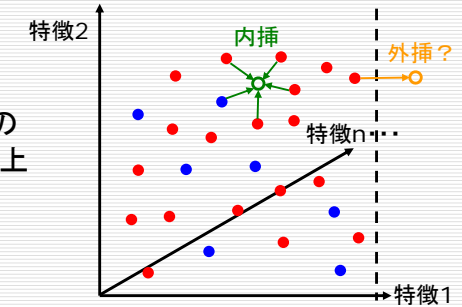


データ同化アプローチの 交通データへの適用可能性

佐々木 邦明 (山梨大学)
布施 孝志 (東京大学)

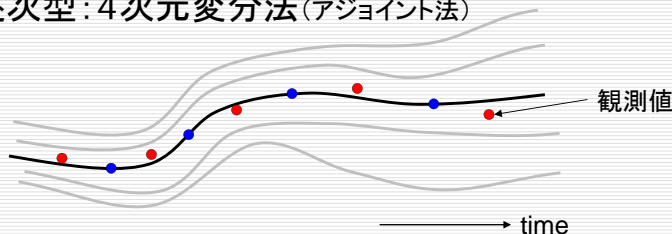
はじめに

- データ取得・解析環境の進展
- データ量・種類の増加は望ましいものであるが...
 - データ空間の次元の膨張: スパースな(疎な)データ
 - 「ビッグデータの価値密度は低い」(統計数理研究所長 樋口知之氏)
 - 内挿の世界(帰納的): 疎な空間では低精度, 外挿は?
- シミュレーション(予測モデル)の進展(演繹的)
 - 内挿・外挿の精度向上への期待
- データ同化
 - 観測データと予測モデルの統合による推定精度の向上



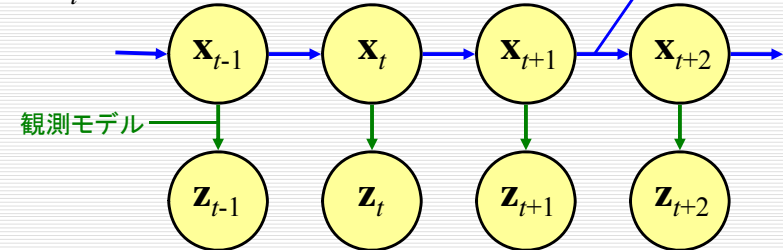
データ同化の意義

- 実測データを用いて数値シミュレーションモデルの精度・性能を改善する
(incl. 初期条件, 境界条件, パラメータ推定, 感度分析)
- シミュレーションモデルを用いることにより観測の不足を補ったり観測誤差を修正する
(再解析データセットの生成, 仮想観測)
- 今回は逐次型のみ対象とする
非逐次型: 4次元変分法(アジョイント法)



基本的な考え方: 一般状態空間モデル

- 時系列の観測モデル
 - x_t : 時刻 t における状態ベクトル(推定したい量): 観測不可
 - z_t : 時刻 t における観測ベクトル: 観測可



システムモデル $x_t = f_t(x_{t-1}, v_t)$ $x_t \square p(x_t | x_{t-1})$

観測モデル $z_t = h_t(x_t, w_t)$ $z_t \square p(z_t | x_t)$

z_t が観測された条件の下で, 最も尤もらしい x_t を推定
事後確率最大化: $p(x_t | z_{1:t}) \rightarrow \max.$

事後確率: ベイズの定理

5

因果関係の推論を行う際に、基礎になっている定理

A(原因)→B(結果) Aの起こる確率
 の確率(尤度) = 事前確率(主観確率)

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

B(結果)が得られた時に
 Aが原因である確率
 = 事後確率

Bの起こる確率

x: 知りたい量, z: データ
 センサ普及 → 尤度値 ストレージ大容量化 → 事前情報

$$p(x|z) = \frac{p(z|x)p(x)}{\int p(z|x)p(x)dx}$$

コンピュータ性能の向上



Thomas Bayes (1702-1761)

ベイズの定理に基づく事後確率

6

時刻 1 ~ t の観測ベクトル $\mathbf{z}_{1:t}$

事後確率 $p(\mathbf{x}_t | \mathbf{z}_{1:t})$

$$p(\mathbf{x}_t | \mathbf{z}_{1:t}) \propto p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{z}_{1:t-1}) = p(\mathbf{z}_t | \mathbf{x}_t) \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{z}_{1:t-1}) d\mathbf{x}_{t-1}$$

観測モデル = 尤度 システムモデル 1期前(前フレーム)の結果

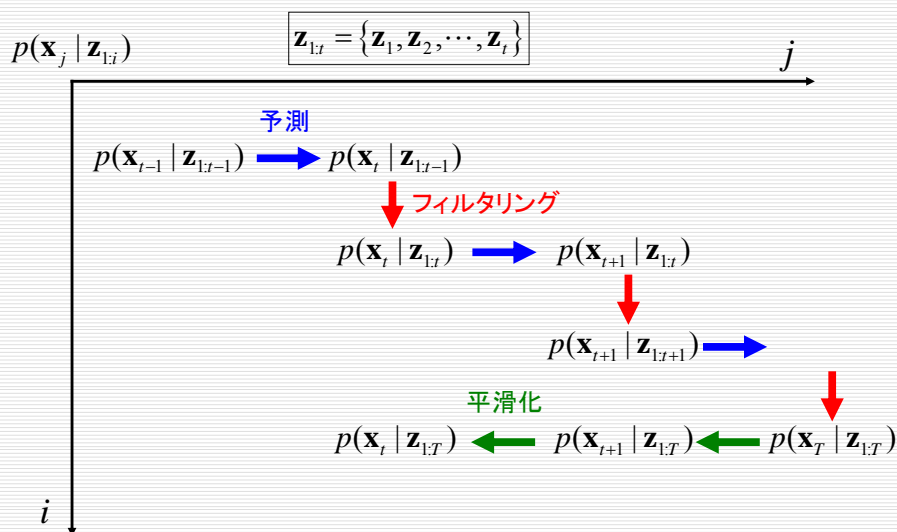
ベイズの定理

モデル化が必要

逐次, MAP推定

逐次ベイズフィルタ

7



線型・ガウス型状態空間モデル

8

事後確率 $p(\mathbf{x}_t | \mathbf{z}_{1:t})$

$$p(\mathbf{z}_t | \mathbf{x}_t) \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{z}_{1:t-1}) d\mathbf{x}_{t-1}$$

観測モデル = 尤度 システムモデル 1期前の結果

観測モデルとシステムモデル

カルマンフィルタ(KF)

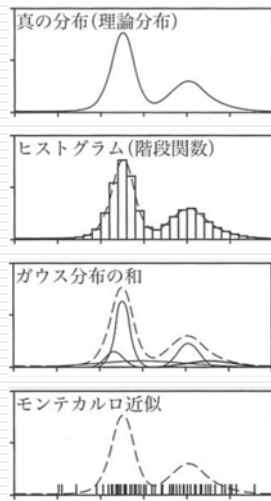
$$\mathbf{x}_t = F_t \mathbf{x}_{t-1} + G_t \mathbf{v}_t \quad \mathbf{v}_t \square N(\mathbf{0}, Q_t)$$

$$\mathbf{z}_t = H_t \mathbf{x}_t + \mathbf{w}_t \quad \mathbf{w}_t \square N(\mathbf{0}, R_t)$$

一般状態空間モデルでは、非線形, 非ガウス性ノイズに対応可能

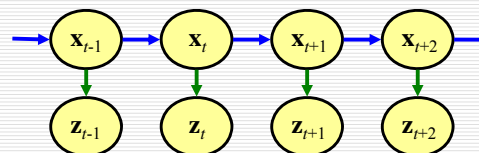
一般状態空間モデル

状態ベクトル:連続	
分布:ガウス型	分布:非ガウス型
カルマンフィルタ (線形・ガウス型)	ヒストグラムフィルタ (HMMの連続版)
拡張KF	バイナリベイズフィルタ
モーメントマッチング (Assumed Density Filter)	ガウス和フィルタ (複数仮説KF)
アンセティッドKF	アンサンブルKF
	パーティクルフィルタ (非線形・非ガウス型)
	アンサンブルベース
状態ベクトル:離散	
隠れマルコフモデル	



応用対象の交通現象

- システムモデル
 - $x_t = f_t(x_{t-1}, v_t)$
- 観測モデル
 - $z_t = h_t(x_t, w_t)$
- 状態変数
 - 密度, 速度, 交通量, 人口
- システムモデル
 - マクロシミュレーションモデル
 - LWR モデル
 - Cell Transmission Model (CTM)
 - 高次項モデル
 - Payne型モデル
 - ミクロシミュレーションモデル



- 観測データ
 - オイリアンデータ
 - 検知器データ
 - メッシュ人口密度データ
 - ラグランジアンデータ
 - プロブカーデータ
 - GPSの移動軌跡

鈴木・中辻(2002): フィードバック原理に基づく交通状態推定手法を応用した高速道路上起終点旅行時間の推定

Nanthawichitら(2003): Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway

- 状態変数
 - 密度, 空間平均速度
- 観測データ
 - 感知器データ(地点交通量, 地点速度)
- モデル
 - 高次項マクロ交通流モデル
- 計算法
 - カルマンフィルタ
- 結果
 - ミクロシミュレーションによるデータを用いた再現性の検討
 - 交通状態の推定を行わないNKFと比較してRMSEが53%開演
- 状態変数
 - 速度
- 観測データ
 - LagrangianデータとしてのProbeデータ
 - Eulerianデータとして検知器データ
- モデル
 - Payne型マクロ交通流モデル
- 計算法
 - カルマンフィルタ
- 結果
 - 1%~3%のProbeデータを混入したとしたときの再現性を検討
 - マクロ交通流モデルのみと比較して70~85%誤差を減少させる

Terpstra et. al(2004): Intelligent Adaptive Traffic Forecasting System using Data Assimilation for use in Traveler Information Systems

Herrera and Bayen (2008): Traffic flow reconstruction using mobile sensors and loop detector data

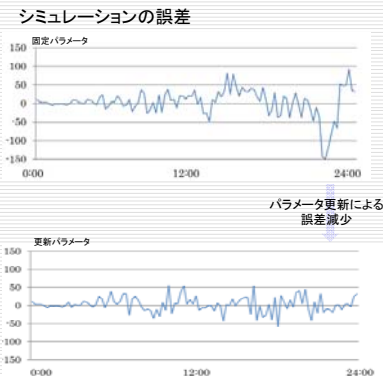
- 状態変数
 - 交通量等
- 観測データ
 - 気象情報, イベント情報
- モデル
 - RAIDO環境シミュレーター
- 結果
 - 天候・イベントなどの外部要因を考慮した交通需要予測体系をデータ同化の枠組みで提案
- 状態変数
 - 交通密度
- 観測データ
 - LagrangianのProbe交通密度
 - Eulerianデータとしての検知器による密度
- モデル
 - LWR-PDEマクロ交通流モデル
- 計算法
 - カルマンフィルタ
- 結果
 - Eulerianデータのみを用いた場合と比較して多大な改善があることを示した

- | | |
|--|--|
| <ul style="list-style-type: none"> □ 状態変数 <ul style="list-style-type: none"> ■ 速度 □ 観測データ <ul style="list-style-type: none"> ■ 携帯電話の速度データを離散的に取得 □ モデル <ul style="list-style-type: none"> ■ CTM-vモデル □ 計算法 <ul style="list-style-type: none"> ■ アンサンブルカルマンフィルタ □ 成果 <ul style="list-style-type: none"> ■ PeMS(速度測定システム)と比較して5%以下の混入で比較的精度よく推測可能 | <ul style="list-style-type: none"> □ 状態変数 <ul style="list-style-type: none"> ■ 密度, 速度 □ 観測データ <ul style="list-style-type: none"> ■ 区間をランダムに選んで観測 □ モデル <ul style="list-style-type: none"> ■ LWR-PDEマクロ交通流モデル □ 計算法 <ul style="list-style-type: none"> ■ パーティクルフィルタ □ 結果 <ul style="list-style-type: none"> ■ 均等に設置するよりも最大53%の誤差改善が可能な戦略を示した |
|--|--|

- | | |
|--|---|
| <ul style="list-style-type: none"> □ 状態変数 <ul style="list-style-type: none"> ■ 交通量 □ 観測データ <ul style="list-style-type: none"> ■ 高速道路上の交通量 □ モデル <ul style="list-style-type: none"> ■ トレンド, 季節調整, 自己回帰モデル □ 計算法 <ul style="list-style-type: none"> ■ カルマンフィルタ □ 結果 <ul style="list-style-type: none"> ■ トレンド・誤差分散から料金の変化による影響を把握 ■ 逐次予測の精度確認 ■ インター間では異なる影響を実証 | <ul style="list-style-type: none"> □ 状態変数 <ul style="list-style-type: none"> ■ 速度 □ 観測データ <ul style="list-style-type: none"> ■ 高速道路上の速度検知器データ □ モデル <ul style="list-style-type: none"> ■ CTM-v □ 計算法 <ul style="list-style-type: none"> ■ パーティクルフィルタ □ 結果 <ul style="list-style-type: none"> ■ プローブカーとの比較で再現性についての検討し検知器の設置間隔の最適性を実証 |
|--|---|

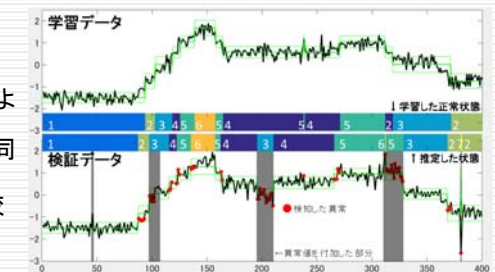
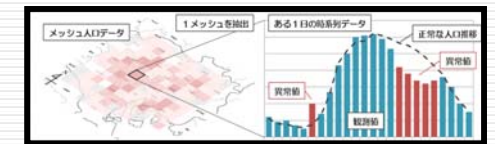
データ同化によるパラメータ推定

- Jackson(2011)
 - アジョイント法による同化
 - 交通流モデルのシステムパラメータを推定
- 広瀬・佐々木(2014)
 - 状態変数
 - 所要時間, 速度
 - モデル
 - マイクロシミュレーション
 - 計算法
 - パーティクルフィルタ
 - 観測変数
 - 区間所要時間
 - 結果
 - シミュレーションの速度パラメータを内生的に求めて誤差の改善



神谷・布施(2014): メッシュ人口データに基づく統計的異常検知

- 状態変数
 - 正常・異常(離散)
- 観測データ
 - メッシュ人口データ(シミュレーションデータ)
- モデル
 - 階層ディリクレ過程
 - 隠れマルコフモデル
- 結果
 - 学習データ(正常データ)による事前学習結果と検証データから状態数および状態を同時推定
 - 学習結果と検証結果の比較により異常地点(状態相違)を自動検知



交通データへの適用可能性

- データを用いたシミュレーションモデルの精度改善
 - ネットワーク上の交通流の状態や交通量, ゾーン別の滞在人数など, 様々な状態を定義可能
 - モデルパラメータを逐次的に更新可能
- シミュレーションモデルを用いた観測データの補完と精度向上
 - 観測データとシミュレーション(モデル)の融合による空間・時間的な補完
 - 各種交通調査の効率化
- 特異事象の検出と対応
 - 突発的なイベント等の検出と, 滞在人数の変化に対する予測
 - シミュレーションによる交通ネットワークの有効活用の検討