 情報量規準 QBIC, BIC, CIC の値 ($QBIC, $BIC, $CIC)

```
> ic <- IC(yuima = yuima, data = Xt, start = init, lower = lower,
+          upper = upper, rcpp = TRUE)
```

```
> ic
$par
    beta1    alpha1
 1.965465 -1.583359

$BIC
[1] -1083.488

$QBIC
[1] -1086.631

$CIC
[1] -1088.698
```

```
> ic$par
    beta1    alpha1
 1.965465 -1.583359
>
> ic$BIC
[1] -1083.488
>
> ic$QBIC
[1] -1086.631
>
> ic$CIC
[1] -1088.698
```

注意

- 変数 start, lower, upper について:

  モデルに含まれているパラメータ全てに関する設定が必須

```
> ymodel <- setModel(drift = "alpha1*x", diffusion = "sqrt(beta1)")
```

× ⇐

```
> init <- list(alpha1 = -1)
> init <- list(gamma1 = -1, beta1 = 0.5)
```

○ ⇐

```
>
> init <- list(alpha1 = -1, alpha2 = -1, beta1 = 0.5, beta2 = 0.5)
```

- QBIC の計算について:

  $\det\left(-\partial_\alpha^2 \mathbb{H}_n(\hat{\theta}_n)\right) \leq 0$ または $\det\left(-\partial_\beta^2 \mathbb{H}_n(\hat{\theta}_n)\right) \leq 0$

  $\Rightarrow$ QBIC は BIC と同じ値をとる ($\mathrm{QBIC}_n = \mathrm{BIC}_n$)

実行例: 一次元データ

- 真のモデル:

$$dX_t = -X_t dt + \exp\left\{\frac{1}{2}(-2\sin X_t + 1)\right\} dw_t, \quad X_0 = 1$$

  - $n = 1000$, $T_n = n^{1/3}$

- 候補モデル:

$$\text{Model 1}: dX_t = \alpha_1 X_t dt + \exp\left\{\frac{1}{2}(\beta_1 \sin X_t + \beta_2)\right\} dw_t$$

$$\text{Model 2}: dX_t = \alpha_1 X_t dt + \exp\left\{\frac{1}{2}(\beta_1 \sin X_t)\right\} dw_t$$

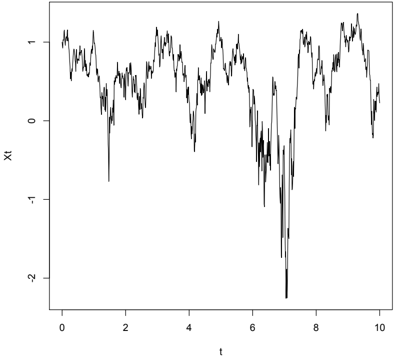
- 真のパラメータ: $(\alpha_{1,0}, \beta_{1,0}, \beta_{2,0}) = (-1, -2, 1)$

```
> set.seed(123)
> model <- setModel(drift = "alpha1*x", diffusion = "exp((beta1*sin(x)+beta2)/2)")
 警告メッセージ:
 yuima.warn("Solution variable (lhs) not specified. Trying to use state variables.") で:

YUIMA: Solution variable (lhs) not specified. Trying to use state variables.

> sample <- setSampling(Terminal = 1000^(1/3), n = 1000)
 警告メッセージ:
 yuima.warn("'delta' (re)defined.") で:
YUIMA: 'delta' (re)defined.

> yuima <- setYuima(model = model, sampling = sample)
> simu <- simulate(yuima, xinit = 1,
+                  true.parameter = list(beta1 = -2, beta2 = 1, alpha1 = -1))
> Xt <- simu@data@original.data
```



- **yuima object 生成:**

Model 1:

```
> model1 <- setModel(drift = "alpha1*x", diffusion = "exp((beta1*sin(x)+beta2)/2)")
 警告メッセージ:
 yuima.warn("Solution variable (lhs) not specified. Trying to use state variables.") で:

YUIMA: Solution variable (lhs) not specified. Trying to use state variables.

> sample1 <- setSampling(Terminal = 1000^(1/3), n = 1000)
 警告メッセージ:
 yuima.warn("'delta' (re)defined.") で:
YUIMA: 'delta' (re)defined.

> yuima1 <- setYuima(model = model1, sampling = sample1)
```

Model 2:

```
> model2 <- setModel(drift = "alpha1*x", diffusion = "exp(beta1*sin(x)/2)")
 警告メッセージ:
 yuima.warn("Solution variable (lhs) not specified. Trying to use state variables.") で:

YUIMA: Solution variable (lhs) not specified. Trying to use state variables.

> sample2 <- setSampling(Terminal = 1000^(1/3), n = 1000)
 警告メッセージ:
 yuima.warn("'delta' (re)defined.") で:
YUIMA: 'delta' (re)defined.

> yuima2 <- setYuima(model = model2, sampling = sample2)
```

```
> Xt <- yuima.simu@data@original.data
```

- **パラメータ設定 $(\alpha_1, \beta_1, \beta_2)$:**

```
> init <- list(alpha1 = -0.5, beta1 = 0.5, beta2 = 0.5)
> low <- list(alpha1 = -5, beta1 = -5, beta2 = -5)
> upp <- list(alpha1 = -0.001, beta1 = 5, beta2 = 5)
```

- **“IC” の使用:**

Model 1:

```
> ic1 <- IC(yuima = yuima1, data = Xt, start = init, lower = low, upper = upp)
```

Model 2:

```
> ic2 <- IC(yuima = yuima2, data = Xt, start = init, lower = low, upper = upp)
```

- **計算結果比較**

Model 1:

```
> ic1
$par
    beta1     beta2    alpha1
-2.082683  1.022628 -1.085644

$BIC
[1] -1684.132

$QBIC
[1] -1687.361

$CIC
[1] -1694.25
```

Model 2:

```
> ic2
$par
    beta1    alpha1
-1.132763 -1.158624

$BIC
[1] -1292.77

$QBIC
[1] -1293.673

$CIC
[1] -1297.981
```

$$\begin{cases} (\text{Model 1 の BIC}) & < & (\text{Model 2 の BIC}) \\ (\text{Model 1 の QBIC}) & < & (\text{Model 2 の QBIC}) \\ (\text{Model 1 の CIC}) & < & (\text{Model 2 の CIC}) \end{cases}$$

実行例: 多次元データ

$$d\begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} -2X_{1,t} \\ -X_{2,t} \end{pmatrix} dt + \begin{pmatrix} -X_{1,t}+2 & -1 \\ 1 & -X_{1,t}+2 \end{pmatrix} d\begin{pmatrix} w_{1,t} \\ w_{2,t} \end{pmatrix}$$

- $X_0 = (1,1)'$
- $n = 2000,\ T_n = n^{1/3}$
- 真のパラメータ: $(\alpha_{1,0}, \alpha_{2,0}, \beta_{1,0}, \beta_{2,0}, \beta_{3,0}) = (-2,-1,-1,0,2)$
- 候補モデル:

$$d\begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} \alpha_1 X_{1,t} \\ \alpha_2 X_{2,t} \end{pmatrix} dt + \begin{pmatrix} b(X_t,\beta) & -1 \\ 1 & b(X_t,\beta) \end{pmatrix} d\begin{pmatrix} w_{1,t} \\ w_{2,t} \end{pmatrix}$$

Model 1 $: b(X_t,\beta) = \beta_1 X_{t,1} + \beta_2 X_{t,2} + \beta_3;$

Model 2 $: b(X_t,\beta) = \beta_1 X_{t,1} + \beta_2 X_{t,2};$

Model 3 $: b(X_t,\beta) = \beta_1 X_{t,1} + \beta_3;$ Model 4 $: b(X_t,\beta) = \beta_2 X_{t,2} + \beta_3;$

Model 5 $: b(X_t,\beta) = \beta_1 X_{t,1};$ Model 6 $: b(X_t,\beta) = \beta_2 X_{t,2};$

Model 7 $: b(X_t,\beta) = \beta_3;$

- 拡散項 (行列) $\begin{pmatrix} -X_{1,t}+2 & -1 \\ 1 & -X_{1,t}+2 \end{pmatrix}$:

```
> diff.coef.matrix <- matrix(c("beta1*x1+beta3", "1", "-1", "beta1*x1+beta3"), 2, 2)
```

- ドリフト項 (ベクトル) $\begin{pmatrix} -2X_{1,t} \\ -X_{2,t} \end{pmatrix}$:

```
> drif.coef.vec <- c("alpha1*x1", "alpha2*x2")
```

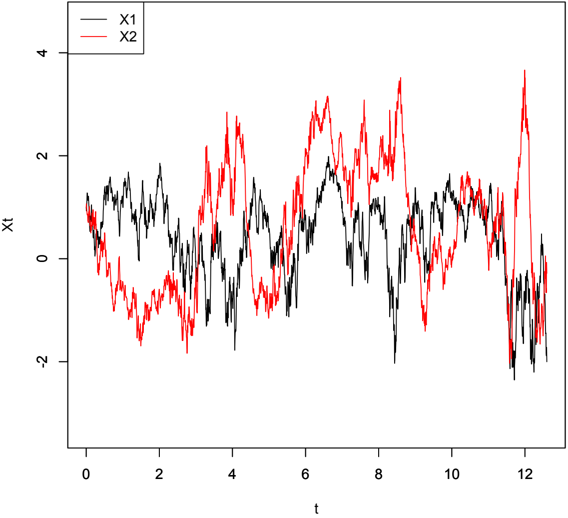
- yuima object:

```
> model <- setModel(drift = drif.coef.vec, diffusion = diff.coef.matrix,
+                   state.variable = c("x1", "x2"), solve.variable = c("x1", "x2"))
> sample <- setSampling(Terminal = 2000^(1/3), n = 2000)
 警告メッセージ:
 yuima.warn("'delta' (re)defined.") で:
YUIMA: 'delta' (re)defined.

> yuima <- setYuima(model = model, sampling = sample)
```

- **データ生成:**

```
> set.seed(200)
> simu <- simulate(yuima, xinit = c(1,1),
+                  true.parameter = list(beta1 = -1, beta3 = 2, alpha1 = -2, alpha2 = -1))
> Xt <- simu@data@original.data
```



- **候補モデルの yuima object:**

```
> cand.diff <- list(matrix(c("beta1*x1+beta2*x2+beta3", "1", "-1", "beta1*x1+beta2*x2+beta3"), 2, 2),
+                   matrix(c("beta1*x1+beta2*x2", "1", "-1", "beta1*x1+beta2*x2"), 2, 2),
+                   matrix(c("beta1*x1+beta3", "1", "-1", "beta1*x1+beta3"), 2, 2),
+                   matrix(c("beta2*x2+beta3", "1", "-1", "beta2*x2+beta3"), 2, 2),
+                   matrix(c("beta1*x1", "1", "-1", "beta1*x1"), 2, 2),
+                   matrix(c("beta2*x2", "1", "-1", "beta2*x2"), 2, 2),
+                   matrix(c("beta3", "1", "-1", "beta3"), 2, 2))
>
> cand.drif <- c("alpha1*x1", "alpha2*x2")
>
> cand.model <- NULL
> for(i in 1:length(cand.diff)){
+   cand <- setModel(drift = cand.drif, diffusion = cand.diff[[i]],
+                    state.variable = c("x1", "x2"), solve.variable = c("x1", "x2"))
+     cand.samp <- setSampling(Terminal = 2000^(1/3), n = 2000)
+     cand.yuima <- setYuima(model = cand, sampling = cand.samp)
+     cand.model <- c(cand.model, list(cand.yuima))
+ }
```

候補拡散項

候補ドリフト項

全候補モデルのyuima object

- **パラメータの設定:**

  最もパラメータが多い候補モデルに合わせて設定する

```
> init <- list(alpha1 = -1, alpha2 = -1, beta1 = 0.5, beta2 = 0.5, beta3 = 0.5)
> low <- list(alpha1 = -3, alpha2 = -3, beta1 = -5, beta2 = -5, beta3 = -5)
> upp <- list(alpha1 = -0.001, alpha2 = -0.001, beta1 = 5, beta2 = 5, beta3 = 5)
```

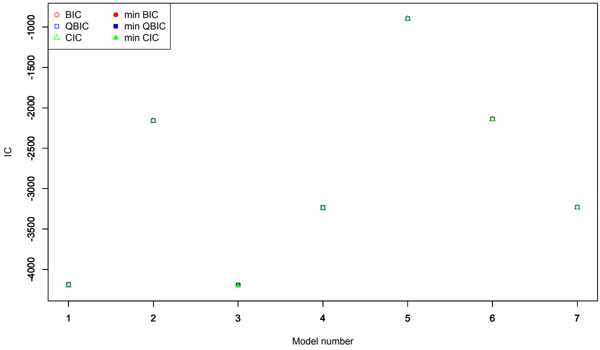
- **“IC”” の使用:**

```
> for(i in 1:7){
+    ic <- IC(yuima = cand.model[[i]], data = Xt, start = init, lower = low, upper = upp)
+ }
```

- **“IC”” の使用 & 結果をまとめる (例):**

```
> paranum <- list(1:5, c(1,2,4,5), c(1,3,4,5), c(2,3,4,5), c(1,4,5), c(2,4,5), c(3,4,5))
> results <- matrix(0, 7, 8)
> for(i in 1:7){
+    ic <- IC(yuima = cand.model[[i]], data = Xt, start = init, lower = low, upper = upp)
+    results[i,paranum[[i]]] <- ic$par
+    results[i,6:8] <- c(ic$BIC, ic$QBIC, ic$CIC)
+ }

> results
              beta1        beta2     beta3    alpha1     alpha2       BIC      QBIC        CIC
Model 1 -1.0583542  0.001811267  2.018310 -1.770173 -0.8747962 -4182.2607 -4184.882 -4200.1307
Model 2 -0.8069696  1.362905768  0.000000 -1.974062 -0.5502677 -2153.1532 -2158.122 -2165.4223
Model 3 -1.0586496  0.000000000  2.019548 -1.769961 -0.8739008 -4189.8550 -4192.581 -4202.1241
Model 4  0.0000000  0.059506439  1.732918 -2.336065 -1.0604995 -3230.4508 -3232.794 -3242.7199
Model 5 -2.1998290  0.000000000  0.000000 -1.851133 -1.4337826  -892.1917  -896.013  -898.8598
Model 6  0.0000000 -1.458342919  0.000000 -1.555513 -0.3980473 -2134.7236 -2137.536 -2141.3918
Model 7  0.0000000  0.000000000 -1.771589 -2.362494 -1.0457829 -3228.2303 -3230.866 -3234.8985
```



真のパラメータ: $(\alpha_{1,0}, \alpha_{2,0}, \beta_{1,0}, \beta_{2,0}, \beta_{3,0}) = (-2, -1, -1, 0, 2)$

	β_1	β_2	β_3	α_1	α_2	BIC	QBIC	CIC
Model 1	-1.058	0.002	2.018	-1.770	-0.875	-4182.261	-4184.882	-4200.131
Model 2	-0.807	1.363	–	-1.974	-0.550	-2153.153	-2158.122	-2165.422
Model 3	-1.059	–	2.020	-1.770	-0.874	-4189.855	-4192.581	-4202.124
Model 4	–	0.060	1.733	-2.336	-1.060	-3230.451	-3232.794	-3242.720
Model 5	-2.200	–	–	-1.851	-1.434	-892.192	-896.013	-898.860
Model 6	–	-1.458	–	-1.556	-0.398	-2134.724	-2137.536	-2141.392
Model 7	–	–	-1.772	-2.362	-1.046	-3228.230	-3230.866	-3234.898

参考文献

- **Specification of the function IC in YUIMA package (https://sites.google.com/site/shoichieguchi821/shi-yang-shu)**

[1] **Eguchi, S. and Masuda, H. (2018). Schwarz type model comparison for LAQ models.** *Bernoulli*, **24(3), 2278–2327.**

[2] **Uchida, M. (2010). Contrast-based information criterion for ergodic diffusion processes from discrete observations.** *Ann. Inst. Statist. Math.*, **62, 161–187.**