

非定常情報源に対する resetting 分布を用いたモデル系列推定

独立行政法人 産業技術総合研究所
サービス工学研究センター
大規模データモデリング研究チーム
櫻井 瑛一

2011/6/22 第2回 Latent Dynamics 研究会
Joint work with 山西 健司(東大)

目次

1. イントロダクション

サービス工学研究センターの紹介+非定常性を考える意義

2. 非定常性への従来の取り組み

従来手法の紹介

3. 動的モデル選択規準

今回モデル系列推定に用いる規準について

4. Resetting 分布と switching 分布

規準に対してどうデータを説明するのか

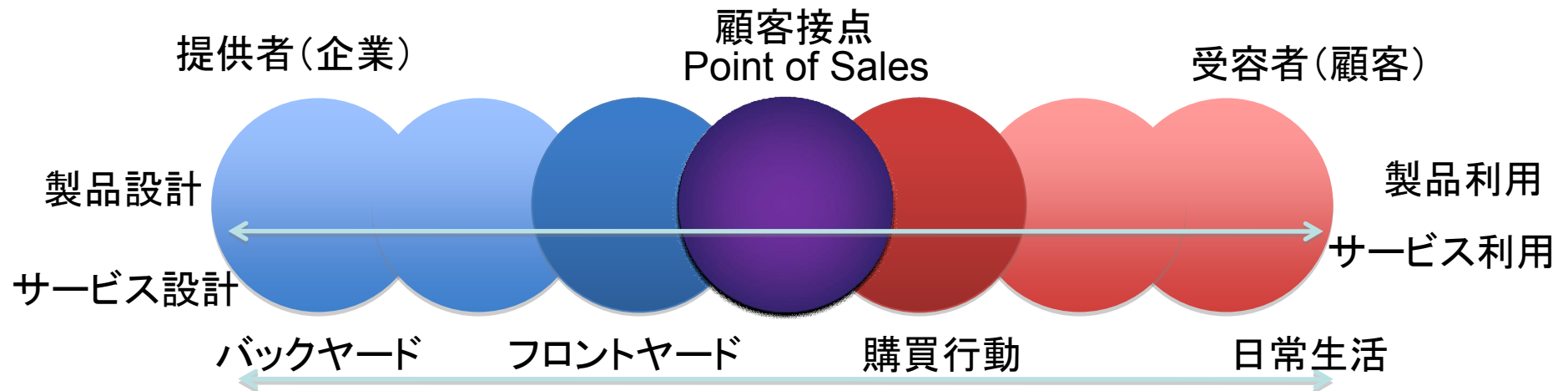
5. 数値実験

6. まとめ

所属センターの紹介(1)

・ サービス工学？

－ サービス場面の 人々の活動 を対象としたシステム工学



- － 観測・モデル化・予測や制御, 最適化を行う
- － それにより「**良いサービス**」を工学的に実現する
- － 最終的な目標はサービスの生産性向上

所属センターの紹介(2)

- 具体的には
 - POS データからの顧客セグメント分類と購買行動分析
 - 飲食店における客数などの需要予測
 - 温泉地における客数移動把握
 - etc.
- 幅広いサービス分野に対して
科学工学的なアプローチによって
サービスの解明と改善の提案を行う

サービスにおける非定常性

- 今までは
定常ではない状態でも定常として解析
 - サービス場面において
 - 顧客の変動
 - 時期変動
 - 景気変動
- などの非定常性が存在
- 隠れた変動に着目する必要がある

目次

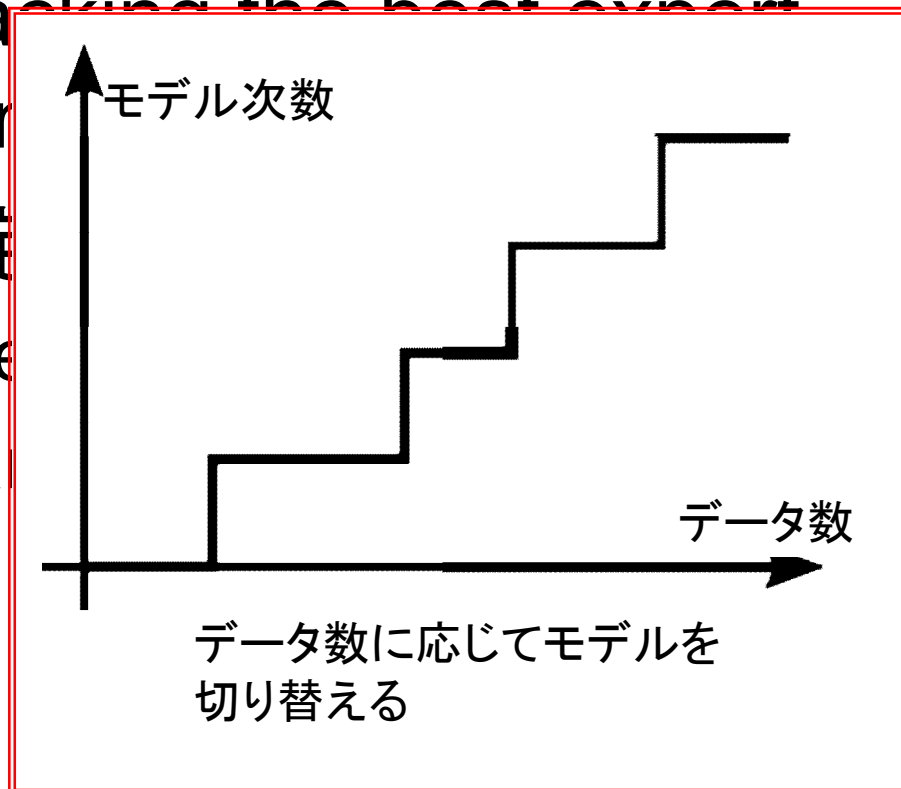
1. イントロダクション
2. 非定常性への従来の取り組み
3. 動的モデル選択規準
4. Resetting 分布と switching 分布
5. 数値実験
6. まとめ

非定常性に対する従来の取り組み

- モデルが切り替わるということ
 - 1) Switching (AIC-BIC dilemma の解決)
 - 2) Tracking the best expert
 - 3) Dynamic model selection (DMS)
- 区間定常情報源のデータ圧縮
 - Willem's 1996
 - Shamir and Merhav 1999

非定常性に対する従来の取り組み

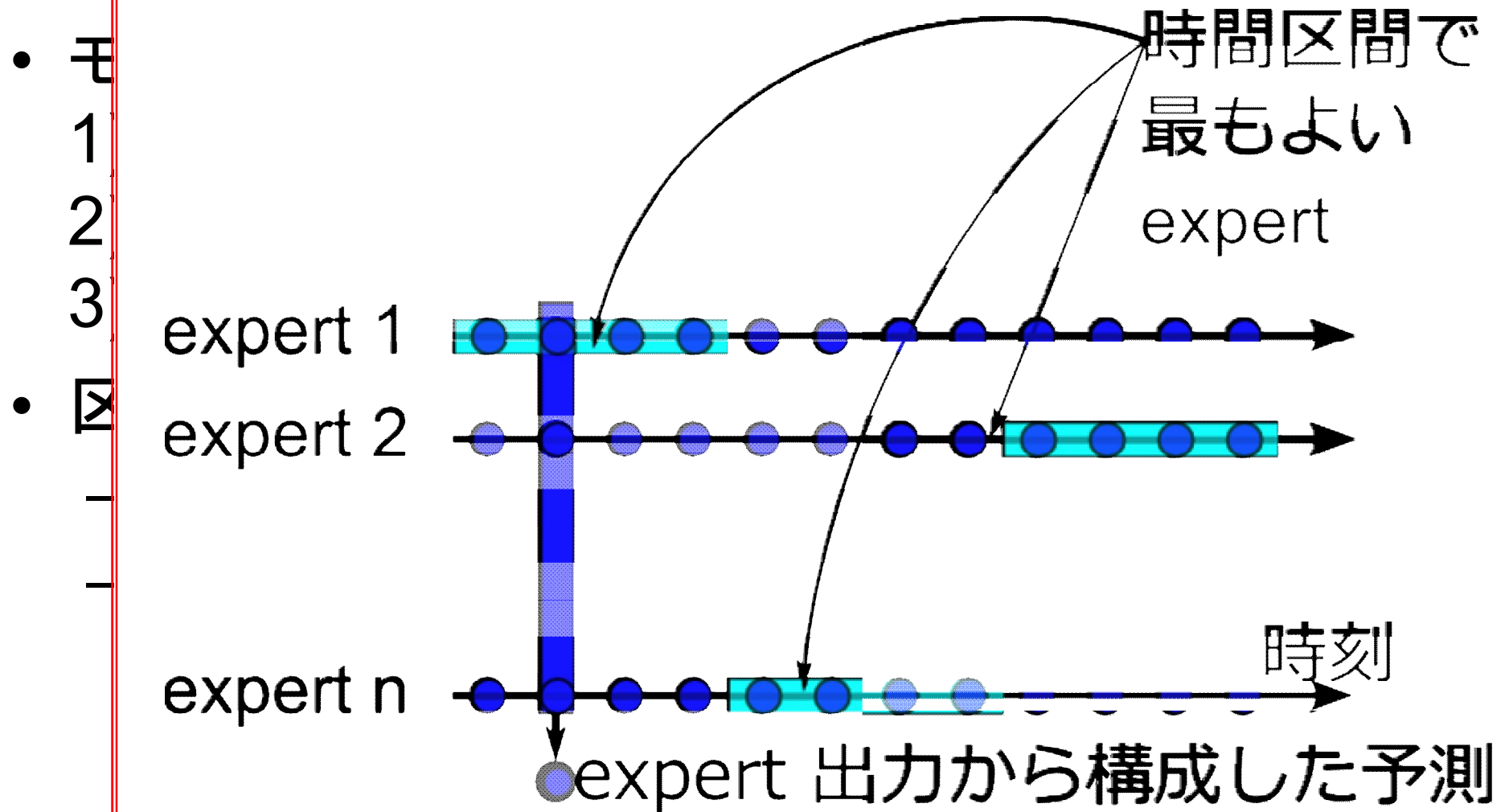
- モデルが切り替わるということ
 - 1) Switching (AIC-BIC dilemma の解決)
 - 2) Tracking the best expert
 - 3) Dynamic (DMS)
- 区間定
 - Wille
 - Sha



非定常性に対する従来の取り組み

- モデルが切り替わるということ
 - 1) Switching (AIC-BIC dilemma の解決)
 - 2) Tracking the best expert
 - 3) Dynamic model selection (DMS)
- 区間定常情報源のデータ圧縮
 - Willem's 1996
 - Shamir and Merhav 1999

非定常性に対する従来の取り組み



非定常性に対する従来の取り組み

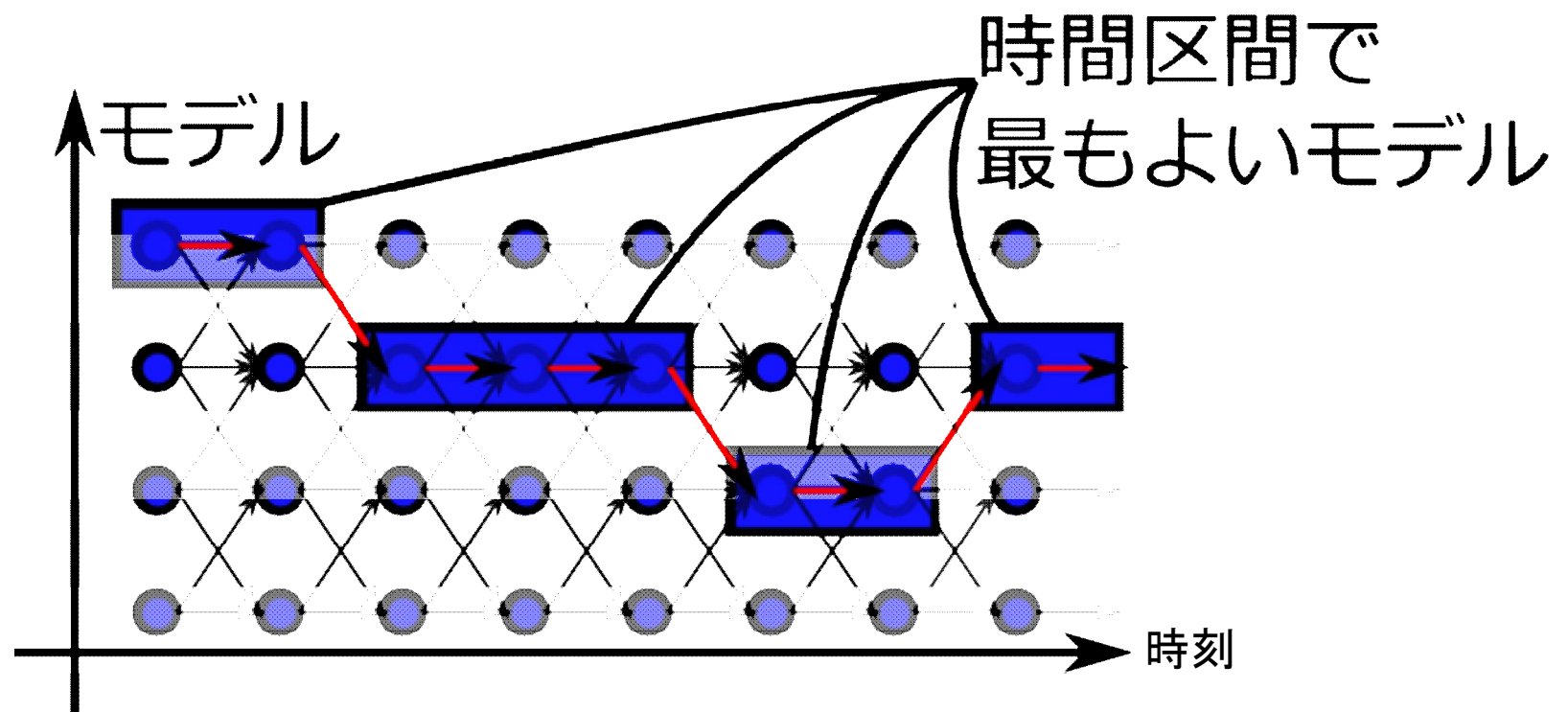
- モデルが切り替わるということ
 - 1) Switching (AIC-BIC dilemma の解決)
 - 2) Tracking the best expert
 - 3) Dynamic model selection (DMS)
- 区間定常情報源のデータ圧縮
 - Willem's 1996
 - Shamir and Merhav 1999

非定常性に対する従来の取り組み

- モデルが変化するシーケンス

- 1)
- 2)
- 3)

- 区間



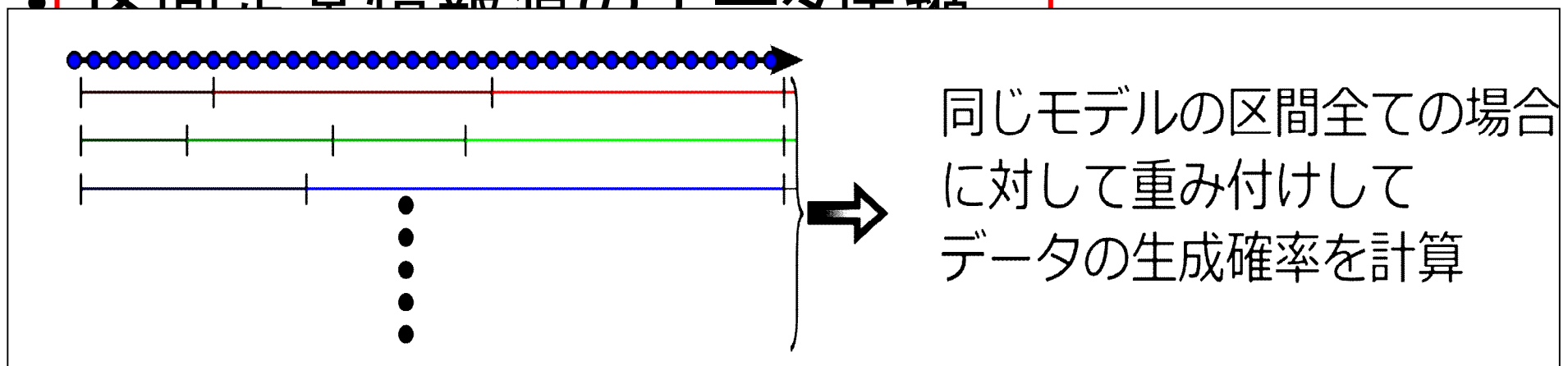
非定常性に対する従来の取り組み

- モデルが切り替わるということ
 - 1) Switching (AIC-BIC dilemma の解決)
 - 2) Tracking the best expert
 - 3) Dynamic model selection (DMS)
- 区間定常情報源のデータ圧縮
 - Willem's 1996
 - Shamir and Merhav 1999

非定常性に対する従来の取り組み

- モデルが切り替わるということ
 - 1) Switching (AIC-BIC dilemma の解決)
 - 2) Tracking the best expert
 - 3) Dynamic model selection (DMS)

区間定常情報のデータ圧縮



本発表の立ち位置

- モデルが切り替わるということ → Switching 分布
 - 1) Switching (AIC-BIC dilemma の解決)
 - 2) Tracking the best expert
 - 3) Dynamic model selection (モデルの系列の選択)
- 区間定常情報源のデータ圧縮 → Resetting 分布
 - Willem's 1996
 - Shamir and Merhav 1999

情報論的評価

計算量

目次

1. イントロダクション
2. 非定常性への従来の取り組み
3. 動的モデル選択規準
4. Resetting 分布と switching 分布
5. 数値実験
6. まとめ

動的モデル選択とは？

- 動的モデル選択はYamanishi and Maruyama 2007
にて提案された非定常な情報源から
モデルの系列を選択する問題
- ここでは非定常性はデータ生成源のモデルが
時間とともに変化することにより現れると設定
- モデル選択におけるMDL 原理
(Minimum Description Length principle)
に基づき規準を定義

動的モデル選択規準の概要

- MDL 原理に基づき

$$\left(\begin{array}{c} \text{モデルの} \textcolor{red}{\text{系列}} \text{自身} \\ \text{の記述長} \end{array} \right) + \left(\begin{array}{c} \text{モデル} \textcolor{red}{\text{系列}} \text{の下での} \\ \text{データの記述長} \end{array} \right)$$

の和を最小にするモデル系列を選ぶ

-
- cf. MDL 原理

$$\left(\begin{array}{c} \text{モデル自身} \\ \text{の記述長} \end{array} \right) + \left(\begin{array}{c} \text{モデルの下での} \\ \text{データの記述長} \end{array} \right)$$

の和を最小にするモデルを選ぶ

動的モデル選択規準の準備

- データを説明するモデル

- パラメトリックな確率モデルクラス

$$\mathcal{C}_k = \{P_k(\cdot|\theta_k)\} (k = 1, \dots, K) \quad \dim(\theta_1) < \dots < \dim(\theta_K)$$

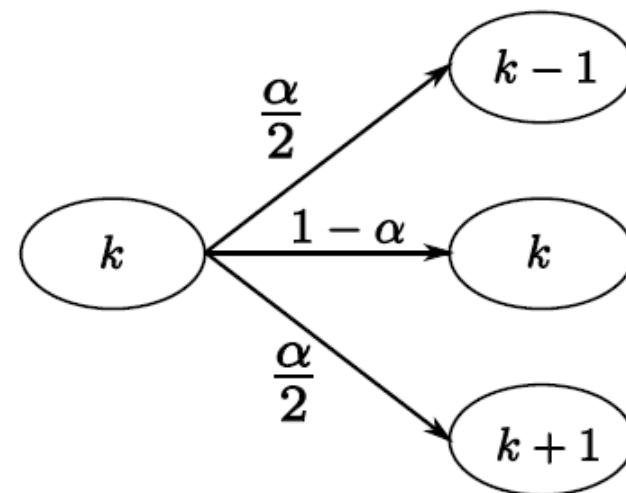
- 逐次的なパラメータ推定 $\theta_k(x^{t-1}) = \theta_k^{(t)}$

- モデルの遷移

マルコフ的に隣のモデルに遷移
 α は Krichevsky–Trofimov 推定
 を使用

$$\hat{\alpha}(k^t) = \frac{n(k^t) + 1/2}{t}$$

$n(k^t)$ = 今までにモデル変化した回数



動的モデル選択規準：DMS 規準

- MDL 原理に基づき

（ モデルの 系列 自身 ）
+
（ モデル 系列 の下での データの記述長 ）

の和を最小にするモデル系列を選ぶ

$$\ell(x^T, k^T) = \underbrace{\sum_{t=1}^T \boxed{-\log P(x_t | \theta_{k_t}^{(t)})}}_{\text{データの記述長}} + \underbrace{\sum_{t=1}^T \boxed{-\log P(k_t | k_{t-1} : \hat{\alpha}(k^{t-1}))}}_{\text{モデル系列の記述長}}$$

予測的記述長として計算 (予測的 MDL)

動的モデル選択規準の特徴

1. 予測的符号長(予測的MDL)で計算できる
2. 定義からモデル系列は動的計画法を用いて探索することができる(計算効率性がよい)
→ 実際 DMS では探索コストは $O(T^2)$
3. 情報論的に評価が可能

$$\ell(x^T, k^T) = \underbrace{\sum_{t=1}^T \boxed{-\log P(x_t | \theta_{k_t}^{(t)})}}_{\text{データの記述長}} + \underbrace{\sum_{t=1}^T \boxed{-\log P(k_t | k_{t-1} : \hat{\alpha}(k^{t-1}))}}_{\text{モデル系列の記述長}}$$

予測的記述長として計算 (予測的 MDL)

情報論的評価

DMS 規準を最小化するモデル系列による総記述長は次のように抑えられる

$$\ell(x^T, k^T) \leq \min_m \left\{ \underbrace{\min_{(t_1, \dots, t_m)} \min_{(k(0), \dots, k(m))} \sum_{j=0}^{m-1} \sum_{t=t_j}^{t_{j+1}-1} -\log P(x_t | \theta_{k(j)}^{(t)})}_{\text{データの記述}} \right. \\ \left. + (T-1)H\left(\frac{m}{T-1}\right) + \frac{1}{2} \log(T-1) \right. \\ \left. + (m+1) \log 2 + \log K \right\} \quad \text{追加損失}$$

ただし, H はエントロピー, m は変数, $\{t_0, \dots, t_{m+1}\}$ はモデルの変化時刻, $k(j)$ は時間区間 $[t_j, t_{j+1} - 1]$ の間のモデル

Yamanishi and Maruyama 2007

本発表のポイント

- 動的モデル選択規準のどこに着目するのか
→モデル系列の下でのデータの記述長をより小さくできるのではないかな？

データの記述長を短くし計算効率性を落とさない
パラメータ推定法は？

$$\ell(x^T, k^T) = \underbrace{\sum_{t=1}^T \boxed{-\log P(x_t | \theta_{k_t}^{(t)})}}_{\text{データの記述長}} + \underbrace{\sum_{t=1}^T \boxed{-\log P(k_t | k_{t-1} : \hat{\alpha}(k^{t-1}))}}_{\text{モデル系列の記述長}}$$

予測的記述長として計算 (予測的 MDL)

目次

1. イントロダクション
2. 非定常性への従来の取り組み
3. 動的モデル選択規準
4. Resetting 分布と switching 分布
5. 数値実験
6. まとめ

概略

データの記述長を短くし計算効率性を落とさない
パラメータ推定法は？



変化するモデル系列に対応する
データを説明する分布を考える

- Resetting 分布
- Switching 分布

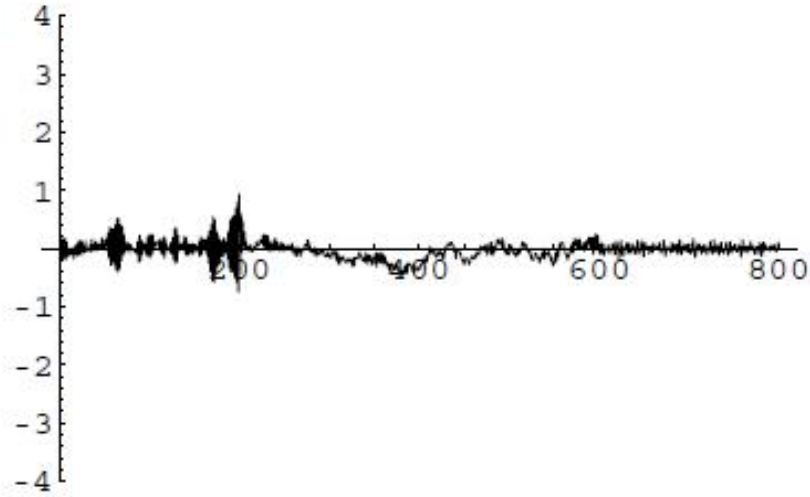
目次

1. イントロダクション
2. 非定常性への従来の取り組み
3. 動的モデル選択規準
4. Resetting 分布と switching 分布
5. 数値実験
6. まとめ

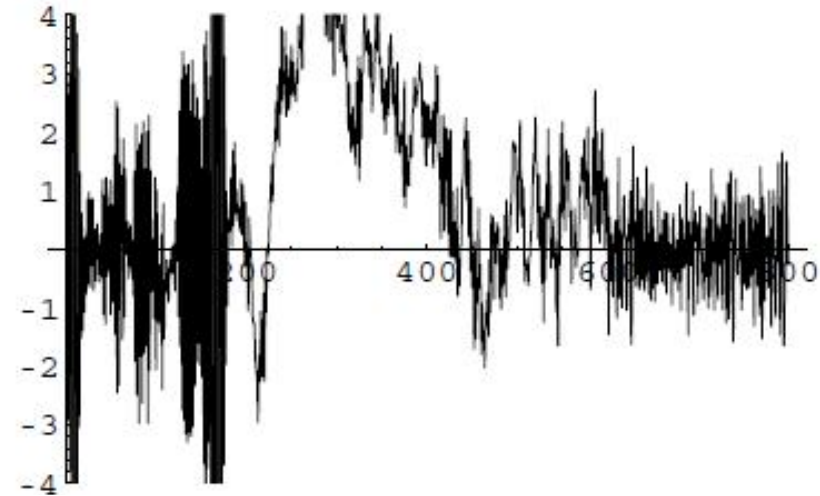
人工データでの実験

- 設定
 - データ
 - 構造が時間変動するAR
 - 設定
 - 生成の構造式が時間区間ごとに変化
 - 時間変化は200 ごと
 - パラメータの変動の種類は4 種作成
 - 誤差項は ± 1 か ± 0.1 の一様分布

データと生成式の例



分散小の場合



分散大の場合

生成式の例:パラメータ推移タイプ C3

$$y_t = \begin{cases} -0.800x_1 - 0.810x_2 + 0.648x_3 + \epsilon_t & ;1 \leq t \leq 200 \\ 0.190x_1 + 0.822x_2 - 0.060x_3 + 0.089x_4 + \epsilon_t & ;201 \leq t \leq 400 \\ 0.400x_1 + 0.770x_2 - 0.360x_3 + \epsilon_t & ;401 \leq t \leq 600 \\ -0.500x_1 - 0.570x_2 - 0.405x_3 - 0.194x_4 + \epsilon_t & ;601 \leq t \leq 800 \end{cases}$$

実験データの評価尺度

- 変化点検出力をbenefit とFAR で評価

- benefit

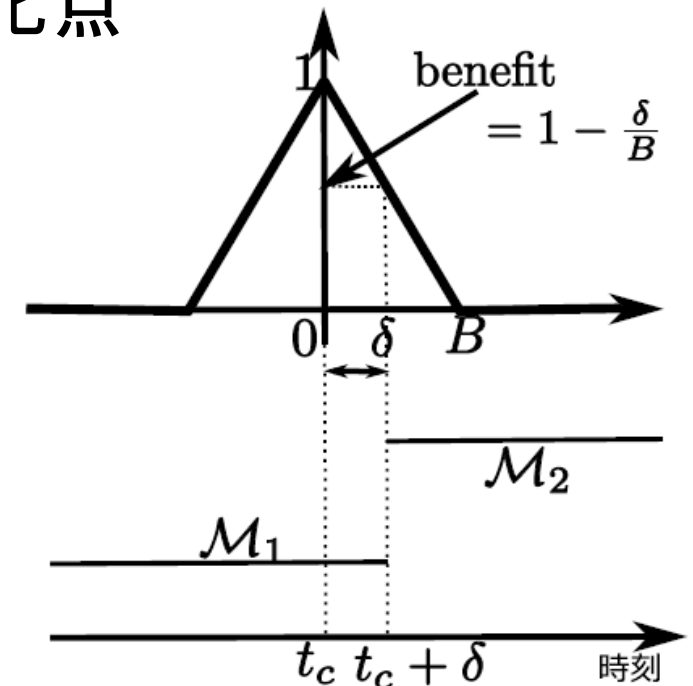
アルゴリズムにより発見された変化点

- FAR (False Alarm Rate)

誤検出率(FAR)=

(誤検出の変化点数)/(データ数)

- benefitの窓サイズB は
所与の定数



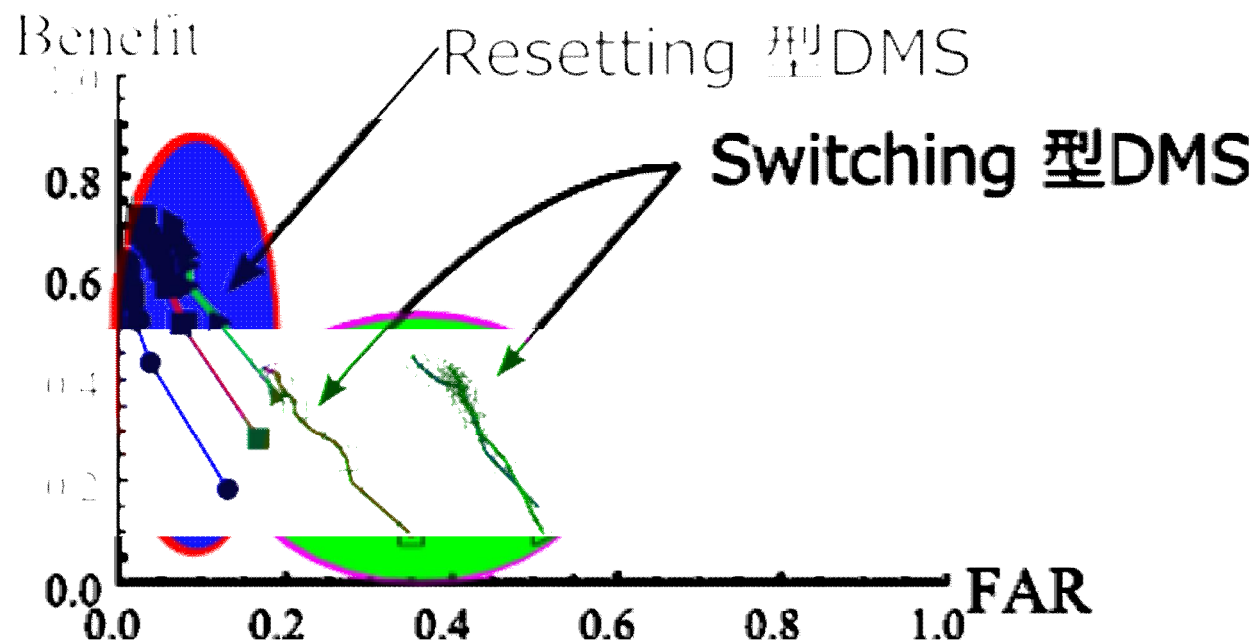
人工データでの結果

3種類のデータでテスト(○□▽が種類に相当)

ウィンドウサイズを変化させ実験

→ウィンドウサイズ大⇒ Benefit ↑・FAR ↓

∴resetting 型のほうがよい結果



目次

1. イントロダクション
2. 非定常性への従来の取り組み
3. 動的モデル選択規準
4. Resetting 分布と switching 分布
5. 数値実験
6. まとめ

まとめ

- 本発表で示したこと:
 - 非定常性に対する考え方の紹介
 - 動的モデル選択の拡張について紹介
 - 理論的数値的評価
- 今後の発展
 - Switching 分布がいい場合？
 - 実データで面白いモデル