Data Assimilation
- A12. EnVAR -

Shunji Kotsuki

Center for Environmental Remote Sensing / Institute of Advanced Academic Research

(shunji.kotsuki@chiba-u.jp)

Today's Goal

- **Lecture**
  - what is the En-Var?
  - what is the cost function of EnVAR?
  - how to employ back propagation?

- **Training Course**
  - to implement EnVAR

Review: 4DVAR

4DVAR Equation

Environmental Prediction Science Laboratory

Cost Function (scalar)

Tangent Linear Model $\mathbf{M}_{i|0} = \mathbf{M}_{i-1} \cdots \mathbf{M}_1 \mathbf{M}_0$

$$J(\delta\mathbf{x}_0) \approx \frac{1}{2}\delta\mathbf{x}_0^T \mathbf{B}_0^{-1}\delta\mathbf{x}_0 + \sum_{i=1}^{k}\frac{1}{2}\left(\mathbf{H}_i\mathbf{M}_{i|0}\delta\mathbf{x}_0 - \mathbf{d}_i^{o-b}\right)^T \mathbf{R}_i^{-1}\left(H_i(M_{i|0}(\mathbf{x}_0)) - \mathbf{y}_i^o\right)$$

Jacobian ($\in \mathbb{R}^n$)

numerical model

obs operator

$$\frac{\partial J}{\partial(\delta\mathbf{x}_0)} \approx \mathbf{B}_0^{-1}(\delta\mathbf{x}_0) + \sum_{i=1}^{k}\mathbf{M}_0^T \cdots \mathbf{M}_{i-2}^T\mathbf{M}_{i-1}^T\mathbf{H}_i^T\mathbf{R}_i^{-1}\left(H_i(M_{i|0}(\mathbf{x}_0)) - \mathbf{y}_i^o\right)$$

backpropagations

Hessian ($\in \mathbb{R}^{n\times n}$)

$$\frac{\partial^2 J}{\partial(\delta\mathbf{x}_0)^2} \approx \mathbf{B}_0^{-1} + \sum_{i=1}^{k}\mathbf{M}_0^T \cdots \mathbf{M}_{i-2}^T\mathbf{M}_{i-1}^T\mathbf{H}_i^T\mathbf{R}_i^{-1}\mathbf{H}_i\,\mathbf{M}_{i-1}\cdots\mathbf{M}_1\mathbf{M}_0 = \mathbf{A}_0^{-1}$$

Problem to be solved

$$\delta\mathbf{x}_0^a = \operatorname{argmin} J(\delta\mathbf{x}_0) \quad \text{subject to} \ \frac{\partial J}{\partial(\delta\mathbf{x}_0)} = 0$$

$$\mathbf{x}_0^a = \mathbf{x}_0^b + \delta\mathbf{x}_0^a$$

$$\mathbf{x}_i^a = M_{i|0}(\mathbf{x}_0^a)$$

Comparison

	4DVAR	4D EnVAR	EnKF
Solver	Iterative (OL & IL)	Iterative (IL only)	Deterministic
Ensemble FCST		☑	☑
TLM & ADJ of M	☑		
TLM & ADJ of H	☑	☑	

Background Knowledge on Minimization

Minimization

Scalar Case

$$J = \frac{1}{2}px^2 + qx + r = \frac{1}{2}p(x + \frac{q}{p})^2 + r'$$

$$J' = px + q \qquad J'' = p$$

$$J' = 0 \Rightarrow x = -\frac{q}{p} = -(J'')^{-1}J'|_{x=0}$$

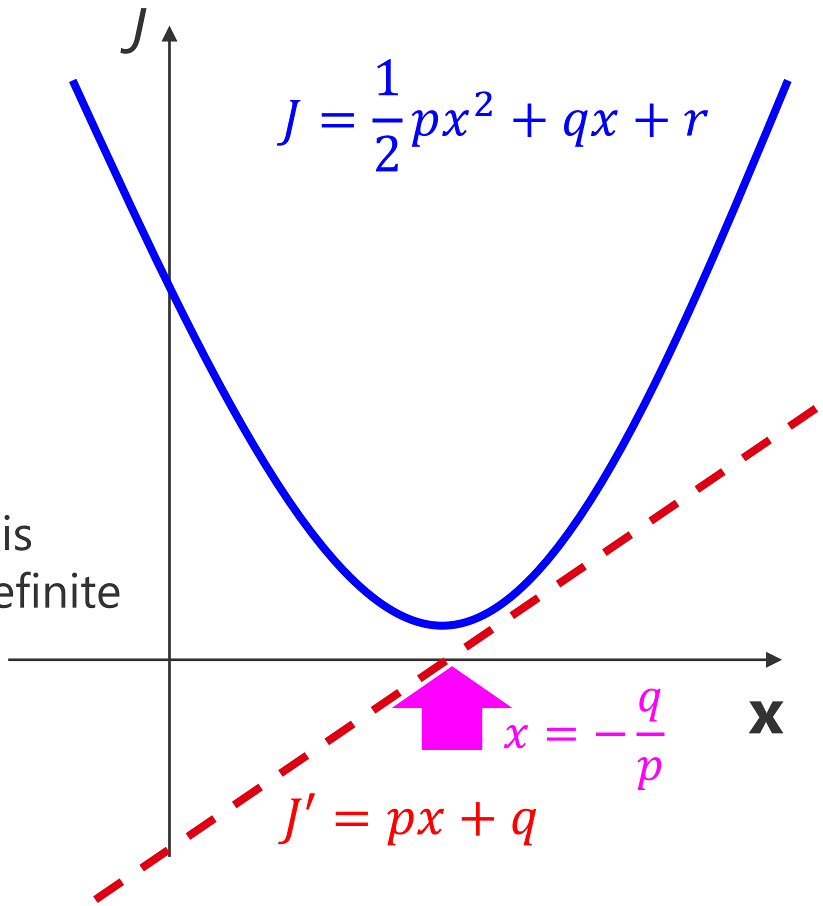
Multi-dimensional Case

$$J = \frac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r$$

suppose **P** is positive-semidefinite

$$J' = \mathbf{P}\mathbf{x} + \mathbf{q} \qquad J'' = \mathbf{P}$$

$$J' = \mathbf{0} \;\Rightarrow\; \mathbf{x} = -\mathbf{P}^{-1}\mathbf{q} = -(J'')^{-1}J'|_{\mathbf{x}=0}$$



Point 1: If inverse of Hessian $(J'')^{-1}$ is available

➔ no iteration is necessary to find $\mathbf{x}$ that minimizes the cost function

Point 2: But computing $(J'')^{-1}$ is expensive $\sim O(n^3)$ where n is the size of $\mathbf{x}$

➔ For EnVAR, computing $(J'')^{-1}$ is affordable $\sim O(m^3)$

To see KF as minimization

Cost Function (scalar), Jacobian ($\in \mathbb{R}^n$), ***Hessian*** ($\in \mathbb{R}^{n\times n}$)

$$J(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^T\mathbf{B}^{-1}(\delta\mathbf{x}) + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}_t^{o-b})^T\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}_t^{o-b})$$

$$J'(\delta\mathbf{x}) = \mathbf{B}^{-1}(\delta\mathbf{x}) + \mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{H}\delta\mathbf{x} - \mathbf{d}_t^{o-b}\right)$$

$$= (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\ \delta\mathbf{x} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{d}_t^{o-b} = \mathbf{A}^{-1}\delta\mathbf{x} - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{d}_t^{o-b}$$

$$J''(\delta\mathbf{x}) = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$$

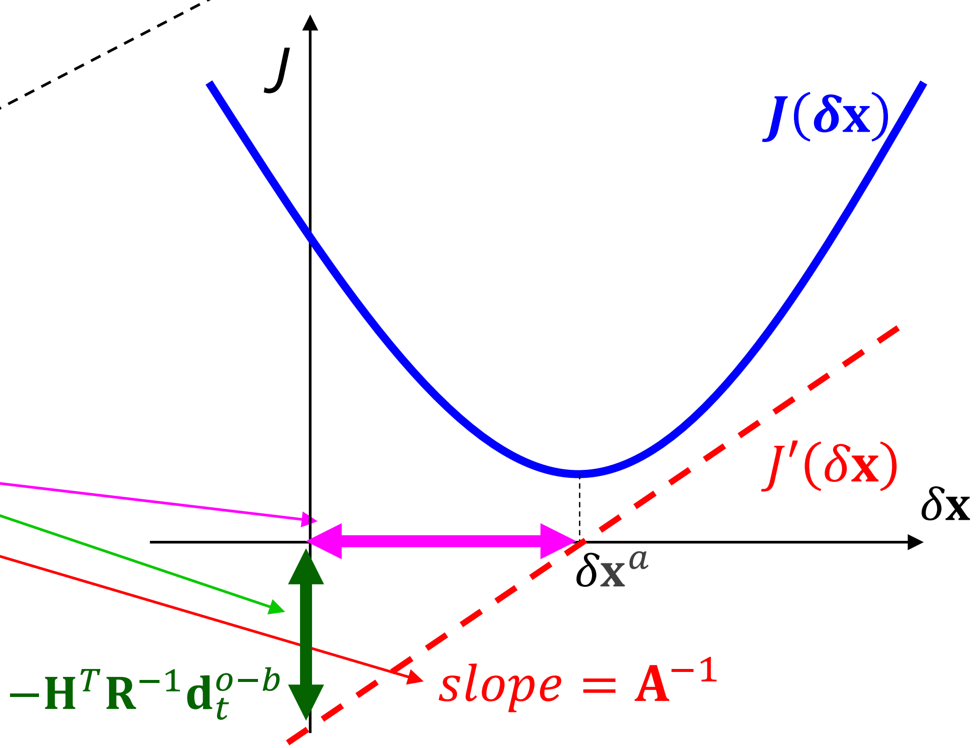
Minimization

$$J'(\delta\mathbf{x}) = 0$$

$$\Rightarrow \delta\mathbf{x}^a = \mathbf{A}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{d}_t^{o-b}$$

where $\mathbf{A} = (J'')^{-1}$

$$\mathbf{H}^T\mathbf{R}^{-1}\mathbf{d}_t^{o-b} = J'|_{\delta\mathbf{x}=0}$$



Pre-conditioning

Cost Function (scalar), Jacobian ($\in \mathbb{R}^n$), ***Hessian*** ($\in \mathbb{R}^{n\times n}$)

$$J(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^T\mathbf{B}^{-1}(\delta\mathbf{x}) + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}_t^{o-b})^T\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}_t^{o-b})$$

$$J'(\delta\mathbf{x}) = \mathbf{A}^{-1}\delta\mathbf{x} - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{d}_t^{o-b}$$

$$J''(\delta\mathbf{x}) = \mathbf{A}^{-1}$$

variable conversion $\mathbf{z} = \mathbf{A}^{-1/2}\delta\mathbf{x}$

$$J(\mathbf{z}) = \frac{1}{2}(\mathbf{z})^T(\mathbf{z}) - (\mathbf{d}_t^{o-b})^T\mathbf{R}^{-1}\mathbf{H}\mathbf{A}^{1/2}\mathbf{z} + \frac{1}{2}(\mathbf{d}_t^{o-b})^T\mathbf{R}^{-1}\mathbf{d}_t^{o-b}$$

$$J'(\mathbf{z}) = \mathbf{z} - (\mathbf{d}_t^{o-b})^T\mathbf{R}^{-1}\mathbf{H}\mathbf{A}^{1/2}$$

$$J''(\mathbf{z}) = \mathbf{I}$$

with pre-conditioning, Hessian becomes to be the identity matrix

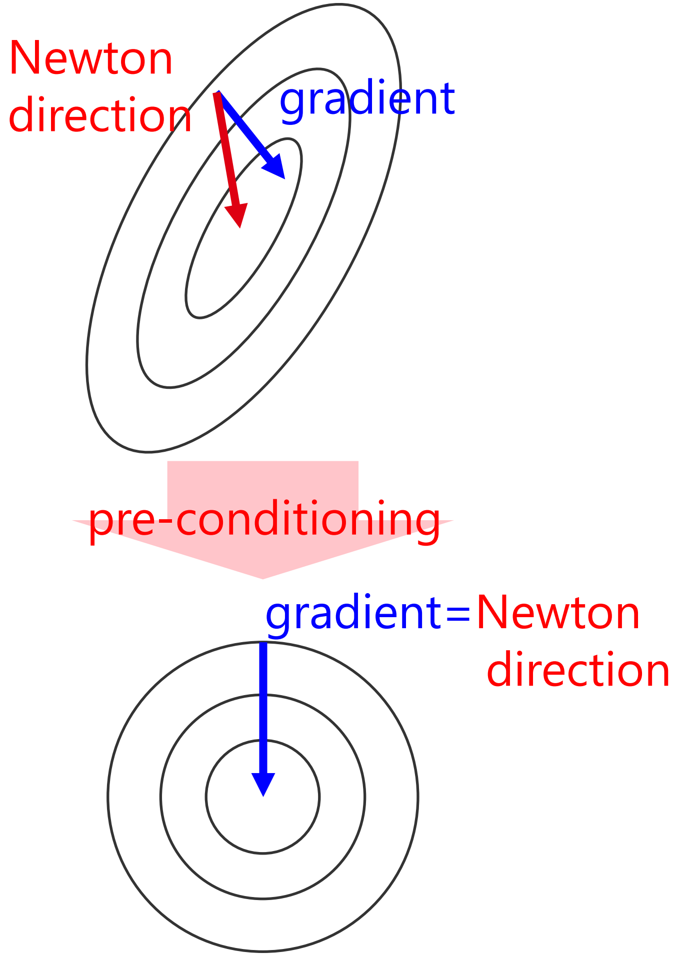
Comparison of minimizations

Gradient and Newton methods

$$J' = \mathbf{B}^{-1}(\delta\mathbf{x}) + \mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{H}\delta\mathbf{x} - \mathbf{d}_t^{o-b}\right)$$

$$\delta\mathbf{x} = -(J'')^{-1}J'$$

to change the gradient toward minima



with pre-conditioning,
gradient gives the way toward minima
because Hessian is the identity matrix

Computational Convergence

	Gradient	Newton
w/o pre-conditioning	slow	good
w/ pre-conditioning	good	/

EnVAR
(Ensemble Variational DA)

Cost Function of EnVAR

Environmental Prediction Science Laboratory

Cost Function of 4DVAR

$$H_i(M_{i|0}(\mathbf{x}_0)) - \mathbf{y}_i^o$$

$$J(\delta\mathbf{x}_0) \approx \frac{1}{2}\delta\mathbf{x}_0^T\mathbf{B}_0^{-1}\delta\mathbf{x}_0 + \sum_{i=1}^{k}\frac{1}{2}\left(H_i(M_{i|0}(\mathbf{x}_0)) - \mathbf{y}_i^o\right)^T\mathbf{R}_i^{-1}\left(H_i(M_{i|0}(\mathbf{x}_0)) - \mathbf{y}_i^o\right)$$

ensemble

$$\mathbf{x}_0 = \bar{\mathbf{x}}_0^b + \delta\mathbf{x}_0 = \bar{\mathbf{x}}_0^b + \delta\mathbf{X}_0\mathbf{w}$$

$$M_{i|0}(\mathbf{x}_0) = M_{i|0}(\bar{\mathbf{x}}_0^b + \delta\mathbf{X}_0\mathbf{w})$$

Cost Function (scalar)

$$J(\mathbf{w}) \approx \frac{1}{2}\mathbf{w}^T\mathbf{w} + \sum_{i=1}^{k}\frac{1}{2}\left(H_i(\bar{\mathbf{x}}_i^b + \delta\mathbf{X}_i\mathbf{w}) - \mathbf{y}_i^o\right)^T\mathbf{R}_i^{-1}\left(H_i(\bar{\mathbf{x}}_i^b + \delta\mathbf{X}_i\mathbf{w}) - \mathbf{y}_i^o\right)$$

Jacobian **($\in \mathbb{R}^m$)**

$$\frac{\partial J}{\partial(\mathbf{w})} \approx \mathbf{w} + \sum_{i=1}^{k}\delta\mathbf{Y}_i^T\mathbf{R}_i^{-1}\left(H_i(\bar{\mathbf{x}}_i^b + \delta\mathbf{X}_i\mathbf{w}) - \mathbf{y}_i^o\right)$$

Hessian **($\in \mathbb{R}^{m\times m}$)**

$$\frac{\partial^2 J}{\partial(\mathbf{w})^2} \approx \mathbf{I} + \sum_{i=1}^{k}\delta\mathbf{Y}_i^T\mathbf{R}_i^{-1}\delta\mathbf{Y}_i = \widetilde{\mathbf{A}}_0^{-1}$$

analysis error covariance in ensemble subspace

Point 1:
$\delta\mathbf{X}$ is not changed over iterations (one ens. fcst)

Point 2:
nonlinearity of obs ope. H should be considered for every iteration.
Otherwise, EnVAR yields the same analysis of ETKF.

Importance of non-linearity

Environmental Prediction Science Laboratory

Cost Function (scalar) and Jacobian ($\in \mathbb{R}^m$) w/ nonlinearity

$$J(\mathbf{w}) \approx \frac{1}{2}\mathbf{w}^T\mathbf{w} + \sum_{i=1}^{k} \frac{1}{2}\left(H_i(\bar{\mathbf{x}}_i^b + \delta\mathbf{X}_i\mathbf{w}) - \mathbf{y}_i^o\right)^T \mathbf{R}_i^{-1}\left(H_i(\bar{\mathbf{x}}_i^b + \delta\mathbf{X}_i\mathbf{w}) - \mathbf{y}_i^o\right)$$

$$\frac{\partial J}{\partial(\mathbf{w})} \approx \mathbf{w} + \sum_{i=1}^{k} \delta\mathbf{Y}_i^T \mathbf{R}_i^{-1}\left(H_i(\bar{\mathbf{x}}_i^b + \delta\mathbf{X}_i\mathbf{w}) - \mathbf{y}_i^o\right)$$

$$H_i\left(\bar{\mathbf{x}}_i^b + \delta\mathbf{X}_i\mathbf{w}\right) - \mathbf{y}_i^o \approx \mathbf{H}\delta\mathbf{X}_i\mathbf{w} - \mathbf{d}_i^o \equiv \delta\mathbf{Y}_i\mathbf{w} - \mathbf{d}_i^o$$

Cost Function (scalar) and Jacobian ($\in \mathbb{R}^m$) w/o nonlinearity

$$J(\mathbf{w}) \approx \frac{1}{2}\mathbf{w}^T\mathbf{w} + \sum_{i=1}^{k} \frac{1}{2}(\delta\mathbf{Y}_i\mathbf{w} - \mathbf{d}_i^o)^T \mathbf{R}_i^{-1}(\delta\mathbf{Y}_i\mathbf{w} - \mathbf{d}_i^o)$$

$$\frac{\partial J}{\partial(\mathbf{w})} \approx \mathbf{w} + \sum_{i=1}^{k} \delta\mathbf{Y}_i^T \mathbf{R}_i^{-1}(\delta\mathbf{Y}_i\mathbf{w} - \mathbf{d}_i^o) \qquad \frac{\partial^2 J}{\partial(\mathbf{w})^2} \approx \mathbf{I} + \sum_{i=1}^{k} \delta\mathbf{Y}_i^T \mathbf{R}_i^{-1}\delta\mathbf{Y}_i = \widetilde{\mathbf{A}}_0^{-1}$$

$$\frac{\partial J}{\partial(\mathbf{w})} = 0 \;\Rightarrow\; \mathbf{w} + \sum_{i=1}^{k} \delta\mathbf{Y}_i^T \mathbf{R}_i^{-1}(\delta\mathbf{Y}_i\mathbf{w} - \mathbf{d}_i^o) = 0$$

Point 1: the weight can be determined deterministically (no iteration is needed).
Point 2: the analysis is equivalent to 4D-LETKF (➔ no reason to use iterative solver)

Thank you for your attention!

Presented by Shunji Kotsuki
(shunji.kotsuki@chiba-u.jp)

Further information is available at
https://kotsuki-lab.com/

Environmental Prediction Science Laboratory