On the Best Choice of Kalman Gain

Yicun Zhen

Department of Mathematics

Question

- Linear filter : $X_i^a = X_i^f + K(y^o - h(X_i^f))$
- What is the optimal matrix K, so that

$$\sum_{i=1}^{N}(X_i^a - X^{true})(X_i^a - X^{true})^{\mathrm{T}}$$

is the minimal SPD matrix

- No assumption on Gaussian distribution or linearity

Answer

- One choice of K is

$$\tilde{K} = C(R + B_h)^{-1}$$

Where

$$C = \frac{1}{N}(X_1^f - X^{true}, \dots, X_N^f - X^{true})(Y_1^f - y^{true}, \dots, Y_N^f - y^{true})^{\mathrm{T}}$$

$$B_h = \frac{1}{N}(Y_1^f - y^{true}, \dots, Y_N^f - y^{true})(Y_1^f - y^{true}, \dots, Y_N^f - y^{true})^{\mathrm{T}}$$

$$Y_i^f = h(X_i^f) \qquad y^{true} = h(X^{true})$$

Implication on the choice of innovation

$$\bar{X}^a = \bar{X}^f + Kv$$

- There was (might be) a debate on using

$$v = y^o - h(\bar{X}^f)$$

or

$$v = y^o - \bar{h}(X^f)$$

- K~ is consistent with the second choice.

The difference with Kalman filter

- Define

perturbation : $$\Delta X^f = (X_1^f - \bar{X}^f, \ldots, X_N^f - \bar{X}^f)$$

bias : $$\delta X^f = \bar{X}^f - X^{true}$$

- Then

B for Kalman filter : $$B_{KF} = \frac{1}{N-1} \Delta X^f (\Delta X^f)^{\mathrm{T}}$$

B for the K~ : $$\tilde{B} = \frac{N-1}{N} B_{KF} + \delta X^f (\delta X^f)^{\mathrm{T}}$$

The similarity with other covariance adjusting methods

- Dick P. Dee (1995) pointed out that

  $\mathrm{E}(v_k v_k^{\mathrm{T}})$ should approximately equal to $H_k B_k H_k^{\mathrm{T}} + R$

  where H is a linearized version of h.

- For K~, by definition $\mathrm{E}(v_k v_k^{\mathrm{T}}) = B_h + R$
- Many adaptive covariance inflation methods are based on Dick P. Dee's criterion (or something similar).

The difference with other covariance adjusting methods

$$\tilde{K} = \left(\frac{1}{N}\Delta X^f \Delta Y^f + \delta X^f \delta Y^f\right)\left(R + \frac{1}{N}\Delta Y^f \Delta Y^f + \delta Y^f \delta Y^f\right)^{-1}$$

Additional terms

$$K = \left(\frac{1}{N}\Delta X^f \Delta Y^f\right)\left(R + \frac{1}{N}\Delta Y^f \Delta Y^f\right)^{-1}$$

Localization and inflation

Adjusting the denominator

$$\tilde{K} = C(R + B_h)^{-1}$$

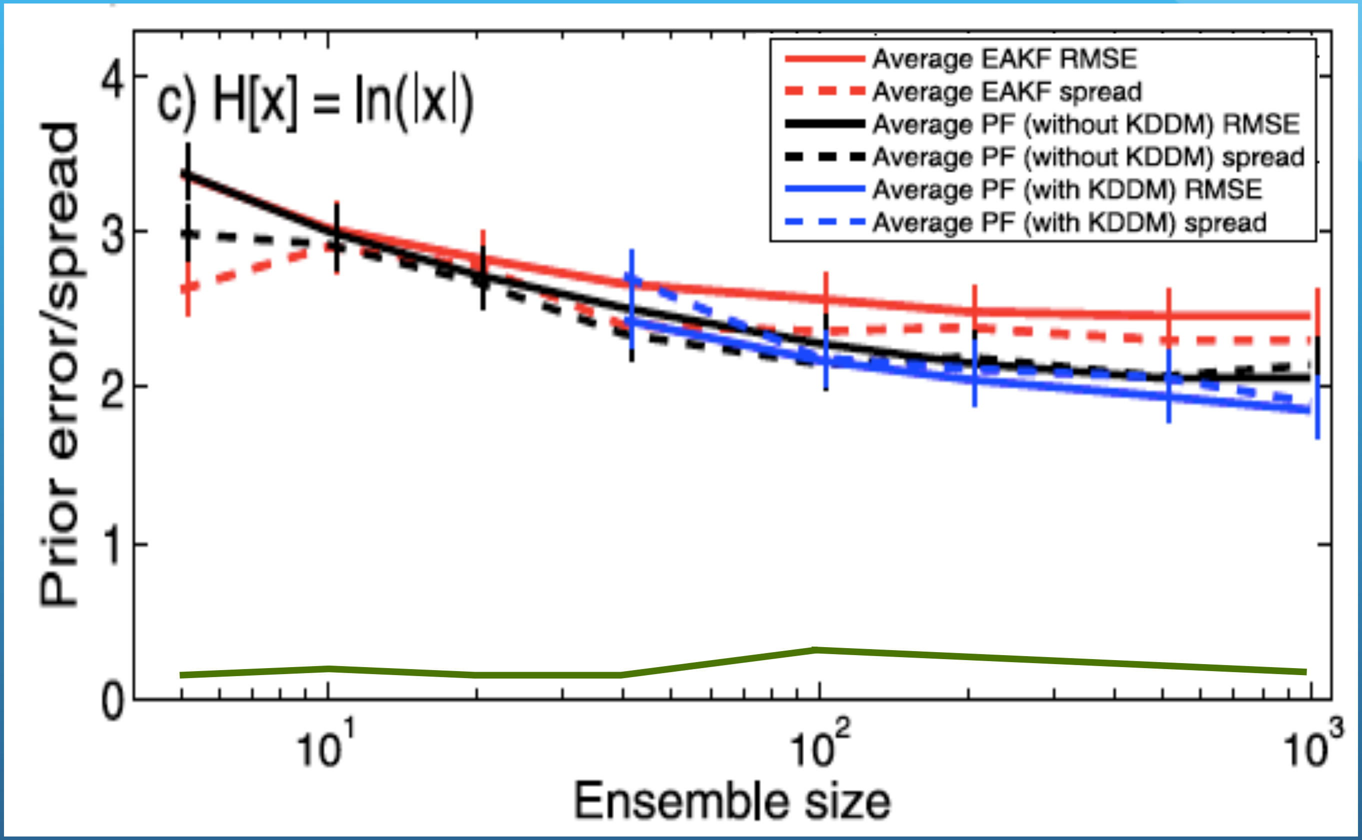
$$\mathrm{E}(VV^{\mathrm{T}}) = R + \frac{1}{N}\Delta Y^f(\Delta Y^f)^{\mathrm{T}} + \delta Y^f(\delta Y^f)^{\mathrm{T}}$$

where

$$V = (y^o - Y_1^f, \ldots, y^o - Y_N^f)$$

- Therefore it is natural to replace B_h+R with VV^T, or (VV^T +R)*(…)
- In AOEI, $\mathrm{E}((y^o - \overline{h}(X^f))^2 - \sigma^2_{h(xb)}) = R + (\delta Y^f)^2 - (\Delta Y^f)^2 / N$

Numerical Results



Conclusions

- Under certain definition, K~ is the best linear regression matrix.
- K~ can not be calculated directly since we do not know the Xtrue.
- It provides the theoretical background to many existing ideas/methods.

Implication

- It suggests that good estimation of the bias, δX^f and δY^f , can lead to a better Kalman gain matrix K.

Future work

- Try to make this idea practical.

$$X_i^a = X_i^f + \tilde{K}(y^o - h(X_i^f))$$

$$\tilde{K} = (\frac{1}{N}\Delta X^f \Delta Y^f + \delta X^f \delta Y^f)(R + \frac{1}{N}\Delta Y^f \Delta Y^f + \delta Y^f \delta Y^f)^{-1}$$