

第43回 TokyoWebMining

ベイズアンモデリングによる マーケティングサイエンス

～状態空間モデルを用いたモデリング～



@sanoche16

名前 : @sanoche16

職業 : コンサルタント

得意 : マーケティングサイエンス

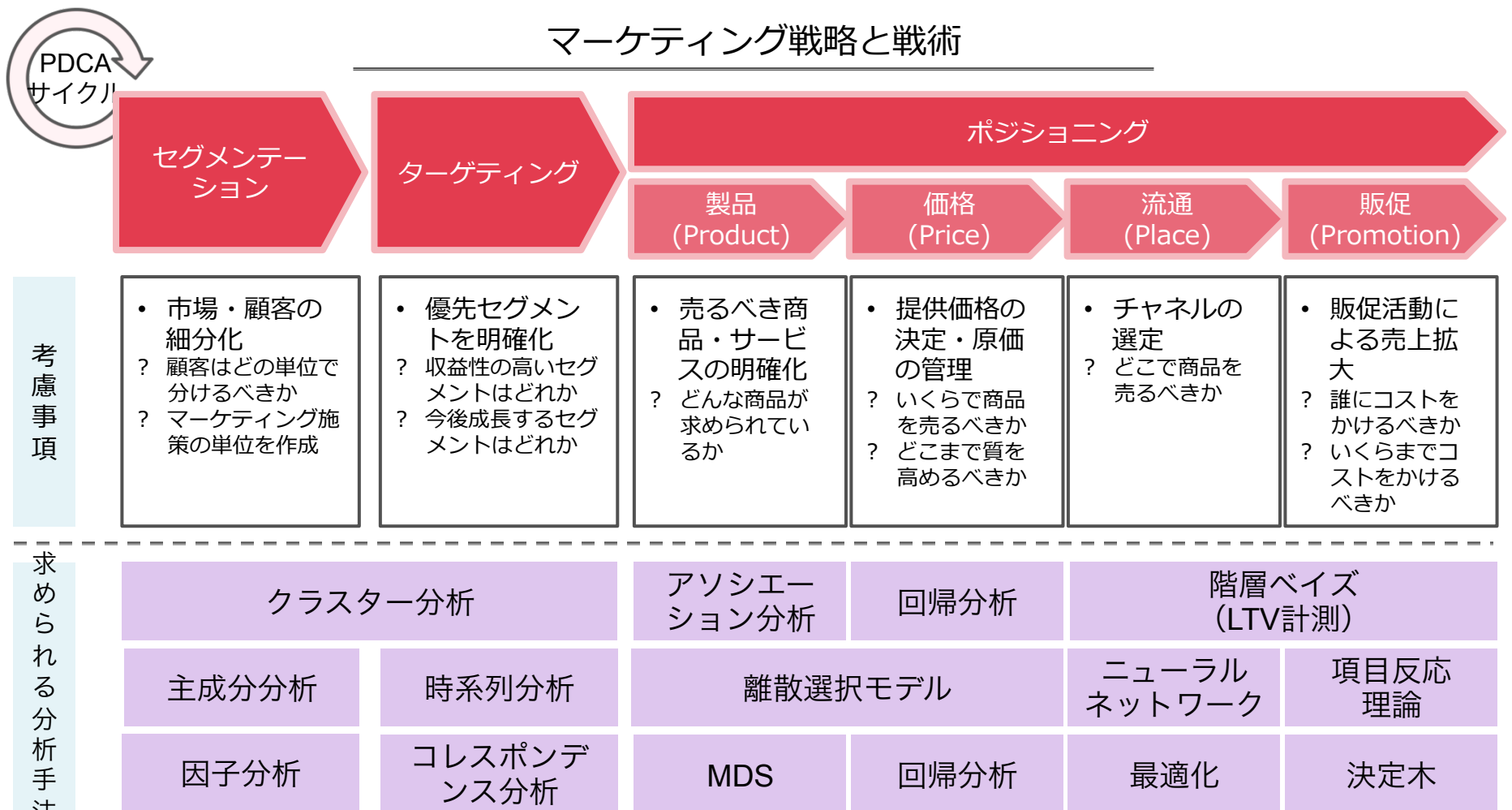
好き : R, Python, Linux, PHP, Ruby

本日の発表に含まれる内容はすべて個人の
見解であり、必ずしも所属団体の見解では
ありません。

- マーケティングにおけるベイジアンモデリングの必要性
- 線形ガウス状態空間モデルの基本概念
- 線形ガウス状態空間モデルへの季節性変数の取込
- 線形ガウス状態空間モデルへのマーケティング変数の取込
- パラメータ推定と実装方法
- おわりに・参考資料

- マーケティングにおけるベイジアンモデリングの必要性
- 線形ガウス状態空間モデルの基本概念
- 線形ガウス状態空間モデルへの季節性変数の取込
- 線形ガウス状態空間モデルへのマーケティング変数の取込
- パラメータ推定と実装方法
- おわりに・参考資料

PDCAサイクルの高速化も相まって、マーケティング戦略/戦術のなかで、統計分析・モデリングが求められる機会はどんどん増えています



※分析手法は様々なシーンで活用できるため、よく利用される観点で一例を示した

➡ 多岐に渡って、より精度の高い分析・予測が求められている

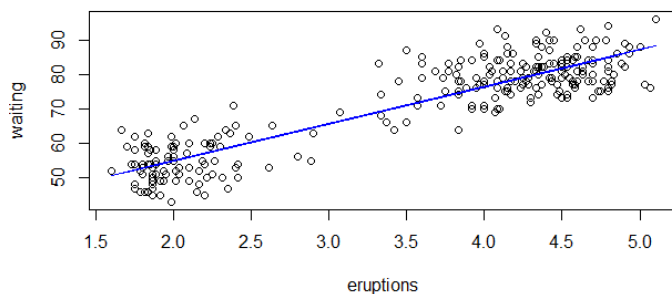
一方で、現実の顧客行動はとても複雑で、簡単な統計モデルが役立たないことも多くあります

単純な手法によるモデリング

顧客	年齢	最終購買日	年間購買回数	新商品購買有無
Aさん	24歳	14/12/01	15階	有
Bさん	42歳	15/02/06	12階	無
Cさん	36歳	15/01/10	9階	有
...

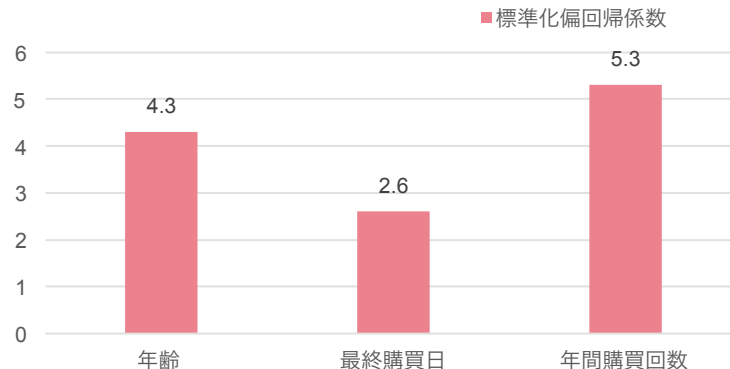


線形回帰分析



- 線形回帰分析により特定の商品の購買確率を予測
- 投入変数として、年齢、最終購買日、年間購買回数を利用

分析結果と問題点



顧客	年齢	最終購買日	年間購買回数	予測購買確率
Xさん	2歳	14/02/27	42階	89%
Yさん	4歳	14/02/25	28階	88%
Zさん	4歳	14/02/26	29階	88%
...

予測確率高

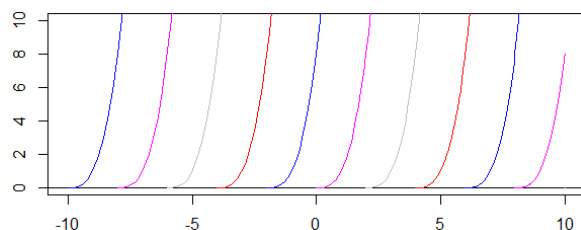


- 良いと予測された顧客は説明変数の最大 (or 最小) の値を持つ顧客になってしまう
- 本来は優先すべきではない顧客を優先してしまう

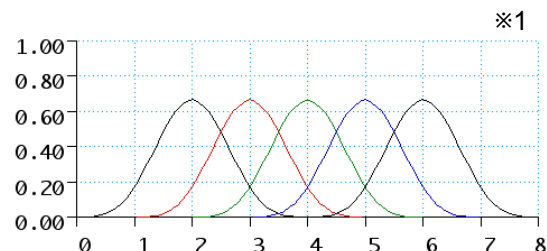
ビッグデータの時代においても、事象を適切にモデル化するために必要なデータ数は多くの場合不足しています。

基底変換を用いたアプローチ

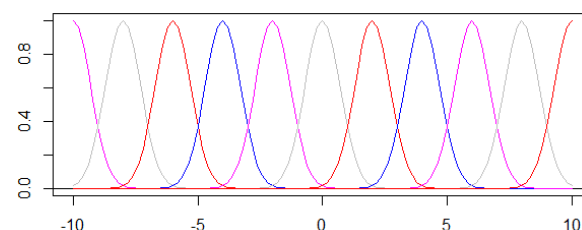
3次スプライン



B-スプライン



ガウス型基底関数



3次スプライン関数の式：
$$u(x; \boldsymbol{\theta}) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{i=1}^m \theta_i (x - t_i)_+^3$$

推定するパラメータベクトル $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \theta_1, \theta_2, \dots, \theta_m)^T$

- 利用する説明変数は1つ
- 推定するパラメータは数十にもなる



説明変数を多い場合はかなり不安定なモデルになってしまう

ベイズ統計を用いて“常識”を教えることで、多くの事象を適切にモデル化できるようになります

頻度主義でのアプローチ

ベイズ統計を用いたアプローチ

データ数 : 0

- 情報は一切なし

- 事前分布を設定
- (無情報事前分布や勘と経験から)

データ数 : 1

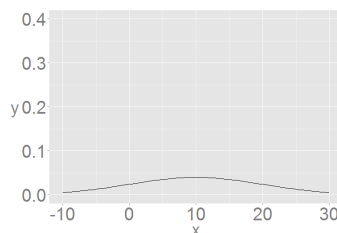
- データはたまたま取れた値のみ利用できる
- データの広がり(=分散)は分からない

- 事前分布を更新
- 分散が小さくなり、推定値の確からしさを分布で表現

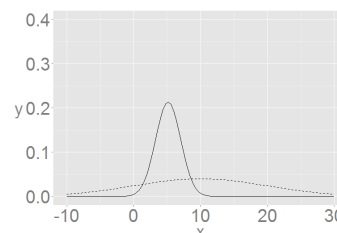
データ数 : 2

- 標本分散を利用して母分散を推定可能
- 但し、2つのデータの近さからのみ推定される(=不安定)

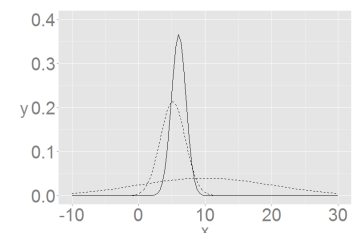
- 事前分布を更新
- 分散が小さくなり、推定値の確からしさを分布で表現



事前分布を更新



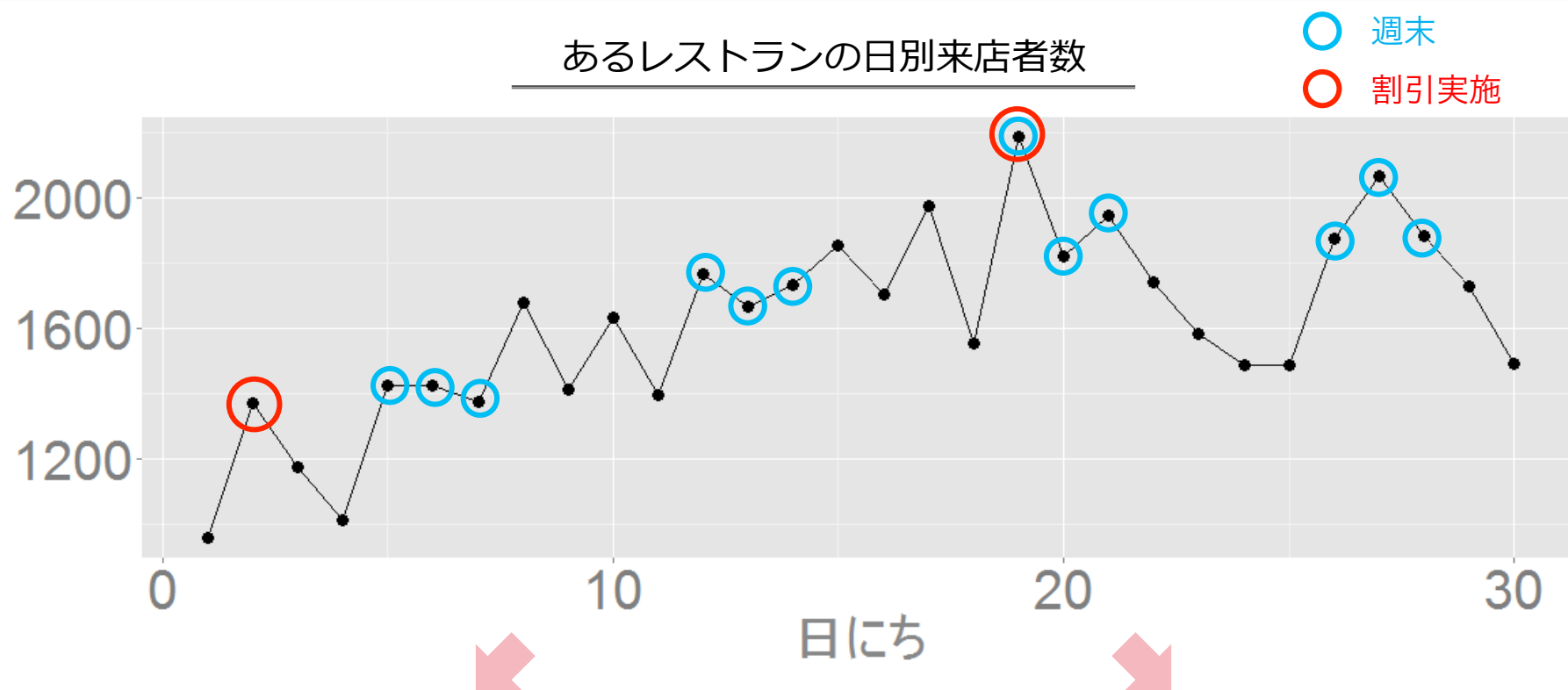
事前分布を更新



- ◆ マーケティングの領域では、現在も非常に多くの統計的手法を用いた分析が求められている
 - きちんとしたモデリングにはそれなりの柔軟性が必要
- ◆ 柔軟なモデリングを行うためには必要なパラメータはかなり多くなってしまふ
 - パラメータが多くても安定的なモデルの構築が必要になる
- ◆ ベイジアンモデリングを利用することで、安定的かつ柔軟なモデリングを行うことが可能になる
 - 事前分布を設定することで、スパースなデータに対しても安定的なモデリングが可能になる

- マーケティングにおけるベイジアンモデリングの必要性
- 線形ガウス状態空間モデルの基本概念
- 線形ガウス状態空間モデルへの季節性変数の取込
- 線形ガウス状態空間モデルへのマーケティング変数の取込
- パラメータ推定と実装方法
- おわりに・参考資料

状態空間モデルとは、時点の変化に関わる影響と、特定の時点で受ける要因の影響をそれぞれモデル化したものです



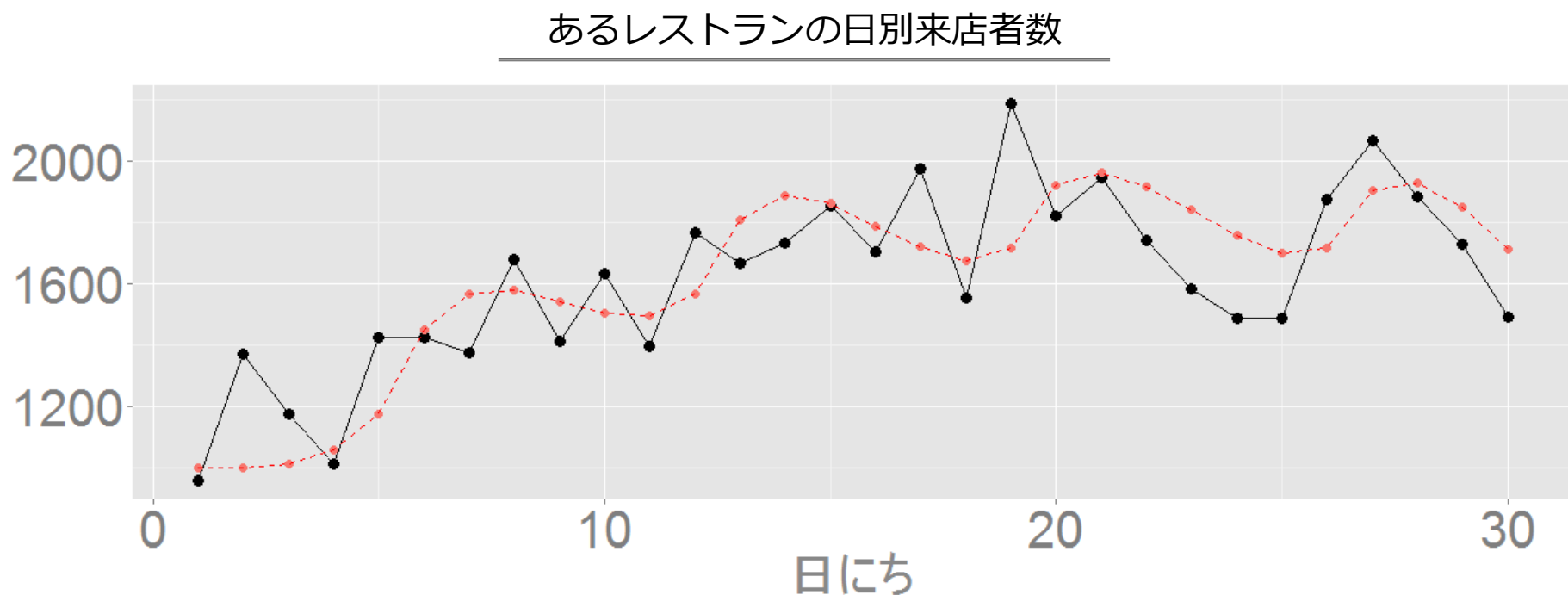
時点の変化に関わる影響

- ✓ 週末になったので、来客数が増加した
- ✓ 夏になったので、来客数が増加した

特定時点での要因による影響

- ✓ 割引を実施したので、来客数が増加した
- ✓ たまたま雨が降ったので、来客数が増加した

観測値から特定時点での要因を排除した“トレンド”を抽出することで
特定時点での要因の分析と時系列の分析を正確に行うことが可能になります



観測値：実際に取得されたデータ



トレンド：観測値から特定時点の要因を排除した潜在的な値



- ・トレンドの変化を抽出することで時点の変化に関わる影響を分析
- ・トレンドと観測値を用いて、特定時点での要因による影響を分析

時点の変化に関わる影響はトレンドの関係のみを考え、特定時点での要因による影響はトレンドと実測値の関係をモデル化します

具体例

- 以下の3つの条件を満たす形で変化するデータについて考える
1. 時点の変化に伴う変化量は、1時点前の変化量と同様となる傾向がある
 2. 時点の変化に伴って一定のノイズが発生する
 3. 実際の観測に伴って一定のノイズが発生する

時点の変化に関わる影響

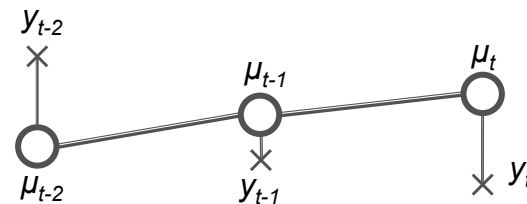
- t時点における変動している変数を μ_t と置く
- (仮定1から) μ_t は、 $\mu_t - \mu_{t-1} = \mu_{t-1} - \mu_{t-2}$ を満たす $\Rightarrow \mu_t = 2\mu_{t-1} - \mu_{t-2}$
- (仮定2から) ノイズを v_t の $N(0, \sigma_1^2)$ に従う正規分布とする (更新ノイズ)

$$\mu_t = 2\mu_{t-1} - \mu_{t-2} + v_t$$

トレンド

特定時点での要因による影響

- t時点における観測値を y_t と置く
- (仮定3から) ノイズを w_t の $N(0, \sigma_2^2)$ に従う正規分布とする (観測ノイズ)



$$y_t = \mu_t + w_t$$

y_t は y_{t-1} からの影響は受けず、 μ_t からの影響のみ受ける

時点の変化に関わる影響をシステムモデルで、特定時点での要因による影響を観測モデルでモデル化します

時点の変化に関わる影響

$$\mu_t = 2\mu_{t-1} - \mu_{t-2} + v_t$$

特定時点での要因による影響

$$y_t = \mu_t + w_t$$

ベクトル・行列で表現

$$\begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \mu_{t-2} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot v_t$$

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix} + w_t$$

$$\mathbf{x}_t = \begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix}$$

$$\mathbf{F}_t = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$$

$$\mathbf{G}_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\mathbf{H}_t = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

状態ベクトル

一般化

線形ガウス状態空間モデル

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{G}_t \mathbf{v}_t$$

システムモデル

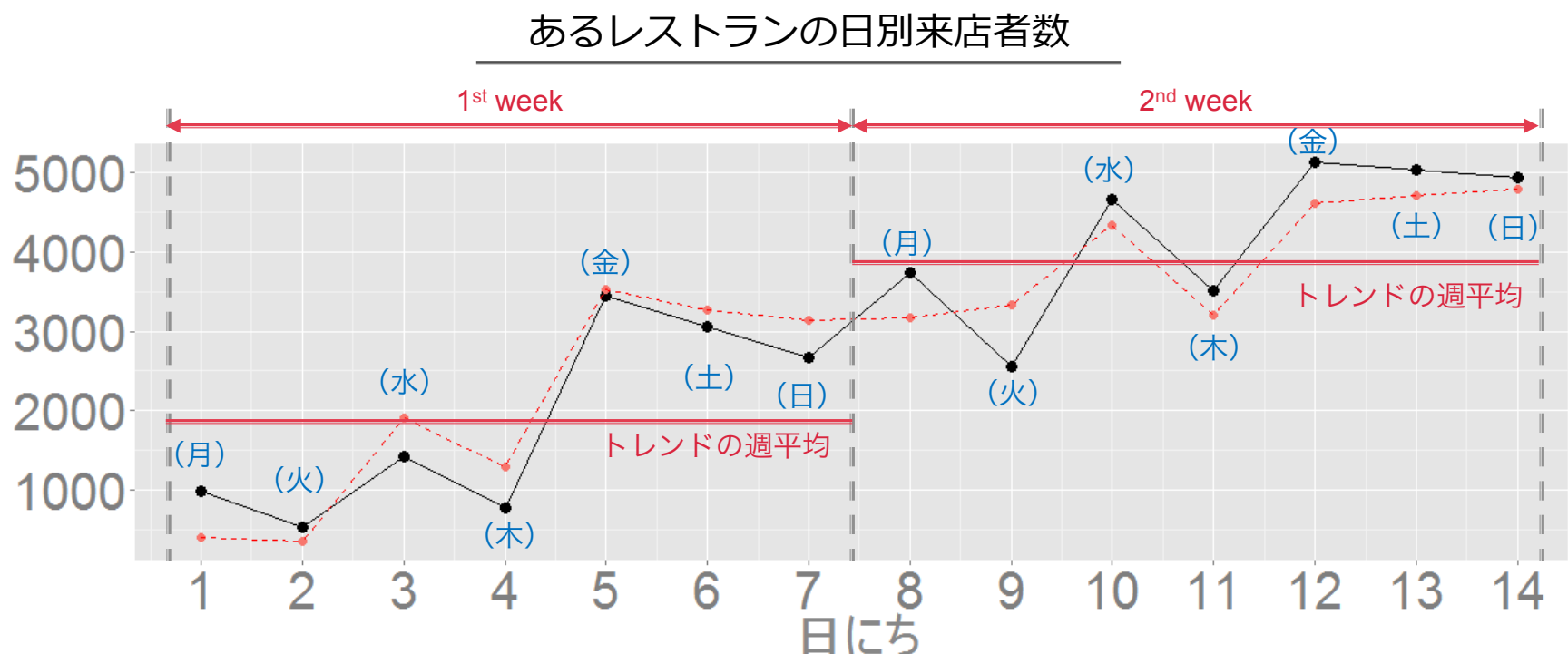
$$y_t = \mathbf{H}_t \mathbf{x}_t + w_t$$

観測モデル

- ◆ 状態空間モデルでは、潜在変数を含む状態ベクトルを用いて時系列データの分析を行う
 - 時点の変化に伴って推移するトレンド成分を抽出することで、トレンドの動きとトレンドと実測値の関係の分析が可能になる
- ◆ 線形ガウス状態空間モデルはシステムモデル・観測モデルの2つにより構成される
 - システムモデルにより時点の変化による影響を分析
 - 観測モデルに特定時点の要因による影響を分析

- マーケティングにおけるベイジアンモデリングの必要性
- 線形ガウス状態空間モデルの基本概念
- 線形ガウス状態空間モデルへの季節性変数の取込
- 線形ガウス状態空間モデルへのマーケティング変数の取込
- パラメータ推定と実装方法
- おわりに・参考資料

曜日による影響の状態空間モデルへの組み込みは、一週間の和がゼロとなる周期性を持つ成分を考えます。



時点の変化に関わる曜日の影響のみをモデルとして表現

$$s_t = - \sum_{i=1}^6 s_{t-i} + v_t \quad v_t \sim N(0, \alpha_s^2 \sigma^2)$$



時点の変化の関係をシステムモデルに組み込んでいく

システムモデルに曜日効果を組込むことで、状態空間モデルに取り込むことが可能になります

時点の変化に関わる影響

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{G}_t \mathbf{v}_t$$

$$\begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \mu_{t-2} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot v_t$$

$$s_t = - \sum_{i=1}^6 s_{t-i} + v_t$$

曜日効果の取込

$$\begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ s_{t-6} \end{pmatrix}$$

特定時点での要因による影響

$$\mathbf{y}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{w}_t$$

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix} + w_t$$

観測モデルへの反映

$$y_t = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \end{pmatrix}$$

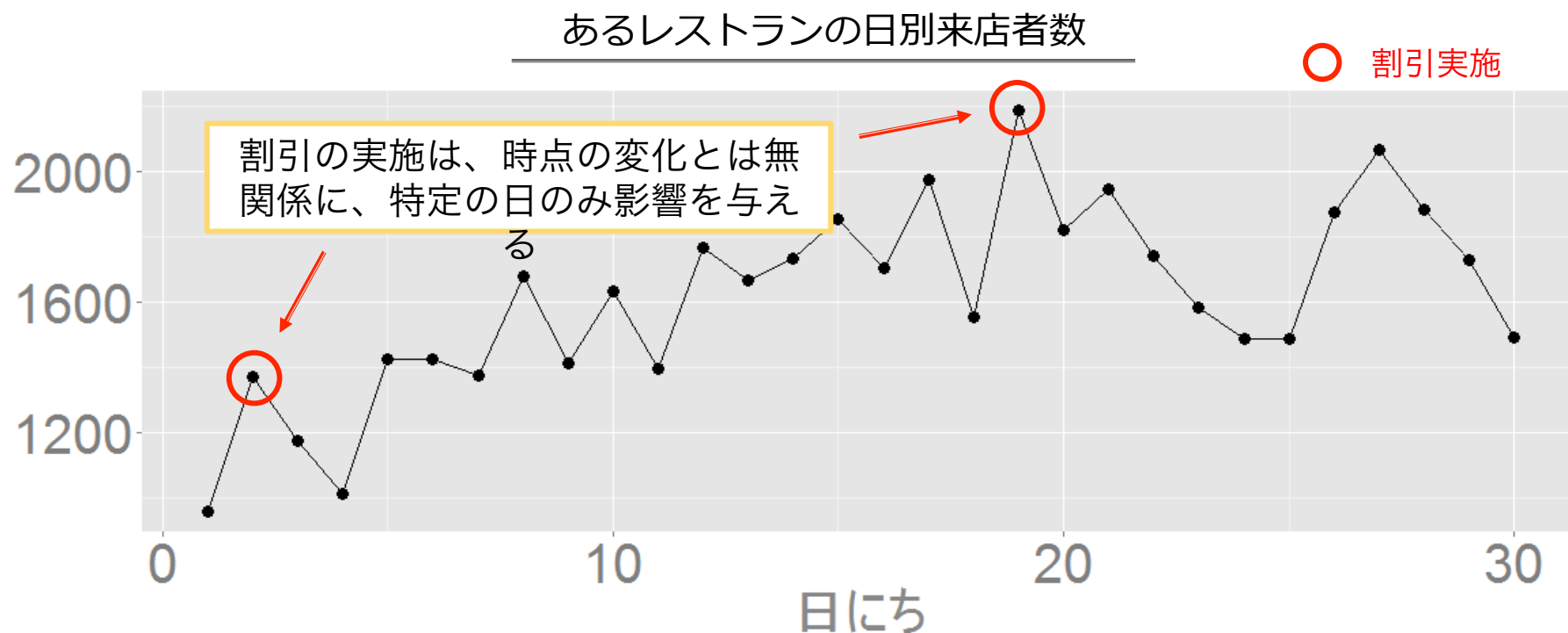
➤ 観測モデルはその時点の成分以外は排除する

状態ベクトルを拡張することで、曜日効果も表現できる

- ◆ 状態ベクトルに対して、1週間の和がゼロとなるような成分を加えることで、曜日効果を組込むことが可能
 - 観測モデルでは、実測値に影響を与える成分をその時点のものに限定することで、時点の変化による影響と区別する

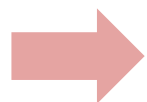
- マーケティングにおけるベイジアンモデリングの必要性
- 線形ガウス状態空間モデルの基本概念
- 線形ガウス状態空間モデルへの季節性変数の取込
- 線形ガウス状態空間モデルへのマーケティング変数の取込
- パラメータ推定と実装方法
- おわりに・参考資料

割引実施効果を測定は、時点に関わらず一定の値を持つ成分を考えます



時点の変化に関わらず一定の値を持つ成分を表現

$$\gamma_t = \gamma_{t-1} + v_t \quad v_t \sim N(0, \alpha_\gamma^2 \sigma^2)$$



一定の値を持つ成分をシステムモデルに組み込み、割引効果を測定

時点に関わらず一定の値を持つ成分をシステムモデルに組み込むことで、値引き効果を状態空間モデルに取り込むことができます

時点の変化に関わる影響

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{G}_t \mathbf{v}_t$$

$$\begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ s_{t-6} \end{pmatrix}$$

$$\gamma_t = \gamma_{t-1} + v_t$$

時点によらない項の組込

$$\begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ \gamma_t \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ s_{t-6} \\ \gamma_{t-1} \end{pmatrix}$$

特定時点での要因による影響

$$\mathbf{y}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{w}_t$$

$$y_t = (1 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0) \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \end{pmatrix}$$

f_t は時点の値引きの有無を表すフラグ

観測モデルへの反映

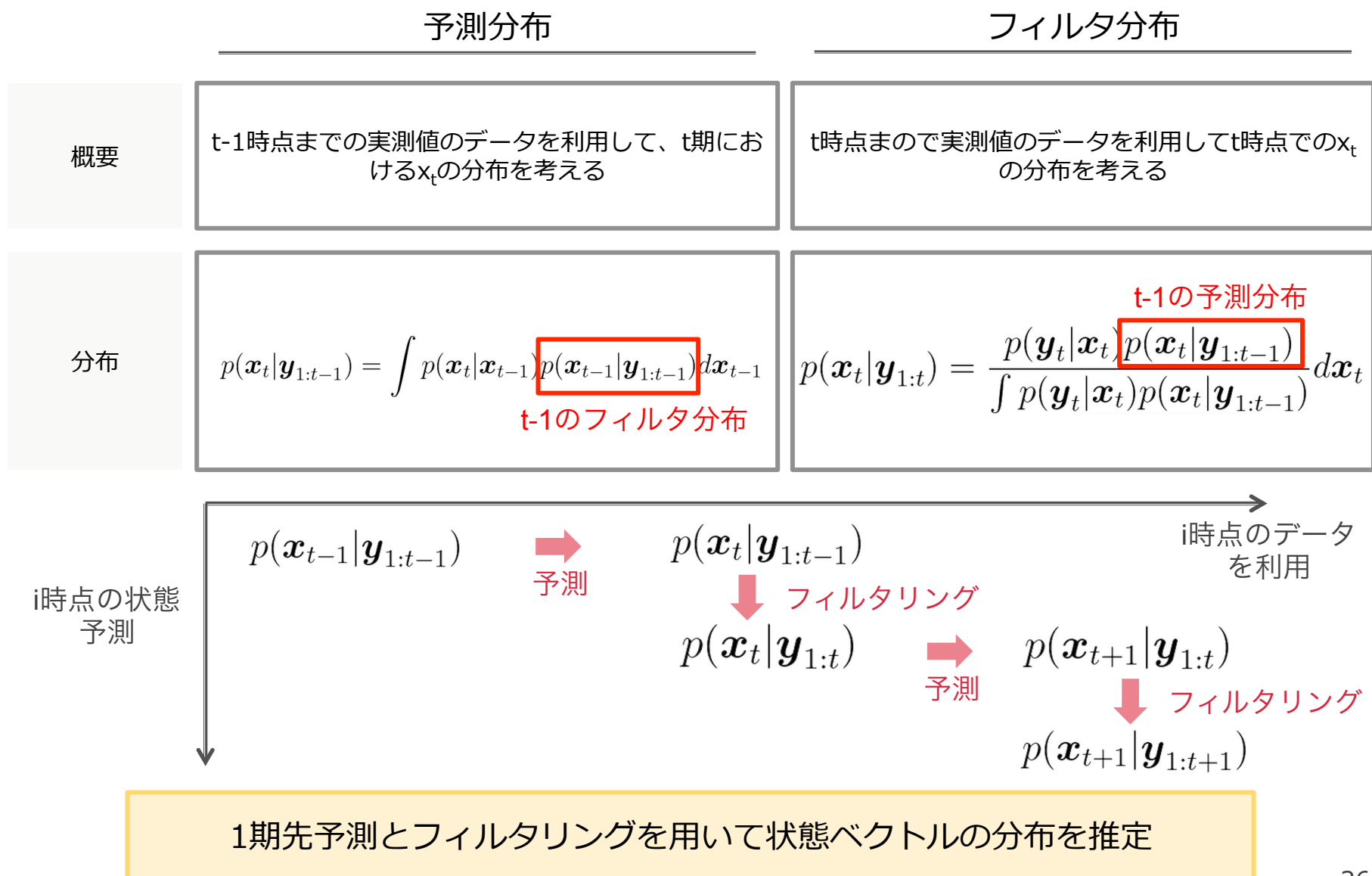
$$y_t = (1 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ f_t) \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ \gamma_t \end{pmatrix}$$

時点によらない成分を組込むことで、値引き効果も表現できる

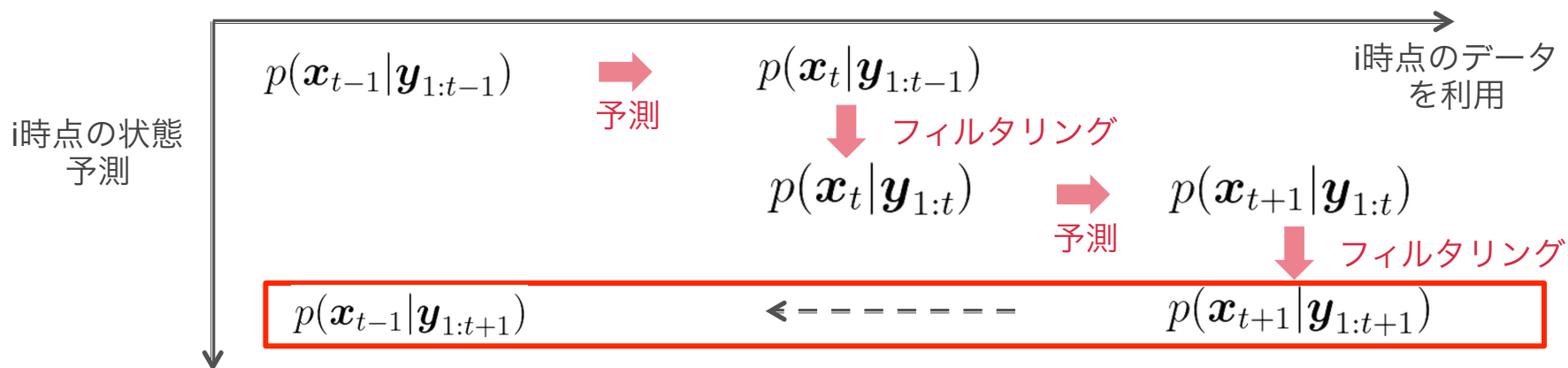
- ◆ マーケティング変数は、システムモデルで時点の変化を排除することでモデルへの組み込みが可能
 - 不定期的な割引など、周期性がないマーケティング施策に関する変数は、時点の変化による影響がないようにモデルに組み込む

- マーケティングにおけるベイジアンモデリングの必要性
- 線形ガウス状態空間モデルの基本概念
- 線形ガウス状態空間モデルへの季節性変数の取込
- 線形ガウス状態空間モデルへのマーケティング変数の取込
- パラメータ推定と実装方法
- おわりに・参考資料

状態空間モデルでは、1期先予測とフィルタリングを繰り返すことで、時点を進めていくことができます



1期前平滑化を繰り返すことで過去の状態ベクトルの推定を修正します



固定区間平滑化アルゴリズム

$$p(\mathbf{x}_t | \mathbf{y}_{1:T}) = \underbrace{p(\mathbf{x}_t | \mathbf{y}_{1:t})}_{\text{フィルタ分布}} \cdot \int \frac{p(\mathbf{x}_{t+1} | \mathbf{x}_t)}{\underbrace{p(\mathbf{x}_{t+1} | \mathbf{y}_{1:t})}_{\text{予測分布}}} \cdot p(\mathbf{x}_{t+1} | \mathbf{y}_{1:T}) d\mathbf{x}_{t+1}$$

- ✓ 予測分布とフィルタ分布によりT時点までのデータを用いて $p(\mathbf{x}_T | \mathbf{y}_T)$ がわかっている
→ T-1から逐次的に状態ベクトルを求められる

過去を振り返って状態ベクトルを表現し、定性的な示唆が得られる

線形ガウス状態空間モデルではカルマンフィルタで逐次推定を行うことで、高速な計算が可能になり、オンライン学習に対応出来ます

フィルタリング

- 逐次学習型の状態空間モデルにおけるパラメータ推定では、予測分布・フィルタリング・平滑化を繰り返すことで、状態ベクトルを求めていく



カルマンフィルタ

- 線形ガウス状態空間モデルにおいて、利用される
- 最小化する推定誤差にガウス分布を仮定
- 計算速度が速い



パーティクルフィルタ

- 非線形の状態空間モデルに対応するフィルタリングアルゴリズム
- 粒子（パーティクル）を発生させてノンパラメトリックに最適解を求める



バッチ学習を行う際にはフィルタリングではなく、MCMCの利用も可能

(ご参考)

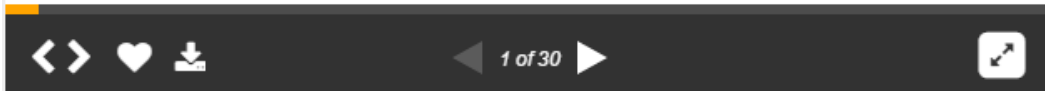
- ◆ Stanを用いたMCMCでのパラメータ推定の例は下記に完全版があります（データもあります）

**『予測にいかす統計モデリングの基本』
の売上データの分析をトレースしてみた**

2013.12.22

@berobero11

BUGS/Stan勉強会 #2 於 ドリコム



- ◆ 状態空間モデルでは、予測分布・フィルタ分布・平滑化分布の3つを利用し、トレンドの推定・予測を行う
 - 予測分布により一期先の予測を行う
 - フィルタ分布に特定時点における状態ベクトルの分布を表現する
 - 平滑化分布により以前の時点における状態ベクトルの分布を表現する
- ◆ カルマンフィルタを利用することで高速な計算・オンライン学習が可能

- マーケティングにおけるベイジアンモデリングの必要性
- 線形ガウス状態空間モデルの基本概念
- 線形ガウス状態空間モデルへの季節性変数の取込
- 線形ガウス状態空間モデルへのマーケティング変数の取込
- パラメータ推定と実装方法
- おわりに・参考資料

- ◆ 状態空間モデルとはシステムモデルと観測モデルを用いて、分析を行うメタモデル
 - システムモデルと観測モデルは自由に設定できる
 - 非線形のモデルを利用した場合は、パーティクルフィルタを利用することで、状態ベクトルを求めることが可能
- ◆ パラメータの推定は必ずしも逐次学習をする必要はない
 - バッチ処理にはMCMCを用いた推定が可能

◆ 状態空間モデルの概要

- ❖ 予測にいかす統計モデリングの基本 樋口知之(2011)
- ❖ データ同化入門 樋口知之(2011)
- ❖ 時系列解析入門 北川源四郎(2005)
- ❖ Rによるベイズアン動的線形モデル G.ペトリス, S.ペトロローネ, P.カンパニョーリ, 和合 肇, 萩原 淳一郎(2013)

◆ 状態空間モデルの活用について

- ❖ 時系列解析入門 北川源四郎(2005)

◆ その他

- ❖ [RStanで『予測にいかす統計モデリングの基本』の売上データの分析をトレースしてみた](#) berobero11(2013)
- ❖ [逐次モンテカルロ/\(粒子|パーティクル|モンテカルロ\)フィルタを実装してみた](#) teramonagi(2014)

Thanks!

ご静聴ありがとうございました。

