An Introduction to Kalman Filter

Overview

- A conceptual view (scalar problem)
- Kalman Filter Formula
- Limitation of traditional Kalman Filter and imperfect solution
- Homework (if you are interested)

Conceptual View (One-dimension)

Where am I? → x^a

Model

Estimate my speed, I estimate I am at position $x^f = 500$km from my origin

Somewhat uncertain, Expressed with standard deviation $\sigma_1 = 50$ km

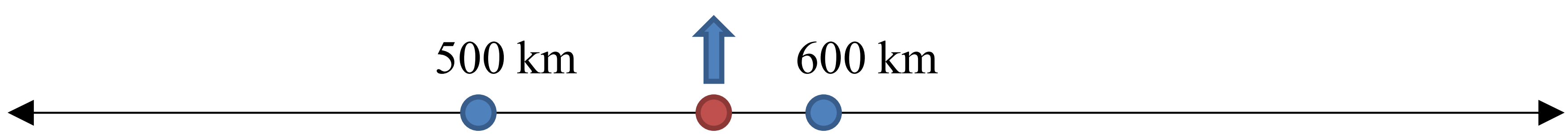
Observation

I also have a very old GPS, it tell me my position at y=600km, with a standard deviation

$\sigma_2 = 20$ km

Combine both pieces of information to get the best estimation of my location.

Somewhere between, but close to GPS location; why?



Mathematical Formulation

$$x^a = ky + (1-k)x^f$$

x^a is our best estimation with uncertainty
k is the unknown coefficients

We want to minimize the uncertainty , i.e. standard deviation σ

Mathematical Formulation

$$\sigma^2 = E[(x^a - x^t)^2] = E[(ky + (1-k)x^f - x^t)^2]$$

$$= E[(k(y - x^t) + (1-k)(x^f - x^t))^2]$$

$$= k^2 E[(y - x^t)^2] + (1-k)^2 E[(x^f - x^t)^2]$$

$$= k^2 \sigma_2^2 + (1-k)^2 \sigma_1^2 = (\sigma_1^2 + \sigma_2^2)k^2 - 2\sigma_1^2 k + \sigma_1^2$$

Here we assume the

$$\mathrm{cov}(\sigma_1, \sigma_2) = 0$$

$$\frac{d\sigma^2}{dk} = 0$$ Give you the minimal σ

$$\frac{d\sigma^2}{dk} = 2(\sigma_1^2 + \sigma_2^2)k - 2\sigma_1^2 = 0$$

$$\Rightarrow k = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

Mathematical Formulation

$$\sigma^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad \text{i.e.} \quad \frac{1}{\sigma^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$$

Uncertainty was reduced !

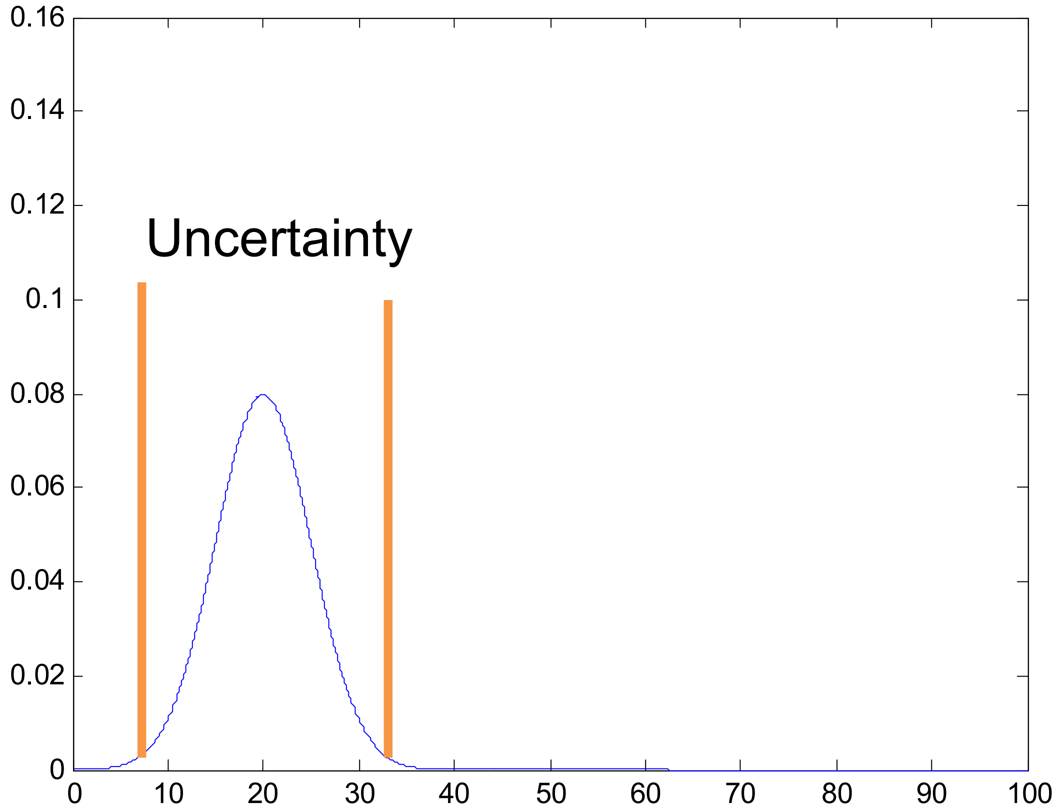
$$x^a = (1-k)x^f + ky = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} x^f + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} y$$

$$x^a = x^f + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}(y - x^f)$$

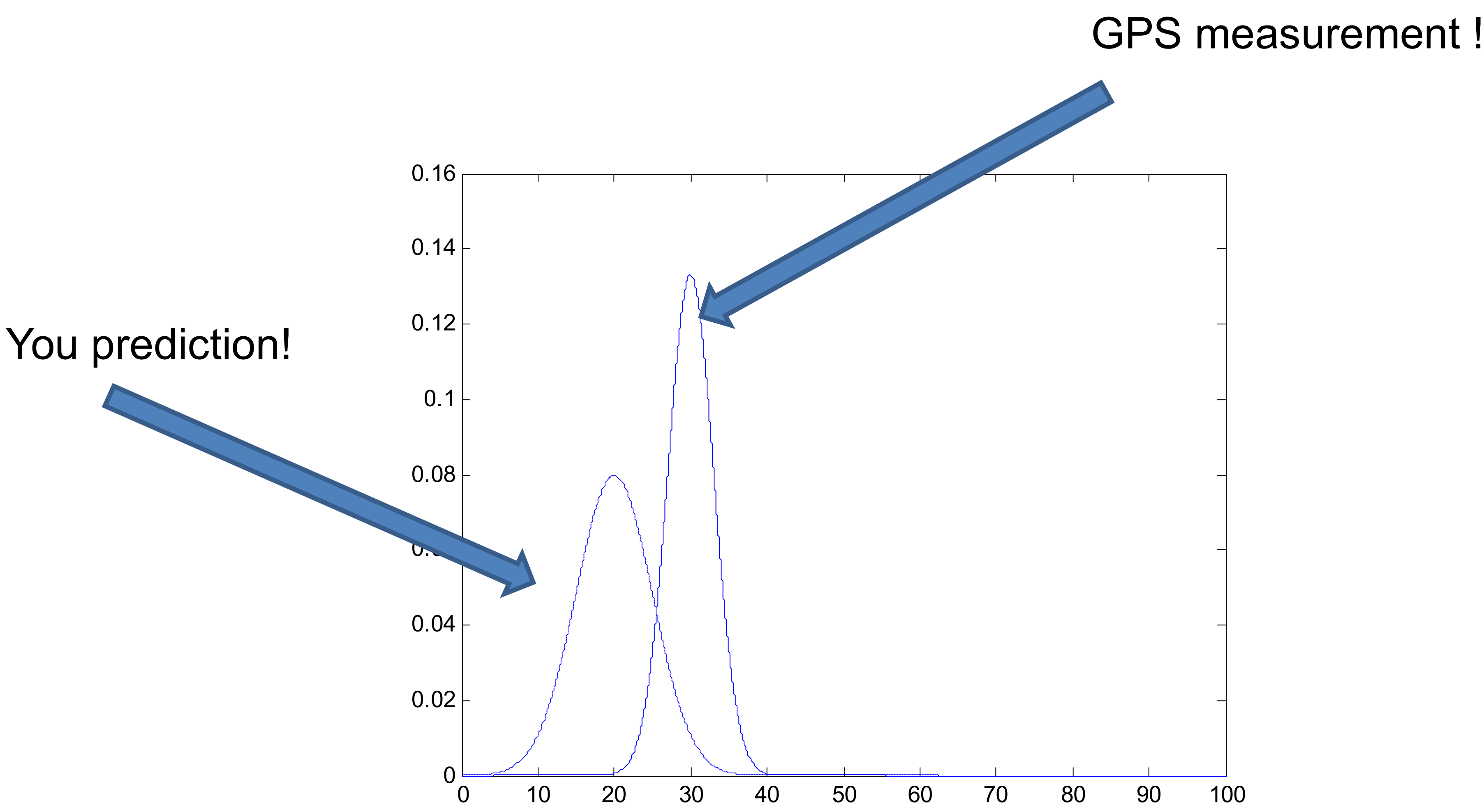
Observation

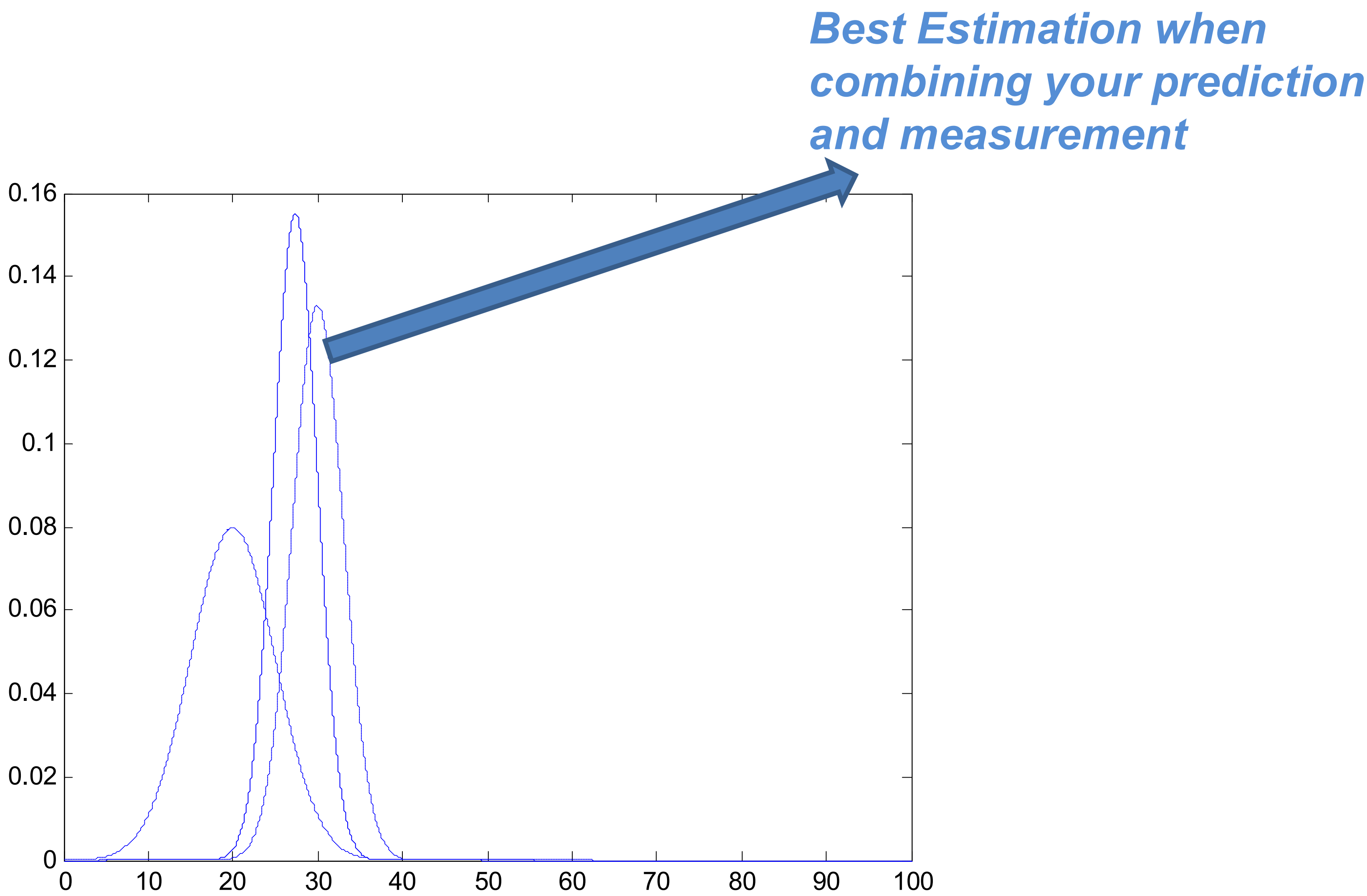
analysis model Gain innovation

Statistical View



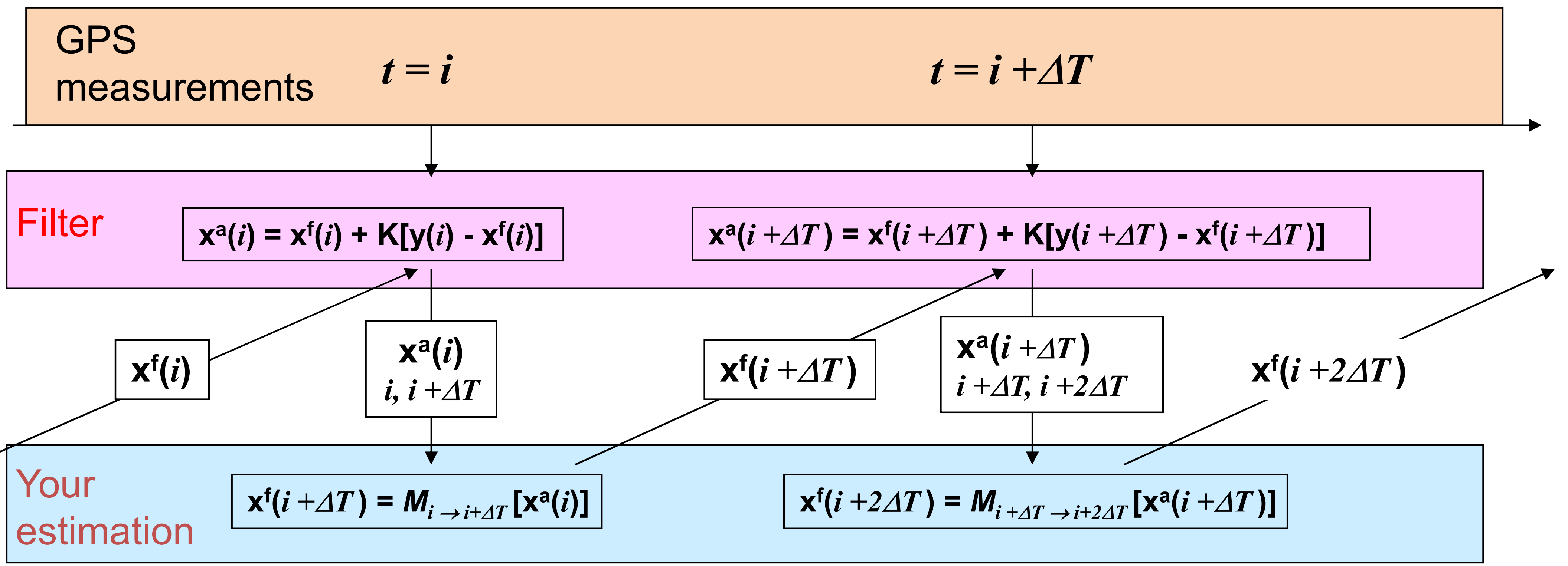
Your prediction!


GPS measurement !
You prediction!
0.16
0.14
0.12
0.1
0.08
0.04
0.02
0
0
10
20
30
40
50
60
70
80
90
100



- Corrected mean is the new optimal estimate of position
- New uncertainty is smaller than either of the previous two variances

Flow chart of the process---doing previously repeatedly



$\mathbf{x}^f$: forecast ;

$\mathbf{x}^a$: improved estimation;

y: measurement;

K: Gain;

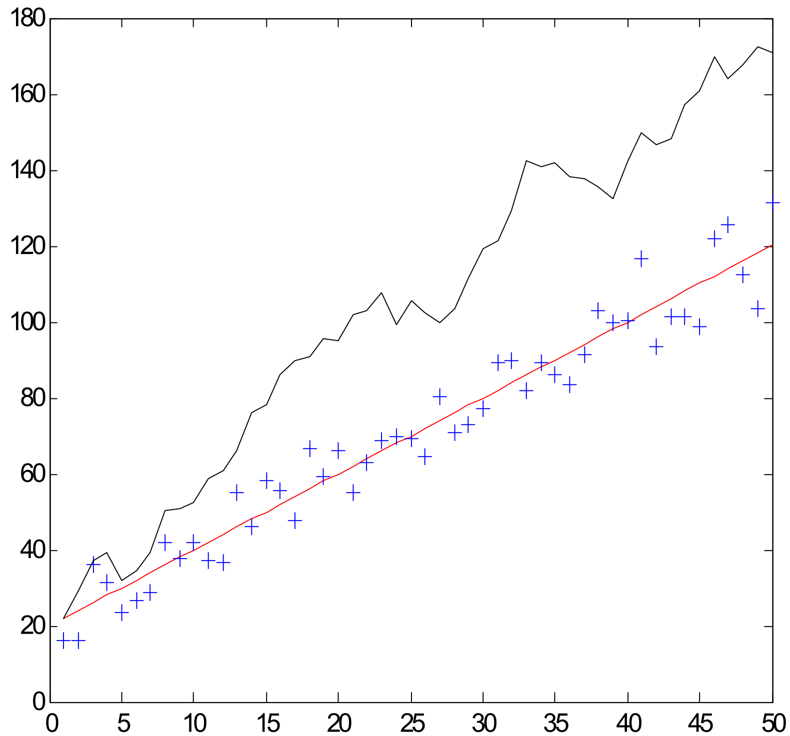
i, $i+\Delta T$: time step

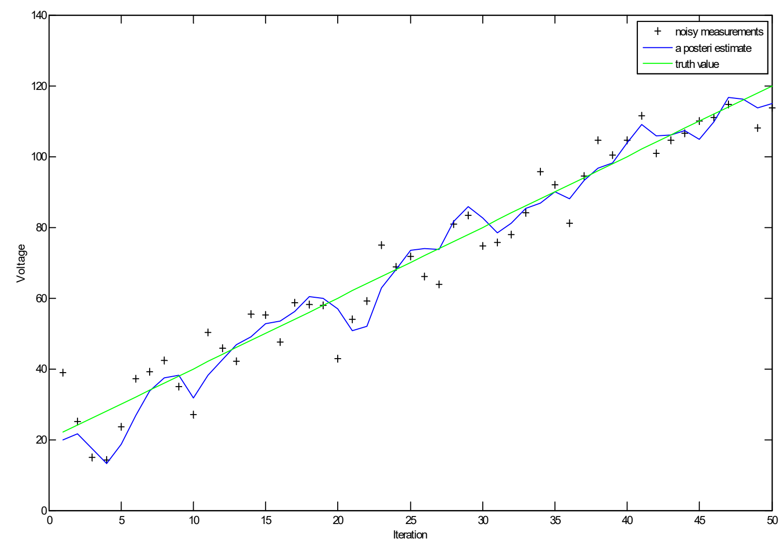
ΔT: assimilation interval

$M_{i\rightarrow i+\Delta T}$: model integration

$$k = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

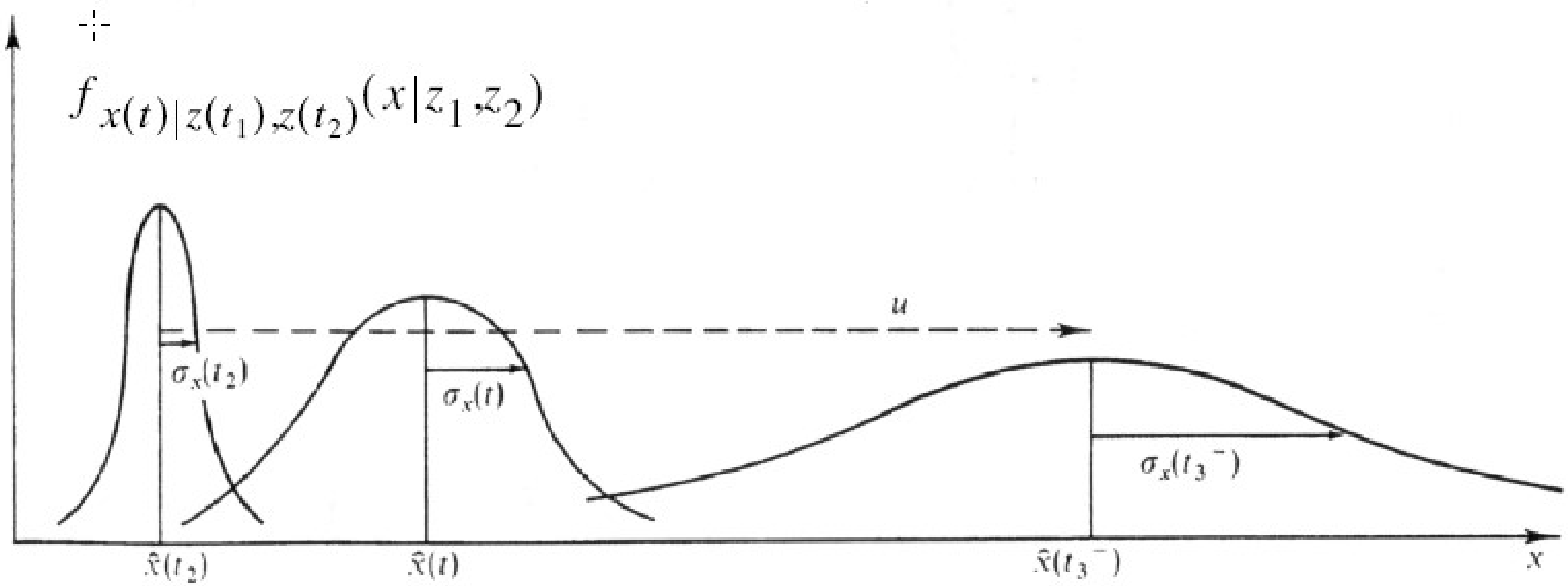
Without and With assimilation





Long-term prediction without assimilation?

Larger and larger error and uncertainty



Maybeck, Peter S., *Stochastic Models, Estimation, and Control*, Vol. 1

In general case: not scalar

Given the linear dynamical system:

$$x_k = M_{k-1} x_{k-1} + B_{k-1} u_{k-1} + v_{k-1}$$

$$y_k = H_k x_k + w_k$$

x_k is the n-dimensional state vector (unknown)

u_k is the m -dimensional input vector (known)

y_k is the p-dimensional output vector (known, measured)

M_k, B_k, H_k are appropriately dimensioned system matrices (known)

v_k, w_k are zero-mean, white Gaussian noise with (known) covariance matrices $Q(k), R(k)$

the Kalman Filter is a recursion that provides the "best" estimate of the state vector *x*.

In general case :not a scalar

- Kalman Filter

Step 1. Model prediction: x^f is estimate based on

$$x_k^f = Mx_{k-1}^a + Bu_k + v_k$$

$$P_k^f = MP_{k-1}^a M^T + Q$$

Noise (v) with covariance Q

this is your estimation of the error propagation.

Step 2. Calculate Kalman Gain:

$$K = \frac{P_k^f H^T}{(HP_k^f H^T + R)}$$

step 3:correction of model state by KF analysis

$$x_k^a = x_k^f + K(y_k - Hx_k^f)$$

$$P_k^a = (I - KH)P_k^f$$

this is your new estimation of the error covariance, reduced from P_k^f to P_k^a

In general case: not a scalar

x^f: forecast [N×1] ; a state vector (a scalar before : your prediction of your location 1-D)
x^a: (analysis) [N×1] ; a state vector (a scalar before: 1-D location)
y: observation [N_o×1] ; a observationa vector (a scalar before: gps measurement location 1-D)
K: Kalman gain [N×N_o] ; a matrix (a scalar before : $k = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$)

$$x^a = x^f + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}(y - x^f)$$

Observation

analysis model Gain innovation

$$x^a = x^f + \frac{P^f H^T}{HP^f H^T + R}(y - Hx^f)$$

H is the observation operator , interpolate the x^f to y ,this is because the observation size are usually smaller than your model state vector

Scalar case $$x^a = x^f + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}(y - x^f)$$ $$K = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

General case $$x^a = x^f + \frac{P^f H^T}{HP^f H^T + R}(y - Hx^f)$$

$$K = \frac{P^f H^T}{HP^f H^T + R}$$

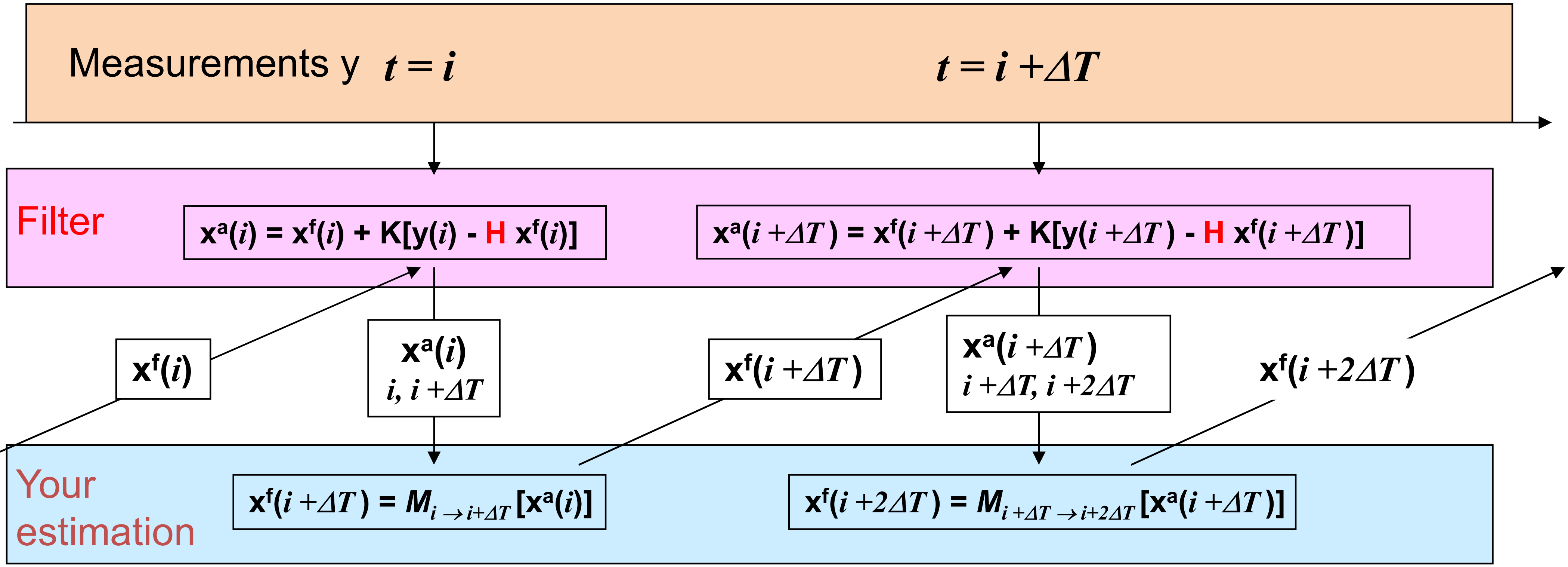
- If we are sure about measurements:
  - Measurement error covariance (R) decreases to zero
  - K decreases and weights residual more heavily than prediction

$$\lim_{R_k \to 0} K_k = H^{-1}.$$

- If we are sure about prediction
  - Prediction error covariance P^f decreases to zero
  - K increases and weights prediction more heavily than residual

$$\lim_{P_k^- \to 0} K_k = 0.$$

Flow chart of the process---doing previously repeatedly



$\mathbf{x}^f$: forecast $[N\times1]$; model prediction

$\mathbf{X}^a$: KF analysis condition $[N\times1]$;

y: observation $[N_o\times1]$; from field measurements

K: Kalman gain

ΔT: assimilation interval

$M_{i \to i+\Delta T}$: model integration from time i to i+ ΔT

Summary

- Recursive data processing algorithm
- Generates optimal estimate of desired quantities given the set of measurements
- Optimal?
  - For linear system and white Gaussian errors, Kalman filter is “best” estimate based on all previous measurements
  - For non-linear system optimality is ‘qualified’
- Recursive?
  - Doesn’t need to store all previous measurements and reprocess all data each time step

Limitation

- Weak nonlinear system (Extended Kalman Filter)
- Computation loads

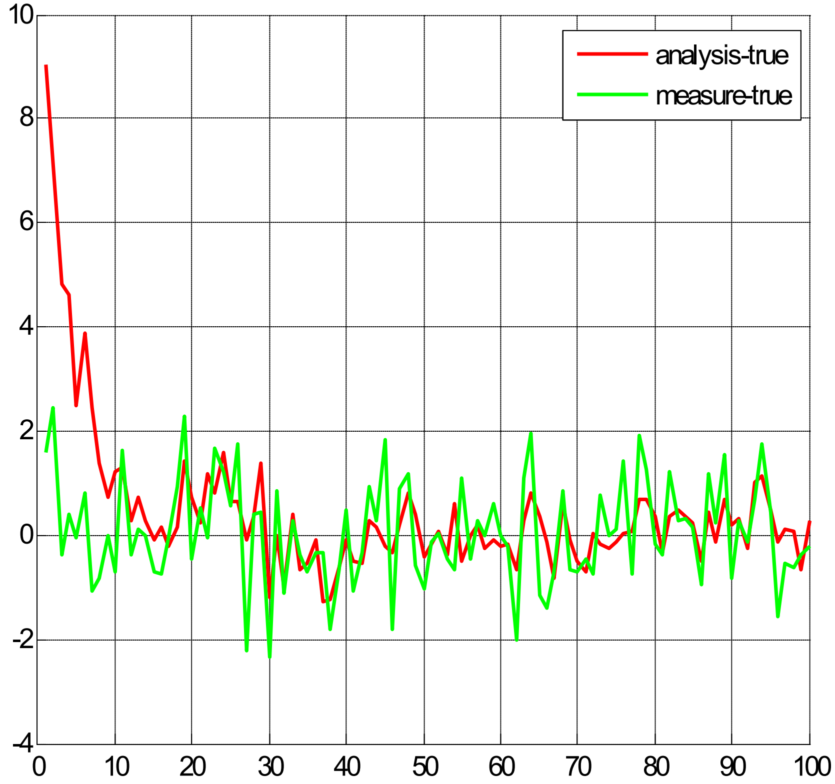
$$K = \frac{P^f H^T}{HP^f H^T + R}$$

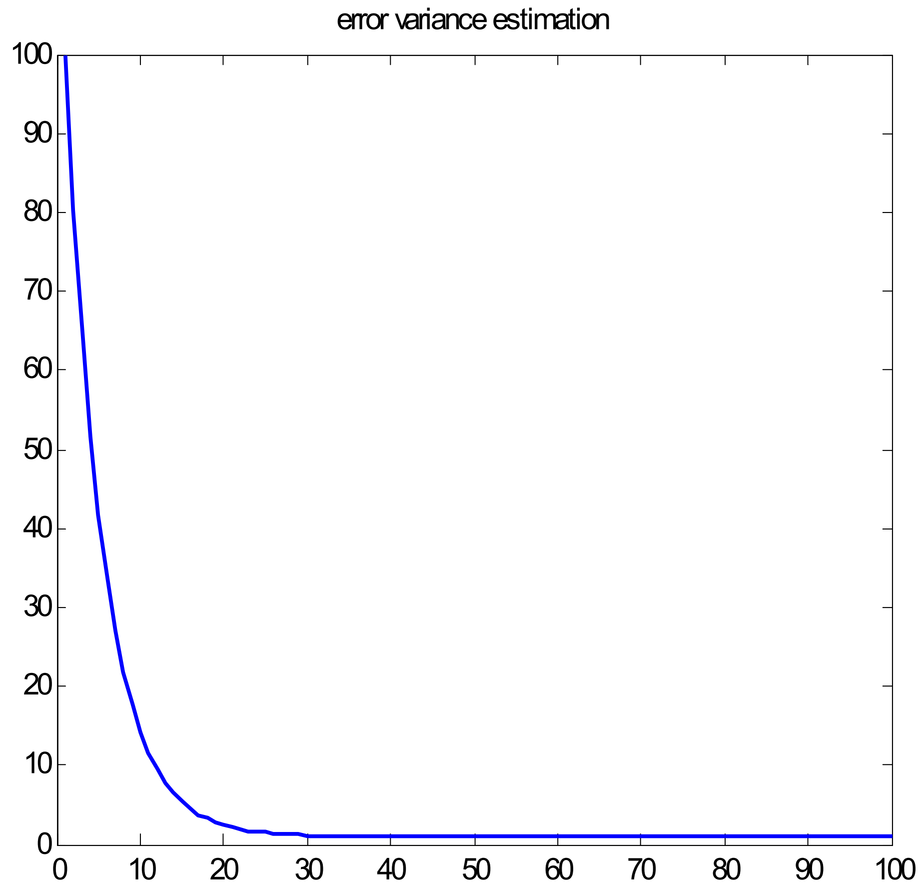
  - P=O(1e6) x O(1e6) matrix
  - Reduced Rank Kalman Filter (project to leading error subspace O(1e2) from EOF analysis and doing KF in model error subspace then project back)
  - Ensemble Kalman Filter (Represents error statistics P^f using an ensemble of model states.)
- (see Chen etc. 2009 for coastal ocean idealized case)

Homework (if you are interested)

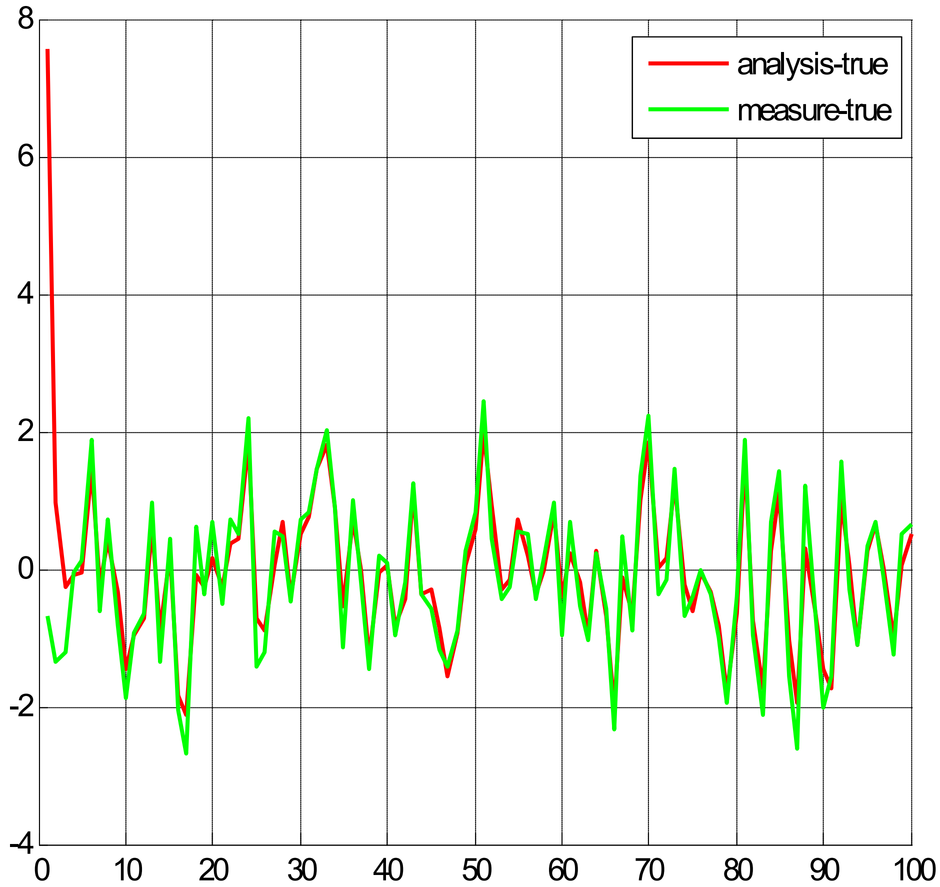
- Consider you are in a room: your estimation is the temperature is constant (you can have your first guess with any temperature)
- You have a thermometer, with a known variance (uncertainty) $\sigma_1^2 = 1$
- We know the true room temperature is 10 °C with some perturbation, variance $\sigma_2^2 = 0.25$
- Using matlab/others to construct a KF model, show your model states analysis and error variance convergence in KF
- *Assuming all error distribution is Gaussian*

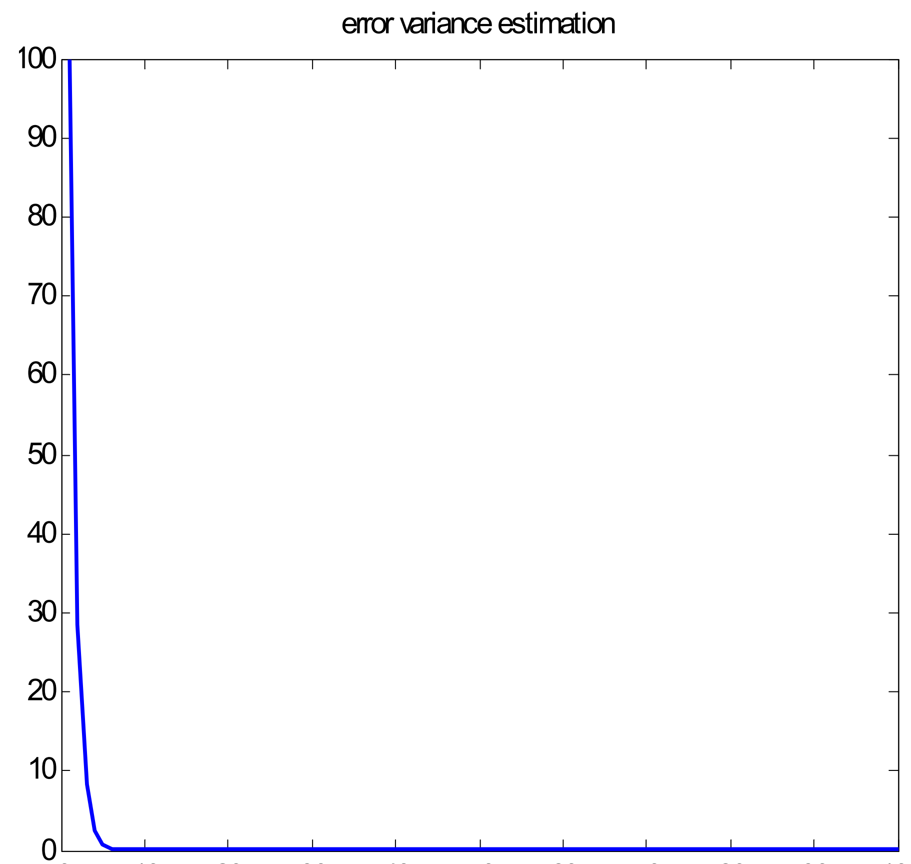
Analysis state is improved, and error variance converged



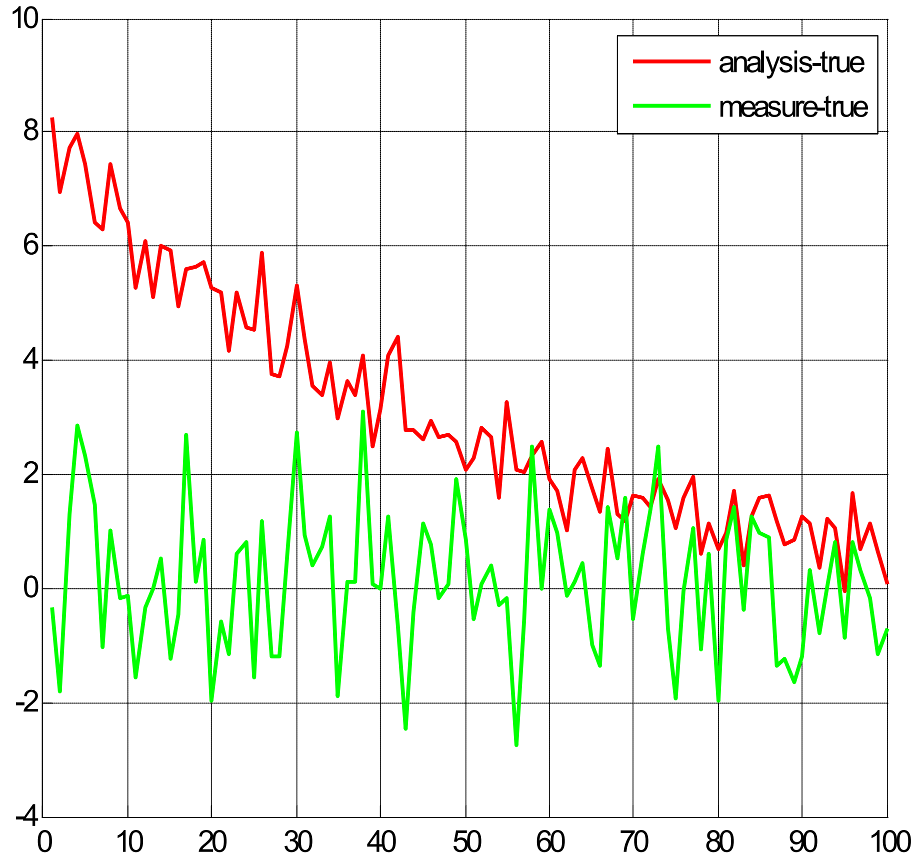


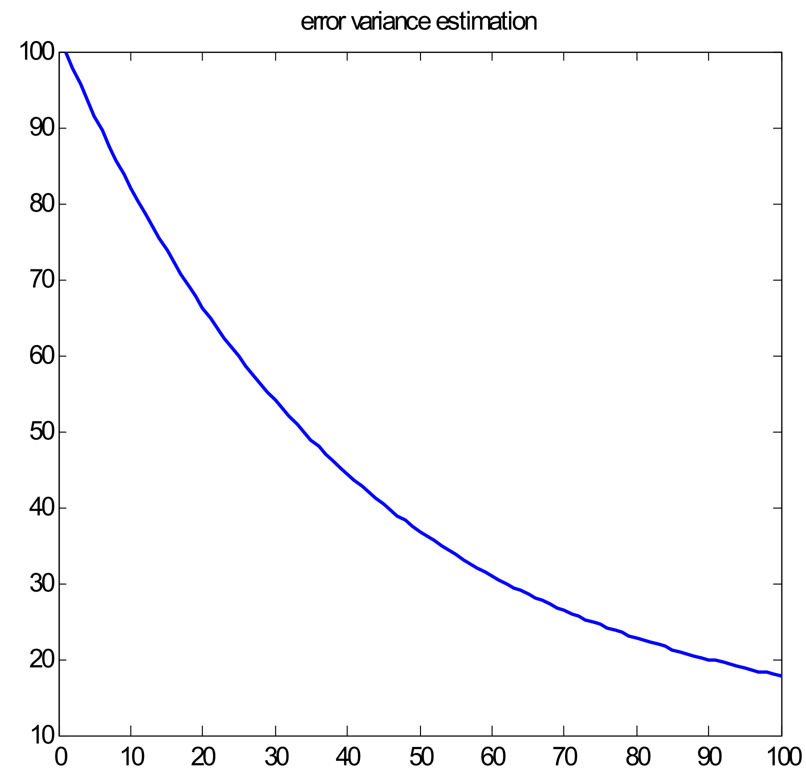
Overestimate the model error variance by a factor of 10





Underestimate the model error variance by a factor of 10





References

Maybeck, Peter S., *Stochastic Models, Estimation, and Control*, Vol. 1

- Greg Welch and Gary Bishop: An Introduction to the Kalman Filter
- Buehner, M., and P. Malanotte-Rizzoli, Reduced-rank Kalman filters applied to an idealized model of the wind-driven ocean circulation, JGR
- Evensen 2003, Ocean Dynamics, Vol 53, No 4 The Ensemble Kalman Filter: Theoretical formulation and practical implementation