5．4次元変分法

MAP推定

Maximum a posteriori (MAP) estimate

変分法によるデータ同化では、事後確率密度関数のモードを解析値とする。

$$\mathbf{x}_k^a := \arg\max_{\mathbf{x}_k} p(\mathbf{x}_k \mid \mathbf{y}_1, \cdots, \mathbf{y}_k)$$

便宜上、次の**評価関数**（cost function）の最小点を求める。

$$J(\mathbf{x}_k) := -\log p(\mathbf{x}_k \mid \mathbf{y}_1, \cdots, \mathbf{y}_k)$$

Bayes の定理を代入すると

$$J(\mathbf{x}_k) := -\log p(\mathbf{y}_k \mid \mathbf{x}_k) - \log p(\mathbf{x}_k \mid \mathbf{y}_1, \cdots, \mathbf{y}_{k-1}) + (\text{terms not dependent on } \mathbf{x}_k)$$

観測値の確率密度関数

$$p(\mathbf{y}_k \mid \mathbf{x}_k) = \frac{1}{\sqrt{(2\pi)^{m_k} |\mathbf{R}_k|}} \exp\left[-\frac{1}{2}(\mathbf{y}_k - H_k(\mathbf{x}_k))^{\mathrm{T}} \mathbf{R}_k^{-1} (\mathbf{y}_k - H_k(\mathbf{x}_k))\right]$$

3次元変分法(1)

仮定

事前確率密度関数は、平均値 $\mathbf{x}_k^f$ 、共分散行列 $\mathbf{P}_k^f$ のガウス分布とする。

$$p(\mathbf{x}_k \mid \mathbf{y}_1, \cdots, \mathbf{y}_{k-1}) = \frac{1}{\sqrt{(2\pi)^n \left|\mathbf{P}_k^f\right|}} \exp\left[-\frac{1}{2}\left(\mathbf{x}_k - \mathbf{x}_k^f\right)^{\mathrm{T}} \left(\mathbf{P}_k^f\right)^{-1} \left(\mathbf{x}_k - \mathbf{x}_k^f\right)\right]$$

評価関数($\mathbf{x}_k$ に依存しない項を省略する)

$$J(\mathbf{x}_k) = \frac{1}{2}\left(\mathbf{x}_k - \mathbf{x}_k^f\right)^{\mathrm{T}} \left(\mathbf{P}_k^f\right)^{-1} \left(\mathbf{x}_k - \mathbf{x}_k^f\right) + \frac{1}{2}\left(H_k(\mathbf{x}_k) - \mathbf{y}_k\right)^{\mathrm{T}} \mathbf{R}_k{}^{-1} \left(H_k(\mathbf{x}_k) - \mathbf{y}_k\right)$$

$$\mathbf{x}_k^a := \arg\min_{\mathbf{x}_k} J(\mathbf{x}_k)$$

評価関数の勾配ベクトル

$$\nabla J(\mathbf{x}_k) = \left(\mathbf{P}_k^f\right)^{-1} \left(\mathbf{x}_k - \mathbf{x}_k^f\right) + \mathbf{H}_k{}^{\mathrm{T}} \mathbf{R}_k{}^{-1} \left(H_k(\mathbf{x}_k) - \mathbf{y}_k\right)$$

$$\left(\mathbf{H}_k\right)_{ij} := \left.\frac{\partial \left(H_k(\mathbf{x})\right)_i}{\partial x_j}\right|_{\mathbf{x}=\mathbf{x}_k}$$

予測値誤差共分散行列 $\mathbf{P}_k^f$ には、長期平均値 $\mathbf{P}^f$ を用いる。

3次元変分法(2)

解析値の計算手順(反復法)

$$k = 0: \quad \mathbf{x}^{(0)} = \mathbf{x}^{f}$$

while not convergent

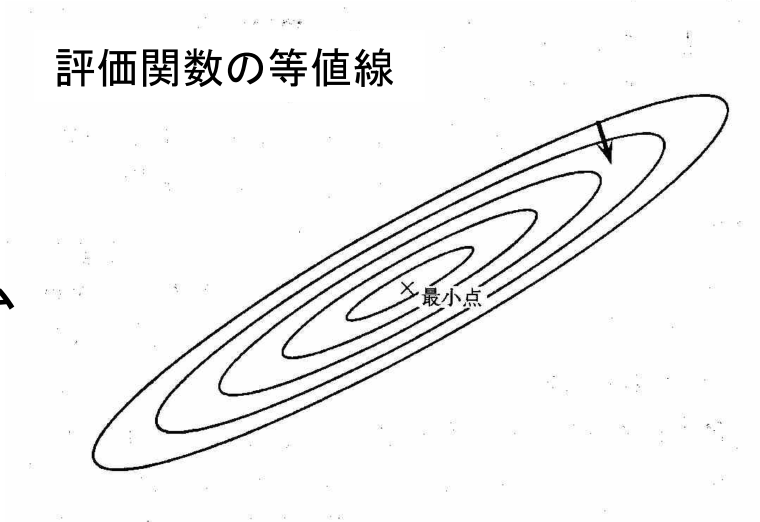
$$\mathbf{x}^{(k)} \Rightarrow J(\mathbf{x}^{(k)}), \nabla J(\mathbf{x}^{(k)})$$

minimization routine $[\mathbf{x}^{(k)}, J(\mathbf{x}^{(k)}), \nabla J(\mathbf{x}^{(k)}) \Rightarrow \mathbf{x}^{(k+1)}]$

$$k \rightarrow k+1$$

end

$$\mathbf{x}^{a} = \mathbf{x}^{(k)}$$

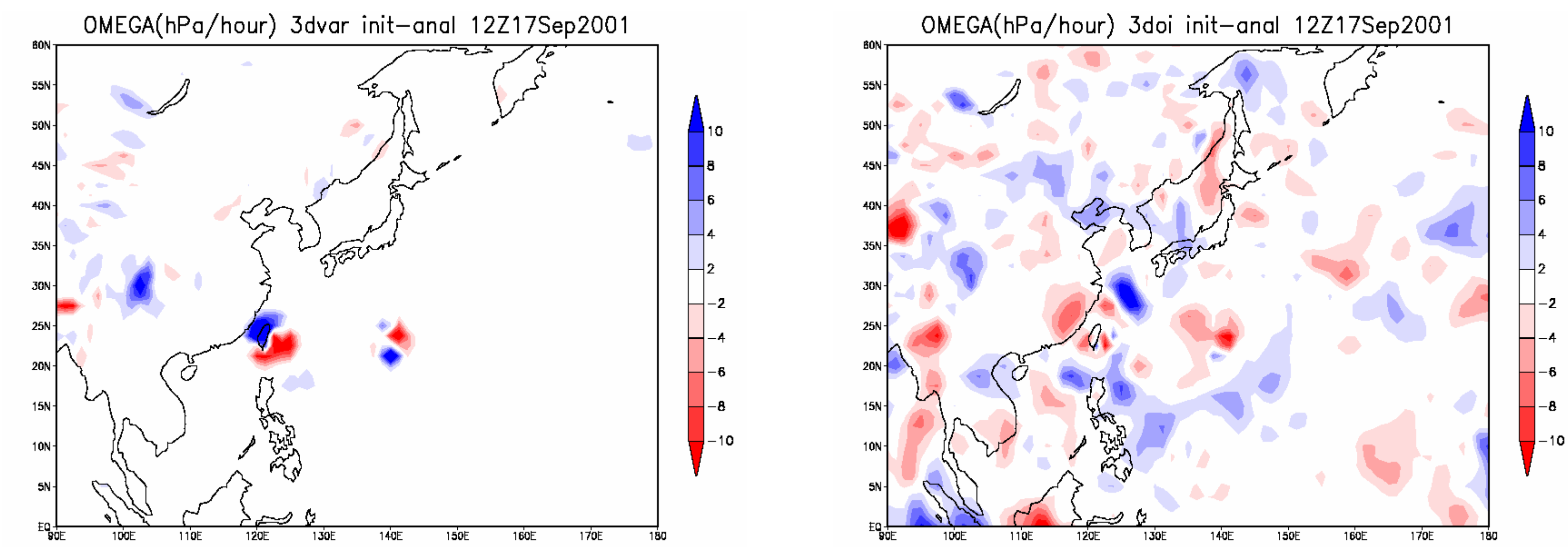


大規模問題に向いた最小値探索アルゴリズム

- 準ニュートン(quasi-Newton)法
- 共役勾配(conjugate gradient)法

3次元変分法（3）

・最適内挿法と比べると、状態変数と異なる物理量の観測データを直接同化できるだけでなく、解析値の力学的バランスに優れている。
・予測誤差共分散行列がflow-dependentでないという問題は、4次元変分法によっておおよそ解決される。



850hPaの鉛直p速度（単位: hPa/hr）の初期値化による変化量。気象庁の3次元変分法（左）と最適内挿法（右）による2001年9月17日12UTCの全球解析値について示す。（Takeuchi and Tsuyuki, 2002）

3次元変分法(4)

観測演算子が線形とすると、解析値は

$$\nabla J(\mathbf{x}_k^a) = \left(\mathbf{P}_k^f\right)^{-1}\left(\mathbf{x}_k^a - \mathbf{x}_k^f\right) + \mathbf{H}_k{}^{\mathrm{T}}\mathbf{R}_k{}^{-1}\left(\mathbf{H}_k\mathbf{x}_k^a - \mathbf{y}\right) = \mathbf{0}$$

$$\therefore \mathbf{x}_k^a = \mathbf{x}_k^f + \left[\left(\mathbf{P}_k^f\right)^{-1} + \mathbf{H}_k{}^{\mathrm{T}}\mathbf{R}_k{}^{-1}\mathbf{H}_k\right]^{-1}\mathbf{H}_k{}^{\mathrm{T}}\mathbf{R}_k{}^{-1}\left(\mathbf{y} - \mathbf{H}_k\mathbf{x}_k^f\right)$$

解析誤差共分散行列

$$\mathit{\Delta}\mathbf{x}_k^a = \left[\left(\mathbf{P}_k^f\right)^{-1} + \mathbf{H}_k{}^{\mathrm{T}}\mathbf{R}_k{}^{-1}\mathbf{H}_k\right]^{-1}\left[\left(\mathbf{P}_k^f\right)^{-1}\mathit{\Delta}\mathbf{x}_k^f + \mathbf{H}_k{}^{\mathrm{T}}\mathbf{R}_k{}^{-1}\mathit{\Delta}\mathbf{y}\right]$$

$$\therefore \mathbf{P}_k^a := \left\langle \mathit{\Delta}\mathbf{x}_k^a\left(\mathit{\Delta}\mathbf{x}_k^a\right)^{\mathrm{T}}\right\rangle = \left[\left(\mathbf{P}_k^f\right)^{-1} + \mathbf{H}_k{}^{\mathrm{T}}\mathbf{R}_k{}^{-1}\mathbf{H}_k\right]^{-1}$$

これらはカルマンフィルタの結果と一致する。

一方、

$$J''(\mathbf{x}_k^a) = \left(\mathbf{P}_k^f\right)^{-1} + \mathbf{H}_k{}^{\mathrm{T}}\mathbf{R}_k{}^{-1}\mathbf{H}_k$$

$$\left(J''(\mathbf{x})\right)_{ij} := \frac{\partial^2 J}{\partial x_i \partial x_j}$$ （Hessian行列）

$$\therefore \mathbf{P}_k^a = J''(\mathbf{x}_k^a)^{-1}$$

問　題　２
（風ベクトルの推定）

問題設定

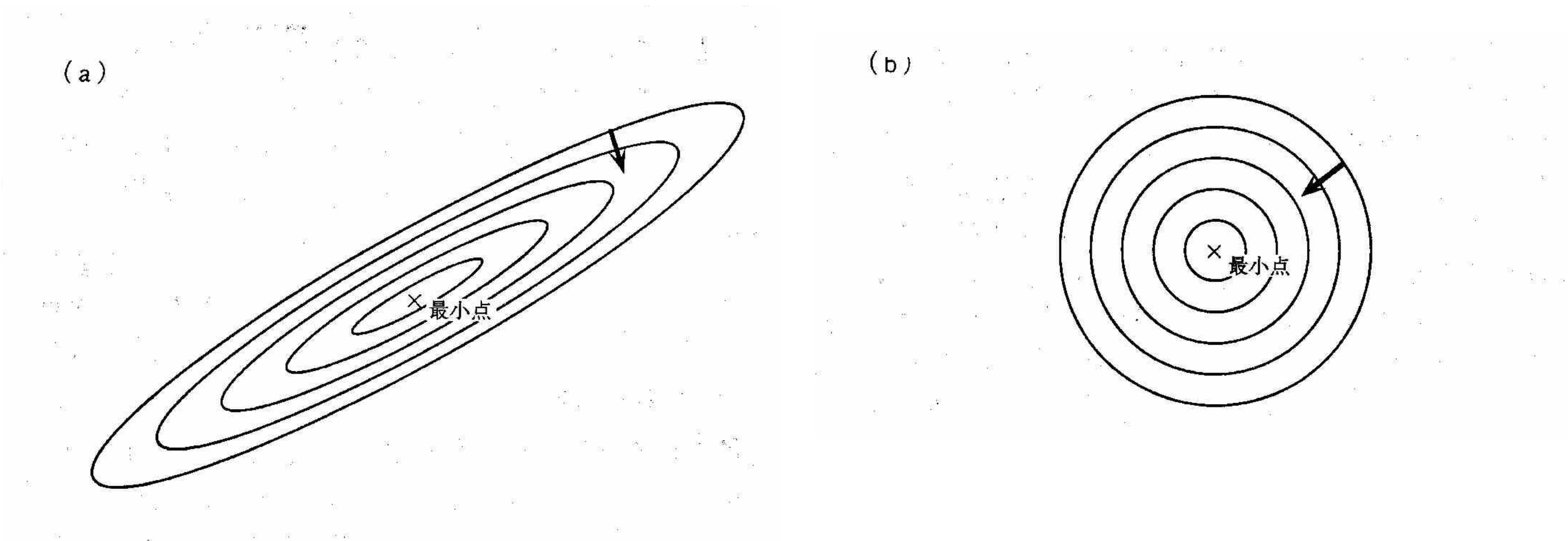
- 数値モデルは1格子で、状態変数は風の水平成分
- 水平風速の観測データが1個
- 予測誤差共分散行列は一様な対角行列

$$\mathbf{x} = \begin{pmatrix} u \\ v \end{pmatrix}, \qquad H(\mathbf{x}) = \sqrt{u^2 + v^2}, \qquad \mathbf{H} = \begin{pmatrix} \dfrac{u}{\sqrt{u^2+v^2}} & \dfrac{v}{\sqrt{u^2+v^2}} \end{pmatrix}$$

$$\mathbf{P}^f = \begin{pmatrix} (\sigma^f)^2 & 0 \\ 0 & (\sigma^f)^2 \end{pmatrix}, \qquad \mathbf{R} = (\sigma^o)^2$$

前処理(1)

最小値探索アルゴリズムの収束速度は、一般にHessian行列J'' の**条件数**(最大固有値と最小固有値の比)が1に近いほど大きい。



評価関数の等値線と勾配ベクトルの負の方向。（a） Hessian行列の条件数が1より十分大きい場合。（b） Hessian行列の条件数が1の場合。

前処理(2)

制御変数を線形変換することにより、ヘッセ行列の固有値の大きさをなるべくそろえる。

予測誤差共分散行列の**Cholesky 分解**

$$\mathbf{P}^{f} = \mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{L}\mathbf{L}^{\mathrm{T}}$$

($\mathbf{U}$: 上三角行列 、$\mathbf{L}$: 下三角行列)

制御変数(最小値探索に用いる変数)を線形変換する。

$$\mathbf{u} := \mathbf{L}^{-1}\left(\mathbf{x}_k - \mathbf{x}_k^f\right)$$

評価関数

$$\tilde{J}(\mathbf{u}) = \frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{u} + \frac{1}{2}\left[H_k(\mathbf{x}_k^f + \mathbf{L}\mathbf{u}) - \mathbf{y}_k\right]^{\mathrm{T}}\mathbf{R}_k{}^{-1}\left[H_k(\mathbf{x}_k^f + \mathbf{L}\mathbf{u}) - \mathbf{y}_k\right]$$

評価関数のHessian行列は、観測演算子が線形の場合

$$\tilde{J}''(\mathbf{u}) = \mathbf{I} + \mathbf{L}^{\mathrm{T}}\mathbf{H}_k{}^{\mathrm{T}}\mathbf{R}_k{}^{-1}\mathbf{H}_k\mathbf{L}$$

一般に観測データの個数は数値モデルの自由度よりかなり小さいので、この変換によってHessian 行列の固有値の多くは1に等しくなる。

予測誤差共分散の推定

NMC法

・ 初期値の時刻が異なる2つの予報値の差の統計から、予測値の誤差相関に関する情報を抽出する。

・ 手軽であるが、誤差分散の大きさを別途推定する必要がある。

解析アンサンブル法

・ 観測データや境界条件などに人工的な誤差を加えた多数のデータ同化サイクルを走らせ、それらの間の予測値の差の統計から、予測値の共分散行列を推定する。

・ 数値モデルの誤差を別途考慮する必要がある。

評価関数の統計的性質（1）

観測演算子が線形とすると、次式が成り立つ。

$$\left\langle J(\mathbf{x}^a) \right\rangle = \frac{m}{2}$$

m は観測データの個数である。さらに

$$\mathbf{d} := \mathbf{y} - H(\mathbf{x}^f), \qquad \left\langle \mathbf{d} \right\rangle = \mathbf{0}$$

がガウス分布するならば、次式が成り立つ。

$$\left\langle \left(J(\mathbf{x}^a) - \left\langle J(\mathbf{x}^a) \right\rangle \right)^2 \right\rangle = \frac{m}{2}$$

したがって、評価関数の値の標準偏差と平均値の比は $\sqrt{\dfrac{2}{m}}$ になる。

これらの統計的性質は、観測誤差分散の推定に有用である。

評価関数の統計的性質(2)

証明1

この場合には、3次元変分法の解析値はカルマンフィルタと同じなので

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{P}^f\,\mathbf{H}^{\mathrm{T}}\left(\mathbf{R} + \mathbf{H}\,\mathbf{P}^f\,\mathbf{H}^{\mathrm{T}}\right)^{-1}\left(\mathbf{y} - H(\mathbf{x}^f)\right)$$

$$\therefore \mathbf{y} - H(\mathbf{x}^a) = \left(\mathbf{y} - H(\mathbf{x}^f)\right) + \mathbf{H}\left(\mathbf{x}^f - \mathbf{x}^a\right)$$

$$= \mathbf{R}\left(\mathbf{R} + \mathbf{H}\,\mathbf{P}^f\,\mathbf{H}^{\mathrm{T}}\right)^{-1}\left(\mathbf{y} - H(\mathbf{x}^f)\right)$$

$$\therefore J(\mathbf{x}^a) = \frac{1}{2}\left(\mathbf{x}^a - \mathbf{x}^f\right)^{\mathrm{T}}\left(\mathbf{P}^f\right)^{-1}\left(\mathbf{x}^a - \mathbf{x}^f\right) + \frac{1}{2}\left(H(\mathbf{x}^a) - \mathbf{y}\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(H(\mathbf{x}^a) - \mathbf{y}\right)$$

$$= \frac{1}{2}\left(\mathbf{y} - H(\mathbf{x}^f)\right)^{\mathrm{T}}\left(\mathbf{R} + \mathbf{H}\,\mathbf{P}^f\,\mathbf{H}^{\mathrm{T}}\right)^{-1}\left(\mathbf{y} - H(\mathbf{x}^f)\right)$$

$$= \frac{1}{2}\mathrm{Tr}\left[\left(\mathbf{R} + \mathbf{H}\,\mathbf{P}^f\,\mathbf{H}^{\mathrm{T}}\right)^{-1}\left(\mathbf{y} - H(\mathbf{x}^f)\right)\left(\mathbf{y} - H(\mathbf{x}^f)\right)^{\mathrm{T}}\right]$$

期待値をとると

$$\left\langle J(\mathbf{x}^a)\right\rangle = \frac{1}{2}\mathrm{Tr}\left[\left(\mathbf{R} + \mathbf{H}\,\mathbf{P}^f\,\mathbf{H}^{\mathrm{T}}\right)^{-1}\left\langle \mathbf{d}\,\mathbf{d}^{\mathrm{T}}\right\rangle\right]$$

$$= \frac{1}{2}\mathrm{Tr}\left[\mathbf{I}_m\right] = \frac{m}{2}$$

評価関数の統計的性質(3)

証明2

前の結果を用いれば

$$\left\langle \left(J(\mathbf{x}^a) - \left\langle J(\mathbf{x}^a) \right\rangle \right)^2 \right\rangle = \left\langle \left(\frac{1}{2} \mathrm{Tr} \left[\left(\mathbf{R} + \mathbf{H} \mathbf{P}^f \mathbf{H}^\mathrm{T} \right)^{-1} \mathbf{d} \, \mathbf{d}^\mathrm{T} \right] \right)^2 \right\rangle - \left(\frac{m}{2} \right)^2$$

$$= \frac{1}{4} \sum_{i,j,k,l=1}^{m} \left(\left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle^{-1} \right)_{ij} \left(\left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle^{-1} \right)_{kl} \left\langle d_i d_j d_k d_l \right\rangle - \frac{m^2}{4}$$

仮定より $\mathbf{d}$ が平均0のガウス分布をするので

$$\left\langle d_i d_j d_k d_l \right\rangle = \left\langle d_i d_j \right\rangle \left\langle d_k d_l \right\rangle + \left\langle d_i d_k \right\rangle \left\langle d_j d_l \right\rangle + \left\langle d_i d_l \right\rangle \left\langle d_j d_k \right\rangle$$

$$\therefore \sum_{i,j,k,l=1}^{m} \left(\left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle^{-1} \right)_{ij} \left(\left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle^{-1} \right)_{kl} \left\langle d_i d_j d_k d_l \right\rangle = \sum_{i,j,k,l=1}^{m} \left(\left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle^{-1} \right)_{ij} \left(\left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle^{-1} \right)_{kl}$$

$$\times \left(\left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle_{ij} \left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle_{kl} + \left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle_{ik} \left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle_{jl} + \left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle_{il} \left\langle \mathbf{d} \, \mathbf{d}^\mathrm{T} \right\rangle_{jk} \right)$$

$$= m^2 + 2m$$

$$\therefore \left\langle \left(J(\mathbf{x}^a) - \left\langle J(\mathbf{x}^a) \right\rangle \right)^2 \right\rangle = \frac{1}{4} \left(m^2 + 2m \right) - \frac{m^2}{4} = \frac{m}{2}$$

解析誤差共分散の近似計算(1)

最小値の探索に準ニュートン法を用いる場合には、反復計算の過程で得られた評価関数の形状に関する情報から、Hessian行列の逆行列を近似しつつ最小点を探索していくので、それを解析誤差共分散行列とすることができる。

ニュートン法の反復計算

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \left[J''(\mathbf{x}^{(k)})\right]^{-1} \nabla J(\mathbf{x}^{(k)})$$

準ニュートン法(**BFGS法**)の反復計算

$$\alpha^{(0)} = 1, \qquad \mathbf{H}_0 = \mathbf{I}$$

$$\alpha^{(k)} = \arg\min_{\alpha} J(\mathbf{x}^{(k)} - \alpha \mathbf{H}^{(k)} \nabla J(\mathbf{x}^{(k)}))$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha^{(k)} \mathbf{H}^{(k)} \nabla J(\mathbf{x}^{(k)})$$

$$\mathbf{H}^{(k+1)} = \left(\mathbf{I} - \frac{\mathbf{p}^{(k)} {\mathbf{q}^{(k)}}^{\mathrm{T}}}{{\mathbf{q}^{(k)}}^{\mathrm{T}} \mathbf{p}^{(k)}} \right) \mathbf{H}^{(k)} \left(\mathbf{I} - \frac{\mathbf{p}^{(k)} {\mathbf{q}^{(k)}}^{\mathrm{T}}}{{\mathbf{q}^{(k)}}^{\mathrm{T}} \mathbf{p}^{(k)}} \right) + \frac{\mathbf{p}^{(k)} {\mathbf{p}^{(k)}}^{\mathrm{T}}}{{\mathbf{q}^{(k)}}^{\mathrm{T}} \mathbf{q}^{(k)}}$$

$$\mathbf{p}^{(k)} := \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}, \qquad \mathbf{q}^{(k)} := \nabla J(\mathbf{x}^{(k+1)}) - \nabla J(\mathbf{x}^{(k)})$$

解析誤差共分散の近似計算（2）

モンテカルロ法によってHessian行列を近似計算することもできる。

$$J''(\mathbf{x}^a) = \left(\mathbf{P}^f\right)^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}$$

計算が必要な右辺第2項は、評価関数の観測項から生じる。

$$J^o(\mathbf{x}) = \frac{1}{2}\left(H(\mathbf{x}) - \mathbf{y}\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(H(\mathbf{x}) - \mathbf{y}\right)$$

人工的な観測データを、正規乱数によるモンテカルロ法で生成する。

$$\mathbf{y} = H(\mathbf{x}^f) + \boldsymbol{\xi}$$

$$\left\langle\boldsymbol{\xi}\right\rangle = \mathbf{0}, \qquad \left\langle\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}}\right\rangle = \mathbf{R}$$

このとき

$$\left\langle \nabla J^o(\mathbf{x}^f)\left(\nabla J^o(\mathbf{x}^f)\right)^{\mathrm{T}} \right\rangle = \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}$$

$$\therefore \left(\mathbf{P}^a\right)^{-1} = \left(\mathbf{P}^f\right)^{-1} + \left\langle \nabla J^o(\mathbf{x}^f)\left(\nabla J^o(\mathbf{x}^f)\right)^{\mathrm{T}} \right\rangle$$

$$\approx \left(\mathbf{P}^f\right)^{-1} + \frac{1}{N-1}\sum_{i=1}^{N} \nabla J^{o(i)}(\mathbf{x}^f)\left(\nabla J^{o(i)}(\mathbf{x}^f)\right)^{\mathrm{T}}$$

解析誤差共分散の近似計算(3)

逆行列を計算する。

$$\mathbf{G}^{(k)} := \left(\mathbf{P}^{f}\right)^{-1} + \frac{1}{N-1}\sum_{i=1}^{k} \nabla J^{o(i)}(\mathbf{x}^{f})\left(\nabla J^{o(i)}(\mathbf{x}^{f})\right)^{\mathrm{T}} \qquad (k = 0, \cdots, N)$$

とすれば

$$\mathbf{P}^{a} \approx \left(\mathbf{G}^{(N)}\right)^{-1}$$

Sherman-Morrison-Woodbury の公式より

$$\left(\mathbf{A} + \boldsymbol{\alpha}\mathbf{b}^{\mathrm{T}}\right)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\boldsymbol{\alpha}\mathbf{b}^{\mathrm{T}}\mathbf{A}^{-1}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{A}^{-1}\boldsymbol{\alpha}}$$

これを使えば

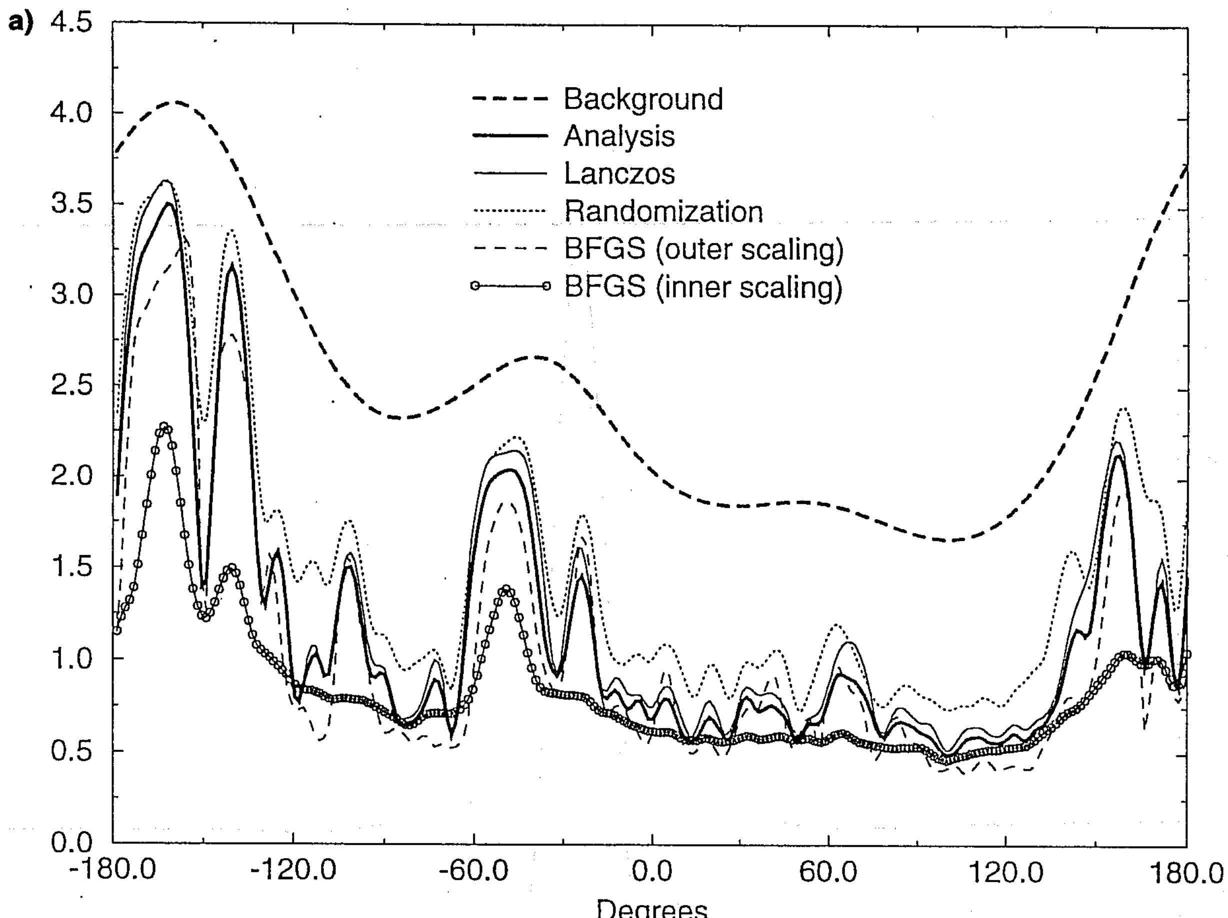
$$\left(\mathbf{G}^{(k)}\right)^{-1} = \left(\mathbf{G}^{(0)}\right)^{-1} - \sum_{i=1}^{k} \frac{\mathbf{s}^{(i)}\mathbf{s}^{(i)\mathrm{T}}}{1 + \frac{1}{\sqrt{N-1}}\left(\nabla J^{o(i)}(\mathbf{x}^{f})\right)^{\mathrm{T}}\mathbf{s}^{(i)}} \qquad (k = 1, \cdots, N)$$

$$\mathbf{s}^{(i)} := \frac{1}{\sqrt{N-1}}\left(\mathbf{G}^{(i-1)}\right)^{-1}\nabla J^{o(i)}(\mathbf{x}^{f})$$

解析誤差共分散の近似計算(4)

1次元分布した観測データに対する3次元変分法による計算例

（格子点数:512個。観測点の分布は中緯度の高層観測に似せてある）



Fisher and Courtier(1995)

変分法 QC

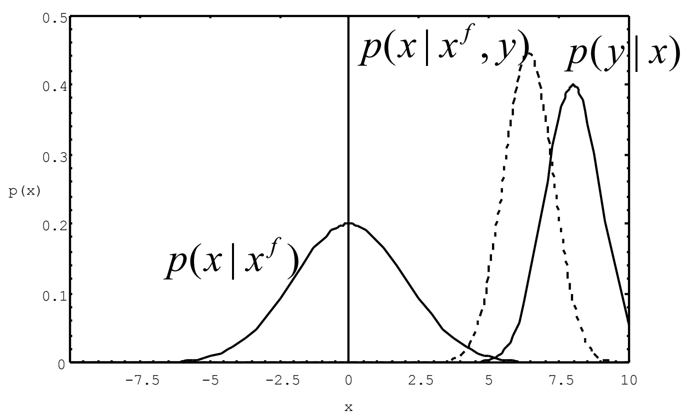
観測データの品質管理（QC）

- 測器の故障などによる誤った観測データをあらかじめ除去する。
- 方法：気候学的チェック、内的整合性チェック、外的整合性チェック（他の観測データや予報値との比較）

変分法では、外的整合性チェックをデータ同化の中で行える。

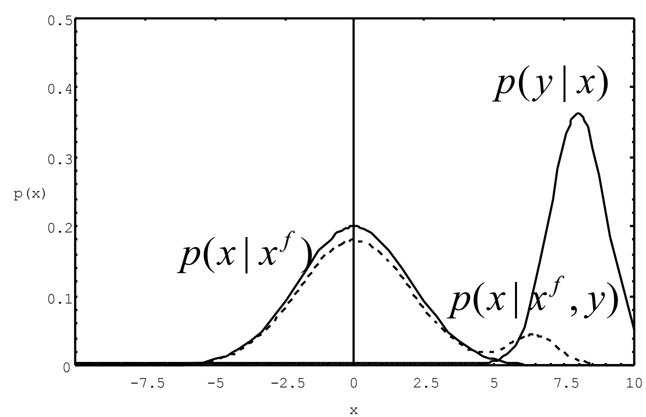
$$p(y\,|\,x)=\frac{1}{\sqrt{2\pi}\,\sigma^o}\exp\left[-\frac{(y-x)^2}{2(\sigma^o)^2}\right]$$

$$p(x\,|\,x^f)=\frac{1}{\sqrt{2\pi}\,\sigma^f}\exp\left[-\frac{(x-x^f)^2}{2(\sigma^f)^2}\right]$$



$$p(y\,|\,x)=\frac{a}{x_2-x_1}+\frac{1-a}{\sqrt{2\pi}\,\sigma^o}\exp\left[-\frac{(y-x)^2}{2(\sigma^o)^2}\right]$$

（a ：誤データ発生確率）



4次元変分法(1)

時刻 t_0 から時刻 t_K までの同化ウィンドウ内の観測データ$\{\mathbf{y}_1,\cdots,\mathbf{y}_K\}$から、同化ウィンドウ内の状態変数$\{\mathbf{x}_0,\cdots,\mathbf{x}_K\}$を一括推定する。

同化ウィンドウの長さが、数値モデルのランダム誤差の影響が無視できる程度に短ければ

$$\mathbf{x}_k = M_k(\mathbf{x}_0) \qquad (k=1,\cdots,K)$$

したがって、$\mathbf{x}_0$ を推定すればよい。異なる時刻の観測データが互いに独立とすれば、Bayesの定理より

$$p(\mathbf{x}_0 \mid \mathbf{y}_1,\cdots,\mathbf{y}_K) = \frac{p(\mathbf{y}_1 \mid \mathbf{x}_1)\cdots p(\mathbf{y}_K \mid \mathbf{x}_K)\,p(\mathbf{x}_0)}{\int p(\mathbf{y}_1 \mid \mathbf{x}_1)\cdots p(\mathbf{y}_K \mid \mathbf{x}_K)\,p(\mathbf{x}_0)\,d^n\mathbf{x}}$$

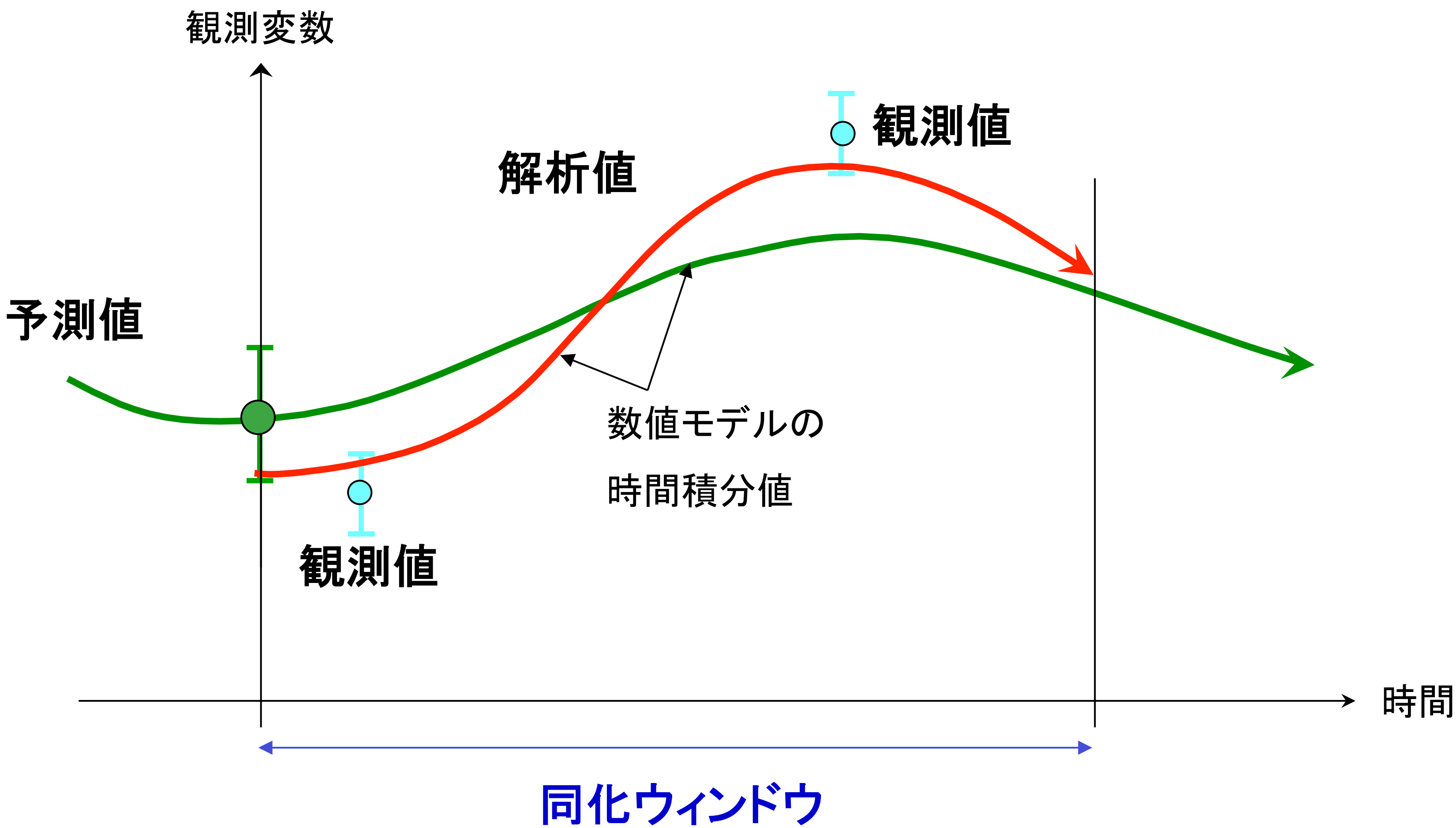
評価関数($\mathbf{x}_0$ に依存しない項を省略する)

$$\begin{aligned} J(\mathbf{x}_0) &:= -\log p(\mathbf{x}_0 \mid \mathbf{y}_1,\cdots,\mathbf{y}_K) \\ &= \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right)^{\mathrm{T}}\left(\mathbf{P}_0^f\right)^{-1}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right) + \sum_{k=1}^{K}\frac{1}{2}\left(H_k(\mathbf{x}_k) - \mathbf{y}_k\right)^{\mathrm{T}}\mathbf{R}_k^{-1}\left(H_k(\mathbf{x}_k) - \mathbf{y}_k\right) \end{aligned}$$

予測値誤差共分散行列$\mathbf{P}_0^f$には、長期平均値$\mathbf{P}^f$を用いる。

4次元変分法(2)

4次元変分法では、同化ウィンドウ内の予測値と観測値の両方に最も近い数値シミュレーションの結果を解析値とする。



4次元変分法(3)

評価関数の勾配ベクトル

$$\nabla J(\mathbf{x}_0) = \left(\mathbf{P}_0^f\right)^{-1}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right) + \sum_{k=1}^{K} \mathbf{M}_k{}^{\mathrm{T}} \mathbf{H}_k{}^{\mathrm{T}} \mathbf{R}_k{}^{-1}\left(H_k(\mathbf{x}_k) - \mathbf{y}_k\right)$$

$$\left(\mathbf{M}_k\right)_{ij} := \left.\frac{\partial\left(M_k(\mathbf{x})\right)_i}{\partial x_j}\right|_{\mathbf{x}=\mathbf{x}_0}$$

観測誤差が力学的バランスを壊すのを抑えるために、評価関数にペナルティ項 $J_c(\mathbf{x}_0)$を追加することが多い。たとえば

$$J_c(\mathbf{x}_0) = \frac{1}{2} r \sum_{k=0}^{K} \left| \mathbf{P}_G \frac{d}{dt}\left(\mathbf{x}_k - M_k(\mathbf{x}_0^f)\right) \right|^2$$

$\mathbf{P}_G$:高周波重力波成分を抽出する射影演算子、r :調節パラメータ

同化ウィンドウ内の観測データなどをまとめて$\tilde{\mathbf{y}}$ などと表せば、評価関数を次のように表記することもできる。

$$J(\mathbf{x}_0) = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right)^{\mathrm{T}}\left(\mathbf{P}_0^f\right)^{-1}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right) + \frac{1}{2}\left[\tilde{H}(\tilde{M}(\mathbf{x}_0)) - \tilde{\mathbf{y}}\right]^{\mathrm{T}} \tilde{\mathbf{R}}^{-1}\left[\tilde{H}(\tilde{M}(\mathbf{x}_0)) - \tilde{\mathbf{y}}\right]$$

4次元変分法(4)

仮定

- 同化ウィンドウの終わりにだけ観測データがある。
- 観測演算子と数値モデルは線形。

評価関数（$\mathbf{P}_0^f$には長期平均値$\mathbf{P}^f$を用いる）

$$J(\mathbf{x}_0) = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right)^{\mathrm{T}}\left(\mathbf{P}^f\right)^{-1}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right) + \frac{1}{2}\left(\mathbf{H}_K\mathbf{x}_K - \mathbf{y}_K\right)^{\mathrm{T}}\mathbf{R}_K^{-1}\left(\mathbf{H}_K\mathbf{x}_K - \mathbf{y}_K\right)$$

$$\mathbf{x}_K = \mathbf{M}_K\mathbf{x}_0, \qquad \mathbf{x}_K^{\{f,a\}} = \mathbf{M}_K\mathbf{x}_0^{\{f,a\}}$$

解析値を求める。

$$\nabla J(\mathbf{x}_0) = \left(\mathbf{P}^f\right)^{-1}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right) + \mathbf{M}_K{}^{\mathrm{T}}\mathbf{H}_K{}^{\mathrm{T}}\mathbf{R}_K{}^{-1}\left(\mathbf{H}_K\mathbf{x}_K - \mathbf{y}_K\right)$$

$$\therefore\ \left(\mathbf{P}^f\right)^{-1}\left(\mathbf{x}_0^a - \mathbf{x}_0^f\right) + \mathbf{M}_K{}^{\mathrm{T}}\mathbf{H}_K{}^{\mathrm{T}}\mathbf{R}_K{}^{-1}\left(\mathbf{H}_K\mathbf{M}_K\mathbf{x}_0^a - \mathbf{y}_K\right) = \mathbf{0}$$

$$\therefore\ \mathbf{x}_0^a = \mathbf{x}_0^f + \left[\left(\mathbf{P}^f\right)^{-1} + \mathbf{M}_K{}^{\mathrm{T}}\mathbf{H}_K{}^{\mathrm{T}}\mathbf{R}_K{}^{-1}\mathbf{H}_K\mathbf{M}_K\right]^{-1}\mathbf{M}_K{}^{\mathrm{T}}\mathbf{H}_K{}^{\mathrm{T}}\mathbf{R}_K{}^{-1}\left(\mathbf{y}_K - \mathbf{H}_K\mathbf{x}_K^f\right)$$

$$= \mathbf{x}_0^f + \mathbf{P}^f\,\mathbf{M}_K{}^{\mathrm{T}}\mathbf{H}_K{}^{\mathrm{T}}\left(\mathbf{R}_K + \mathbf{H}_K\mathbf{M}_K\,\mathbf{P}^f\,\mathbf{M}_K{}^{\mathrm{T}}\mathbf{H}_K{}^{\mathrm{T}}\right)^{-1}\left(\mathbf{y}_K - \mathbf{H}_K\mathbf{x}_K^f\right)$$

4次元変分法(5)

4次元変分法による同化ウィンドウの終わりにおける解析値

$$\mathbf{x}_K^a = \mathbf{M}_K \mathbf{x}_0^f + \mathbf{M}_K \mathbf{P}^f \mathbf{M}_K{}^{\mathrm{T}} \mathbf{H}_K{}^{\mathrm{T}} \left(\mathbf{R}_K + \mathbf{H}_K \mathbf{M}_K \mathbf{P}^f \mathbf{M}_K{}^{\mathrm{T}} \mathbf{H}_K{}^{\mathrm{T}} \right)^{-1} \left(\mathbf{y}_K - \mathbf{H}_K \mathbf{x}_K^f \right)$$

$$= \mathbf{x}_K^f + \underline{\underline{\mathbf{M}_K \mathbf{P}^f \mathbf{M}_K}}{}^{\mathrm{T}} \mathbf{H}_K{}^{\mathrm{T}} \left(\mathbf{R}_K + \mathbf{H}_K \underline{\underline{\mathbf{M}_K \mathbf{P}^f \mathbf{M}_K}}{}^{\mathrm{T}} \mathbf{H}_K{}^{\mathrm{T}} \right)^{-1} \left(\mathbf{y}_K - \mathbf{H}_K \mathbf{x}_K^f \right)$$

3次元変分法による同化ウィンドウの終わりにおける解析値

$$\mathbf{x}_K^a = \mathbf{x}_K^f + \underline{\underline{\mathbf{P}^f}} \mathbf{H}_K{}^{\mathrm{T}} \left(\mathbf{R}_K + \mathbf{H}_K \underline{\underline{\mathbf{P}^f}} \mathbf{H}_K{}^{\mathrm{T}} \right)^{-1} \left(\mathbf{y}_K - \mathbf{H}_K \mathbf{x}_K^f \right)$$

予測誤差共分散行列の時間発展方程式

$$\mathbf{P}_K^f := \left\langle \Delta \mathbf{x}_K^f \left(\Delta \mathbf{x}_K^f \right)^{\mathrm{T}} \right\rangle = \mathbf{M}_K \left\langle \Delta \mathbf{x}_0^f \left(\Delta \mathbf{x}_0^f \right)^{\mathrm{T}} \right\rangle \mathbf{M}_K{}^{\mathrm{T}}$$

$$= \mathbf{M}_K \mathbf{P}_0^f \mathbf{M}_K{}^{\mathrm{T}} \rightarrow \mathbf{M}_K \mathbf{P}^f \mathbf{M}_K{}^{\mathrm{T}}$$

4次元変分法では、予測誤差共分散行列を時間発展させることによってflow-dependent にしている。

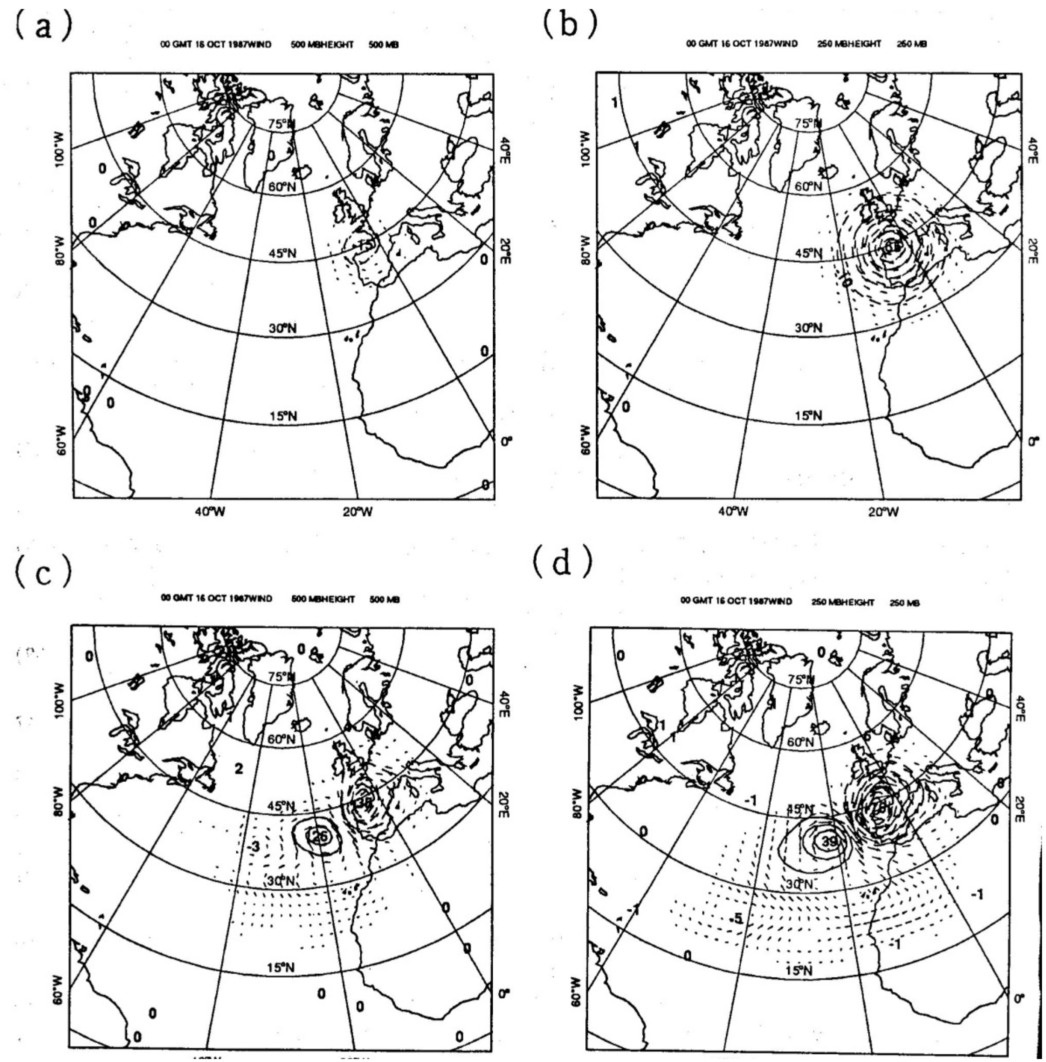
4次元変分法(6)

修正量の比較

500hPa　　250hPa

3次元変分法

4次元変分法



（Thepaut et al., 1993）

Adjoint（随伴）方程式

数値モデル

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}(t), t)$$

線形方程式

$$\frac{d\delta\mathbf{x}}{dt} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}}\delta\mathbf{x}, \qquad \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right)_{ij} := \frac{\partial F_i}{\partial x_j}$$

（Jacobi行列）

$\mathbf{p}(t)$ と $\mathbf{q}(t)$ を $\mathbf{x}(t)$ と同じ次元の任意のベクトル関数とすると、次の恒等式が成り立つ。

$$\int_0^T \mathbf{p}(t)^{\mathrm{T}}\left(\frac{d\mathbf{q}}{dt} - \frac{\partial \mathbf{F}}{\partial \mathbf{x}}\mathbf{q}(t)\right)dt = \left[\mathbf{p}(t)^{\mathrm{T}}\mathbf{q}(t)\right]_0^T - \int_0^T \left[\frac{d\mathbf{p}}{dt} + \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right)^{\mathrm{T}}\mathbf{p}(t)\right]^{\mathrm{T}}\mathbf{q}(t)\,dt$$

数値モデルのAdjoint方程式

$$\boxed{-\frac{d\mathbf{p}}{dt} = \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right)^{\mathrm{T}}\mathbf{p}(t)}$$

Adjoint 法による感度解析（1）

数値モデル（ $\boldsymbol{\alpha}$ ：パラメータからなるベクトル）

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(t, \mathbf{x}(t); \boldsymbol{\alpha})$$

数値モデルの解から構成される任意の汎関数

$$G\left[\mathbf{x}(0), \boldsymbol{\alpha}\right] = \int_0^T g(t, \mathbf{x}(t); \boldsymbol{\alpha})\, dt$$

初期値とパラメータに対する汎関数の感度を調べる。そのために、汎関数の第１変分をとると

$$\delta G\left[\mathbf{x}(0), \boldsymbol{\alpha}\right] = \int_0^T \left(\delta\mathbf{x}(t)^{\mathrm{T}} \frac{\partial g}{\partial \mathbf{x}} + \delta\boldsymbol{\alpha}^{\mathrm{T}} \frac{\partial g}{\partial \boldsymbol{\alpha}} \right) dt$$

数値モデルの解の摂動は、次の線形方程式を満たす。

$$\frac{d\delta\mathbf{x}}{dt} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \delta\mathbf{x}(t) + \frac{\partial \mathbf{F}}{\partial \boldsymbol{\alpha}} \delta\boldsymbol{\alpha}$$

Adjoint 法による感度解析（2）

強制項のあるAdjoint方程式

$$-\frac{d\mathbf{p}}{dt} = \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right)^{\mathrm{T}} \mathbf{p}(t) + \frac{\partial g}{\partial \mathbf{x}}$$

汎関数の第1変分にこれらを代入して部分積分すると

$$\delta G\left[\mathbf{x}(0), \boldsymbol{\alpha}\right] = \delta \mathbf{x}(0)^{\mathrm{T}} \mathbf{p}(0) - \delta \mathbf{x}(T)^{\mathrm{T}} \mathbf{p}(T) + \delta \boldsymbol{\alpha}^{\mathrm{T}} \int_0^T \left(\left(\frac{\partial \mathbf{F}}{\partial \boldsymbol{\alpha}} \right)^{\mathrm{T}} \mathbf{p}(t) + \frac{\partial g}{\partial \boldsymbol{\alpha}} \right) dt$$

Adjoint 方程式の解のうち、終期条件

$$\mathbf{p}(T) = \mathbf{0}$$

を満足する解を採用すれば

$$\nabla_{\mathbf{x}(0)} G = \mathbf{p}(0)$$

$$\nabla_{\boldsymbol{\alpha}} G = \int_0^T \left(\left(\frac{\partial \mathbf{F}}{\partial \boldsymbol{\alpha}} \right)^{\mathrm{T}} \mathbf{p}(t) + \frac{\partial g}{\partial \boldsymbol{\alpha}} \right) dt$$

評価関数の勾配計算（1）

数値モデル

$$\mathbf{x}_k = M_k(\mathbf{x}_k) \qquad (k = 1, \cdots, K)$$

線形モデル

$$\delta\mathbf{x}_k = \mathbf{M}_k \delta\mathbf{x}_{k-1} \qquad (k = 1, \cdots, K)$$

Adjoint 方程式を定義するための恒等式

$$\sum_{k=1}^{K} \mathbf{p}_k{}^{\mathrm{T}} \left(\delta\mathbf{x}_k - \mathbf{M}_k \delta\mathbf{x}_{k-1} \right) = \left(\mathbf{p}_K{}^{\mathrm{T}} \delta\mathbf{x}_K - \mathbf{p}_0{}^{\mathrm{T}} \delta\mathbf{x}_0 \right) - \sum_{k=0}^{K-1} \left(\mathbf{p}_k - \mathbf{M}_{k+1}{}^{\mathrm{T}} \mathbf{p}_{k+1} \right)^{\mathrm{T}} \delta\mathbf{x}_k$$

$$\mathbf{p}_0 := \mathbf{M}_1{}^{\mathrm{T}} \mathbf{p}_1$$

Adjoint 方程式

$$\mathbf{p}_k = \mathbf{M}_{k+1}{}^{\mathrm{T}} \mathbf{p}_{k+1} \qquad (k = 0, \cdots, K-1)$$

評価関数の勾配計算（2）

評価関数の第1変分

$$\delta J = \sum_{k=0}^{K} \delta \mathbf{x}_k{}^{\mathrm{T}} \frac{\partial J}{\partial \mathbf{x}_k}$$

$$\frac{\partial J}{\partial \mathbf{x}_0} = \left(\mathbf{P}_0^f\right)^{-1}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right)$$

$$\frac{\partial J}{\partial \mathbf{x}_k} = \mathbf{H}_k{}^{\mathrm{T}} \mathbf{R}_k{}^{-1}\left(H_k(\mathbf{x}_k) - \mathbf{y}_k\right) \qquad (k = 1, \cdots, K)$$

強制項のあるAdjoint 方程式

$$\mathbf{p}_k = \mathbf{M}_{k+1}{}^{\mathrm{T}} \mathbf{p}_{k+1} + \frac{\partial J}{\partial \mathbf{x}_k} \qquad (k = 0, \cdots, K)$$

$$\mathbf{p}_{K+1} = \mathbf{0}$$

の解より

$$\nabla J = \mathbf{p}_0$$

アジョイントコードの書き方（1）

線形モデル: $\delta\mathbf{x}_{k+1} = \mathbf{M}_k \delta\mathbf{x}_k$

アジョイントモデル: $\mathbf{p}_k = \mathbf{M}_k^{\mathrm{T}} \mathbf{p}_{k+1}$

$$\mathbf{A} = \mathbf{A}_r \cdots \mathbf{A}_2 \mathbf{A}_1 \quad \rightarrow \quad \mathbf{A}^{\mathrm{T}} = \mathbf{A}_1^{\mathrm{T}} \mathbf{A}_2^{\mathrm{T}} \cdots \mathbf{A}_r^{\mathrm{T}}$$

1. 数値モデルのプログラムを任意の基本場の周りで線形化することによって、線形モデルのプログラムを書く。その際、条件分岐文の分岐条件には摂動を考慮しない。

2. 線形モデルの摂動変数の処理部分を、後ろから機械的に変換すれば、アジョイントモデルのプログラムが得られる。

3. 線形かつ可逆な定型処理のアジョイントコードは、元のプログラムをそのまま利用して作成できる。

TAMC: アジョイントコード自動生成ソフトウェア

http://www.autodiff.com/tamc/

アジョイントコードの書き方(2)

線形モデルの代表的な処理の変換規則

$$z = a*x + b*y$$

xとyを以後は使わない場合	xとyを以後も使う場合
1. 処理を行列形式で表す $z = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$	$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ a & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$
2. 右辺の係数行列を転置し、入出力ベクトルを交換する $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} z$	$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$
3. これを展開すれば、アジョイントコードが得られる $x = a*z$, $y = b*z$	$x = x + a*z$, $y = y + b*z$

アジョイントコードの書き方(3)

波・格子変換のアジョイントコード

(格子→波)変換と(波→格子)変換

$$\hat{\mathbf{f}} = \mathbf{S}\,\mathbf{f}, \qquad \mathbf{f} = \mathrm{Inv}\left[\mathbf{S}\right]\hat{\mathbf{f}}$$

$$\mathbf{f} := \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{pmatrix}, \qquad \hat{\mathbf{f}} := \begin{pmatrix} \hat{f}_1 \\ \vdots \\ \hat{f}_M \end{pmatrix} \qquad (M \le N)$$

基底関数系$\{u_k(x)\}$の直交条件

$$\sum_{j=1}^{N} w_j\, u_k(x_j) u_l(x_j) = c_k \delta_{kl} \qquad \left(k, l = 1, \cdots, M\right)$$

アジョイントコードは、以下の関係を用いて作成できる。

$$\mathbf{S}^{\mathrm{T}} = \mathbf{W}\,\mathrm{Inv}\left[\mathbf{S}\right]\mathbf{C}^{-1}, \qquad \mathrm{Inv}\left[\mathbf{S}\right]^{\mathrm{T}} = \mathbf{C}\,\mathbf{S}\,\mathbf{W}^{-1}$$

$$\mathbf{W} := \begin{pmatrix} w_1 & & 0 \\ & \ddots & \\ 0 & & w_N \end{pmatrix}, \qquad \mathbf{C} := \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_M \end{pmatrix}$$

ソースコードの検証法

線形コード：　$\mathbf{M}$

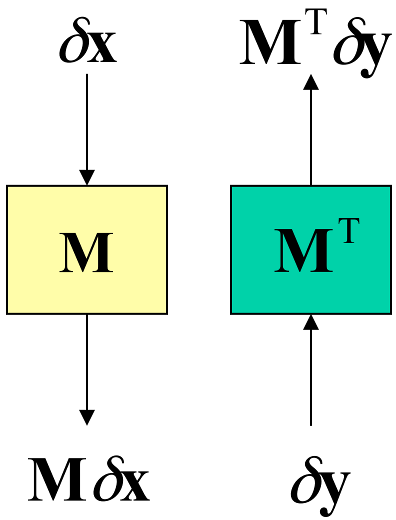
$$\frac{\left\| M(\mathbf{x}+\alpha\delta\mathbf{x}) - M\mathbf{x} \right\|}{|\alpha|\,\|\mathbf{M}\delta\mathbf{x}\|} = 1 + O(\alpha), \qquad |\alpha| << 1$$

アジョイント・コード：　$\mathbf{M}^{\mathrm{T}}$

$$(\mathbf{M}\delta\mathbf{x})^{\mathrm{T}}\mathbf{M}\delta\mathbf{x} = \delta\mathbf{x}^{\mathrm{T}}[\mathbf{M}^{\mathrm{T}}(\mathbf{M}\delta\mathbf{x})]$$

より一般的には

$$(\mathbf{M}\delta\mathbf{x})^{\mathrm{T}}\delta\mathbf{y} = \delta\mathbf{x}^{\mathrm{T}}(\mathbf{M}^{\mathrm{T}}\delta\mathbf{y})$$



評価関数の勾配計算コード

$$\frac{J(\mathbf{x}+\alpha\nabla J(\mathbf{x})) - J(\mathbf{x})}{\alpha\nabla J(\mathbf{x})^{\mathrm{T}}\nabla J(\mathbf{x})} = 1 + O(\alpha), \qquad |\alpha| << 1$$

問　題　3

平均値を計算する、次のサブルーチンのアジョイントコードを書くこと（入力変数xは、この処理のあとも使用されるものとする）。

```
subroutine average (x, n, xmean)

integer, intent (in)::n
real, intent (in)::x(n)
real, intent (out)::xmean

sum = 0.
do i = 1, n
   sum = sum + x(i)
end do
xmean = sum/real(n)

return
end
```

4次元変分法の計算手順(1)

$k=0: \quad \mathbf{x}_0{}^{(0)} = \mathbf{x}_0{}^f$

while not convergent

forward integration of numerical model

$\Rightarrow (\mathbf{x}_0{}^{(k)}, \cdots, \mathbf{x}_K{}^{(k)}), J(\mathbf{x}_0{}^{(k)})$

backward integration of adjoint model $\Rightarrow \nabla J(\mathbf{x}_0{}^{(k)})$

minimization routine $[\mathbf{x}_0{}^{(k)}, J(\mathbf{x}_0{}^{(k)}), \nabla J(\mathbf{x}_0{}^{(k)}) \Rightarrow \mathbf{x}_0{}^{(k+1)}]$

$k \rightarrow k+1$

end

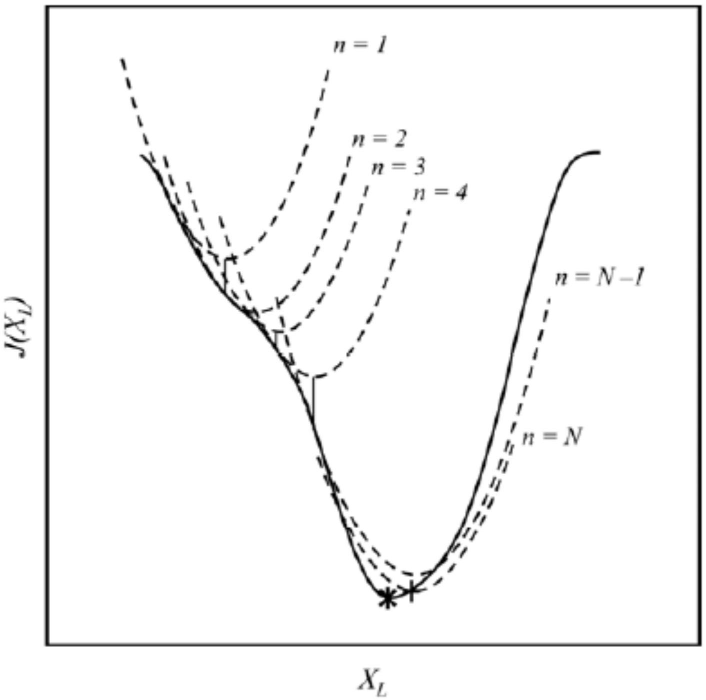
$\mathbf{x}_0{}^a = \mathbf{x}_0{}^{(k)}$

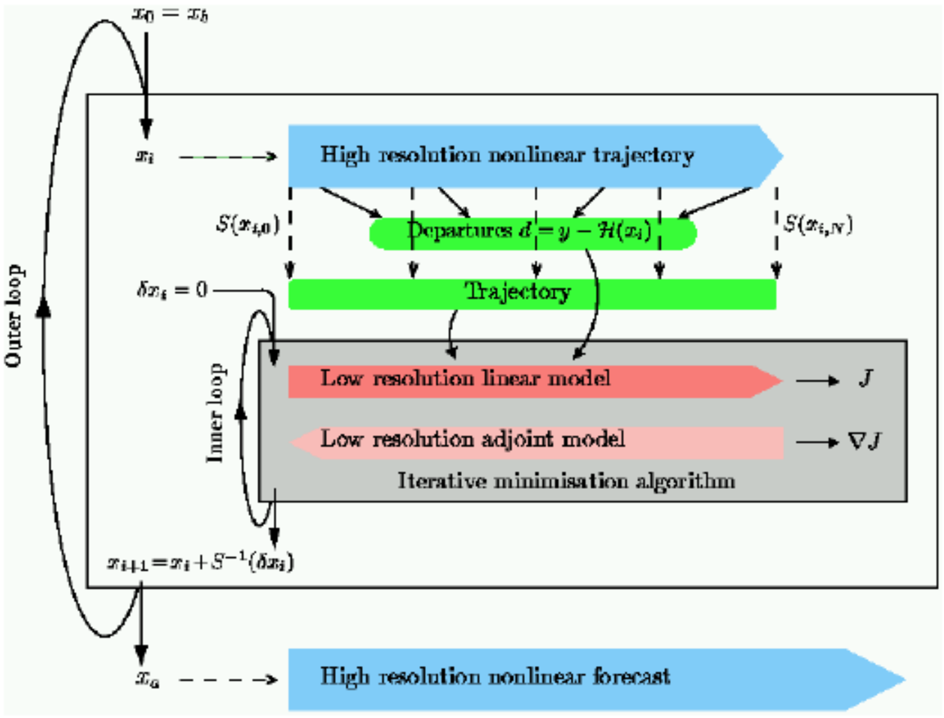
forward integration of numerical model $\Rightarrow (\mathbf{x}_0{}^a, \cdots, \mathbf{x}_K{}^a)$

4次元変分法の計算手順（2）

計算量を大幅に減らすために･･･

- 評価関数を局所的に2次関数で近似して反復的に解く。
- 局所的探索では、低解像度かつ単純化した線形数値モデルを用いる。





ただし、非線形性が強いとうまくいかない

Andersson et al.(2005)

ローレンツモデル（1）

運動方程式

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}), \qquad \mathbf{x} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}, \qquad \mathbf{F}(\mathbf{x}) = \begin{pmatrix} -\sigma X_1 + \sigma X_2 \\ -X_1 X_3 + r X_1 - X_2 \\ X_1 X_2 - b X_3 \end{pmatrix}$$

線形方程式

$$\frac{d\delta\mathbf{x}}{dt} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \delta\mathbf{x}, \qquad \mathbf{x} = \begin{pmatrix} \delta X_1 \\ \delta X_2 \\ \delta X_3 \end{pmatrix}, \qquad \frac{\partial F}{\partial \mathbf{x}} = \begin{pmatrix} -\sigma & \sigma & 0 \\ r - X_3 & -1 & -X_1 \\ X_2 & X_1 & -b \end{pmatrix}$$

アジョイント方程式

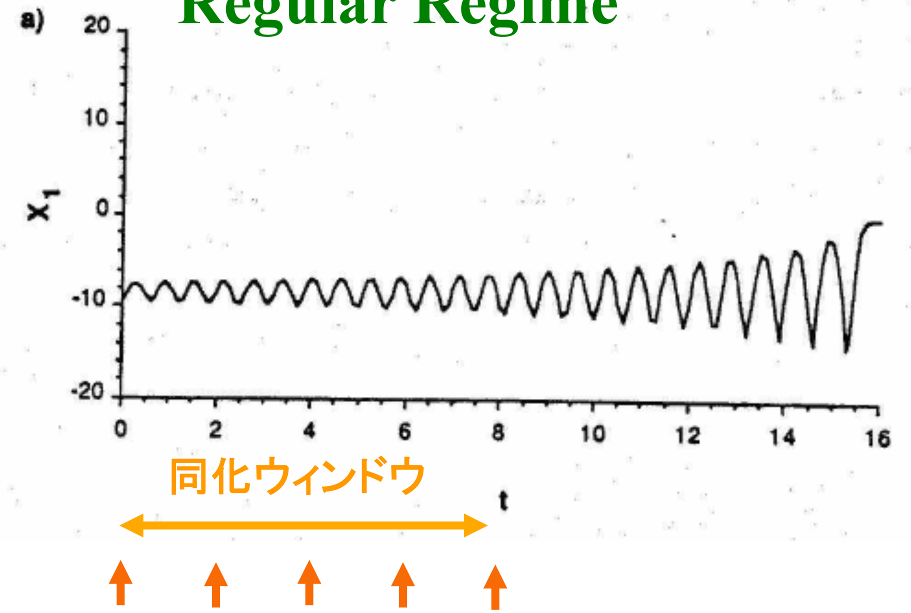
$$-\frac{d\mathbf{p}}{dt} = \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right)^{\mathrm{T}} \mathbf{p}, \qquad \mathbf{p} = \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix}, \qquad \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right)^{\mathrm{T}} = \begin{pmatrix} -\sigma & r - X_3 & X_2 \\ \sigma & -1 & X_1 \\ 0 & -X_1 & -b \end{pmatrix}$$

ローレンツモデル(2)

評価関数

$$J = \frac{1}{K+1}\sum_{k=0}^{K}\sum_{i=1}^{3}\left(X_i - X_i^{\,o}\right)^2 \qquad \left(\sigma = 10,\quad b = \frac{8}{3},\quad r = 28\right)$$

Regular Regime



Transition Regime

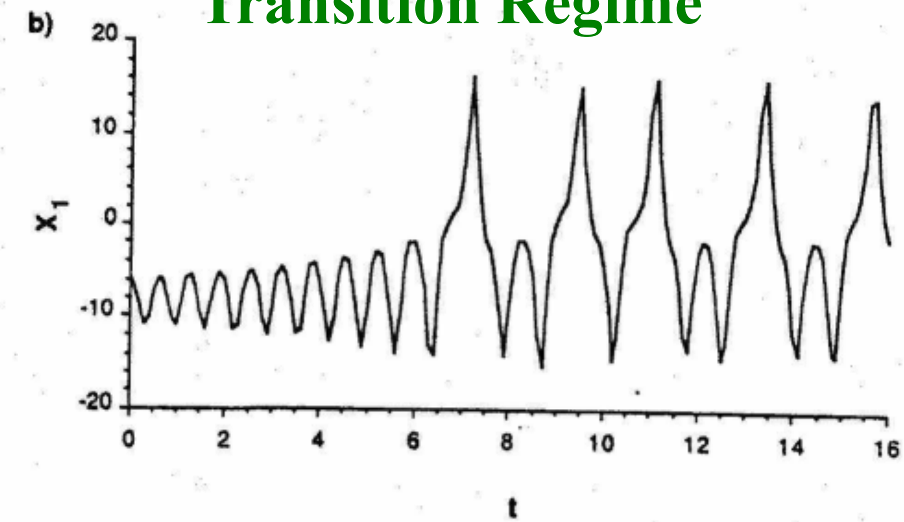


Fig. 1. Representation of the component $X_1(t)$ as a function of time: (a) Regular case generated from the initial conditions $Z_0 = (-9.42,\ -9.34,\ 28.3)$, (b) Case with transition generated from $Z_0 = (-5.92,\ -5.90,\ 24.0)$. Data assimilation is performed over the time interval $0 < t < 8$ and a forecast is made from $t = 8$ up to $t = 16$.

(Gauther, 1992)

ローレンツモデル（3）

評価関数の形状１：Regular regime

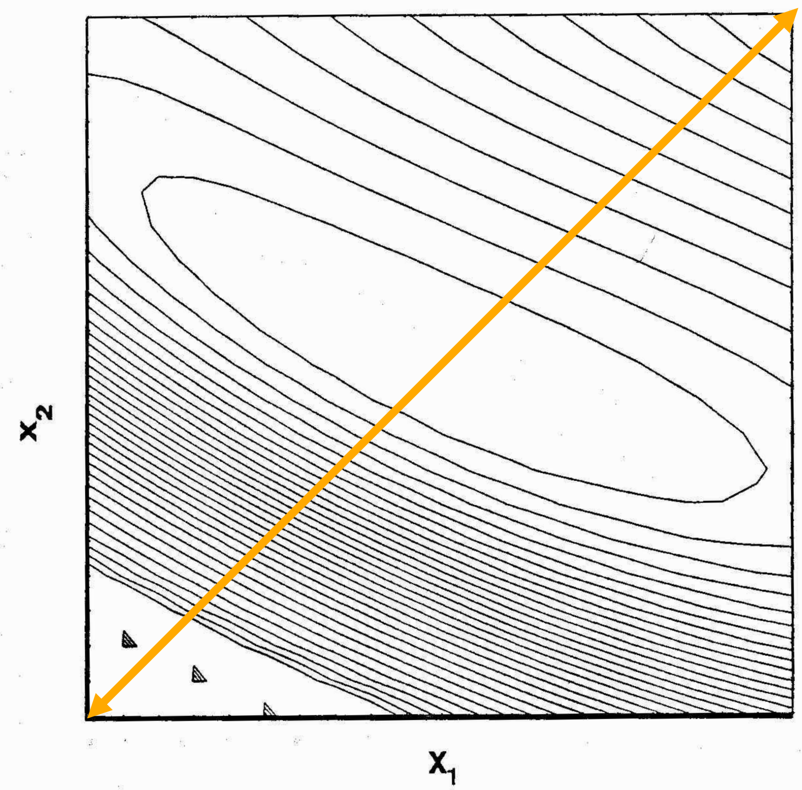


Fig. 4. Representation of the functional $J(Z_0)$ for the regular case: $Z_0 = (X_1, X_2, X_3)$ was varied such that:

$X_1^* - 2 < X_1 < X_1^* + 2,$

$X_2^* - 2 < X_2 < X_2^* + 2,$

with $X_3 = X_3^*$ and $(X_1^*, X_2^*, X_3^*) = (-9.42, -9.34, 28.3)$. Contour intervals are evenly spaced for values of $J(Z_0)$ between 0 and 50.

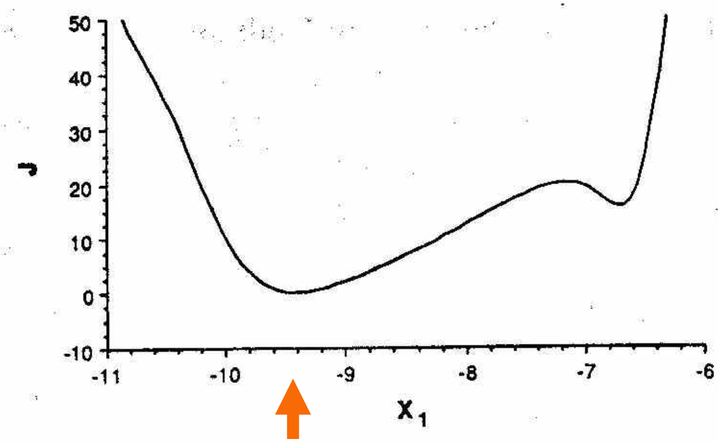


Fig. 5. Values of $J(Z_0)$ versus X_1 along a straight line going through the points

$P_{\pm} = (X_1^* \pm 2, X_2^* \pm 2, X_3^*).$

（Gauther，1992）

ローレンツモデル（4）

評価関数の形状2：Transition regime

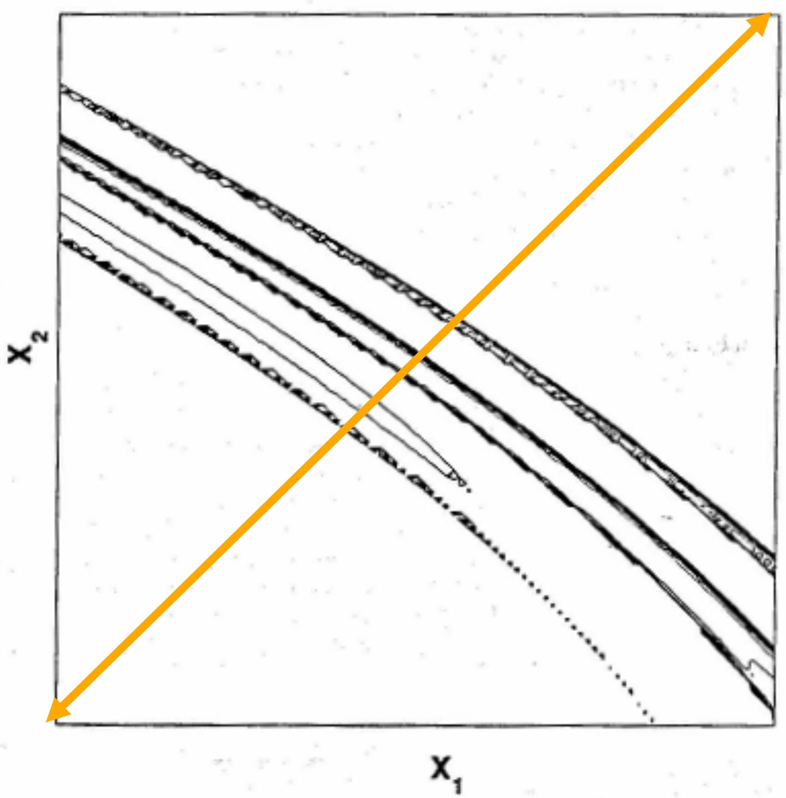


Fig. 8. Same as Fig. 4 but for the case with transition with

$(X_1^*, X_2^*, X_3^*) = (-5.92, -5.90, 24.0)$.

Contour intervals are evenly spaced for values of $J(Z_0)$ between 0 and 50. Values of $J(Z_0)$ above 50 are not shown.

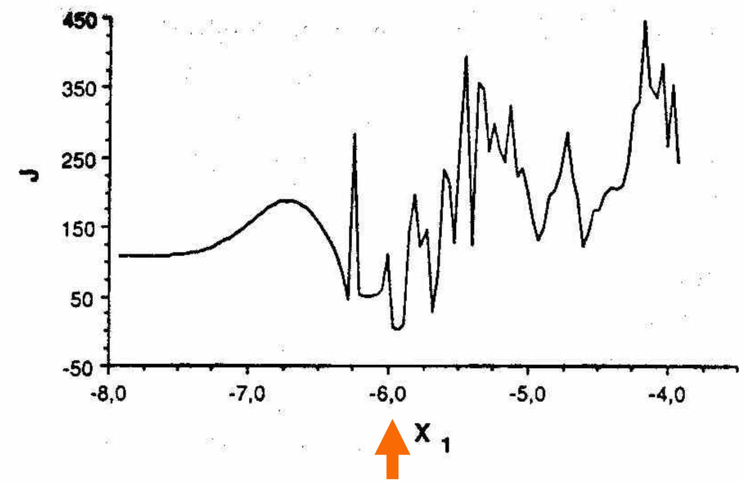


Fig. 9. Same as Fig. 5 but for the case with transition with

$(X_1^*, X_2^*, X_3^*) = (-5.92, -5.90, 24.0)$.

（Gauther, 1992）

ローレンツモデル（5）

評価関数の形状3：同化ウィンドウの長さの影響

T=8

T=10

T=15

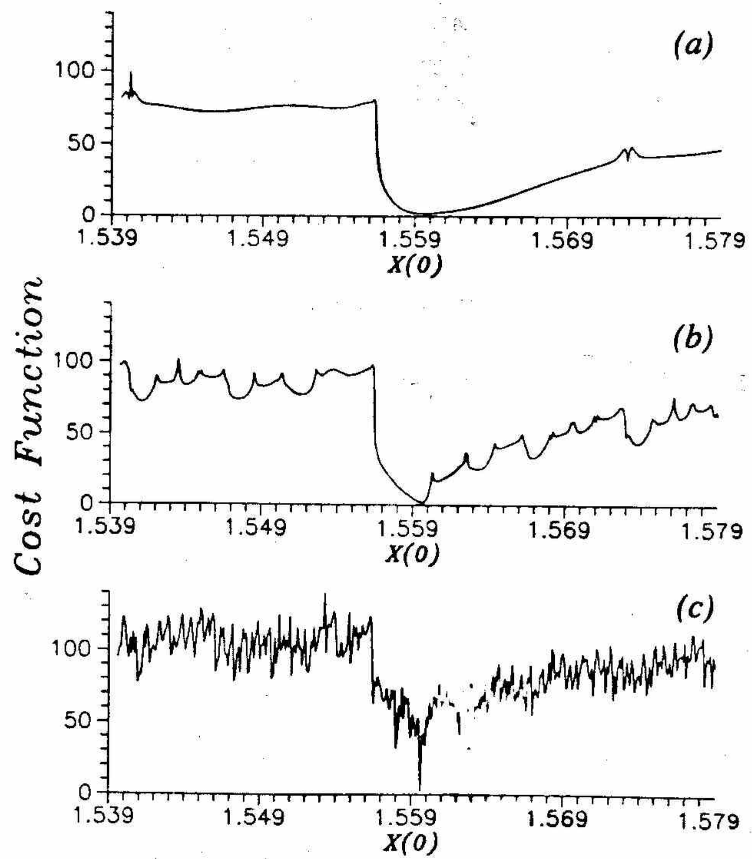


FIG. 6. Values of cost function as a funct: : initial X, with initial Y and Z held constant, in the neighborhood of the initial values used in calculating the reference solution. (a) Cost function, i.e., mean-square deviation of model solution with given initial data from "observed" values, where observations up to $t = 8$ are considered. (b) As in (a) but for observations up to $t = 10$. (c) As in (a) but for observations up to $t = 15$.

（Miller et al.，1994）

解析力学による導出(1)

観測データ $\mathbf{y}$ は時間的に連続的に得られるものとし、簡単のため評価関数 $J[\mathbf{x}]$ には予測値からの寄与がないものとする。

$$J[\mathbf{x}] = \int_0^T \frac{1}{2}\left(H(\mathbf{x}) - \mathbf{y}\right)^{\mathrm{T}} \mathbf{R}^{-1}\left(H(\mathbf{x}) - \mathbf{y}\right) dt$$

$$\dot{\mathbf{x}} := \frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}, t)$$

ラグランジアン($\mathbf{p}$:Lagrangeの未定乗数)

$$L(\mathbf{x}, \dot{\mathbf{x}}, \mathbf{p}, \dot{\mathbf{p}}, t) = -\frac{1}{2}\left(H(\mathbf{x}) - \mathbf{y}\right)^{\mathrm{T}} \mathbf{R}^{-1}\left(H(\mathbf{x}) - \mathbf{y}\right) + \mathbf{p}^{\mathrm{T}}\left(\dot{\mathbf{x}} - \mathbf{F}(\mathbf{x}, t)\right)$$

一般化運動量

$$\frac{\partial L}{\partial \dot{\mathbf{x}}} = \mathbf{p}, \qquad \frac{\partial L}{\partial \dot{\mathbf{p}}} = \mathbf{0}$$

(特異ラグランジアン)

ハミルトニアン

$$H = \dot{\mathbf{x}}^{\mathrm{T}} \frac{\partial L}{\partial \dot{\mathbf{x}}} + \dot{\mathbf{p}}^{\mathrm{T}} \frac{\partial L}{\partial \dot{\mathbf{p}}} - L$$

$$= \frac{1}{2}\left(H(\mathbf{x}) - \mathbf{y}\right)^{\mathrm{T}} \mathbf{R}^{-1}\left(H(\mathbf{x}) - \mathbf{y}\right) + \mathbf{p}^{\mathrm{T}}\mathbf{F}(\mathbf{x}, t) = H(\mathbf{x}, \mathbf{p}, t)$$

解析力学による導出（2）

ハミルトン方程式

$$\frac{d\mathbf{x}}{dt} = \frac{\partial H}{\partial \mathbf{p}} = \mathbf{F}(\mathbf{x}, t)$$ （物理の世界）

$$-\frac{d\mathbf{p}}{dt} = \frac{\partial H}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right)^{\mathrm{T}} \mathbf{p} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\big(H(\mathbf{x}) - \mathbf{y}\big)$$ （情報の世界）

4次元変分法の最小値探索のためのアジョイント方程式が得られた。

$\mathbf{p}$ は情報の世界の状態変数なので、同化ウィンドウ$[0, T]$ の外では $\mathbf{p} = \mathbf{0}$ と仮定するのが自然。そこで、連続性から次の境界条件を課す。

$$\mathbf{p}(0) = \mathbf{p}(T) = \mathbf{0}$$

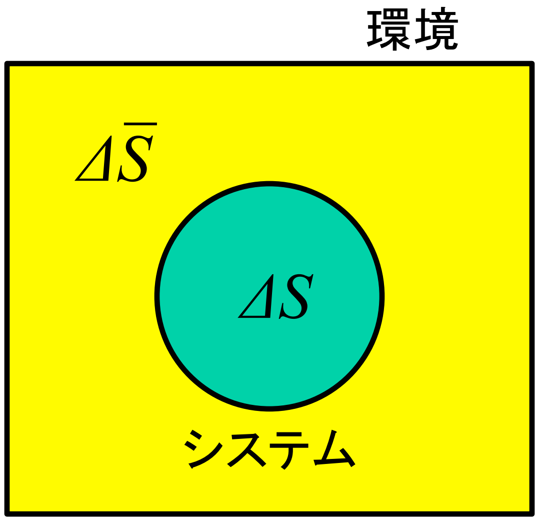
この条件の下での解 $\big(\mathbf{x}(t), \mathbf{p}(t)\big)$ は、評価関数 J を最小にする。なぜなら、アジョイント方程式を終期条件 $\mathbf{p}(T) = \mathbf{0}$ から過去に向かって時間積分した解を$\mathbf{p}(t)$ とすると

$$\nabla_{\mathbf{x}(0)} J = \mathbf{p}(0)$$

が成り立つから。

物理と情報（1）

ΔS ：システムのエントロピーの変化
$\Delta \bar{S}$ ：環境のエントロピーの変化



熱力学第2法則より

$$\Delta S + \Delta \bar{S} \geq 0$$

ΔI ：システムに関して取得した情報量

$$\Delta I = -\Delta S, \qquad \therefore \Delta \bar{S} \geq \Delta I$$

① 1ビットの情報量をエントロピーに換算すると、k_B をBoltzmann定数（$1.3806 \times 10^{-23}\,\mathrm{J\,K^{-1}}$）として $k_B \ln 2$ 。

したがって、一般的に、マクロなシステムに関する情報の取得によるエントロピーの減少は無視できる。

（Laplaceの魔にとって、世界のエントロピーは0か？）

物理と情報（2）

② 熱力学的なエントロピーは、熱平衡状態に対して定義される。局所熱力学平衡が成り立っていれば、エントロピーを局所的に定義できる。

マクロなシステムについて部分的な情報を取得しても、情報の得られていない他の部分との相互作用のために、情報の取得によるエントロピーの減少は短時間で失われる。

（Maxwellの魔は、なぜシステムのエントロピーを減らせるのか？）

③ 情報は何らかの物理系に記録されるので、システムのエントロピーではなく、その物理系のエントロピーが減る。ただし、ノイズに対して情報を一定期間保持するためには、エントロピーの減少量を保持する情報量より大きくする必要がある。

マクロなシステムでは、物理と情報を独立に扱える。

モデル誤差の考慮(1)

数値モデルのランダム誤差を考慮する($\boldsymbol{\varepsilon}_k$:白色ガウス過程)。

$$\mathbf{x}_k = M_k(\mathbf{x}_{k-1}) + \boldsymbol{\varepsilon}_k \qquad (k = 1, \cdots, K)$$

$$\langle \boldsymbol{\varepsilon}_k \rangle = \mathbf{0}, \qquad \delta_{kl} \langle \boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_l^{\mathrm{T}} \rangle = \mathbf{Q}_k \delta_{kl}$$

Bayes の定理より

$$p(\mathbf{x}_0, \cdots, \mathbf{x}_K \mid \mathbf{y}_1, \cdots, \mathbf{y}_K)$$

$$= \frac{p(\mathbf{y}_1 \mid \mathbf{x}_1) \cdots p(\mathbf{y}_K \mid \mathbf{x}_K) p(\mathbf{x}_0, \cdots, \mathbf{x}_K)}{\int \cdots \int p(\mathbf{y}_1 \mid \mathbf{x}_1) \cdots p(\mathbf{y}_K \mid \mathbf{x}_K) p(\mathbf{x}_0, \cdots, \mathbf{x}_K) d^n \mathbf{x}_0 \cdots d^n \mathbf{x}_K}$$

$$= \frac{p(\mathbf{y}_1 \mid \mathbf{x}_1) \cdots p(\mathbf{y}_K \mid \mathbf{x}_K) p(\mathbf{x}_0) p(\mathbf{x}_1 \mid \mathbf{x}_0) \cdots p(\mathbf{x}_K \mid \mathbf{x}_{K-1})}{\int \cdots \int p(\mathbf{y}_1 \mid \mathbf{x}_1) \cdots p(\mathbf{y}_K \mid \mathbf{x}_K) p(\mathbf{x}_0) p(\mathbf{x}_1 \mid \mathbf{x}_0) \cdots p(\mathbf{x}_K \mid \mathbf{x}_{K-1}) d^n \mathbf{x}_0 \cdots d^n \mathbf{x}_K}$$

評価関数

$$J(\mathbf{x}_0, \cdots, \mathbf{x}_K) = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right)^{\mathrm{T}} \left(\mathbf{P}_0^f\right)^{-1} \left(\mathbf{x}_0 - \mathbf{x}_0^f\right) + \sum_{k=1}^{K} \frac{1}{2}\left(H_k(\mathbf{x}_k) - \mathbf{y}_k\right)^{\mathrm{T}} \mathbf{R}_k^{-1} \left(H_k(\mathbf{x}_k) - \mathbf{y}_k\right)$$

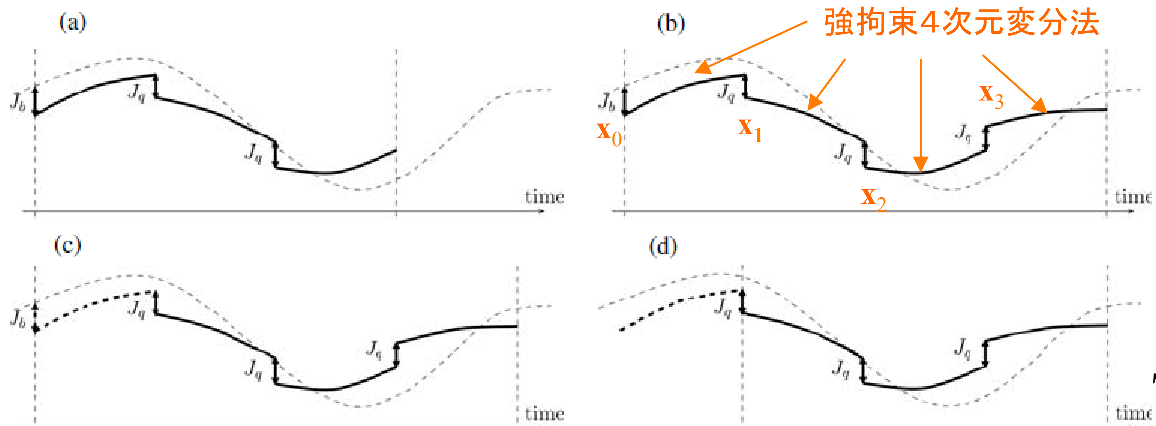
$$+ \sum_{k=1}^{K} \frac{1}{2}\left(\mathbf{x}_k - M_k(\mathbf{x}_{k-1})\right)^{\mathrm{T}} \mathbf{Q}_k^{-1} \left(\mathbf{x}_k - M_k(\mathbf{x}_{k-1})\right)$$

モデル誤差の考慮(2)

弱拘束4次元変分法

・モデル誤差を考慮することによって、長い同化ウィンドウを採用できる。

・モデル誤差の導入を間欠的に行えば、制御変数の個数を減らせる。



Tremolet (2006)

評価関数(同化ウィンドウが十分長いので予測値による項を無視する)

$$J(\mathbf{x}_0,\cdots,\mathbf{x}_N) = J_1(\mathbf{x}_0)+\cdots+J_N(\mathbf{x}_{N-1})+J_q(\mathbf{x}_1,\mathbf{x}_2)+\cdots+J_q(\mathbf{x}_{N-1},\mathbf{x}_N)$$

$$J_i(\mathbf{x}_{i-1}) := \sum_{k=1}^{K_i}\frac{1}{2}\left(H_k^{(i)}(\mathbf{x}_k^{(i)})-\mathbf{y}_k^{(i)}\right)^{\mathrm{T}}\mathbf{R}_k^{(i)-1}\left(H_k^{(i)}(\mathbf{x}_k^{(i)})-\mathbf{y}_k^{(i)}\right)$$

$$J_q(\mathbf{x}_{i-1},\mathbf{x}_i) := \frac{1}{2}\left(\mathbf{x}_i - M_i(\mathbf{x}_{i-1})\right)^{\mathrm{T}}\mathbf{Q}_i^{-1}\left(\mathbf{x}_i - M_i(\mathbf{x}_{i-1})\right)$$

$J_i(\mathbf{x}_{i-1}), \nabla J_i(\mathbf{x}_{i-1}) \quad (i=1,\cdots,N)$ をすべて並列的に計算できる。

アンサンブル4次元変分法（1）

En4DVar: Liu et al (2008, 2009)

$\mathbf{x}_0^f$ と$\mathbf{P}_0^f$ に加えて予測値のアンサンブル$\left\{\mathbf{x}_0^{f(1)},\cdots,\mathbf{x}_0^{f(N)}\right\}$があるとする。

$$\mathbf{X}_0^f := \frac{1}{\sqrt{N-1}}\left(\mathbf{x}_0^{f(1)}-\mathbf{x}_0^f,\cdots,\mathbf{x}_0^{f(N)}-\mathbf{x}_0^f\right)$$

$$\mathbf{P}_0^f \approx \mathbf{X}_0^f\left(\mathbf{X}_0^f\right)^{\mathrm{T}}$$

同化ウィンドウの初めにおける状態変数を、アンサンブル偏差の線形結合を用いて近似する（$\mathbf{w}$:N 次元）。

$$\mathbf{x}_0 = \mathbf{x}_0^f + \mathbf{X}_0^f\mathbf{w}$$

事前確率密度関数による期待値をとると

$$\left\langle \mathbf{x}_0 \right\rangle = \mathbf{x}_0^f, \qquad \left\langle \left(\mathbf{x}_0-\mathbf{x}_0^f\right)\left(\mathbf{x}_0-\mathbf{x}_0^f\right)^{\mathrm{T}} \right\rangle = \mathbf{X}_0^f\left(\mathbf{X}_0^f\right)^{\mathrm{T}}$$

このためには$\mathbf{w}$ が次の条件を満たせばよい。

$$\left\langle \mathbf{w} \right\rangle = \mathbf{0}, \qquad \left\langle \mathbf{w}\,\mathbf{w}^{\mathrm{T}} \right\rangle = \mathbf{I}$$

アンサンブル4次元変分法（2）

評価関数（制御変数：$\mathbf{w}$ ）

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + \frac{1}{2}\left[\tilde{H}(\tilde{M}(\mathbf{x}_0^f + \mathbf{X}_0^f\mathbf{w})) - \tilde{\mathbf{y}}\right]^{\mathrm{T}} \tilde{\mathbf{R}}^{-1}\left[\tilde{H}(\tilde{M}(\mathbf{x}_0^f + \mathbf{X}_0^f\mathbf{w})) - \tilde{\mathbf{y}}\right]$$

評価関数の勾配

$$\nabla J(\mathbf{w}) = \mathbf{w} + \left(\mathbf{X}_0^f\right)^{\mathrm{T}} \tilde{\mathbf{M}}^{\mathrm{T}} \tilde{\mathbf{H}}^{\mathrm{T}} \tilde{\mathbf{R}}^{-1}\left[\tilde{H}(\tilde{M}(\mathbf{x}_0^f + \mathbf{X}_0^f\mathbf{w})) - \tilde{\mathbf{y}}\right]$$

$$= \mathbf{w} + \left(\tilde{\mathbf{H}}\tilde{\mathbf{M}}\mathbf{X}_0^f\right)^{\mathrm{T}} \tilde{\mathbf{R}}^{-1}\left[\tilde{H}(\tilde{M}(\mathbf{x}_0^f + \mathbf{X}_0^f\mathbf{w})) - \tilde{\mathbf{y}}\right]$$

アンサンブル予報を使えば、勾配をアジョイントモデルなしで近似計算できる。

$$\tilde{\mathbf{H}}\tilde{\mathbf{M}}\mathbf{X}_0^f \approx \frac{1}{\sqrt{N-1}}\Big[\tilde{H}(\tilde{M}(\mathbf{x}_0^{f(1)} + \mathbf{X}_0^f\mathbf{w})) - \tilde{H}(\tilde{M}(\mathbf{x}_0^f + \mathbf{X}_0^f\mathbf{w})), \cdots, \\ \tilde{H}(\tilde{M}(\mathbf{x}_0^{f(N)} + \mathbf{X}_0^f\mathbf{w})) - \tilde{H}(\tilde{M}(\mathbf{x}_0^f + \mathbf{X}_0^f\mathbf{w}))\Big]$$

評価関数を逐次的に2次関数で近似する場合には

$$\tilde{H}(\tilde{M}(\mathbf{x}_0^{f(i)} + \mathbf{X}_0^f\mathbf{w})) \approx \tilde{H}(\tilde{M}(\mathbf{x}_0^{f(i)})) + \tilde{\mathbf{H}}\tilde{\mathbf{M}}\mathbf{X}_0^f\mathbf{w}$$

アンサンブル4次元変分法(3)

アンサンブルカルマンフィルタと同様に、サンプリングエラーを抑えるために共分散を局所化する。そのための相関行列$\mathbf{C}$ を次のように分解する。

$$\mathbf{C} = \mathbf{C}'\left(\mathbf{C}'\right)^{\mathrm{T}} \qquad （\mathbf{C}' : n \times r \text{ 行列}）$$

局所化を行わない場合は$\mathbf{C}' = \mathbf{1}$ である。

同化ウィンドウの初めの状態変数を次のように表す（$\mathbf{w}$ ：Nr 次元）。

$$\mathbf{x}_0 = \mathbf{x}_0^f + \left(\mathrm{diag}\left[\left(\mathbf{X}_0^f\right)_1\right]\mathbf{C}', \cdots, \mathrm{diag}\left[\left(\mathbf{X}_0^f\right)_N\right]\mathbf{C}'\right)\mathbf{w}$$

$$\mathrm{diag}\left[\left(\mathbf{X}_0^f\right)_i\right] := \frac{1}{\sqrt{N-1}}\begin{pmatrix} \left(\mathbf{x}_0^{f(i)} - \mathbf{x}_0^f\right)_1 & & 0 \\ & \ddots & \\ 0 & & \left(\mathbf{x}_0^{f(i)} - \mathbf{x}_0^f\right)_n \end{pmatrix} \qquad (i = 1, \cdots, N)$$

このとき予測誤差共分散行列は以下のように局所化される。

$$\left\langle \left(\mathbf{x}_0^f - \mathbf{x}_0^f\right)\left(\mathbf{x}_0^f - \mathbf{x}_0^f\right)^{\mathrm{T}} \right\rangle = \sum_{i=1}^{N} \mathrm{diag}\left[\left(\mathbf{X}_0^f\right)_1\right]\mathbf{C}\,\mathrm{diag}\left[\left(\mathbf{X}_0^f\right)_1\right]$$

$$= \mathbf{C} \circ \mathbf{X}_0^f\left(\mathbf{X}_0^f\right)^{\mathrm{T}}$$