Some Issues in Reduced Model Problem

Yicun Zhen and John Harlim

Group Meeting
May 9, 2014

Outline

We assume that the true model is

$$dx = F_1(x, y)dt + \sigma_x dW_x \quad (1)$$
$$dy = \frac{1}{\epsilon} F_2(x, y)dt + \frac{\sigma_y}{\sqrt{\epsilon}} dW_y \quad (2)$$
$$y_n^o = h_1(x_n) + h_2(y_n) + \xi_n \quad (3)$$

where the equations for x and y form the prior model. Equation (3) is the true observational model.

The role of ϵ

- x is the slow process that we can build a model to resolve;
- y is the fast process that we can not model.

Goal:

- build a model for x without explicitly modeling for y;
- build an observational model that only involves x.

$$dx = F(x)dt + \widetilde{\sigma}_x d\widetilde{W}_x \quad (4)$$
$$y^o = h(x) + \widetilde{\xi}_n \quad (5)$$

Question:

How to choose $F, \widetilde{\sigma}_x, h$ and $\widetilde{\xi}_n$, so that the analysis/short forecast of the state x using reduced models is as close as possible to those of the true models.

Some partial results:

$$\begin{aligned} dx &= F_1(x,y)dt + \sigma_x dW_x \\ dy &= \frac{1}{\epsilon}F_2(x,y)dt + \frac{\sigma_y}{\sqrt{\epsilon}}dW_y \\ y_n^o &= h_1(x_n) + h_2(y_n) + \xi_n \end{aligned}$$

$$\begin{aligned} dx &= F(x)dt + \widetilde{\sigma}_x d\widetilde{W}_x \\ y^o &= h(x) + \widetilde{\xi}_n \end{aligned}$$

1, We need to significantly modify $F, \widetilde{\sigma}_x, h$ and $\widetilde{\xi}_n$. Inparticular, h could be very different from h_1, $Var(\widetilde{\xi}_n)$ is no longer the instrumental error variance;
2, We may need to assume some nonzero correlation between the system noise $d\widetilde{W}_x$ and the observational noise $\widetilde{\xi}_n$;
3, When observations are frequently taken, we may need to assume an extremely large $Var(\widetilde{\xi}_n)$.

Linear case

We consider the following system (true model):

$$\begin{aligned} dx &= a_{11}xdt + a_{12}ydt + \sigma_x dW_x && (6)\\ dy &= \frac{a_{21}}{\epsilon}xdt + \frac{a_{22}}{\epsilon}ydt + \sigma_y dW_y && (7)\\ y_n^o &= h_1 x_n + h_2 y_n + \sqrt{R}v_n && (8) \end{aligned}$$

We also asumme a linear reduced model:

$$\begin{aligned} dx &= \alpha xdt + \widetilde{\sigma}_x d\widetilde{W}_x && (9)\\ y_n^o &= hx_n + \sqrt{r}\widetilde{v}_n && (10) \end{aligned}$$

First we convert the prior model to a discrete time system:

$$x_{n+1} = F_{11}x_n + F_{12}y_n + \sqrt{q_x}W_x \quad (11)$$

$$y_{n+1} = F_{21}x_n + F_{22}y_n + \sqrt{q_y}W_y \quad (12)$$

$$y_n^o = h_1x_n + h_2y_n + \sqrt{R}v_n \quad (13)$$

and the reduced model:

$$x_{n+1} = \widetilde{F}x_n + \sqrt{\widetilde{q}}\widetilde{W_x} \quad (14)$$

$$y_n^o = hx_n + \sqrt{r}\widetilde{v}_n \quad (15)$$

Question: How to find those parameters in the reduced model?

Write

$$y_n = l_n x_n + \eta_n, \tag{16}$$

for some constant l_n, such that $\eta_n = y_n - l_n x_n$ and x_n are uncorrelated. $l_n = \frac{Cov(x_n, y_n)}{Var(x_n)}$.
Then

$$\begin{aligned} x_{n+1} &= F_{11}x_n + F_{12}y_n + \sqrt{q_x}W_x \\ &= F_{11}x_n + F_{12}(l_n x_n + \eta_n) + \sqrt{q_x}W_x \\ &= (F_{11} + F_{12}l_n)x_n + (F_{12}\eta_n + \sqrt{q_x}W_x) \end{aligned} \tag{17}$$

This suggests $\widetilde{F} = F_{11} + F_{12}l_n$ and $\widetilde{q}_x = F_{12}^2 Var(\eta_n) + q_x$.

Similarly

$$
\begin{aligned}
y_n^o &= h_1 x_n + h_2 y_x + \sqrt{R} v_n \\
&= h_1 x + h_2 (l_n x_n + \eta_n) + \sqrt{R} v_n \\
&= (h_1 + h_2 l_n) x_n + (h_2 \eta_n + \sqrt{R} v_n) \qquad (18)
\end{aligned}
$$

This suggests $h = h_1 + h_2 l_n$ and $r = h_2^2 Var(\eta_n) + R$.

In sum, we rewrite the dynamics of x and the observational model as :

$$\begin{aligned} x_{n+1} &= (F_{11} + F_{12} I_n) x_n + (F_{12}\eta_n + \sqrt{q_x} W_x) \qquad (19) \\ y_n^o &= (h_1 + h_2 I_n) x_n + (h_2 \eta_n + \sqrt{R} v_n) \qquad (20) \end{aligned}$$

where the "reduced" system noise is $(F_{12}\eta_n + \sqrt{q_x} W_x)$, and the "reduced" observational noise is $(h_2\eta_n + \sqrt{R} v_n)$. They are correlated! The correlation is $F_{12} h_2 Var(\eta_n)$.

The only shortcoming of the decomposition above is that the "reduced" system noise and observational noise are no longer white. This effect is significant when observations are frequently taken.

Numerical results

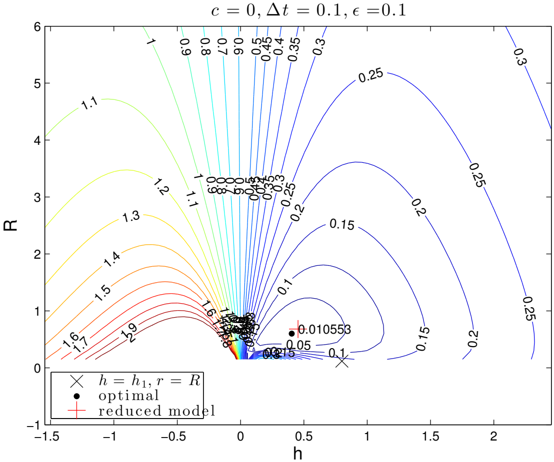
Consider the following system (true model):

$$
\begin{aligned}
dx &= -xdt + ydt + 2dW_x && (21)\\
dy &= -\frac{1}{\epsilon}xdt - \frac{1}{\epsilon}ydt + \frac{2}{\sqrt{\epsilon}}dW_y && (22)\\
y_n^o &= 0.8x_n + 0.5y_n + \sqrt{R}v_n && (23)
\end{aligned}
$$

where $R = 0.1\,Var(x)$. Observations are taken for every $\Delta t = 0.1$ time units. ($x_n = x(0.1n)$, $y_n = y(0.1n)$).

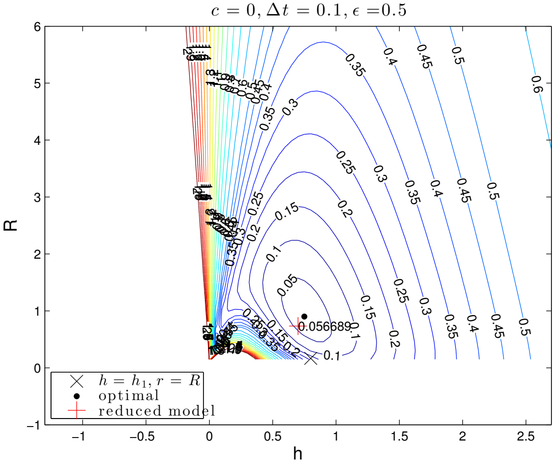
h-r plot

Contour plot of the error in percentage for varying h and r and fixed $\widetilde{F}$, $\widetilde{q}_x$ and $c = 0$.



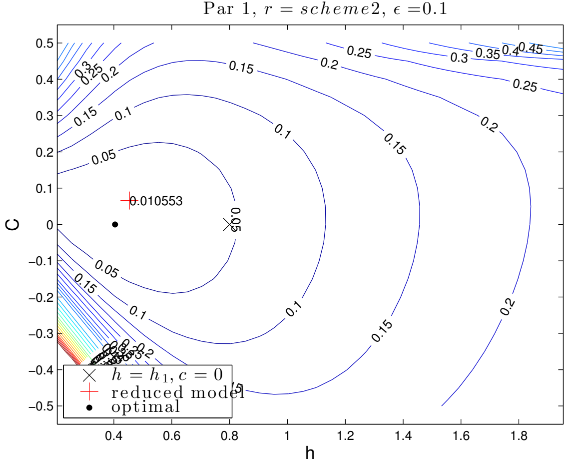
h-r plot

Contour plot of the error in percentage for varying h and r and fixed $\widetilde{F}$, $\widetilde{q}_x$ and $c = 0$.



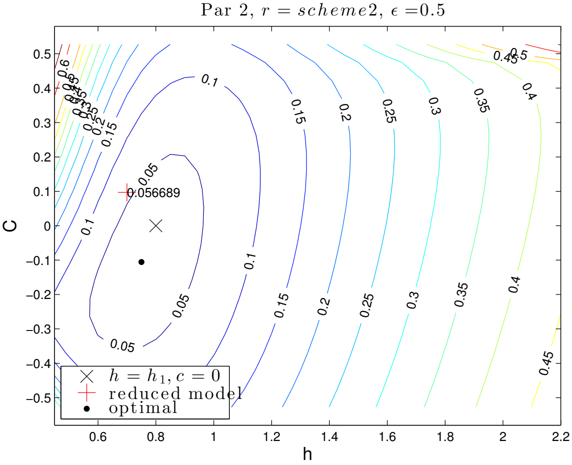
h-c plot

Contour plot of the error in percentage for varying h and c and fixed $\widetilde{F}, \widetilde{q}_x$ and r.



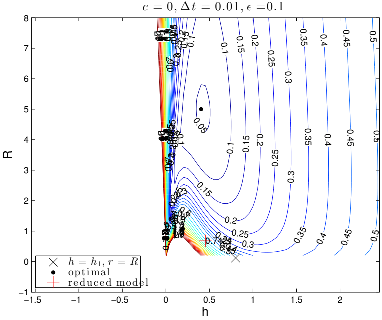
h-c plot

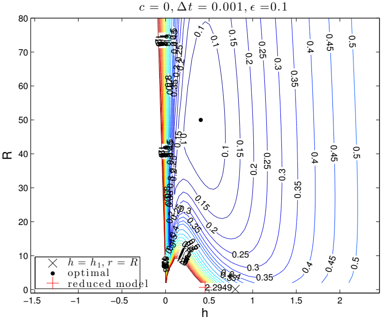
Contour plot of the error in percentage for varying h and c and fixed $\widetilde{F}$, $\widetilde{q}_x$ and r.



h-r plot for $\Delta t = 0.01$ and $\Delta t = 0.001$.

Contour plot of the error in percentage for varying h and r and fixed $\widetilde{F}, \widetilde{q}_x$ and $c = 0$.





Question: why do we need a large r when Δt is small?

Recall the reduced model:

$$y_n^o = (h_1 + h_2 l_n)x_n + (h_2\eta_n + \sqrt{R}v_n) \qquad (24)$$

where $\eta_n = y_n - l_n x_n$.
η_n is not a white noise!
When $\Delta t = 0.01$ or 0.001, η_n and η_{n+1} are highly correlated.

A mathematical justification

Suppose at a fixed time point t, we have two observations $y^o_{n,1}$ and $y^o_{n,2}$ which are drawn from the same observational model. Now assume that they are the same (i.e. they have correlation 1). Two ways to assimilate these two observations:

1, only assimilate one of $y^o_{n,1}$ and $y^o_{n,2}$ and use the instrumental R;
2, assimilate these two observations one by one but use an inflated r.

Lemma

Using $r = 2R$ in the second way we can get the same result by the first way.

Consider the system:

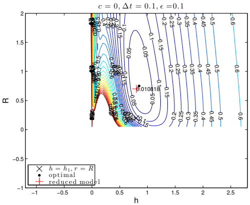
$$\begin{aligned} dx &= -xdt + 2dW_x && (25) \\ dy &= -ydt + 2dW_y && (26) \\ y_n^o &= 0.8x_n + 0.5y_n + \sqrt{R}v_n. && (27) \end{aligned}$$

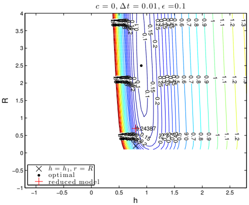
There is a natural way of choosing the reduced prior model:

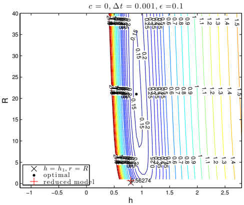
$$dx = -xdt + 2dW_x. \qquad (28)$$

And it is natural to assume for this model that the reduced observational noise and the system noise are uncorrelated. Hence the only parameters left to determine are h and r.

h-r plot for $\Delta t = 0.1, 0.01, 0.001$.







Summary

1, We need to significantly modify $F, \widetilde{\sigma}_x, h$ and $\widetilde{\xi}_n$. Inparticular, h could be very different from h_1, $Var(\widetilde{\xi}_n)$ is no longer the instrumental error variance;
2, We may need to assume some nonzero correlation between the system noise $d\widetilde{W}_x$ and the observational noise $\widetilde{\xi}_n$;
3, When observations are frequently taken, we may need to assume an extremely large $Var(\widetilde{\xi}_n)$.