

データ同化夏の学校
オンライン開催
2020/8/5
(2020/8/12一部訂正)

データ同化の基礎

伊藤耕介(琉球大学)

謝辞：コメントを頂いた藤井陽介さん、中野慎也さんに感謝申し上げます。

事務連絡

- これから3本の動画に分けて解説を行います。
- p p t 資料の二次配布に当たっては事前に伊藤にご連絡ください。

全体の流れ

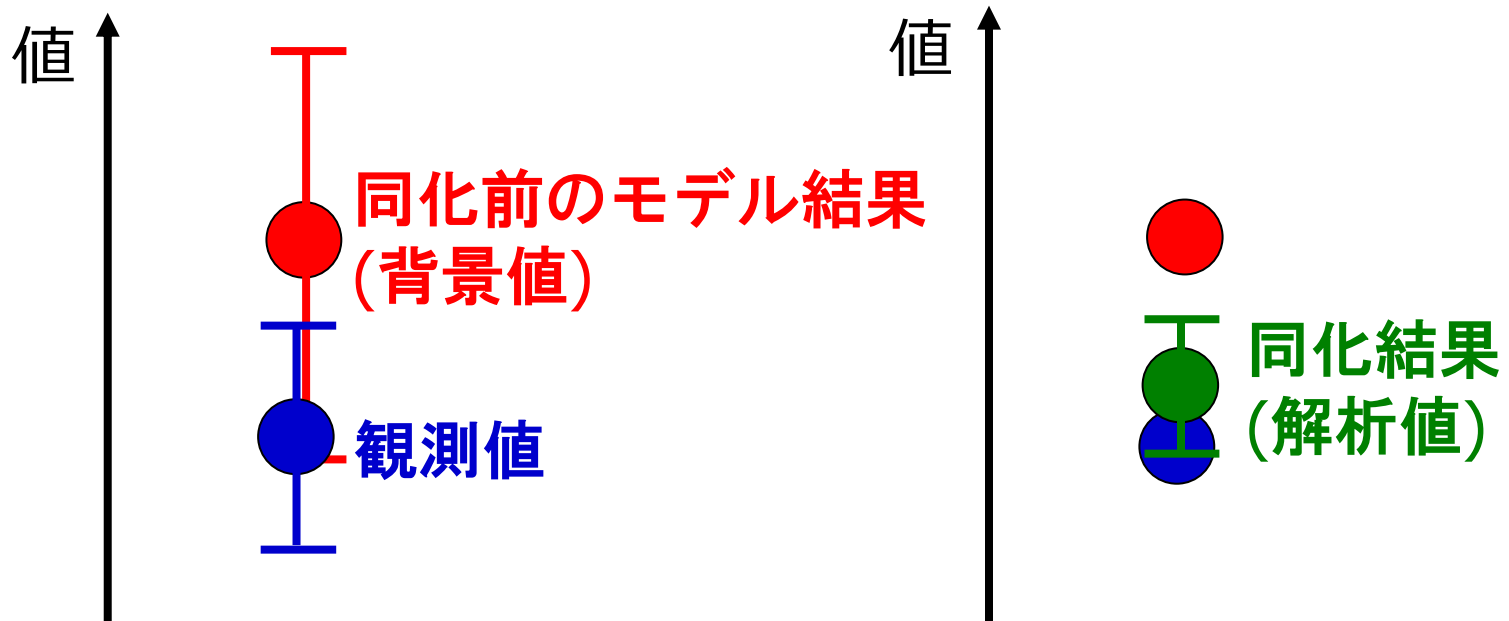
- データ同化とは何か？
- 1変数のデータ同化(時間発展なし)
- 多変数のデータ同化(時間発展なし)
 - 変数の定義・式の導出
 - 誤差の概念
 - 最適内挿法
 - 3次元変分法 (3DVar)
- 多変数のデータ同化(時間発展あり)
 - 変数と誤差の時間発展
 - カルマンフィルタ(KF)
 - 4次元変分法 (4DVar; アジョイント法)

データ同化とは何か？

データ同化とは何か？

- 観測値と同化前のモデル結果（第一推定値；背景値；予報値）の双方から、統計的な最適値を得ること。
- 同化結果（解析値）は、観測値と背景値の信頼度（誤差共分散に代表される誤差統計）に応じて値が決まる。

観測値 + 背景値 + 観測誤差統計 + 背景誤差統計
⇒ 解析値 + 解析誤差統計



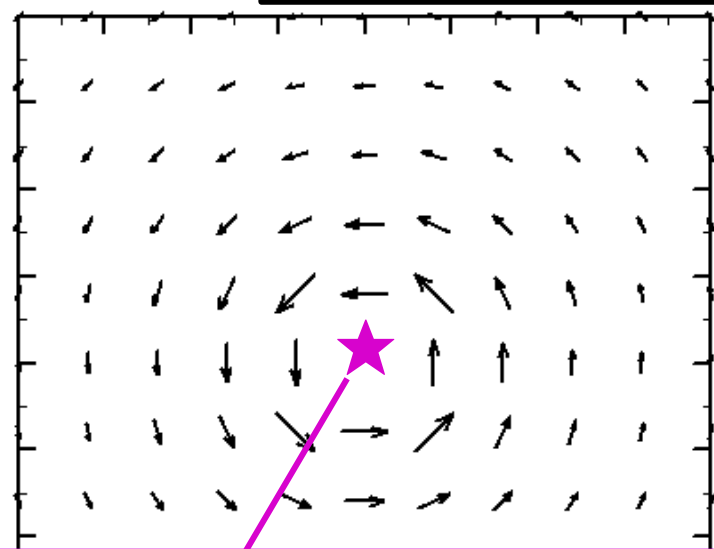
データ同化とは何か？（複数の変数）

イノベーション
(観測値 - 背景値)

誤差共分散の
推定値

解析インクリメント
(解析値 - 背景値)

例：台風の中で中心気圧の観測だけが得られた



観測: 960hPa, 背景値: 980hPa
→ イノベーション: -20hPa

中心気圧と風速の
共分散情報を追加



解析値：更新された気圧

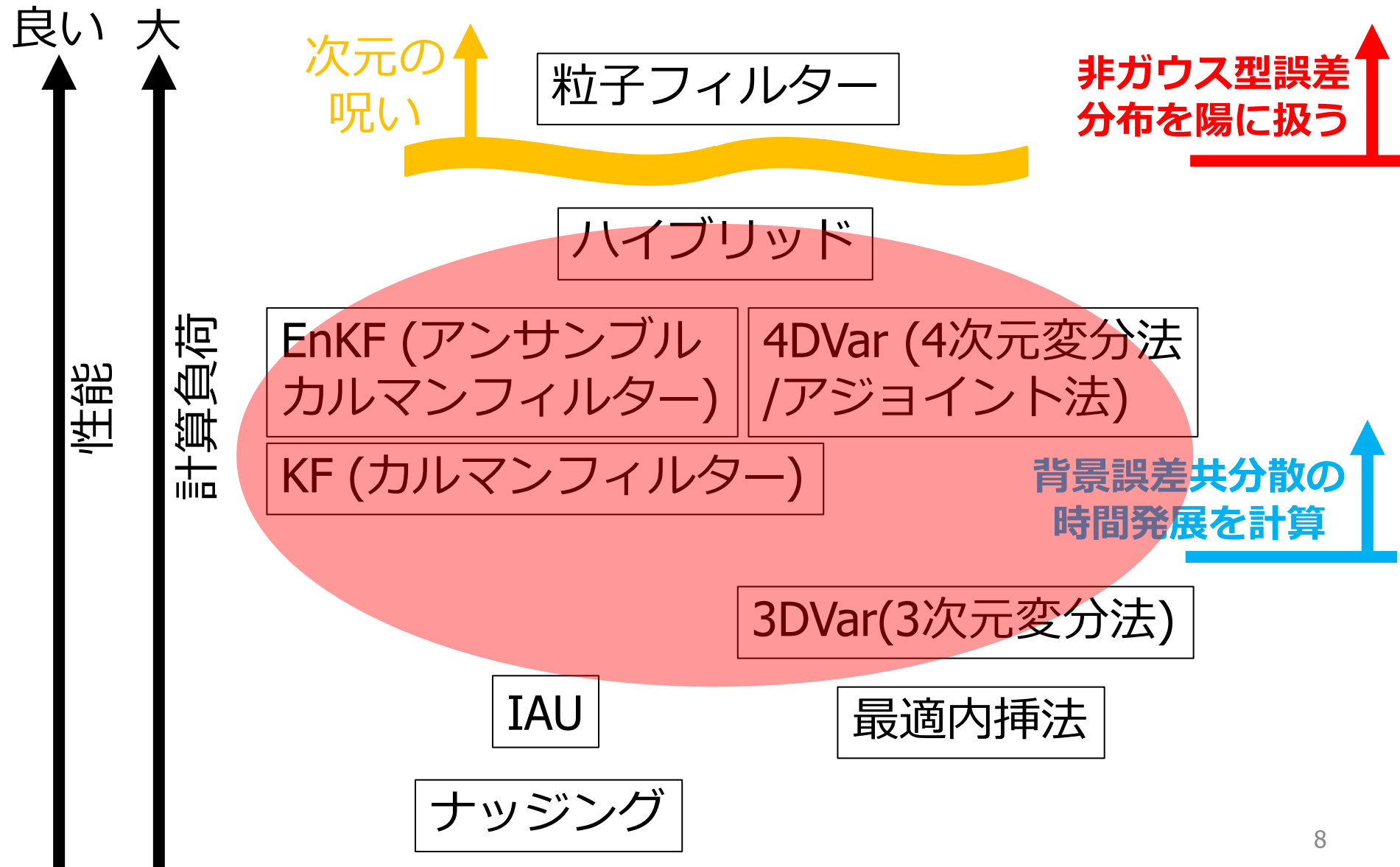
解析値：更新された風

データ同化の活用法

- 将来予測の初期値：品質の良い解析値を将来予測モデルの初期値として入力することで正確な将来予測ができる
- 観測値に近く稠密なデータセットの生成
 - 日々の大気・海洋の解析データの構築
 - 100年以上前の上空の大気の状態を地上気圧データと物理モデルだけから復元する試み
- モデルで用いられているパラメータの値を（物理モデルと観測の精度を組み合わせた精度の範囲で）推定できる。
- 観測システムの設計

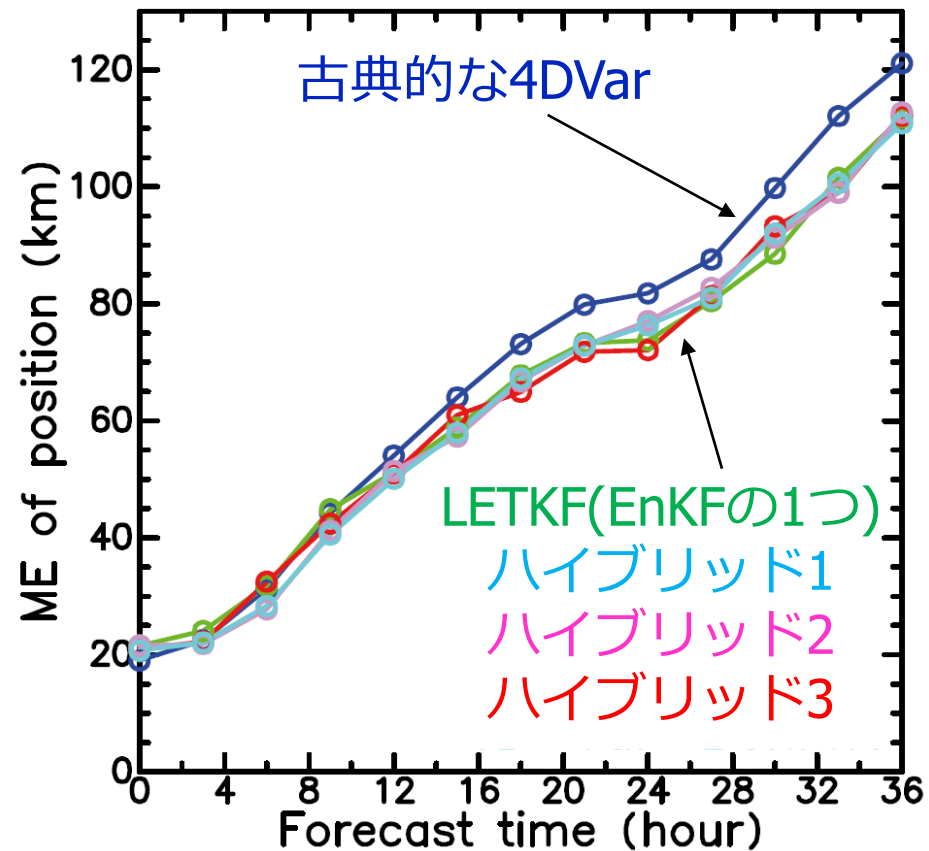
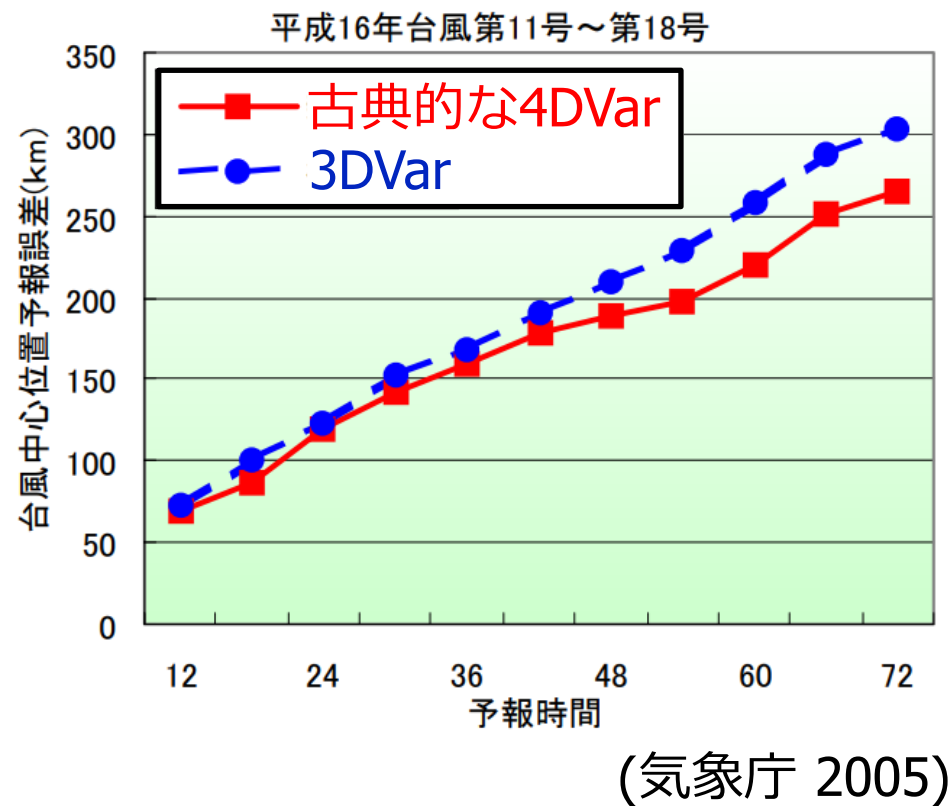
種々の同化手法

対比は一般論。性能や計算負荷は実装や解く問題の性質にも依存。



台風進路予報の誤差

- 高度なデータ同化手法を用いて推定したモデル変数を初期値として将来予測を行った方が予測誤差が小さい。
(「良い」同化手法は実装や解く問題にも依存する)



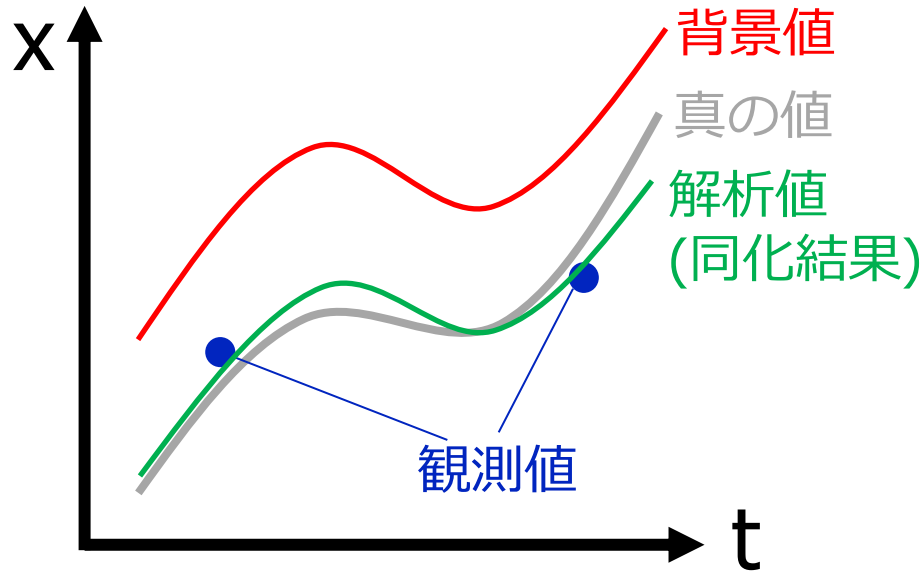
(Ito et al. 2016)

様々なバックグラウンドを お持ちの皆さんへ

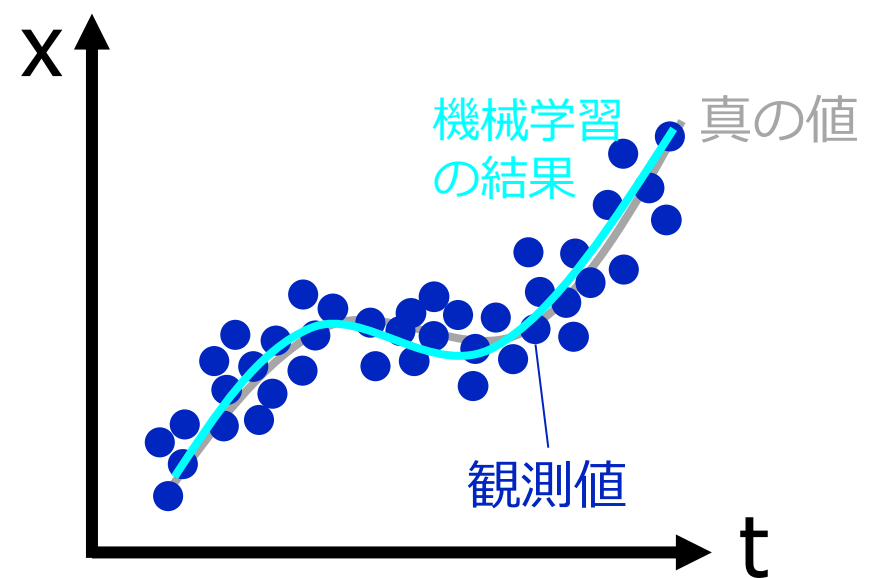
- データ同化と逆問題
 - 線形データ同化はイノベーションから解析インクリメントをダンプ付き最小二乗解として求める逆問題とみなすことができる（伊藤・藤井, 2020, ながれ）
- データ同化と機械学習
 - 「データサイエンスにおける遠い親戚」
 - プロセス駆動型 VS データ駆動型
 - 観測データの誤差に対する考え方

プロセス駆動型 VS データ駆動型

プロセス駆動型：データ同化
(良いモデル & 観測不足なら優位)



データ駆動型：機械学習
(良いモデル無し & 観測大量なら優位)



- 良いモデルの有無と稠密な観測データの有無に依存して、データ同化と機械学習の優位性は異なる。
- 気象学・海洋学：物理学を根拠にそこそこのモデルを作成でき、観測量が不足していたことからデータ同化が発展してきたと考えられる（機械学習も必要に応じて）¹¹。

1 変数のデータ同化 (時間発展を考えない場合)

ある時刻に1つの物理量に関する
背景値と**観測値**の組が得られた
質問「現在の気温は何℃ですか？」

モデルさん $24.0 \pm 5.0^{\circ}\text{C}$

24℃ぐらいじゃないですか？？？

21℃くらいではないでしょうか？

観測さん $21.0 \pm 2.0^{\circ}\text{C}$

観測値と**背景値**の両方に誤差がある。

⇒そもそも、ここに出てきた ± 2.0 や ± 5.0 の意味は？

誤差とは何か？(詳細な説明はまた別途行う)

- 不可知な真の気温を T_{true} とする
- モデルに基づく気温の背景値 T_b との差を $\varepsilon_b = T_b - T_{true}$
- 観測された気温の観測値 T_o との差を $\varepsilon_o = T_o - T_{true}$
- ε_b や ε_o は真値からのずれという意味の「誤差」だが、これらは不可知である。
- 全く同じ状況で複数回観測でき、観測誤差の分散 σ_o^2 を推定できたとする。

$$\sigma_o^2 = \langle \varepsilon_o^2 \rangle \quad (< > \text{は期待値})$$

- 同様に何らかの方法で σ_b を推定できたとする。
- 重要な仮定：バイアスは無い。 $\langle \varepsilon_b \rangle = \langle \varepsilon_o \rangle = 0$

- σ_b^2 は背景誤差分散、 σ_o^2 は観測誤差分散、 σ_b (σ_o) は背景(観測)誤差標準偏差

データ同化分野では
これら誤差統計量も
単に「誤差」と呼ぶ
ことが多い

同化結果とは何か？

下記条件を満たすものを解析値(同化結果)とする

(1) 背景値と観測値の加重平均

(2) (1)のうち誤差標準偏差が最も小さいもの

$$T_a = (1-\alpha)T_b + \alpha T_o = T_b + \alpha(T_o - T_b)$$

↓ 両辺から T_{true} を引く。

$$\boxed{T_a - T_{true}} = (1-\alpha) \boxed{T_b - T_{true}} + \alpha \boxed{T_o - T_{true}}$$

$= \varepsilon_a$ $= \varepsilon_b$ $= \varepsilon_o$

↓ 2乗して複数回の試行で期待値を計算する

$$\boxed{\langle \varepsilon_a^2 \rangle} = (1-\alpha)^2 \boxed{\langle \varepsilon_b^2 \rangle} + \alpha^2 \boxed{\langle \varepsilon_o^2 \rangle} + \cancel{2\alpha(1-\alpha) \langle \varepsilon_o \varepsilon_b \rangle}$$

$= \sigma_a^2$ $= \sigma_b^2$ $= \sigma_o^2$

観測誤差とモデル結果の誤差が無相関なら0

解析値： $T_a = (1-\alpha)T_b + \alpha T_o$

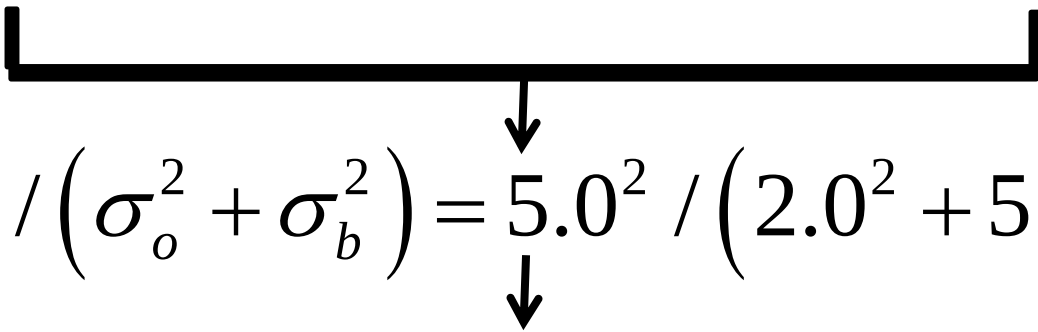
解析誤差分散： $\sigma_a^2 = (1-\alpha)^2 \sigma_b^2 + \alpha^2 \sigma_o^2$

σ_a が最小となる条件を探索

$$\rightarrow \frac{d\sigma_a^2}{d\alpha} = 2(\alpha-1)\sigma_b^2 + 2\alpha\sigma_o^2 = 0 \rightarrow \boxed{\alpha = \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2}}$$

モデルさん：24.0±5.0°C

観測さん：21.0±2.0°C


$$\alpha = \sigma_b^2 / (\sigma_o^2 + \sigma_b^2) = 5.0^2 / (2.0^2 + 5.0^2) = 0.86$$

解析値：0.14×24.0+ 0.86×21.0=21.4°C

解析誤差標準偏差は1.9°C。すなわち21.4±1.9°C

解析値と解析誤差分散

- 重み $\alpha = \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2}$

- 解析値

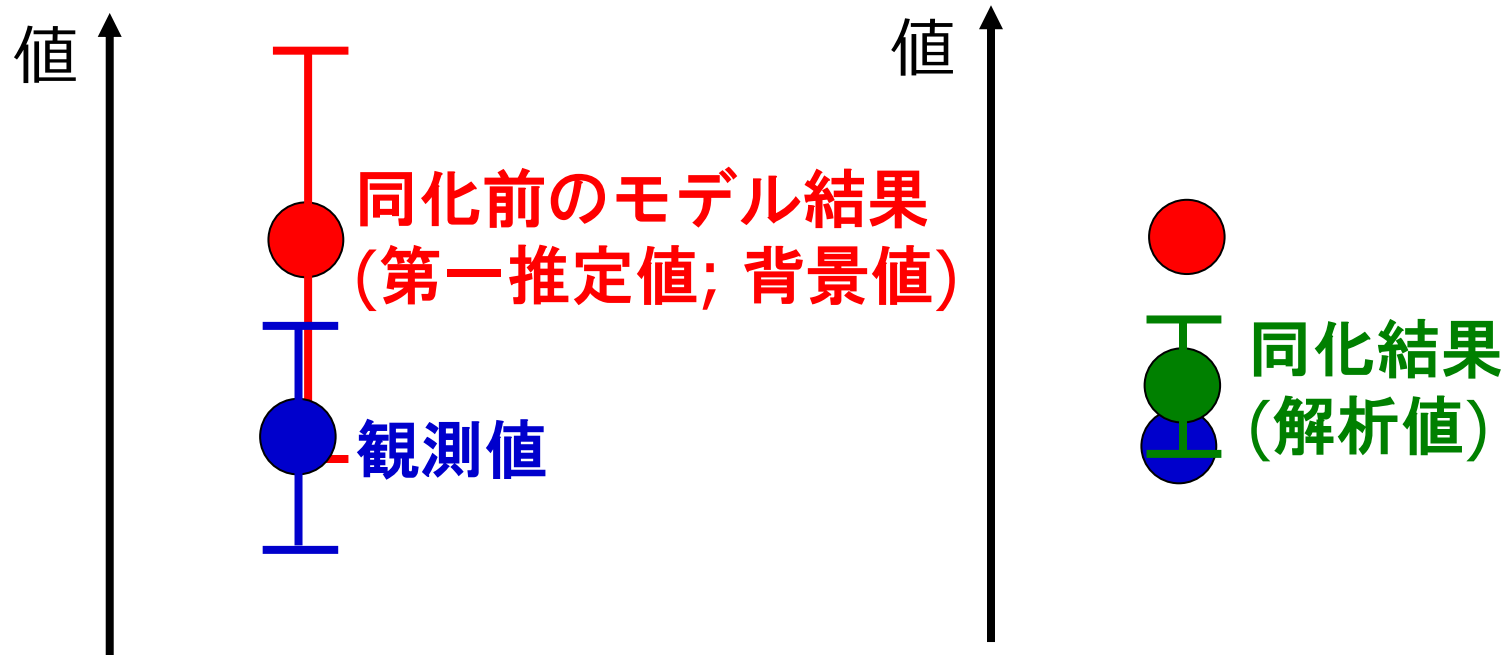
$$T_a = \frac{\sigma_o^2}{\sigma_b^2 + \sigma_o^2} T_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2} T_o = T_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2} (T_o - T_b)$$

解析インクリメント
イノベーション

- 解析誤差分散

$$\sigma_a^2 = (1 - \alpha)^2 \sigma_b^2 + \alpha^2 \sigma_o^2 = \frac{\sigma_o^2 \sigma_b^2}{\sigma_o^2 + \sigma_b^2} \leq \min(\sigma_o^2, \sigma_b^2)$$

観測値+背景値+観測誤差分散+背景誤差分散
⇒解析値+解析誤差分散



- 観測誤差分散 σ_o^2 と背景誤差分散 σ_b^2 に応じて、重み付けを行う。
- 誤差分散が小さいほうに近い値になる。
- 解析誤差分散 σ_a^2 は σ_o^2 や σ_b^2 より小さくなる。

もう一つの解析値の導出方法

評価関数と呼ばれる以下の関数 $J(T)$ を定義し、これが最小となる T を解析値 $T = T_a$ とする。

$$J(T) = \frac{(T - T_b)^2}{2\sigma_b^2} + \frac{(T - T_o)^2}{2\sigma_o^2}$$

解析値の
背景値からの
離れ具合 解析値の
観測値からの
離れ具合

$$\frac{dJ}{dT} = \frac{T - T_b}{\sigma_b^2} + \frac{T - T_o}{\sigma_o^2} = 0 \rightarrow T_a = \frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2} T_b + \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} T_o$$

ベイズの定理：評価関数最小化の根拠

- 条件付確率 $p(x|y)$
ある事象 y が起こる条件のもとでの別の事象 x の確率
 - 馬 x が1着になる確率 $p(x)$
 - 騎手 y が馬 x に騎乗する場合に馬 x が1着になる確率 $p(x|y)$
 - 以下の同時確率 $p(x, y)$ とは異なることに注意
これは騎手 y が馬 x に騎乗し、かつ、馬 x が1着になる確率
- ベイズの定理

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$

 - 証明：同時確率は条件付確率を用いて以下の2通りで表せる。
$$p(x, y) = p(x|y)p(y) = p(y|x)p(x)$$
 - $p(x|y)p(y)$ は「騎手 y が馬 x に騎乗する場合に馬 x が1着になる確率×騎手 y が馬 x に騎乗する確率」だから同時確率に等しい。

背景値 T_b と観測値 T_o の確率が
与えられた $\Rightarrow$ 変数 T の確率

ベイズの定理 $p(T|T_o) = \frac{p(T_o|T)p(T)}{p(T_o)}$

事前確率 $p(T)$ が以下の正規分布で与えられている

$$p(T) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left[-\frac{(T-T_b)^2}{2\sigma_b^2}\right]$$

尤度 $p(T_o|T)$ が以下の正規分布で与えられている

$$p(T_o|T) = \frac{1}{\sqrt{2\pi\sigma_o^2}} \exp\left[-\frac{(T_o-T)^2}{2\sigma_o^2}\right]$$

評価関数

$$p(T|T_o) = \frac{1}{2\pi\sigma_o\sigma_b p(T_o)} \exp\left[-\left(\frac{(T-T_o)^2}{2\sigma_o^2} + \frac{(T-T_b)^2}{2\sigma_b^2}\right)\right]$$

まとめ：1変数のデータ同化

- 観測値+背景値+観測誤差統計+背景誤差統計
⇒解析値+解析誤差統計
- 2つの方法による解析値の導出
 - 解析値は「観測値と背景値の加重平均」で「誤差分散が最小となるもの」（最小分散推定）
 - 解析値は「ベイズの定理を原理とし、観測値と解析値の差と背景値と解析値の差に関し正規分布を仮定したときの確率を最大とするもの」（最尤値推定）
 - 正規分布であれば両者は一致する。

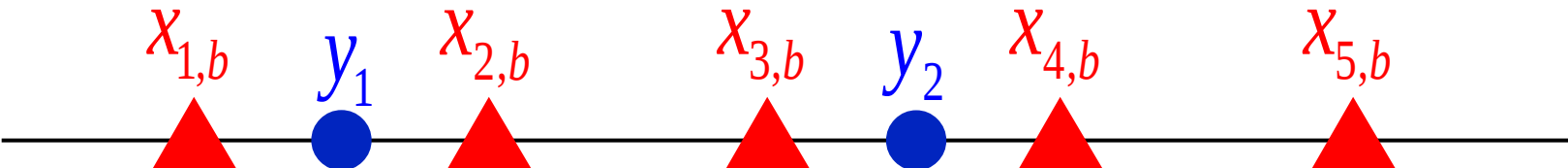
- 解析値
$$T_a = \frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2} T_b + \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} T_o = T_b + \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} (T_o - T_b)$$

- 解析誤差分散
$$\sigma_a^2 = \frac{\sigma_o^2 \sigma_b^2}{\sigma_o^2 + \sigma_b^2} \leq \min(\sigma_o^2, \sigma_b^2)$$

多変数のデータ同化 (時間発展を考えない)

変数の定義

- 解析値ベクトル $\mathbf{x}_a$, 背景値ベクトル $\mathbf{x}_b$, 観測値ベクトル $\mathbf{y}$ ($\mathbf{x}_a$ と $\mathbf{x}_b$ の長さは格子点数×変数の種類)
- 背景誤差共分散行列 $\mathbf{B}$ (または $\mathbf{P}^f$) , 観測誤差共分散行列 $\mathbf{R}$, 解析誤差共分散行列 $\mathbf{P}_a$ (または $\mathbf{P}^a$)
- 観測行列 $\mathbf{H}$: モデルの値を観測値と比較できるようにするための演算子(格子点内挿・変数変換など)。



$$\mathbf{x}_b = \begin{pmatrix} x_{1,b} \\ x_{2,b} \\ x_{3,b} \\ x_{4,b} \\ x_{5,b} \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \mathbf{H} = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 \end{pmatrix} \quad \mathbf{H}\mathbf{x}_b = \begin{pmatrix} 0.5x_{1,b} + 0.5x_{2,b} \\ 0.5x_{3,b} + 0.5x_{4,b} \end{pmatrix}$$

最小分散推定

解析値は線形の加重平均で得られるとする $\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$

両辺からモデル空間の真値 $\mathbf{x}_{true}$ を引き「誤差」を計算

$$\boxed{\mathbf{x}_a - \mathbf{x}_{true}} = \boxed{\mathbf{x}_b - \mathbf{x}_{true}} + \mathbf{K} \left(\boxed{\mathbf{y} - \mathbf{H}\mathbf{x}_{true}} - \mathbf{H} \left(\boxed{\mathbf{x}_b - \mathbf{x}_{true}} \right) \right)$$

$= \boldsymbol{\varepsilon}_a$
 $= \boldsymbol{\varepsilon}_b$
 $= \boldsymbol{\varepsilon}_o$
 $= \boldsymbol{\varepsilon}_b$

複数回の試行について期待値を取り誤差共分散を計算する

$$\mathbf{P}_a = \langle \boldsymbol{\varepsilon}_a \boldsymbol{\varepsilon}_a^T \rangle = \left\langle \left(\boldsymbol{\varepsilon}_b + \mathbf{K}(\boldsymbol{\varepsilon}_o - \mathbf{H}\boldsymbol{\varepsilon}_b) \right) \left(\boldsymbol{\varepsilon}_b + \mathbf{K}(\boldsymbol{\varepsilon}_o - \mathbf{H}\boldsymbol{\varepsilon}_b) \right)^T \right\rangle$$

無バイアスの仮定 $\langle \epsilon_b \rangle = \langle \epsilon_o \rangle = 0$

背景誤差共分散の定義 $\mathbf{B} = \langle \boldsymbol{\varepsilon}_b \boldsymbol{\varepsilon}_b^T \rangle$

観測誤差共分散の定義 $\mathbf{R} = \langle \boldsymbol{\varepsilon}_0 \boldsymbol{\varepsilon}_0^T \rangle$

$\boldsymbol{\varepsilon}_b$ と $\boldsymbol{\varepsilon}_o$ の無相関の仮定 $\langle (\mathbf{H}\boldsymbol{\varepsilon}_b) \boldsymbol{\varepsilon}_o^T \rangle = \mathbf{0}$

$$\mathbf{P}_q = \mathbf{B} - \mathbf{B}\mathbf{H}^T\mathbf{K}^T + \mathbf{K}\mathbf{R}\mathbf{K}^T - \mathbf{K}\mathbf{H}\mathbf{B} + \mathbf{K}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{K}^T$$

$$\mathbf{P}_a = \mathbf{B} - \mathbf{B}\mathbf{H}^T\mathbf{K}^T + \mathbf{K}\mathbf{R}\mathbf{K}^T - \mathbf{K}\mathbf{H}\mathbf{B} + \mathbf{K}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{K}^T$$

対角成分 $\text{trace}(\mathbf{P}_a)$ の最小化：重み行列 $\mathbf{K}$ で微分した値が0

$$\frac{\partial}{\partial \mathbf{X}} \text{trace}(\mathbf{X}\mathbf{Y}) = \mathbf{Y}^T, \quad \frac{\partial}{\partial \mathbf{X}} \text{trace}(\mathbf{X}\mathbf{Y}\mathbf{X}^T) = \mathbf{X}(\mathbf{Y} + \mathbf{Y}^T)$$

$$\frac{\partial}{\partial \mathbf{K}} \text{trace}(\mathbf{P}_a) = -2\mathbf{B}\mathbf{H}^T + 2\mathbf{K}\mathbf{R} + 2\mathbf{K}\mathbf{H}\mathbf{B}\mathbf{H}^T = 0$$

1変数の場合と比べてみよう

$$\alpha = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}$$

数式の整理

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}$$

1変数の場合と比べてみよう

$$x_a = x_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2} (x_o - x_b)$$

$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$ に代入

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

具体的な2変数の問題で考えよう

- 台風における中心気圧 p と最大風速 v の2つのモデル変数があり、中心気圧のみ観測されたとする
 - 中心気圧の背景誤差分散 $\sigma_{p,b}^2$
 - 最大風速の背景誤差分散 $\sigma_{v,b}^2$
 - 中心気圧と最大風速の背景誤差共分散 $r_{pv}\sigma_{p,b}\sigma_{v,b}$
 - 中心気圧の観測誤差分散 $\sigma_{p,o}^2$
- 解析値は「モデル結果と観測結果の線形和」かつ「 $\sigma_{p,a}^2 + \sigma_{v,a}^2$ を最小にするもの」と仮定。

$$\begin{pmatrix} p_a \\ v_a \end{pmatrix} = \begin{pmatrix} p_b \\ v_b \end{pmatrix} + \mathbf{K} \left(\begin{pmatrix} p_o \end{pmatrix} - \underbrace{\begin{pmatrix} 1 & 0 \end{pmatrix}}_{\text{観測行列H}} \begin{pmatrix} p_b \\ v_b \end{pmatrix} \right)$$

モデル結果が2変数、観測値が1変数なので、
観測行列を導入して引き算可能にしている

解析値(同化結果)

- 先ほどの計算結果に従うと

$$\mathbf{B} = \begin{pmatrix} \sigma_{p,b}^2 & r_{pv} \sigma_{p,b} \sigma_{v,b} \\ r_{pv} \sigma_{p,b} \sigma_{v,b} & \sigma_{v,b}^2 \end{pmatrix}, \mathbf{R} = (\sigma_{p,o}^2), \mathbf{H} = (1 \ 0) \text{ として}$$

$$\mathbf{K} = \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} = \begin{pmatrix} \sigma_{p,b}^2 / (\sigma_{p,b}^2 + \sigma_{p,o}^2) \\ r_{pv} \sigma_{p,b} \sigma_{v,b} / (\sigma_{p,b}^2 + \sigma_{p,o}^2) \end{pmatrix}$$

が $\sigma_{p,a}^2 + \sigma_{v,a}^2$ の最小値を与える「重み」となる。

- 陽に書くと
$$\begin{cases} p_a = p_b + \frac{\sigma_{p,b}^2}{\sigma_{p,b}^2 + \sigma_{p,o}^2} (p_o - p_b) \\ v_a = v_b + \frac{r_{pv} \sigma_{p,b} \sigma_{v,b}}{\sigma_{p,b}^2 + \sigma_{p,o}^2} (p_o - p_b) \end{cases}$$

観測値 p_o が
相関を通じ
 v を修正

最尤値推定

背景値に対するベイズの定理の適用
& 観測誤差と背景誤差の無相関の仮定

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}$$

以下の確率分布が正規分布に従うと仮定

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^m |\mathbf{B}|}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) \right]$$

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{R}|}} \exp \left[-\frac{1}{2} (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \right]$$

($|\cdot|$ は行列式を表すものとする)

$$p(\mathbf{x}|\mathbf{y}) \propto \exp \left[-\left(\frac{1}{2} (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \right) \right]$$

$$p(\mathbf{x}|\mathbf{y}) \propto \exp \left[- \left(\frac{1}{2} (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \right) \right]$$

$p(\mathbf{x}|\mathbf{y})$ の最大化問題 ↓

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \text{ の最小化問題}$$

$$\frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) - \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) = \mathbf{0}$$

整理 ↓

$$\mathbf{x} = \mathbf{x}_b + (\mathbf{B}^{-1} + \mathbf{H}\mathbf{R}^{-1}\mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

$$\mathbf{y} - \mathbf{H}\mathbf{x}_b = (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_b) \text{ を代入して整理}$$

$$\boxed{\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_b)}$$

データ同化の具体的なアルゴリズム(1)

- 最小分散推定・最尤値推定ともに線形では同じ解。

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T \left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T \right)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

$(J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}))$ の最小化問題と等価)

(1) $\mathbf{B}$, $\mathbf{R}$ を何らかの方法で設定

1-a. $\mathbf{B}$ を時間変化しない静的なものとして扱う

1-b. $\mathbf{B}$ を時間変化する動的なものとして扱う

(2) 下記のいずれかの問題を解く

2-a. $\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T \left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T \right)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_b)$ を陽に計算

2-b. $J(\mathbf{x})$ の最小化問題を解いて $\mathbf{x} = \mathbf{x}_a$ を決定

データ同化の具体的なアルゴリズム(2)

- 主要な静的手法(教科書2章)

- 最適内挿法 : $1a+2a$

- $\mathbf{B}$, $\mathbf{R}$ を何らかの方法で設定し、 $\mathbf{x}_a$ を陽に計算

- 3次元変分法 (3DVar) (*): $1a+2b$

- $\mathbf{B}$, $\mathbf{R}$ を何らかの方法で設定し、 $J(\mathbf{x})$ の最小値問題を(通常は) $J(\mathbf{x})$ の勾配を計算して $\mathbf{x} = \mathbf{x}_a$ を得る。

- 主要な動的手法(詳細はのちほど)

- カルマンフィルタ(KF; 教科書3章) : $1b+2a$

- 4次元変分法 (4DVar; アジョイント法) (教科書4章) : $1b+2b$

(*) 1次元・2次元の格子点上に定義された変数に対するデータ同化でも、時間軸が含まれていなければ3次元変分法と呼ぶことがある。逆に、時間軸が入っていれば、4次元変分法と呼ばれる。

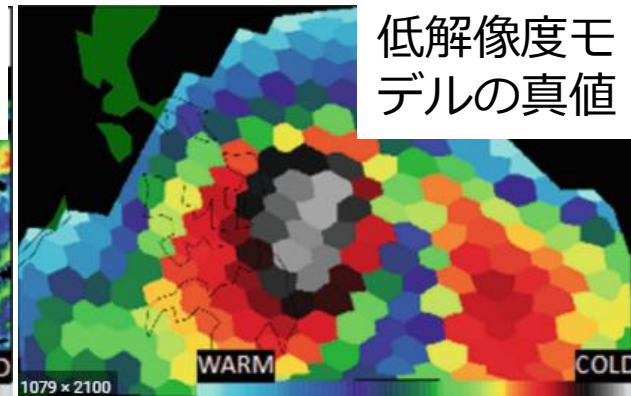
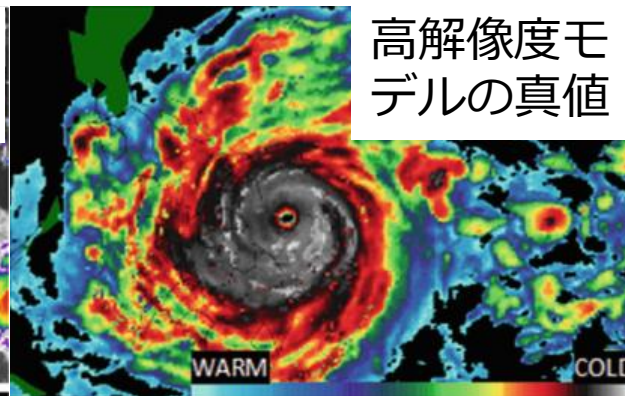
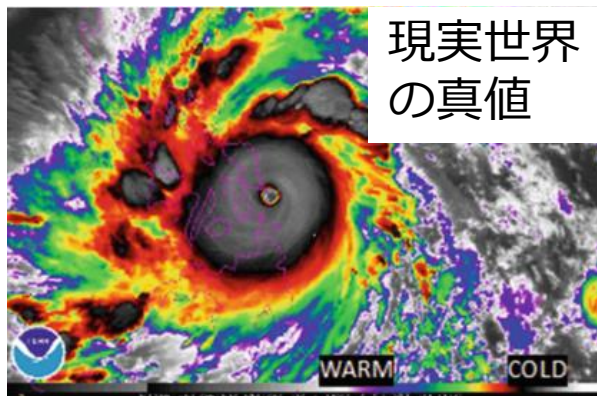
B, Rの設定・・・の前に「誤差の概念」(1)

- 古典的な統計論の立場で考える。
- 背景誤差共分散行列**B**と観測誤差共分散行列**R**
⇒背景誤差や観測誤差は現実世界の真値からの距離ではなく、モデルの世界の真値からの距離に対応することに注意。

$$\boxed{\mathbf{x}_a - \mathbf{x}_{true}} = \boxed{\mathbf{x}_b - \mathbf{x}_{true}} + \mathbf{K} \left(\boxed{\mathbf{y} - \mathbf{H}\mathbf{x}_{true}} - \mathbf{H} \left(\boxed{\mathbf{x}_b - \mathbf{x}_{true}} \right) \right)$$

$= \boldsymbol{\varepsilon}_a \qquad \qquad = \boldsymbol{\varepsilon}_b \qquad \qquad = \boldsymbol{\varepsilon}_o \qquad \qquad = \boldsymbol{\varepsilon}_b$

- モデルの世界の真値は実在しない。誰かが「この値をモデルの世界の真値とする」と宣言した場合、それが真値となる。
- モデルの世界の真値と実世界の真値の差を表現誤差と呼ぶ。



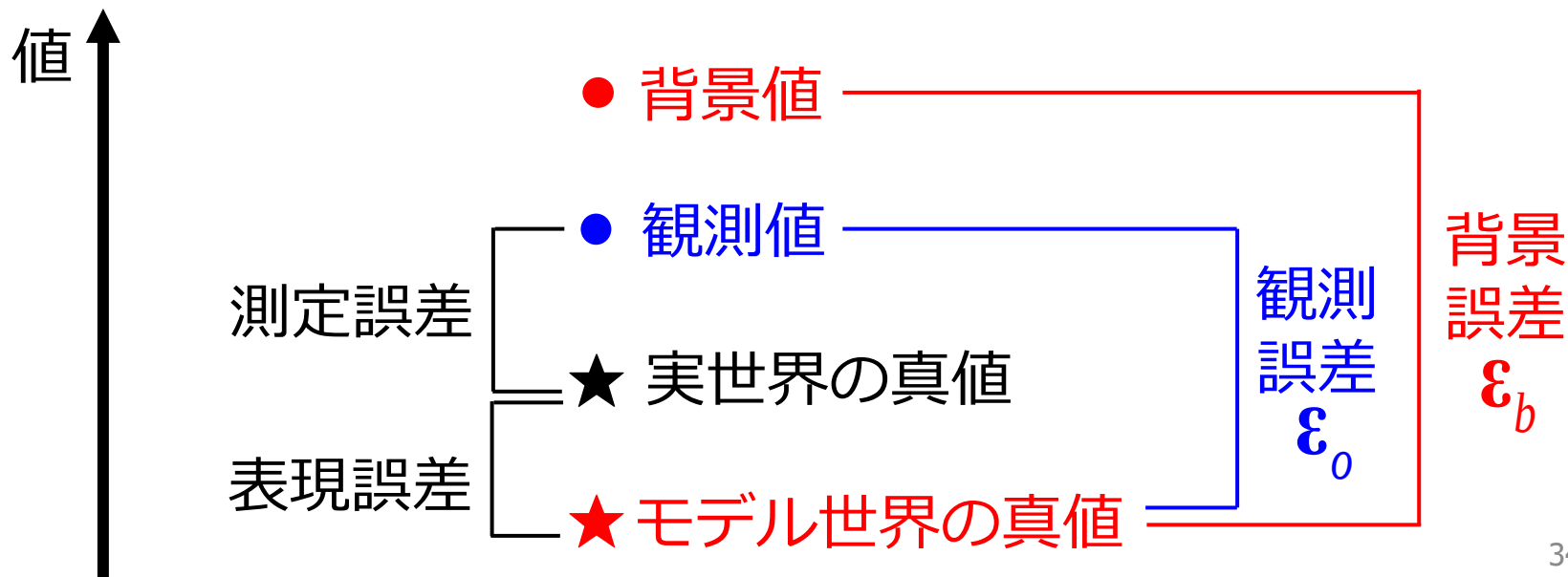
(現実世界の真値は実際には不可知だがWehner et al. (2019)の図を用い模式的に説明)

B, Rの設定・・・の前に「誤差の概念」(2)

- 観測誤差が測器誤差（実世界の真値からの距離）ではなく、モデルの世界の真値の内挿値からの距離であることに注意。

$$\underbrace{\mathbf{x}_a - \mathbf{x}_{true}}_{=\boldsymbol{\varepsilon}_a} = \underbrace{\mathbf{x}_b - \mathbf{x}_{true}}_{=\boldsymbol{\varepsilon}_b} + \mathbf{K} \left(\underbrace{\mathbf{y} - \mathbf{H}\mathbf{x}_{true}}_{=\boldsymbol{\varepsilon}_o} - \mathbf{H} \left(\underbrace{\mathbf{x}_b - \mathbf{x}_{true}}_{=\boldsymbol{\varepsilon}_b} \right) \right)$$

- 仮に、観測行列 $\mathbf{H}$ による射影(内挿や変数変換)を無視すると
⇒ 観測誤差 = 観測値 - モデルの世界の真値
= 測定誤差 + 表現誤差



「誤差の概念」に関し 伊藤が個人的に思うこと

- 「観測誤差」という用語を，観測値とモデルの世界の真値の差に用いるため，誤解を招くことが多い。
- とはいえ，実際に我々が知ることができるのは，せいぜい，モデルの世界と観測値だけなので，それを使ってどうにかこうにか，問題を設定し解いていくしかない。
(モデルを改良し，現実世界の振舞いをより精度よく表現できるなら，勿論素晴らしい！)

B, Rの設定: 単純な方法

- 観測誤差共分散行列**R**は観測値どうしに相関がないとして、対角行列とすることがほとんどである。
- 背景誤差共分散行列**B**は、単純には、以下のような設定の方法がある。
 - 対角行列であると仮定し、典型的な変動幅を入れる
 - 各変数が別々の位置に配置されている場合、距離の関数として設定する


$$(\mathbf{B})_{ij} = \sigma_{i,b} \sigma_{j,b} \exp \left[-\frac{d_{ij}^2}{2L^2} \right]$$

- 物理法則から異なる物理量間の共分散の近似式を陽に決められる場合もある（気象学や海洋物理学では、流速と海面高度などの関係は地衡流として近似できる）

Bの設定: コバリアンス・マッチング

(このスライドでは $\mathbf{x}_{true}$ を $\mathbf{x}_t$ と表記している)

$$\mathbf{x}_b = \mathbf{x}_t + \boldsymbol{\varepsilon}_b$$

観測行列 $\mathbf{H}$ を作用

$$\mathbf{H}\mathbf{x}_b = \mathbf{H}\mathbf{x}_t + \mathbf{H}\boldsymbol{\varepsilon}_b$$

複数回試行をもとに共分散を計算

$$\langle \mathbf{H}\mathbf{x}_b \mathbf{x}_b^T \mathbf{H}^T \rangle = \langle (\mathbf{H}\mathbf{x}_t + \mathbf{H}\boldsymbol{\varepsilon}_b)(\mathbf{H}\mathbf{x}_t + \mathbf{H}\boldsymbol{\varepsilon}_b)^T \rangle$$

真値と背景誤差の
間に無相関を仮定 $\langle \mathbf{x}_t \boldsymbol{\varepsilon}_b^T \rangle = \mathbf{0}$

背景誤差共分
散行列の定義 $\langle \boldsymbol{\varepsilon}_b \boldsymbol{\varepsilon}_b^T \rangle = \mathbf{B}$

$$\langle \mathbf{H}\mathbf{x}_b \mathbf{x}_b^T \mathbf{H}^T \rangle = \langle \mathbf{H}\mathbf{x}_t \mathbf{x}_t^T \mathbf{H}^T \rangle + \mathbf{H}\mathbf{B}\mathbf{H}^T$$

$$\mathbf{y} - \mathbf{H}\mathbf{x}_b = \boldsymbol{\varepsilon}_o - \mathbf{H}\boldsymbol{\varepsilon}_b$$

$$\mathbf{y} = \mathbf{H}\mathbf{x}_t + \boldsymbol{\varepsilon}_o$$

複数回試行をもとに共分散を計算

$$\langle \mathbf{y}\mathbf{y}^T \rangle = \langle (\mathbf{H}\mathbf{x}_t + \boldsymbol{\varepsilon}_o)(\mathbf{H}\mathbf{x}_t + \boldsymbol{\varepsilon}_o)^T \rangle$$

真値と観測誤差の
間に無相関を仮定 $\langle (\mathbf{H}\mathbf{x}_t) \boldsymbol{\varepsilon}_o^T \rangle = \mathbf{0}$

観測誤差共分
散行列の定義 $\langle \boldsymbol{\varepsilon}_o \boldsymbol{\varepsilon}_o^T \rangle = \mathbf{R}$

$$\langle \mathbf{y}\mathbf{y}^T \rangle = \langle \mathbf{H}\mathbf{x}_t \mathbf{x}_t^T \mathbf{H}^T \rangle + \mathbf{R}$$

複数回試行をもとに共分散を計算

$$\langle (\mathbf{y} - \mathbf{H}\mathbf{x}_b)(\mathbf{y} - \mathbf{H}\mathbf{x}_b)^T \rangle = \langle (\boldsymbol{\varepsilon}_o - \mathbf{H}\boldsymbol{\varepsilon}_b)(\boldsymbol{\varepsilon}_o - \mathbf{H}\boldsymbol{\varepsilon}_b)^T \rangle$$

観測誤差と背景誤差の
間に無相関を仮定 $\langle \boldsymbol{\varepsilon}_o (\mathbf{H}\boldsymbol{\varepsilon}_b)^T \rangle = \mathbf{0}$

BとRの定義

$$\langle (\mathbf{y} - \mathbf{H}\mathbf{x}_b)(\mathbf{y} - \mathbf{H}\mathbf{x}_b)^T \rangle = \mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T$$

整理

$$\mathbf{R} = \left\{ \langle (\mathbf{y} - \mathbf{H}\mathbf{x}_b)(\mathbf{y} - \mathbf{H}\mathbf{x}_b)^T \rangle - \langle \mathbf{H}\mathbf{x}_b\mathbf{x}_b^T\mathbf{H}^T \rangle + \langle \mathbf{y}\mathbf{y}^T \rangle \right\} / 2$$

$$\mathbf{H}\mathbf{B}\mathbf{H}^T = \left\{ \langle (\mathbf{y} - \mathbf{H}\mathbf{x}_b)(\mathbf{y} - \mathbf{H}\mathbf{x}_b)^T \rangle + \langle \mathbf{H}\mathbf{x}_b\mathbf{x}_b^T\mathbf{H}^T \rangle - \langle \mathbf{y}\mathbf{y}^T \rangle \right\} / 2$$

$\mathbf{y}$ と $\mathbf{x}_b$ を集めて来て統計処理をすることで推定できる。

Bの設定：コバリアンスマッチング

- 観測値と背景値をたくさん集めてきて統計を取ることに
より, $\mathbf{HBH}^T$ や $\mathbf{R}$ を推定できる。
- Fu et al. (1993)

$$\mathbf{R} = \left\{ \left\langle (\mathbf{y} - \mathbf{H}\mathbf{x}_b)(\mathbf{y} - \mathbf{H}\mathbf{x}_b)^T \right\rangle - \left\langle \mathbf{H}\mathbf{x}_b\mathbf{x}_b^T\mathbf{H}^T \right\rangle + \left\langle \mathbf{y}\mathbf{y}^T \right\rangle \right\} / 2$$

$$\mathbf{HBH}^T = \left\{ \left\langle (\mathbf{y} - \mathbf{H}\mathbf{x}_b)(\mathbf{y} - \mathbf{H}\mathbf{x}_b)^T \right\rangle + \left\langle \mathbf{H}\mathbf{x}_b\mathbf{x}_b^T\mathbf{H}^T \right\rangle - \left\langle \mathbf{y}\mathbf{y}^T \right\rangle \right\} / 2$$

- Fu et al. (1993)の簡略化

$$\mathbf{R} = \left\langle \mathbf{y}\mathbf{y}^T \right\rangle - \left\langle \mathbf{y}(\mathbf{H}\mathbf{x}_b)^T \right\rangle$$

$$\mathbf{HBH}^T = \left\langle \mathbf{H}\mathbf{x}_b\mathbf{x}_b^T\mathbf{H}^T \right\rangle - \left\langle \mathbf{y}(\mathbf{H}\mathbf{x}_b)^T \right\rangle$$

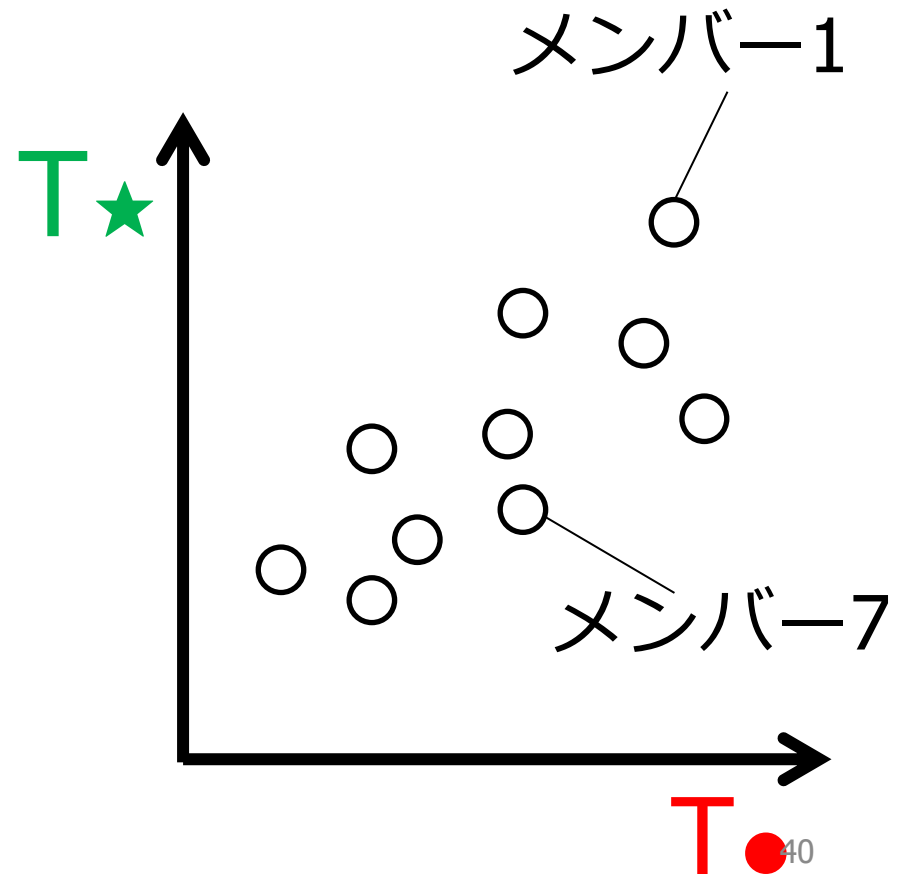
- Desroziers et al. (2005)

$$\mathbf{R} = \left\langle (\mathbf{y} - \mathbf{H}\mathbf{x}_a)(\mathbf{y} - \mathbf{H}\mathbf{x}_b)^T \right\rangle$$

$$\mathbf{HBH}^T = \left\langle (\mathbf{H}\mathbf{x}_a - \mathbf{H}\mathbf{x}_b)(\mathbf{y} - \mathbf{H}\mathbf{x}_a)^T \right\rangle$$

Bの設定：アンサンブル計算

- アンサンブル予報(モデルの初期値を何らかの方法で散らばしたあとに予報計算)を実施。
→アンサンブルのばらつきから、●**地点の気温**と★**地点の気温**の背景誤差分散共分散を推定。



データ同化アルゴリズム：最適内挿法

＜基礎となる式＞

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T \left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T \right)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

- 事前に何らかの方法で $\mathbf{B}$, $\mathbf{R}$ を推定しておく。
- 観測値 $\mathbf{y}$ と背景値 $\mathbf{x}_b$ が与えられたら, 式を計算して解析値 $\mathbf{x}_a$ を求める。
- 計算のコツ
 - $\mathbf{K} = \mathbf{B}\mathbf{H}^T \left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T \right)^{-1}$ を陽に解くには逆行列の計算（大自由度だと大変！）が必要となるので、数値的に $\mathbf{K} \left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T \right) = \mathbf{B}\mathbf{H}^T$ を解いて $\mathbf{K}$ を求めたりする。
 - ある距離内にある観測値のみ使用⇒効率的

データ同化アルゴリズム：3次元変分法

<基礎となる式>

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) \text{ の最小化}$$
$$\frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) - \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x})$$

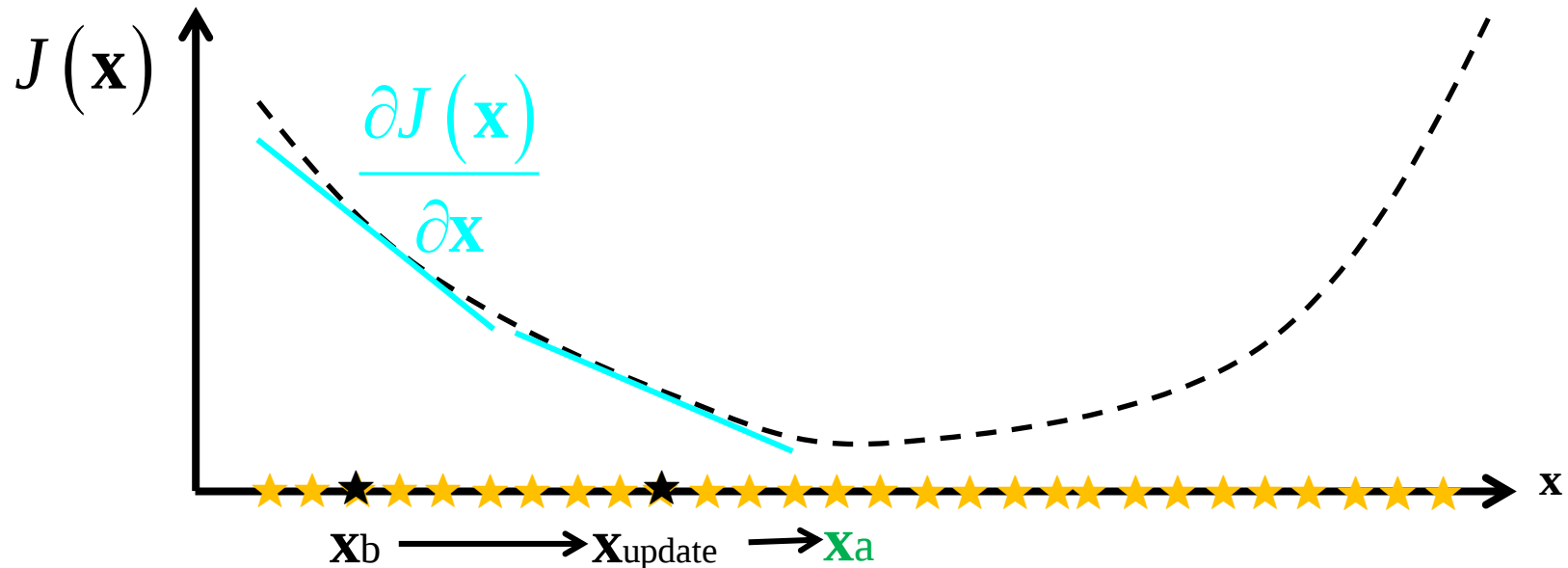
- 事前に何らかの方法で $\mathbf{B}$, $\mathbf{R}$ を推定しておく。
- $\partial J(\mathbf{x}) / \partial \mathbf{x}$ を使い, $J(\mathbf{x})$ を最小とする $\mathbf{x} = \mathbf{x}_a$ を探索する
(次のスライドを参照)。
- 計算のコツ
 - $\mathbf{x} = \mathbf{B}^{1/2} \mathbf{v}$ となる $\mathbf{v}$ を定義すると, $\partial J / \partial \mathbf{v} = \mathbf{B}^{1/2} \partial J / \partial \mathbf{x}$ となるので, これを最小値探索に用いると J を最小とする $\mathbf{v}$ が求まる。ここから $\mathbf{x}$ を求めると, 結果として, 計算が大変な $\mathbf{B}^{-1}$ を使わずに済む (平方根行列 $\mathbf{B}^{1/2}$ を計算する必要はあるが、これは固有値分解で容易に求まる)

最小値探索

$$J(\mathbf{x}) = (1/2)(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + (1/2)(\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x})$$

$$\partial J(\mathbf{x}) / \partial \mathbf{x} = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{y})$$

- どの $\mathbf{x}$ で $J(\mathbf{x})$ が最小となるかは事前にはわからない
- 解法1： $\mathbf{x}$ の値を変えて実験を繰り返す。
- 解法2：まずは $\mathbf{x} = \mathbf{x}_b$ と想定し, $\partial J(\mathbf{x}) / \partial \mathbf{x}$ を計算。その結果から $J(\mathbf{x})$ が小さくなるように $\mathbf{x}$ を更新。この操作を複数回繰り返し, 落ち着いたところを $\mathbf{x} = \mathbf{x}_a$ とみなす。



拡張：非線形な観測演算子への対応

- モデル変数と観測値が非線形的に対応している場合 (すなわち $\mathbf{H}\mathbf{x}$ の代わりに非線形な関数 $H(\mathbf{x})$ を使う) にも、3次元変分法(最尤値推定)では対応可能

$$p(\mathbf{y} | \mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{R}|}} \exp \left[-\frac{1}{2} (\mathbf{y} - \underbrace{H(\mathbf{x})}_{\text{red wavy line}})^T \mathbf{R}^{-1} (\mathbf{y} - \underbrace{H(\mathbf{x})}_{\text{red wavy line}}) \right]$$

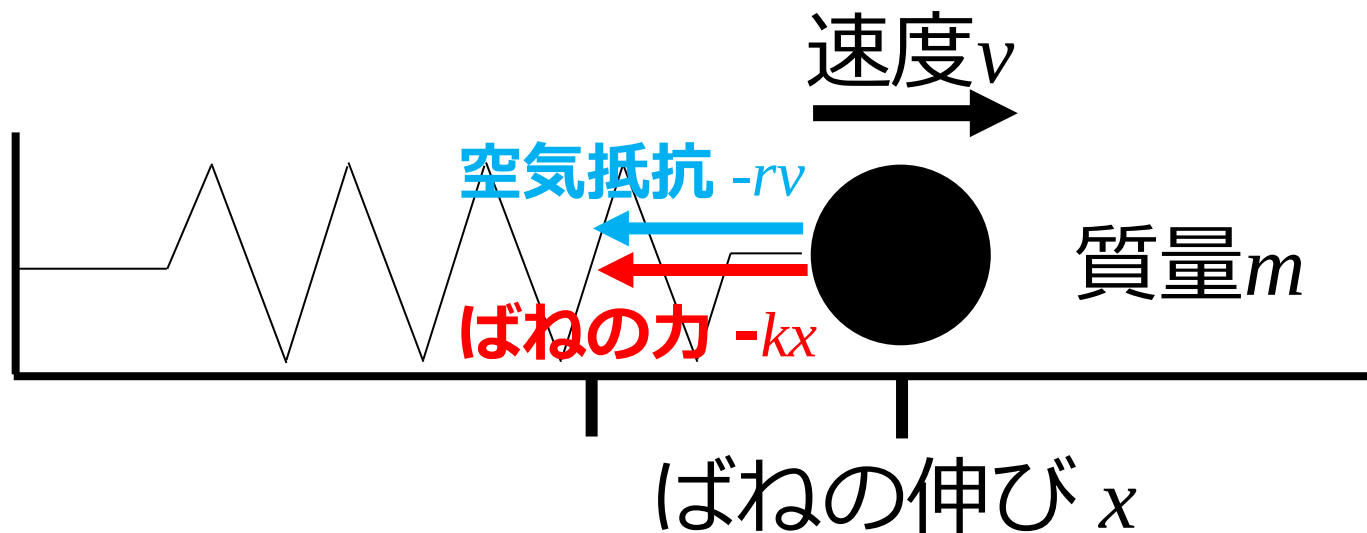
- 評価関数の勾配は下記のように近似されるので、これを用いて評価関数の極小値を探索すればよい。

$$\frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) - \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y} - \underbrace{H(\mathbf{x})}_{\text{red wavy line}})$$

多変数のデータ同化 (時間発展を考える)

ばね運動の時間発展の行列形式での表現

- ばねの運動について考える (x と v の2変数)。



速度の定義

$$\frac{dx}{dt} = v$$

前方差分で離散化

$$\frac{x_i - x_{i-1}}{\Delta t} = v_{i-1}$$

運動方程式

$$m \frac{dv}{dt} = -kx - rv$$

$$m \frac{v_i - v_{i-1}}{\Delta t} = -kx_{i-1} - rv_{i-1}$$

$$\frac{x_i - x_{i-1}}{\Delta t} = v_{i-1}, \quad m \frac{v_i - v_{i-1}}{\Delta t} = -kx_{i-1} - rv_{i-1}$$

整理 ↓

$$x_i = x_{i-1} + \Delta t v_{i-1}, \quad v_i = -\frac{k\Delta t}{m} x_{i-1} + \left(1 - \frac{r\Delta t}{m}\right) v_{i-1}$$

行列形式で書き直す ↓

$$\begin{pmatrix} x_i \\ v_i \end{pmatrix} = \begin{pmatrix} 1 & \Delta t \\ -k\Delta t / m & 1 - r\Delta t / m \end{pmatrix} \begin{pmatrix} x_{i-1} \\ v_{i-1} \end{pmatrix}$$

ばね運動の「誤差」の時間発展の 行列形式での表現

- 位置 x と速度 v に微小な違い $\delta x_i, \delta v_i$ があるとする（加わってもモデルに変化はない）。

$$\frac{d(x + \delta x)}{dt} = v + \delta v, \quad m \frac{d(v + \delta v)}{dt} = -k(x + \delta x) - r(v + \delta v)$$

- 違いが加わる前の式を引き算 $\Rightarrow$ この場合は同じ式,

$$\frac{d\delta x}{dt} = \delta v, \quad m \frac{d\delta v}{dt} = -k\delta x - r\delta v$$

となるので、誤差の時間発展も行列で同様に書ける。

$$\begin{pmatrix} \delta x_i \\ \delta v_i \end{pmatrix} = \begin{pmatrix} 1 & \Delta t \\ -k\Delta t / m & 1 - r\Delta t / m \end{pmatrix} \begin{pmatrix} \delta x_{i-1} \\ \delta v_{i-1} \end{pmatrix}$$

$= \mathbf{\epsilon}_i$

$= \mathbf{M}$ (状態遷移行列；時間推進演算子；
誤差伝播行列；接線形演算子など)

非線形モデルでの誤差の時間発展

- 変数の時間発展が非線形的な場合でも，微小変化量の時間発展は2乗以上の項を無視すれば，線形で近似できる。

$$\frac{dx}{dt} = -\sigma(x - y)$$

$$\frac{d(x + \delta x)}{dt} = -\sigma(x + \delta x - y - \delta y)$$

$$\frac{dy}{dt} = -y - xz + rx$$

$$\frac{d(y + \delta y)}{dt} = -(y + \delta y) - (x + \delta x)(z + \delta z) + r(x + \delta x)$$

$$\frac{dz}{dt} = xy - bz$$

$$\frac{d(z + \delta z)}{dt} = (x + \delta x)(y + \delta y) - b(z + \delta z)$$

ただし， $\mathbf{M}$ は線形の場合と違い，時刻ごとの状態にも依存する。

$$\frac{d\delta x}{dt} = -\sigma(\delta x - \delta y)$$

$$\frac{d\delta y}{dt} = -\delta y - x\delta z - z\delta x + r\delta x$$

$$\frac{d\delta z}{dt} = x\delta y + y\delta x - b\delta z$$

$$\mathbf{M}_i = \begin{pmatrix} 1 - \Delta t \sigma & \Delta t \sigma & 0 \\ -\Delta t z_i + \Delta t r & 1 - \Delta t & -\Delta t x_i \\ \Delta t y_i & \Delta t x_i & 1 - \Delta t b \end{pmatrix}$$

$$\boldsymbol{\varepsilon}_i = \mathbf{M}_{i-1} \boldsymbol{\varepsilon}_{i-1}$$

非線形モデルでの誤差の時間発展

- モデルの状態変数が以下の式に従って時間変化するとする。(下付き文字の0やtは時間を表す)

$$\mathbf{x}_t = M(\mathbf{x}_0)$$

- モデル変数の誤差 $\boldsymbol{\varepsilon}_b$ の時間発展は

$$\mathbf{x}_t + \boldsymbol{\varepsilon}_{b,t} = M(\mathbf{x}_0 + \boldsymbol{\varepsilon}_{b,0})$$

から引き算して、テイラー展開の1次の近似を適用し

$$\boldsymbol{\varepsilon}_{b,t} = M(\mathbf{x}_0 + \boldsymbol{\varepsilon}_{b,0}) - M(\mathbf{x}_0) \cong \frac{\partial M}{\partial \mathbf{x}_0} \boldsymbol{\varepsilon}_{b,0}$$

- $\mathbf{M} \equiv \frac{\partial M}{\partial \mathbf{x}_0}$ が状態遷移行列(接線形演算子)である。

- 時刻0~Tが複数タイムステップなら $\mathbf{M} = \mathbf{M}_{T-1} \cdots \mathbf{M}_1 \mathbf{M}_0$

背景誤差共分散の時間発展

- 誤差の時間発展は $\boldsymbol{\varepsilon}_{b,t} \cong \frac{\partial M}{\partial \mathbf{x}_0} \boldsymbol{\varepsilon}_{b,0} = \mathbf{M} \boldsymbol{\varepsilon}_{b,0}$

- 背景誤差共分散は

$$\mathbf{B}_0 = \langle \boldsymbol{\varepsilon}_{b,0} \boldsymbol{\varepsilon}_{b,0}^T \rangle$$

2変数のときの誤差の時間発展が

$$\mathbf{B} = \begin{pmatrix} \sigma_{p,b}^2 & r_{pv} \sigma_{p,b} \sigma_{v,b} \\ r_{pv} \sigma_{p,b} \sigma_{v,b} & \sigma_{v,b}^2 \end{pmatrix} = \begin{pmatrix} \langle \varepsilon_p^2 \rangle & \langle \varepsilon_p \varepsilon_v \rangle \\ \langle \varepsilon_p \varepsilon_v \rangle & \langle \varepsilon_v^2 \rangle \end{pmatrix}$$

であったことを思い出そう。

$$\mathbf{B}_t = \langle \boldsymbol{\varepsilon}_{b,t} \boldsymbol{\varepsilon}_{b,t}^T \rangle = \langle \mathbf{M} \boldsymbol{\varepsilon}_{b,0} \boldsymbol{\varepsilon}_{b,0}^T \mathbf{M}^T \rangle = \mathbf{M} \langle \boldsymbol{\varepsilon}_{b,0} \boldsymbol{\varepsilon}_{b,0}^T \rangle \mathbf{M}^T = \mathbf{M} \mathbf{B}_0 \mathbf{M}^T$$

- 初期の $\mathbf{B}_0$ が与えれば、摂線形モデル $\mathbf{M}$ と随伴(アジョイント)モデル $\mathbf{M}^T$ を使って原理的には $\mathbf{B}_t$ も計算可能となる。

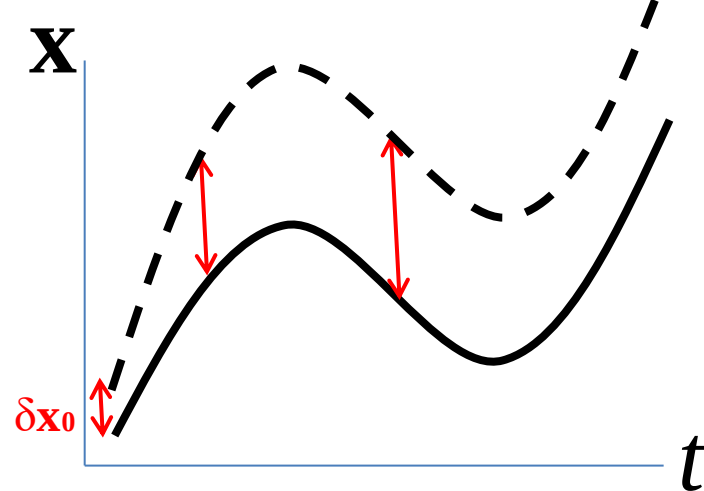
非線形モデル

$$\begin{aligned}\frac{dx}{dt} &= -\sigma(x-y) \\ \frac{dy}{dt} &= -y - xz + rx \\ \frac{dz}{dt} &= xy - bz\end{aligned}$$

$$\mathbf{x}_i \equiv \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}$$

接線形モデル

$$\begin{aligned}\frac{d\delta x}{dt} &= -\sigma(\delta x - \delta y) \\ \frac{d\delta y}{dt} &= -\delta y - x\delta z - z\delta x + r\delta x \\ \frac{d\delta z}{dt} &= x\delta y + y\delta x - b\delta z\end{aligned}$$



$$\delta \mathbf{x}_i \simeq \frac{\partial \mathbf{M}(\mathbf{x}_{i-1})}{\partial \mathbf{x}_{i-1}} \delta \mathbf{x}_{i-1} = \mathbf{M}_{i-1} \delta \mathbf{x}_{i-1}$$

$$\mathbf{M}_i = \begin{pmatrix} 1 - \Delta t \sigma & \Delta t \sigma & 0 \\ -\Delta t z_i + \Delta t r & 1 - \Delta t & -\Delta t x_i \\ \Delta t y_i & \Delta t x_i & 1 - \Delta t b \end{pmatrix}$$

転置(厳密に言うと, 内積の定義にも依存するが . . .)

随伴(アジョイント)モデル $\mathbf{M}_i^T = \begin{pmatrix} 1 - \Delta t \sigma & -\Delta t z_i + \Delta t r & \Delta t y_i \\ \Delta t \sigma & 1 - \Delta t & \Delta t x_i \\ 0 & -\Delta t x_i & 1 - \Delta t b \end{pmatrix}$

背景誤差の時間発展に関する補足

- ここまでは $\mathbf{x}_t = M(\mathbf{x}_0)$ をもとに $\mathbf{B}_t = \mathbf{M}\mathbf{B}_0\mathbf{M}^T$ を導出したがホワイトノイズ $\boldsymbol{\eta}$ が存在する状況

$$\mathbf{x}_t = M(\mathbf{x}_0) + \boldsymbol{\eta}$$

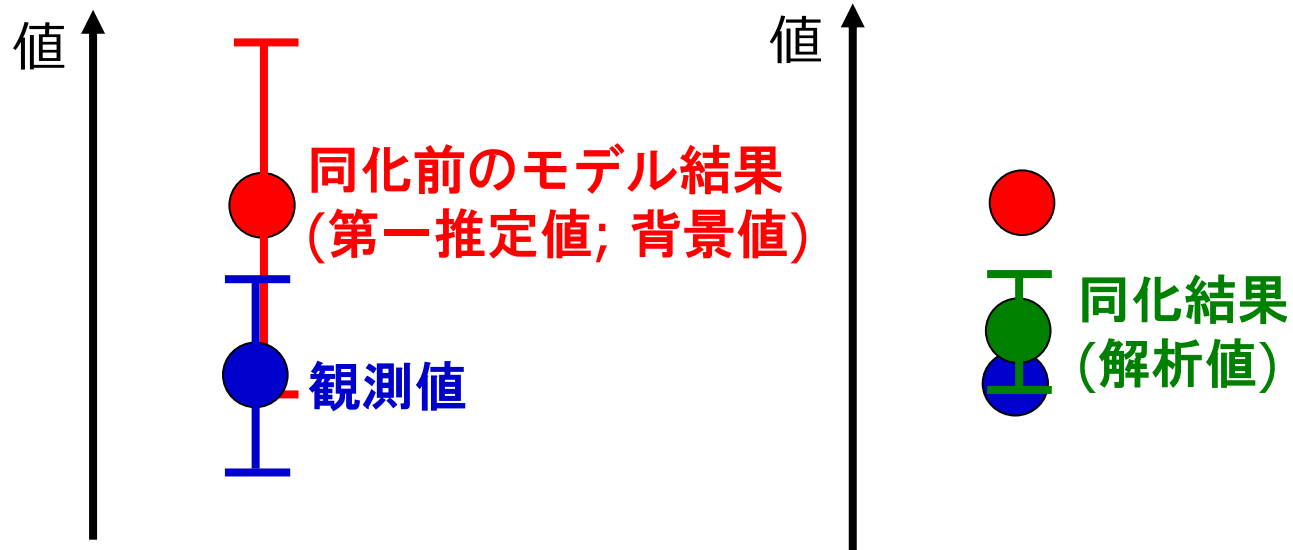
においては,

$$\mathbf{B}_t = \mathbf{M}\mathbf{B}_0\mathbf{M}^T + \langle \boldsymbol{\eta}\boldsymbol{\eta}^T \rangle$$

- ホワイトノイズが存在する場合だけでなく, 背景誤差共分散の時間発展 $\mathbf{M}\mathbf{B}_0\mathbf{M}^T$ が近似により過小評価されている場合にも, $\mathbf{B}$ を拡大する操作 (インフレーション) が必要になる。
⇒ $\mathbf{B}$ に縮小近似を適用した場合やアンサンブルカルマンフィルタのメンバー数が少ないときなど。

最小分散推定の復習

観測値 + 背景値 + 観測誤差分散 + 背景誤差分散
 $\Rightarrow$ 解析値 + 解析誤差分散



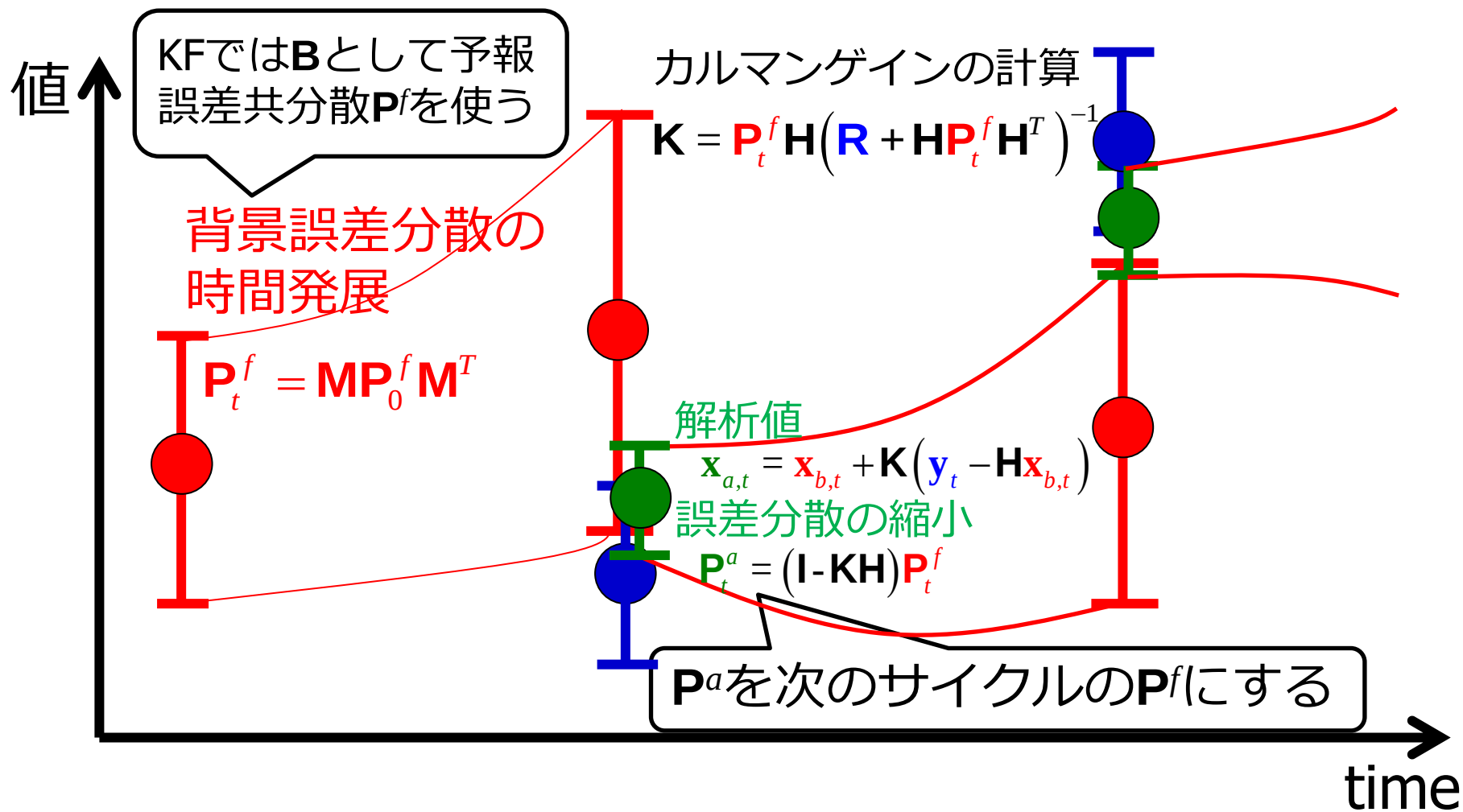
- 1変数の場合

$$x_a = x_b + \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} (x_o - x_b)$$

- 多変数の場合 カルマンゲイン K

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T \left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T \right)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

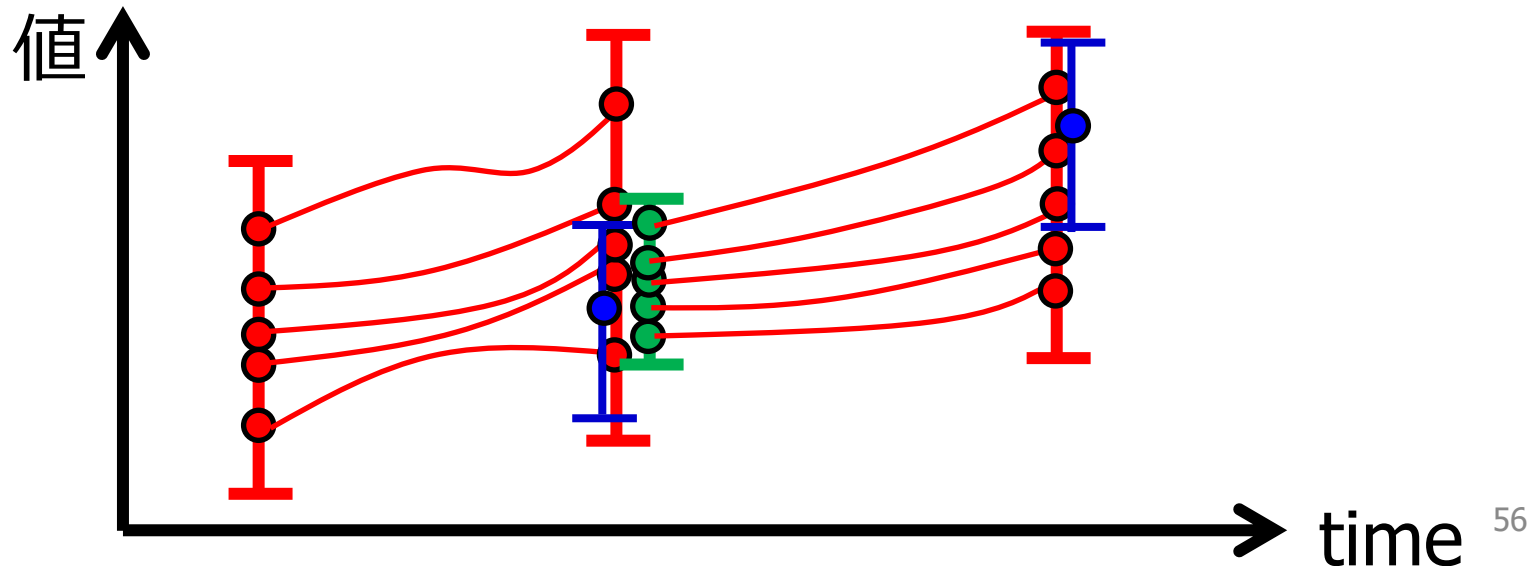
カルマンフィルター(KF)系列の手法の概念図



- 背景(予報)誤差分散の時間発展を計算し、観測値が得られたら、逐次的に重みづけをして修正する。
- 解析誤差分散を新たな背景誤差分散に置き換え計算を続ける

KF系列の手法

- モデルが線形で摂動も線形時間発展する
⇒カルマンフィルター
- モデルが非線形で摂動の線形時間発展を仮定
⇒拡張カルマンフィルター
(演算回数は状態変数の長さ n に対し $O(n^3)$)
- 背景誤差共分散の時間発展をアンサンブルメンバーのばらつきで表現する
⇒アンサンブル・カルマンフィルター (次ページでも説明)



アンサンブルカルマンフィルター

- m メンバーのアンサンブル予報を行い, 時刻 t における各メンバーの状態変数を $\mathbf{x}_{b,t}^{(1)}, \mathbf{x}_{b,t}^{(2)}, \dots, \mathbf{x}_{b,t}^{(m)}$ とし, その平均を $\bar{\mathbf{x}}_b$ と書くことにすると, 分散を以下の式で表せる。

$$\mathbf{P}_t^f = \frac{1}{m-1} \sum_{i=1}^m \left(\mathbf{x}_{b,t}^{(i)} - \bar{\mathbf{x}}_{b,t} \right) \left(\mathbf{x}_{b,t}^{(i)} - \bar{\mathbf{x}}_{b,t} \right)^T = \frac{1}{m-1} \overset{\text{摂動をまとめた行列}}{\boxed{\delta \mathbf{X}_b}} \delta \mathbf{X}_b^T$$

- 具体的なアルゴリズムは多数存在する。
 - 観測摂動法：観測値にノイズを加えたうえで各アンサンブルメンバーについて独立に解析値を求める。

$$\mathbf{x}_{a,t}^{(i)} = \mathbf{x}_{b,t}^{(i)} + \mathbf{P}_t^f \mathbf{H} \left(\mathbf{R} + \mathbf{H} \mathbf{P}_t^f \mathbf{H}^T \right)^{-1} \left(\mathbf{y}_t^{(i)} - \mathbf{H} \mathbf{x}_{b,t}^{(i)} \right)$$

- 平方根フィルタ：アンサンブル平均を解析値として求め, 摂動部分は以下の式で求める。

$$\mathbf{P}_t^a = (\mathbf{I} - \mathbf{K} \mathbf{H}) \mathbf{P}_t^f \xrightarrow{\text{アンサンブルで表現}} \delta \mathbf{X}_t^a \left(\delta \mathbf{X}_t^a \right)^T = (\mathbf{I} - \mathbf{K} \mathbf{H}) \delta \mathbf{X}_t^f \left(\delta \mathbf{X}_t^f \right)^T$$

$$\delta \mathbf{X}_t^a = \delta \mathbf{X}_t^f \mathbf{T} \text{ を仮定 } \longrightarrow \delta \mathbf{X}_t^f \mathbf{T} \mathbf{T}^T \left(\delta \mathbf{X}_t^f \right)^T = (\mathbf{I} - \mathbf{K} \mathbf{H}) \delta \mathbf{X}_t^f \left(\delta \mathbf{X}_t^f \right)^T \text{ から } \delta \mathbf{X}_t^f$$

アンサンブルカルマンフィルタについて、色々

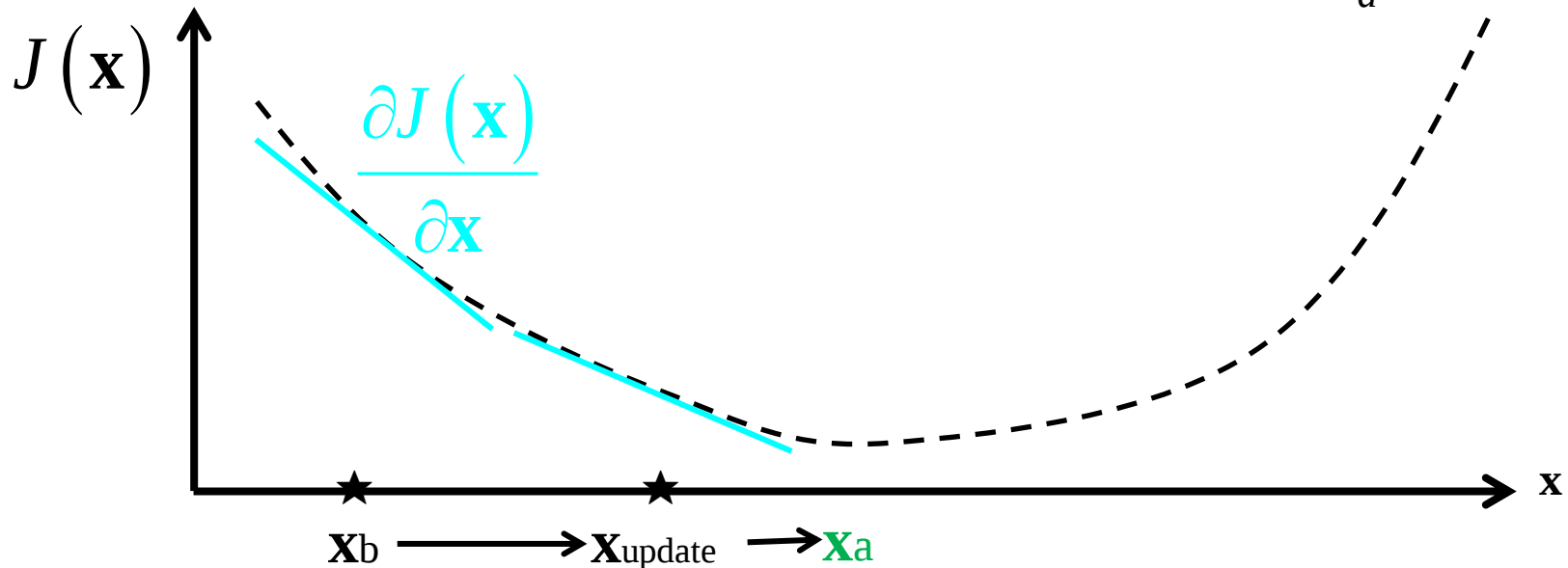
- 大自由度系に適用することに関して
 - 気象分野であれば数百万自由度の系に対して数十メンバーの計算で実用に耐えうる解析値が得られる。
(オリジナルの拡張カルマンフィルタでは計算困難)
 - ただし、メンバー数不足によりいくつかの問題が生じる
 - ◆ サンプルングノイズが生じる
⇒十分に離れた場所に誤差相関が無いと仮定する(局所化)などして、サンプルングノイズを低減する。
 - ◆ $\mathbf{P}^f = \mathbf{M}\mathbf{P}_0^f\mathbf{M}^T$ よりも空間の広がりが過小になる
⇒誤差共分散を疑似的に大きくする操作(インフレーション)などを適用するなどして対処
 - $\mathbf{M}$ の固有値が小さいものばかりだと、効率が悪い(?)。
- 接線形モデルやアジョイントモデルを作る必要が無いため、実装は比較的容易で、一度作ってしまえば使いまわしが効く。
- 実は背景誤差共分散に正規分布を仮定しなくても導出できる
(Evensen 1994; 教科書3.4節)

最尤推定法⇒最小値探索の復習

$$J(\mathbf{x}) = (1/2)(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + (1/2)(\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x})$$

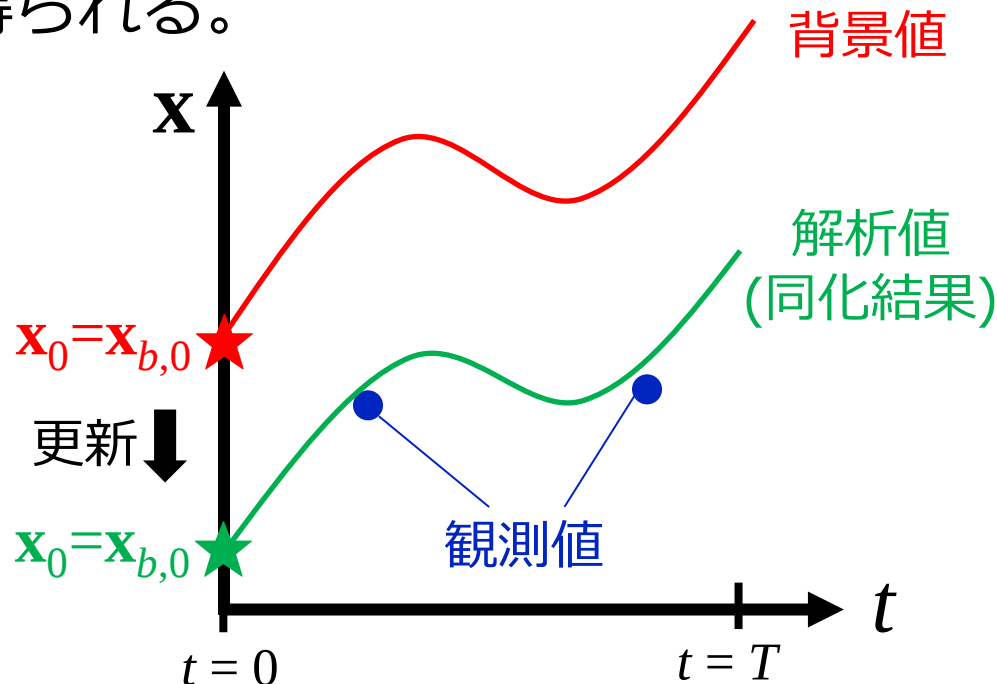
$$\partial J(\mathbf{x}) / \partial \mathbf{x} = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{y})$$

- どの $\mathbf{x}$ で $J(\mathbf{x})$ が最小となるかは事前にはわからない
 - まずは $\mathbf{x} = \mathbf{x}_b$ と想定し, $\partial J(\mathbf{x}) / \partial \mathbf{x}$ を計算。
 - $\partial J(\mathbf{x}) / \partial \mathbf{x}$ の値を見て, $J(\mathbf{x})$ が小さくなる方向に $\mathbf{x}$ を更新。
 - 再度, $\partial J(\mathbf{x}) / \partial \mathbf{x}$ を計算して, $\mathbf{x}$ を更新。
⇒複数回繰り返し, 落ち着いたところを $\mathbf{x} = \mathbf{x}_a$ とみなす。



4次元変分法(アジョイント法)

- カルマンフィルタが時間方向に逐次的な修正であったのに対し、4次元変分法ではある固定区間内全体(時刻0~T)の観測値をまとめて同化する。
- 基本的には、初期場が決まれば時刻0~Tの値はモデルにより決まると考えて、背景値の初期値を最適な解析値の初期値へと更新する問題を解く(★から★への更新)。
- 同化が終われば、固定区間内全体でモデル力学に整合的な解析値が得られる。



4次元変分法の定式化

- 評価関数の定義

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_{b,0})^T \mathbf{B}_0^{-1}(\mathbf{x}_0 - \mathbf{x}_{b,0}) + \frac{1}{2} \sum_{i=0}^T (\mathbf{y}_i - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i(\mathbf{x}_i))$$

モデルで時刻*i*の
 $\mathbf{x}_i$ を計算する

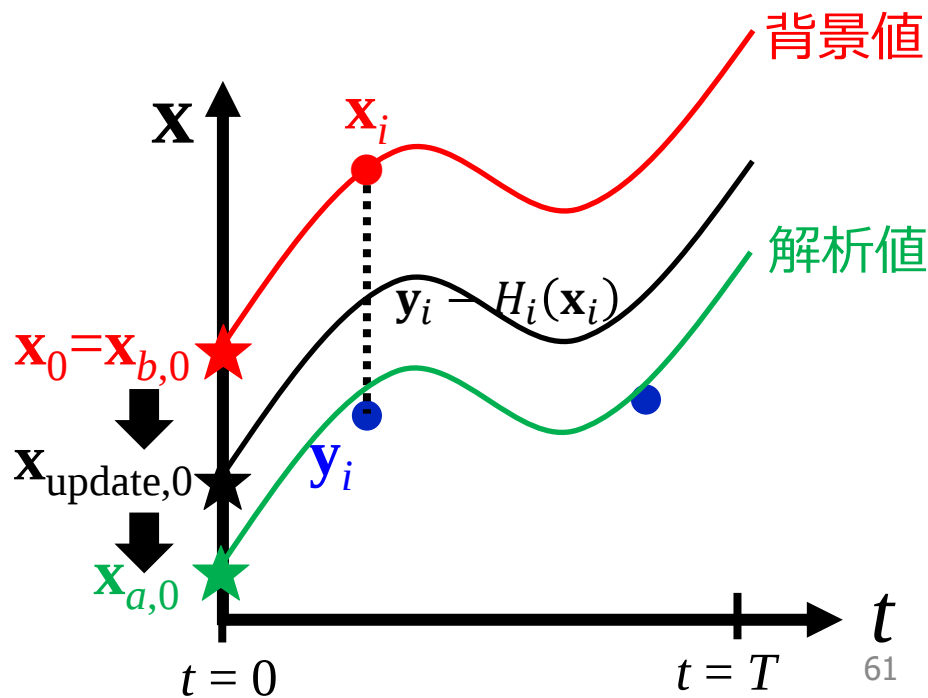
- 評価関数の勾配（証明は4.2.1節；次ページにも説明有り）

$$\partial J(\mathbf{x}_0) / \partial \mathbf{x}_0 = \mathbf{B}_0^{-1}(\mathbf{x}_0 - \mathbf{x}_{b,0}) + \sum_{i=0}^T \mathbf{M}_0^T \mathbf{M}_1^T \cdots \mathbf{M}_{i-1}^T \mathbf{H}_i^T \mathbf{R}_i^{-1} (H_i(\mathbf{x}_i) - \mathbf{y}_i)$$

タイムステップ0から1に進める演算

- 手続き

- $\mathbf{x}_0 = \mathbf{x}_{b,0}$ を初期値としてモデルを走らせ、評価関数の勾配を計算する。
- $J(\mathbf{x}_0)$ が小さくなる方向に、 $\mathbf{x}_0$ を動かす。
- 更新された $\mathbf{x}_0$ を初期値として改めてモデルを走らせる。



評価関数の勾配の計算方法(教科書4.2節)

$$\frac{\partial J(\mathbf{x}_0)}{\partial \mathbf{x}_0} = \mathbf{B}_0^{-1}(\mathbf{x}_0 - \mathbf{x}_{b,0}) + \sum_{i=0}^T \mathbf{M}_0^T \mathbf{M}_1^T \cdots \mathbf{M}_{i-1}^T \mathbf{H}_i^T \mathbf{R}_i^{-1} (H_i(\mathbf{x}_i) - \mathbf{y}_i)$$

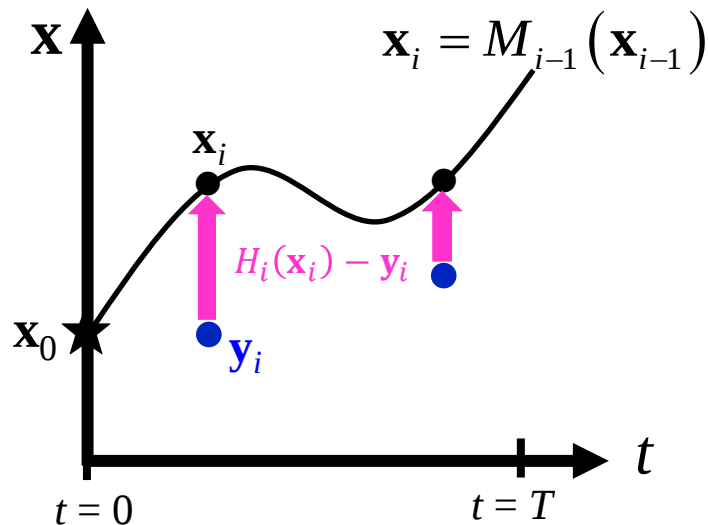
アジョイント変数 $\lambda_i = \frac{\partial J(\mathbf{x}_i)}{\partial \mathbf{x}_i}$ (勾配の情報)を定義して帰納的な式に書換

$$\lambda_T = \mathbf{H}_T^T \mathbf{R}_T^{-1} (H_T(\mathbf{x}_T) - \mathbf{y}_T)$$

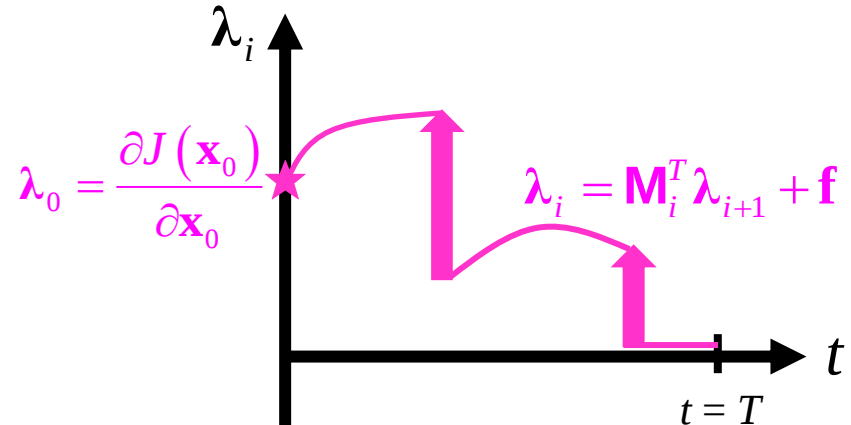
$$\lambda_i = \mathbf{M}_i^T \lambda_{i+1} + \mathbf{H}_i^T \mathbf{R}_i^{-1} (H_i(\mathbf{x}_i) - \mathbf{y}_i) \quad (i = 1, 2, \dots, T-1)$$

$$\lambda_0 = \mathbf{M}_0^T \lambda_1 + \mathbf{B}_0^{-1}(\mathbf{x}_0 - \mathbf{x}_{b,0}) + \mathbf{H}_0^T \mathbf{R}_0^{-1} (H_0(\mathbf{x}_0) - \mathbf{y}_0)$$

状態変数の時間発展(前方積分)



アジョイント変数の時間発展(後方積分)



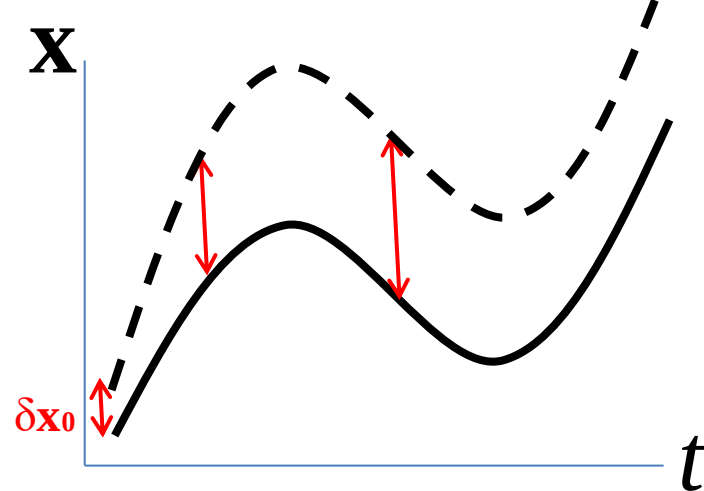
非線形モデル

$$\begin{aligned}\frac{dx}{dt} &= -\sigma(x-y) \\ \frac{dy}{dt} &= -y - xz + rx \\ \frac{dz}{dt} &= xy - bz\end{aligned}$$

$$\mathbf{x}_i \equiv \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}$$

接線形モデル

$$\begin{aligned}\frac{d\delta x}{dt} &= -\sigma(\delta x - \delta y) \\ \frac{d\delta y}{dt} &= -\delta y - x\delta z - z\delta x + r\delta x \\ \frac{d\delta z}{dt} &= x\delta y + y\delta x - b\delta z\end{aligned}$$



$$\delta \mathbf{x}_i \simeq \frac{\partial M(\mathbf{x}_{i-1})}{\partial \mathbf{x}_{i-1}} \delta \mathbf{x}_{i-1} = \mathbf{M}_{i-1} \delta \mathbf{x}_{i-1}$$

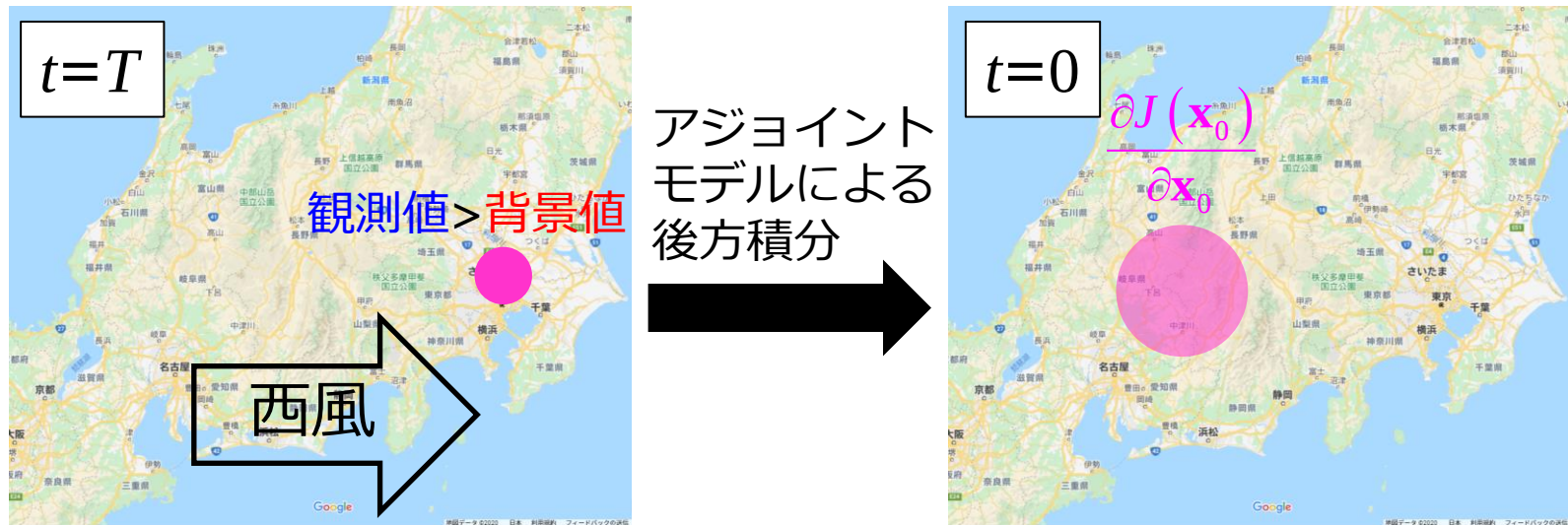
$$\mathbf{M}_i = \begin{pmatrix} 1 - \Delta t \sigma & \Delta t \sigma & 0 \\ -\Delta t z_i + \Delta t r & 1 - \Delta t & -\Delta t x_i \\ \Delta t y_i & \Delta t x_i & 1 - \Delta t b \end{pmatrix}$$

転置(厳密に言うと, 内積の定義にも依存するが . . .)

随伴(アジョイント)モデル $\mathbf{M}_i^T = \begin{pmatrix} 1 - \Delta t \sigma & -\Delta t z_i + \Delta t r & \Delta t y_i \\ \Delta t \sigma & 1 - \Delta t & \Delta t x_i \\ 0 & -\Delta t x_i & 1 - \Delta t b \end{pmatrix}$

4次元変分法について、色々

- アジョイント変数の直観的理解(例：移流拡散方程式)



- 4次元変分法では固定区間内ではモデルの力学に従うので収支解析などに向いているが、ある固定区間から次の固定区間への継ぎ目では、データ同化による補正がかかる。
- 解析誤差共分散を陽に計算しないので、単純な定式化では、次の固定区間に更新された $\mathbf{B}_t$ を渡すことはできない。
- 非線形な観測演算子についても対応することができる。
- アジョイントモデルのコードの自動生成ツールが便利。

カルマンフィルタ系列の手法と4次元変分法

- 4次元変分法の一つの固定区間について考え、カルマンフィルタ系列の手法のノイズの項を無視すると、両手法は誤差の時間発展と観測行列が線形の枠組みで、最終時刻において理論的には同じ解をもたらす。

(教科書4.5節；簡単のため時刻 t にのみ観測があると仮定)

$$\mathbf{x}_{a,t} = \mathbf{x}_{b,t} + \mathbf{M}\mathbf{B}_0\mathbf{M}^T\mathbf{H}_t^T \left(\mathbf{R}_t + \mathbf{H}_t\mathbf{M}\mathbf{B}_0\mathbf{M}^T\mathbf{H}_t^T \right)^{-1} \left(\mathbf{y}_t - \mathbf{H}_t\mathbf{x}_{b,t} \right)$$

- 最終時刻以外の固定区間では解は一致しないが、固定区間スムーザー(教科書3.3.3節)を用いると理論的には一致する。
- 一方で、現実的なデータ同化システムの実装にあたっては、近似が入ったり、複数の同化サイクルをまたぐようになるため、通常は解は一致しない。
- 近年は、カルマンフィルタ系列の手法により得られた背景誤差共分散を4次元変分法に渡す手法(私などは「ハイブリッド」と呼ぶがそう呼ぶべきでないという人もいる・・・)

3次元変分法, 4次元変分法, EnKF, Hybridの違い

- 簡単のため、観測誤差共分散**R**は与えられたとして、線形的な誤差の時間発展について考える。
- 手法によって、**B**と**M**の与え方が異なる。
- 「アンサンブル近似と過去の統計のどちらがよいか」という問いは解くべき問題やアンサンブルメンバー数によって異なるだろう。
- 気象分野では一般にHybrid>4DVar~EnKF>3DVar

同化手法	背景誤差共分散 B	時間発展 M
3DVar	過去の統計	計算しない
古典的4DVar	過去の統計	(陰に)厳密に解く
EnKF	アンサンブルで近似	アンサンブルで近似
Hybrid	アンサンブルで近似	(陰に)厳密に解く

おわり

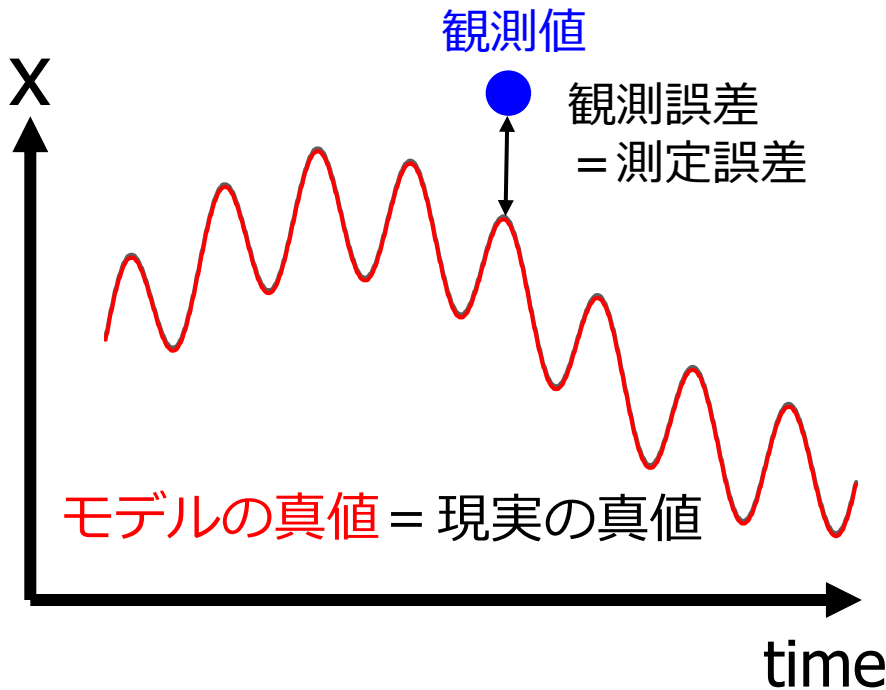
皆さん、お疲れさまでした。

(補足) 「モデルの真値」とは何か？

(Thanks: 藤井さん、中野さん)

- モデルにバイアスは無いことを前提とする。
(既知のバイアスがある場合は、単純には、補正を掛ける)
- 例：海面の変動を対象としたデータ同化に関し、パーフェクトなモデルと、長周期成分しか計算しないモデルを考える。

パーフェクトなモデル



長周期のみ計算するモデル

