# 強拘束下でのランダム部分空間アンサンブルによる推定

中野 慎也 モデリング研究系 准教授

## Iterative ensemble variational method

We consider the following strong constraint data assimilation problem:

$$J(\boldsymbol{x}_0)=\frac{1}{2}(\boldsymbol{x}_0-\bar{\boldsymbol{x}}_{0,b})^T\mathsf{P}_{0,b}^{-1}(\boldsymbol{x}_0-\bar{\boldsymbol{x}}_{0,b})+\frac{1}{2}\sum_{k=1}^{K}\left(\boldsymbol{y}_k-\boldsymbol{h}_k(\boldsymbol{x}_k)\right)^T\mathsf{R}^{-1}\left(\boldsymbol{y}_k-\boldsymbol{h}_k(\boldsymbol{x}_k)\right),$$

where $\boldsymbol{x}_k=\boldsymbol{f}_k\circ\boldsymbol{f}_{k-1}\circ\cdots\circ\boldsymbol{f}_1(\boldsymbol{x}_0)$.

Defining a function $\boldsymbol{g}_k$ as $\boldsymbol{g}_k(\boldsymbol{x}_0)=\boldsymbol{h}_k\circ\boldsymbol{f}_k\circ\boldsymbol{f}_{k-1}\circ\cdots\circ\boldsymbol{f}_1(\boldsymbol{x}_0)$,

$$J(\boldsymbol{x}_0)=\frac{1}{2}(\boldsymbol{x}_0-\bar{\boldsymbol{x}}_{0,b})^T\mathsf{P}_{0,b}^{-1}(\boldsymbol{x}_0-\bar{\boldsymbol{x}}_{0,b})+\frac{1}{2}\sum_{k=1}^{K}\left(\boldsymbol{y}_k-\boldsymbol{g}_k(\boldsymbol{x}_0)\right)^T\mathsf{R}^{-1}\left(\boldsymbol{y}_k-\boldsymbol{g}_k(\boldsymbol{x}_0)\right).$$

At each iteration, we consider matrices $\mathsf{X}_{0,m}$ and $\Gamma_m$ as follows:

$$\mathsf{X}_{0,m}=\frac{1}{\sqrt{N}}\left(\boldsymbol{x}_{0,m}^{(1)}-\bar{\boldsymbol{x}}_{0,m-1}\quad\cdots\quad\boldsymbol{x}_{0,m}^{(N)}-\bar{\boldsymbol{x}}_{0,m-1}\right),$$

$$\Gamma_{0,m}=\frac{1}{\sqrt{N}}\left(\boldsymbol{g}(\boldsymbol{x}_{0,m}^{(1)})-\boldsymbol{g}(\bar{\boldsymbol{x}}_{0,m-1})\quad\cdots\quad\boldsymbol{g}(\boldsymbol{x}_{0,m}^{(N)})-\boldsymbol{g}(\bar{\boldsymbol{x}}_{0,m-1})\right).$$

where $\{\boldsymbol{x}_{0,m}^{(i)}\}_{i=1}^{N}$ is the ensemble spanned in a random subspace.

We then minimize the following surrogate cost function:

$$J_m(\boldsymbol{w}_m\,|\,\bar{\boldsymbol{x}}_m)=\frac{\sigma_m^2}{2}\boldsymbol{w}_m^T\boldsymbol{w}_m+\frac{1}{2}\sum_{k=1}^{K}\left(\boldsymbol{y}_k-\boldsymbol{g}_k(\bar{\boldsymbol{x}}_{0,m})-\Gamma_{k,m}\boldsymbol{w}_m\right)^T\mathsf{R}^{-1}\left(\boldsymbol{y}_k-\boldsymbol{g}_k(\bar{\boldsymbol{x}}_{0,m})-\Gamma_{k,m}\boldsymbol{w}_m\right).$$

The (m+1)-th estimate is obtained by $\bar{\boldsymbol{x}}_{0,m+1}=\bar{\boldsymbol{x}}_{0,m}+\mathsf{X}_{0,m}\hat{\boldsymbol{w}}_m$.

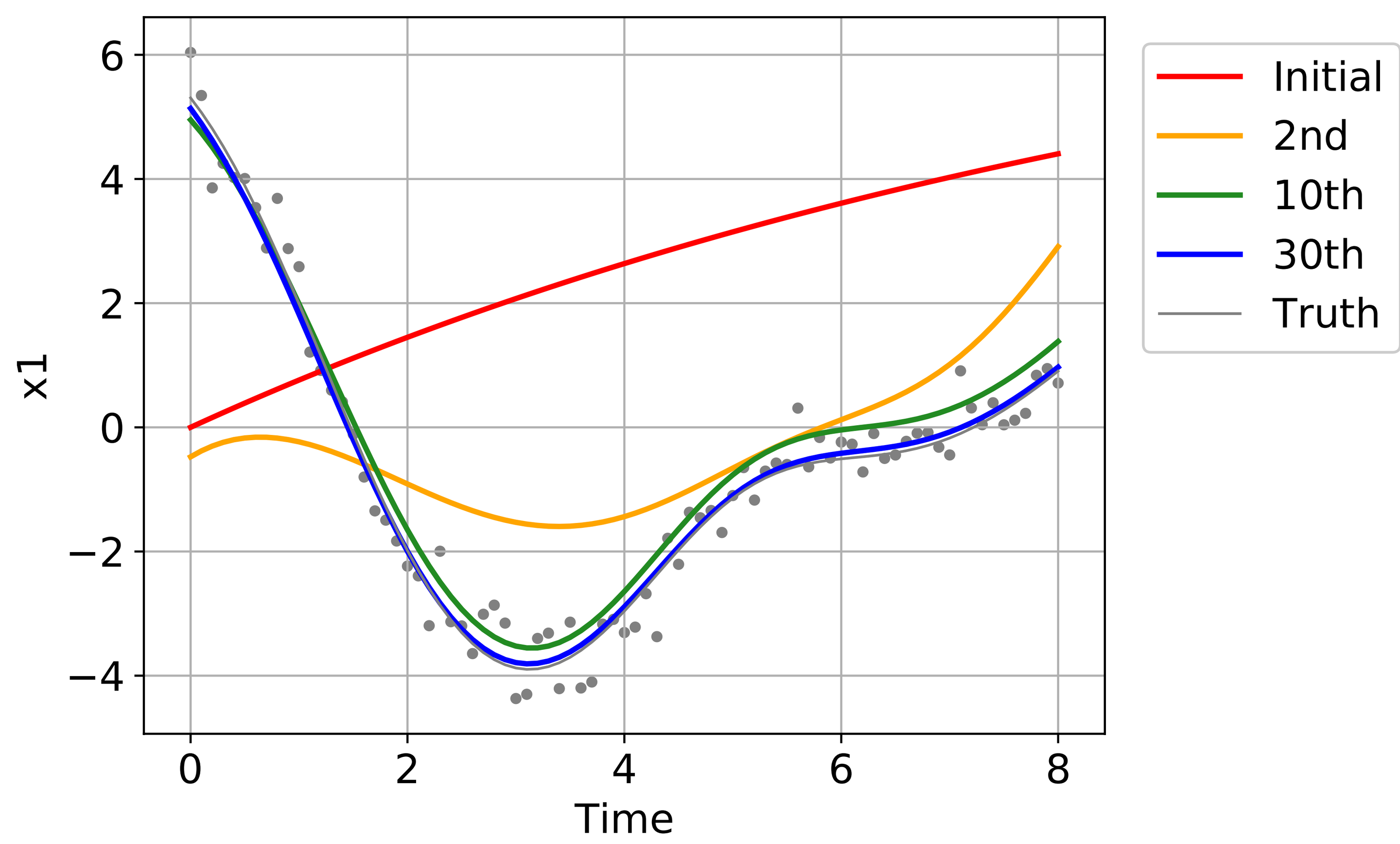
*See Nakano (2021) for details.*

## Application to Lorenz 96 model

We applied the Lorenz 96 model (Lorenz and Emanuel, 1998) with 400 variables:

$$\frac{dx_i}{dt}=(x_{i+1}-x_{i-2})-x_i+f,\qquad (i=1,\ldots,400).$$

200 ensemble members were used at each iteration.




Temporal evolution of one of the 400 variables at the initial guess, 2nd, 10th, and 30th iterations.

## Echo state network

We employ an echo state network for approximating the relationship between the solar wind variables and the *AL* index which is a geomagnetic index representing auroral activity.

At each time step, we update each state variable as follows:

$$x_{k,i}=0.1x_{k-1,i}+0.9\tanh(\gamma_i[\boldsymbol{w}_i^T(\boldsymbol{x}_{k-1}+\boldsymbol{x}_{\mathrm{in},k}+\alpha_i]),$$

where the parameters $\gamma_i$, $\boldsymbol{w}_i$, and $\alpha_i$ were randomly determined in advance and fixed.

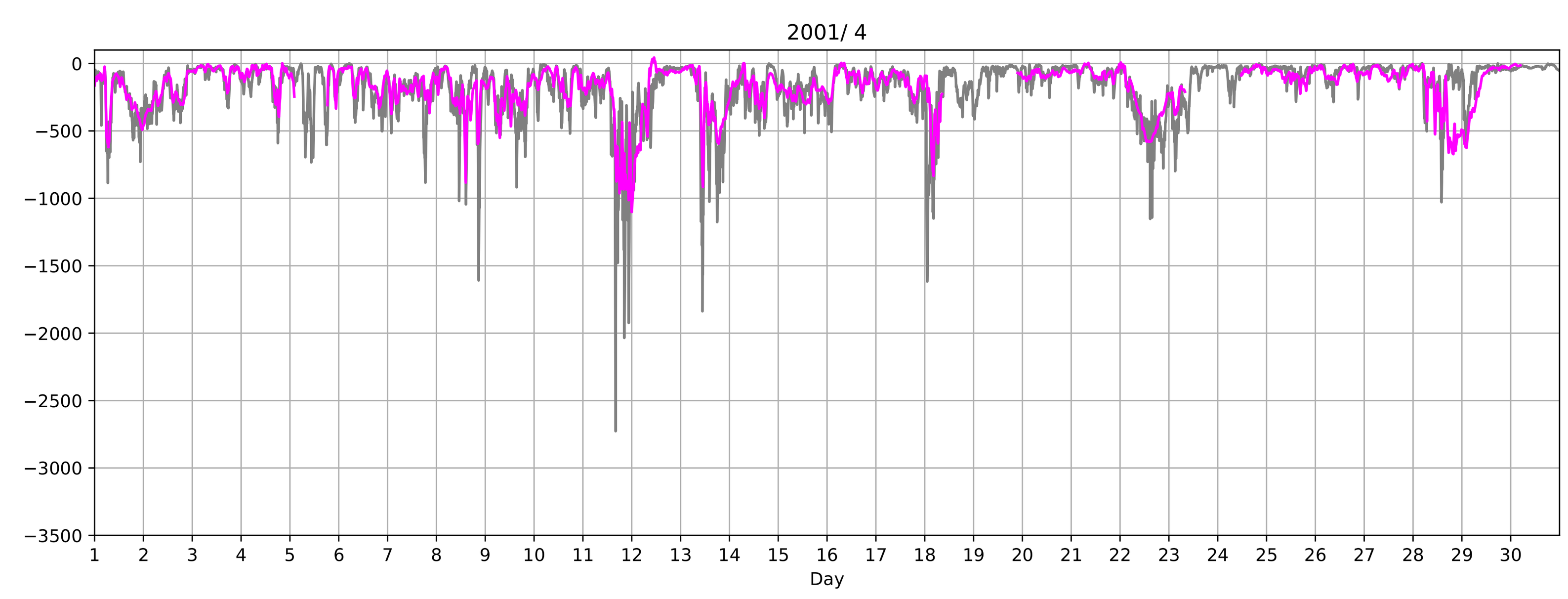
The output for the time step $k$ is then obtained as follows:

$$y_k=\boldsymbol{\beta}^T\boldsymbol{x}_k.$$

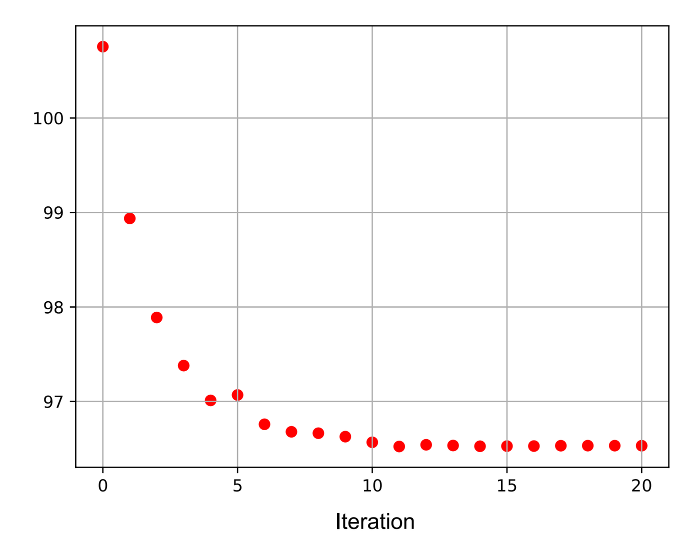
The weight $\boldsymbol{\beta}$ is usually determined by simple linear regression. Although linear regression provides a good solution, we do not necessarily obtain an optimal solution when we consider a feedback of the (unknown) output. We improve the weight $\boldsymbol{\beta}$ by the iterative ensemble variational method.

## Time series analysis with an echo state network

The weight $\boldsymbol{\beta}$ has been determined by using the data for three years from 1998 to 2000. We set the number of state variables of the echo state network to be 600, while 64 ensemble members were used at each iteration.




The time series of the *AL* index (gray) and the prediction by the echo state network with a given solar wind input which was optimized by the iterative method (magenta).




The root-mean-square error for one year (2001) as a function of the iteration.

**References**

S. Nakano: "Behavior of the iterative ensemble-based variational method in nonlinear problems", *Nonlin. Process. Geophys.*, v. 28, pp. 93-109, https://doi.org/10.5194/npg-28-93-2021, 2021.

E. N. Lorenz and K. A. Emanuel: "Optimal sites for supplementary weather observations: Simulations with a small model", *J. Atmos. Sci.*, v. 55, pp, 399-414, https://doi.org/10.1175/1520-0469(1998)055<0399:OSFSWO>2.0.CO;2, 1998.

World Data Center for Geomagnetism, Kyoto, M. Nose, T. Iyemori, M. Sugiura, T. Kamei (2015), Geomagnetic *AE* index, https://doi.org/10.17593/15031-54800, 2015.

**Acknowledgements**

The AL index was provided by the WDC for Geomagnetism, Kyoto (http://wdc.kugi.kyoto-u.ac.jp/wdc/Sec3.html).