A Simple Introduction to Kalman Filter

Duy-Ky Nguyen, PhD 2002-Nov-22

1. Random Variable

1.1. Probability Axioms

Given an event E in a sample space S which is either finite with N elements or countaby infinite with $N = \infty$ elements, then we can write

$$S \equiv \left(\bigcup_{i=1}^{N} E_i\right)$$

and a quality $P(E_i)$, called the *probability* of event E_i, is defined such that

1. $0 \le P(E_i) \le 1$,
2. $\sum_{i=1}^{N} P(E_i) = 1$
3. Additivity $P(E_1 \cup E_2) = P(E_1) + P(E_2)$, where E_1 and E_2 are mutually exclusive, *ie* $E_1 \bigcap E_2 = \varnothing$

1.2. Conditional Probability

The conditional probability of an event A assuming that B has occurred, defined as

$$P(A \cap B) = P(A|B)P(B)$$

For *independent* events, we have

$$P(A|B) = P(A)$$

so

$$P(A \cap B) = P(A)\,P(B)$$

1.3. Expectation

Expectation, also known as *average* or *mean*, is defined as

$$\overline{X} = E(x) = \mu = \langle x \rangle = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Recursive expectation

$$\mu_{n+1} = \frac{1}{n+1}\sum_{i=1}^{n+1} x_i = \frac{1}{n+1}\left(x_{n+1} + \sum_{i=1}^{n} x_i\right) = \frac{x_{n+1} + n\mu_n}{n+1}$$

1.4. Variance and Deviation

Variance is a measure of variation around the mean and defined as

$$\sigma^2 = E(x-\mu)^2 = \left\langle (x-\mu)^2 \right\rangle = \mathrm{var}(x) = E[x - E(x)]^2 = \left\langle (x - \langle x \rangle)^2 \right\rangle$$

and *Deviation* is the square root of *Variance*, *ie.* σ

1.5. Covariance

Covariance provides a measure of how strong relation between 2 variables, and defined as

$$\operatorname{cov}(x_i, x_j) = \langle (x_i - \mu_i)(x_j - \mu_j) \rangle = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle$$

as

$$\langle (x_i - \mu_i)(x_j - \mu_j) \rangle = \langle x_i x_j + \mu_i \mu_j - \mu_i x_j - \mu_j x_i \rangle = \langle x_i x_j \rangle + \mu_i \mu_j - \mu_i \langle x_j \rangle - \mu_j \langle x_i \rangle$$

For independent variables

$$\langle x_i x_j \rangle = \langle x_i \rangle \langle x_j \rangle$$

so covariance of independent variables equals to zero, as expected.

Covariance of the same variable is its variance

$$\operatorname{cov}(x_i, x_i) = \langle (x - \mu_i)(x - \mu_i) \rangle = \langle (x - \mu_i)^2 \rangle = \operatorname{var}(x)$$

For 2 variables, the covariance is related to the variance by

$$\operatorname{var}(x + y) = \langle (x - \mu_x + y - \mu_y)^2 \rangle = \langle (x - \mu_x)^2 + (y - \mu_y)^2 + 2(x - \mu_x)(y - \mu_y) \rangle = \operatorname{var}(x) + \operatorname{var}(y) + 2\operatorname{cov}(x, y)$$

For 2 independent variable

$$\operatorname{var}(x + y) = \operatorname{var}(x) + \operatorname{var}(y)$$

2. Derivation of Kalman Gain

Discrete time linear systems with noise are presented in the state equation below

$$x_j = a\,x_{j-1} + b\,u_j + w_j \tag{1}$$

and its measurable output, also with noise

$$y_j = h\,x_j + v_j \tag{2}$$

Because noises are unknown, so we need to define 2 new noise-free states

- $\bar{x}_j$: priori state as predictor
- $\tilde{x}_j$: posteriori state as corrector

we then have predicted estimate system equation (without noises)

$$\bar{x}_j = a\,\tilde{x}_{j-1} + b\,u_j \tag{3}$$

$$\tilde{y}_j = h\,\bar{x}_j \tag{4}$$

Based on priori estimate, we have a corrected estimate below

$$\tilde{x}_j = \bar{x}_j + K\left(y_j - \tilde{y}_j\right) \tag{5}$$

where

$(y_j - \tilde{y}_j)$: known as **residual**

K is Kalman gain to be determined to minimize error variance below

$$\tilde{e}_j = x_j - \tilde{x}_j = x_j - \left[\bar{x}_j + K(y_j - \tilde{y}_j)\right] = x_j - \bar{x}_j - K(hx_j + v_j - h\bar{x}_j) = (1 - hK)(x_j - \bar{x}_j) - Kv_j$$

$$\tilde{e}_j = x_j - \tilde{x}_j = (1 - hK)\bar{e}_j - Kv_j \tag{6}$$

where

$$\bar{e}_j = x_j - \bar{x}_j \tag{7}$$

so its variance is

$$\tilde{p}_j = E\left(\tilde{e}_j^2\right) = E\left[(1-hK)\bar{e}_j - Kv_j\right]^2 = E\left[(1-hK)^2\bar{e}_j^2 + K^2v_j^2 - 2\bar{e}_jKv_j(1-hK)\right]$$

$$\tilde{p}_j = E\left(\tilde{e}_j^2\right) = (1-hK)^2\,\bar{p}_j + K^2R \qquad (8)$$

where

$E\left(\bar{e}_jv_j\right) = E\left(\bar{e}_j\right)E\left(v_j\right) = 0$: as v_j uncorrelated to x_j and $E\left(v_j\right) = 0$

$\bar{p}_j = E\left(\bar{e}_j^2\right)$: variance of priori error $\bar{e}_j$, known as **priori covariance**

$R = E\left(v_j^2\right)$: variance of noise at measured output

> Remark 1
> By Kalman, and for reason of mutual interaction, we have used priori and posteriori in the same equation, *e.g.* $\bar{x}_j$ and $\tilde{x}_{j-1}$ in Eq.(3), $\tilde{y}_j$ and $\bar{x}_j$ in Eq.(4), $\tilde{x}_j$ and $\bar{x}_j$ in Eq.(5).

Eq.(8) gives

$$\frac{\partial p_j}{\partial K} = 2(hK-1)h\bar{p}_j + 2RK = 2\left[\left(h^2\bar{p}_j + R\right)K - h\bar{p}_j\right]$$

equating to 0 to minimize the estimate error, and the optimal Kalman gain K is given below

$$K_j = \frac{h\bar{p}_j}{h^2\bar{p}_j + R} \qquad (9)$$

Note we change K to K_j as it's varied with j. We will use Eq.(9) to eliminate R in Eq.(8). So Eq.(9) is rewritten as

$$R = \frac{h\bar{p}_j}{K_j} - h^2\bar{p}_j = \frac{h\bar{p}_j}{K_j}\left(1-hK_j\right) \qquad (10)$$

Substituting into Eq.(8) after changing K to K_j

$$\tilde{p}_j = E\left(\tilde{e}_j^2\right) = \left(1-hK_j\right)^2\bar{p}_j + K_jh\bar{p}_j\left(1-hK_j\right) = \bar{p}_j\left(1-hK_j\right)\left(1-hK_j+hK_j\right)$$

$$\tilde{p}_j = E\left(\tilde{e}_j^2\right) = \bar{p}_j\left(1-hK_j\right) \qquad (11)$$

However, it's pointed out by Peter Joseph that Eq.(8) is numerically stable, while its simplified Eq.(11) is not due to round-off computation

We next to compute priori covariance

3. Derivation of Priori Covariance

By Eqs.(1) & (3), the priori covariance is given by

$$\bar{p}_j = E\left(\bar{e}_j^2\right) = E\left(x_j - \bar{x}_j\right)^2 = E\left[\left(a\,x_{j-1} + bu_j + w_j\right) - \left(a\,\tilde{x}_{j-1} + bu_j\right)\right]^2 = E\left(a\tilde{e}_{j-1} + w_j\right)^2 = E\left(a^2\tilde{e}_{j-1}^2 + w_j^2 + 2a\tilde{e}_{j-1}w_j\right)$$

$$\bar{p}_j = a^2\tilde{p}_{j-1} + Q \qquad (12)$$

where

$E\left(\tilde{e}_{j-1}w_j\right) = E\left(\tilde{e}_{j-1}\right)E\left(w_j\right) = 0$: as w_j uncorrelated to x_j, and $E\left(w_j\right) = 0$

$\tilde{p}_{j-1} = E\left(\tilde{e}_{j-1}^2\right)$: variance of error $\bar{e}_{j-1}$,

$Q = E\left(w_j^2\right)$: variance of noise at input

4. Scalar Kalman Filter Algorithm

System equations

$$\begin{cases} x_j = a\,x_{j-1} + b u_j + w_j, & E\left(w_j^2\right) = Q \\ y_j = h\,x_j + v_j, & E\left(v_j^2\right) = R \end{cases}$$

Init State

$$\tilde{x}_0,\ \tilde{p}_0$$

Predictor updates (Priori)

$$\bar{p}_j = a^2 \tilde{p}_{j-1} + Q$$
$$\bar{x}_j = a\,\tilde{x}_{j-1} + b u_j$$

Corrector updates (Poteriori)

$$K_j = \frac{h\bar{p}_j}{h^2 \bar{p}_j + R}$$
$$\tilde{p}_j = E\left(\tilde{e}_j^2\right) = \underline{\underline{\left(1 - hK_j\right)^2 \bar{p}_j + K_j^2 R}} = \bar{p}_j\left(1 - hK_j\right)$$
$$\tilde{x}_j = \bar{x}_j + K_j\left(y_j - h\,\bar{x}_j\right)$$

5. Vector Kalman Filter Algorithm

System equations

$$\begin{cases} \mathbf{x}_j = \mathbf{A}\,\mathbf{x}_{j-1} + \mathbf{B}\mathbf{u}_j + \mathbf{w}_j, & E\left(\mathbf{w}_j \mathbf{w}_j^T\right) = \mathbf{Q} \in \Re^{n\times n} \\ \mathbf{y}_j = \mathbf{H}\,\mathbf{x}_j + \mathbf{v}_j, & E\left(\mathbf{v}_j \mathbf{v}_j^T\right) = \mathbf{R} \in \Re^{m\times m} \end{cases}$$

Init State

$$\tilde{\mathbf{x}}_0,\ \tilde{\mathbf{P}}_0$$

Predictor updates (Priori)

$$\bar{\mathbf{P}}_j = \mathbf{A}\,\tilde{\mathbf{P}}_{j-1}\mathbf{A}^T + \mathbf{Q}$$
$$\bar{\mathbf{x}}_j = \mathbf{A}\,\tilde{\mathbf{x}}_{j-1} + \mathbf{B}\mathbf{u}_j$$

Corrector updates (Posteriori)

$$\mathbf{K}_j = \bar{\mathbf{P}}_j \mathbf{H}^T\left(\mathbf{H}\bar{\mathbf{P}}_j\mathbf{H}^T + \mathbf{R}\right)^{-1}$$
$$\tilde{\mathbf{P}}_j = E\left(\tilde{\mathbf{e}}_j \tilde{\mathbf{e}}_j^T\right) = \underline{\underline{\left(\mathbf{I} - \mathbf{K}_j\mathbf{H}\right)\bar{\mathbf{P}}_j\left(\mathbf{I} - \mathbf{K}_j\mathbf{H}\right)^T + \mathbf{K}_j\mathbf{R}\mathbf{K}_j^T}} = \left(\mathbf{I} - \mathbf{K}_j\mathbf{H}\right)\bar{\mathbf{P}}_j$$
$$\tilde{\mathbf{x}}_j = \bar{\mathbf{x}}_j + \mathbf{K}_j\left(\mathbf{y}_j - \mathbf{H}\,\bar{\mathbf{x}}_j\right)$$

where

$$\mathbf{x}_j, \mathbf{w}_j \in \Re^{n\times 1};\quad \mathbf{y}_j, \mathbf{v}_j \in \Re^{m\times 1};\quad \mathbf{A}, \mathbf{Q}, \mathbf{R}, \mathbf{P}_j \in \Re^{n\times n};\quad \mathbf{B}, \mathbf{u}_j, \mathbf{K}_j \in \Re^{n\times m};\quad \mathbf{H} \in \Re^{m\times n}$$

6. Kalman Filter Examples

For this paper to be self-contained, I'll use my approximate discretization, An accurate discretization can be found in Ref[1].

6.1. Scalar Kalman filter

Estimate a scalar constant x¸ a voltage for example. Let's assume that we have the ability to take measurements of the constant, but the measurements are corrupted by a 0.1 volt RMS white measurement noise, thus $R = 0.1^2 = 0.01$

The scalar euqations describing this situation are

$$x_{j+1} = x_j + w_j$$

for the system and

$$y_j = x_j + v_j$$

for the measurement, where

$$E(v_j) = 0, \quad E(v_j^2) = R = 0.01$$

Prersuming a very small process variance, we let $E(w_j^2) = Q = 10^{-5}$, thus we have a computational algorithm below

- System equations

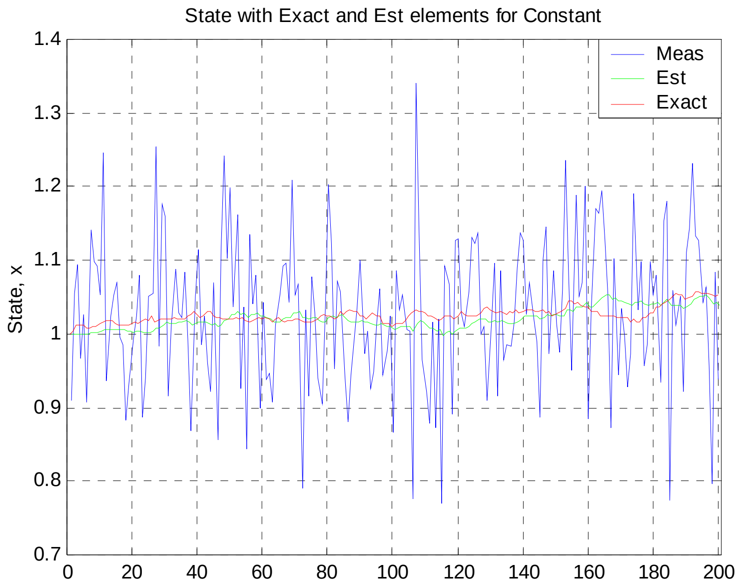
$$x_{j+1} = x_j + w_j$$
$$y_j = x_j + v_j$$

- Predictor equations (Priori)

$$\bar{p}_j = \tilde{p}_{j-1} + Q, \quad Q = 10^{-5}$$
$$\bar{x}_j = \tilde{x}_{j-1}$$

- Corrector equations (Posteriori)

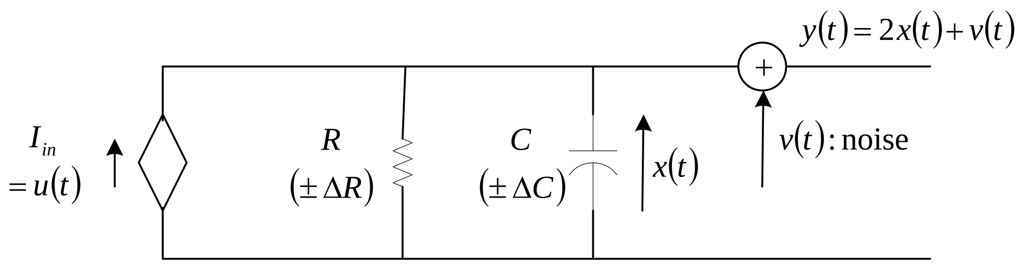
$$K_j = \frac{\bar{p}_j}{\bar{p}_j + R}, \quad R = 0.01$$
$$\tilde{p}_j = \underline{\underline{(1 - K_j)^2 \bar{p}_j + K_j^2 R}} = \bar{p}_j (1 - K_j)$$
$$\tilde{x}_j = \bar{x}_j + K_j (y_j - \bar{x}_j)$$



Remark 2
Note the green-colored estimate tracks the red-colored exact signal quite well even with the blue-colored noisy measurement.

6.2. Scalar Kalman filter formulation for RC circuit

We consider the voltage measurement at the output of the RC circuit in the figure below, using a high-impedance voltmeter. Because these measurements are noisy, and also the component values imprecise ($\pm\Delta R, \pm\Delta C$), we require an improved estimate of the output voltage. For this purpose we want to use a Kalman filter for which we develop the system and measurement models as follows.



The Kirchoff nodal equation is

$$u(t) = \frac{x(t)}{R} + C\frac{dx}{dt}$$

By definition of derivative, with sampling T¸we have

$$u_j = \frac{x_j}{R} + C\frac{x_{j+1} - x_j}{T}$$

or

$$x_{j+1} = \left(1 - \frac{T}{RC}\right)x_j + \frac{T}{C}u_j$$

Assuming the measurement to have gain of 2, the circuit elements to have values $R = 3.3k\Omega, C = 1000\mu F$ and sampling period $T = 0.1s$, the input to be step function of $300\mu A$, our signal and measurement equations are

$$x_{j+1} = 0.97x_k + 0.03 + w_j$$
$$y_j = x_j + v_j$$

Assuming the model parameter uncertainty $E(w_j^2) = Q = 10^{-4}$, and the measurement error $E(v_j^2) = R = 0.01$, thus we have a computational algorithm below

- System equations

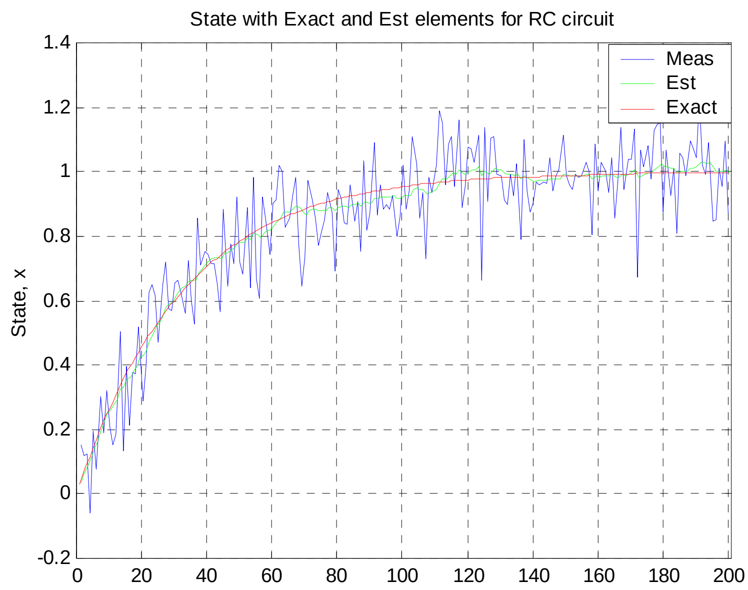
$$x_{j+1} = 0.97x_k + 0.03 + w_j$$
$$y_j = x_j + v_j$$

- Predictor equations (Priori)

$$\overline{p}_j = a^2 \tilde{p}_{j-1} + Q, \quad a = 0.97, \; Q = 0.01$$
$$\overline{x}_j = a\,\tilde{x}_{j-1} + 0.03$$

- Corrector equations (Posteriori)

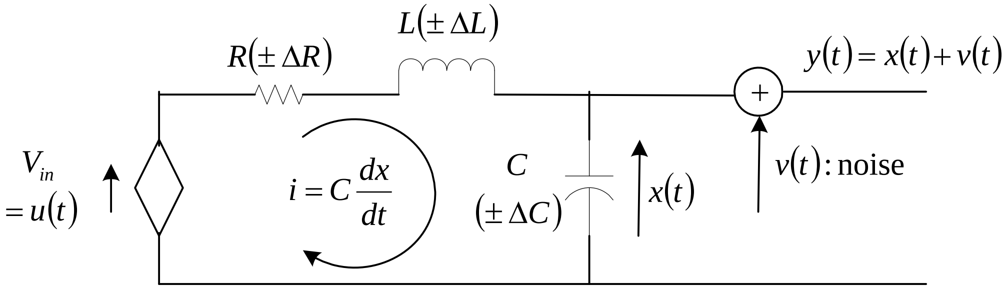
$$K_j = \frac{h\overline{p}_j}{h^2\overline{p}_j + R}, \quad h = 1, \quad R = 0.01$$
$$\tilde{p}_j = (1 - hK_j)^2 \overline{p}_j + K_j^2 R = \overline{p}_j (1 - hK_j)$$
$$\tilde{x}_j = \overline{x}_j + K(y_j - h\overline{x}_j)$$



Remark 3
Note the green-colored estimate tracks the red-colored exact signal quite well even with the blue-colored noisy measurement.

6.3. Vector Kalman filter formulation for RLC circuit

We consider a design of an estimator for a second-order system consisting of *R*, *L*, *C* elements



The loop equation for this circuit is

$$u(t) = Ri + L\frac{di}{dt} + x(t), \quad i = C\frac{dx}{dt}$$

so

$$u(t) = RC\frac{dx}{dt} + LC\frac{d^2x}{dt^2} + x(t)$$

or

$$\frac{d^2x}{dt^2} + \frac{R}{L}\frac{dx}{dt} + \frac{1}{LC}x(t) = \frac{1}{LC}u(t)$$

Assuming $R = 5k\Omega$, $L = 2.5H$, and $C = 0.1\mu F$, we have

$$\frac{d^2x}{dt^2} + 2\times10^3\frac{dx}{dt} + 4\times10^6 x(t) = 4\times10^6 u(t)$$

If we scale time from seconds to milliseconds, *ie.* $t \rightarrow 10^{-3}t$, we obtain

$$\frac{d^2x}{10^{-6}dt^2} + 2\times10^3\frac{dx}{10^{-3}dt} + 4\times10^6 x(t) = 4\times10^6 u(t)$$

or

$$\frac{d^2x}{dt^2} + 2\frac{dx}{dt} + 4x(t) = 4u(t)$$

To get a state-space form, we let

$$x_1 = x, \quad x_2 = \frac{dx_1}{dt}$$

and get

$$\frac{dx_2}{dt} + 2x_2 + 4x_1 = 4u$$

thus the state-space equation is

$$\left.\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -4x_1 - 2x_2 + 4u \end{aligned}\right\}$$

By definition of derivative, with sampling time *T*, we have

$$\left.\begin{aligned} x_1(j+1) - x_1(j) &= Tx_2(j) \\ x_2(j+1) - x_2(j) &= -4Tx_1(j) - 2Tx_2(j) + 4Tu(j) \end{aligned}\right\}$$

or

$$\left.\begin{aligned} x_1(j+1) &= x_1(j) + Tx_2(j) \\ x_2(j+1) &= -4Tx_1(j) + (1-2T)x_2(j) + 4Tu(j) \end{aligned}\right\}$$

We thus have system equation below

$$\left.\begin{aligned} \mathbf{x}_{j+1} &= \mathbf{A}\mathbf{x}_j + \mathbf{B}u_j + \mathbf{w}_j \\ y_j &= \mathbf{H}\mathbf{x}_j + v_j \end{aligned}\right\}$$

where, assuming $T = 0.1mS$

$$\mathbf{A} = \begin{bmatrix} 1 & 0.1 \\ -0.4 & 0.8 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0.4 \end{bmatrix}, \quad \mathbf{w}_j = \begin{bmatrix} w_j \\ 0 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

Assuming the model parameter uncertainty $E\left(w_j^2\right) = Q = 10^{-4}$, and the measurement error $E\left(v_j^2\right) = R = 0.01$, thus we have a computational algorithm below

- System equations

$$\left.\begin{aligned} \mathbf{x}_{j+1} &= \mathbf{A}\mathbf{x}_j + \mathbf{B}u_j + \mathbf{w}_j \\ y_j &= \mathbf{H}\mathbf{x}_j + v_j \end{aligned}\right\}$$

- Predictor equations (Priori)

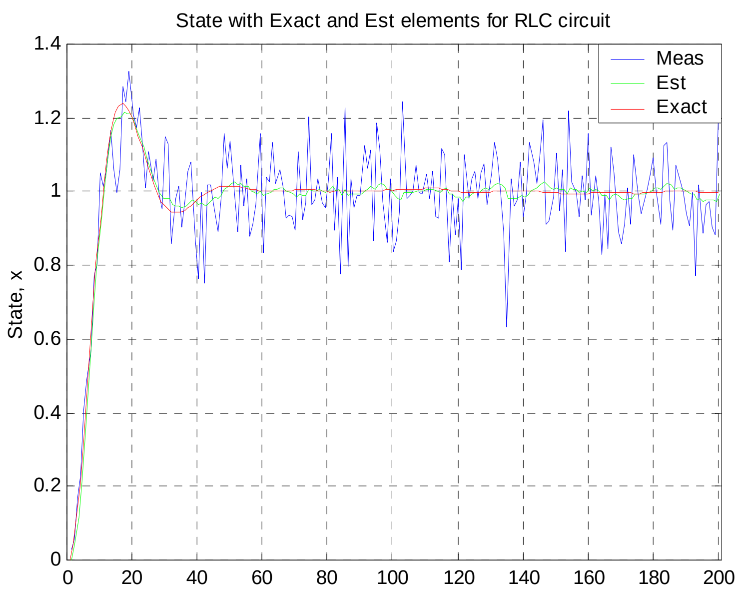
$$\begin{aligned} \overline{\mathbf{P}}_j &= \mathbf{A}\widetilde{\mathbf{P}}_{j-1}\mathbf{A}^T + \mathbf{Q} \\ \overline{\mathbf{x}}_j &= \mathbf{A}\widetilde{\mathbf{x}}_{j-1} + \mathbf{B}\mathbf{u}_j \end{aligned}$$

- Corrector equations (Posteriori)

$$\begin{aligned} \mathbf{K}_j &= \overline{\mathbf{P}}_j\mathbf{H}^T\left(\mathbf{H}\overline{\mathbf{P}}_j\mathbf{H}^T + \mathbf{R}\right)^{-1} \\ \widetilde{\mathbf{P}}_j &= \underline{\underline{\left(\mathbf{I} - \mathbf{K}_j\mathbf{H}\right)\overline{\mathbf{P}}_j\left(\mathbf{I} - \mathbf{K}_j\mathbf{H}\right)^T + \mathbf{K}_j\mathbf{R}\mathbf{K}_j^T}} = \left(\mathbf{I} - \mathbf{K}_j\mathbf{H}\right)\overline{\mathbf{P}}_j \\ \widetilde{\mathbf{x}}_j &= \overline{\mathbf{x}}_j + \mathbf{K}_j\left(\mathbf{y}_j - \mathbf{H}\overline{\mathbf{x}}_j\right) \end{aligned}$$

where

$$\mathbf{x}_j, \mathbf{w}_j \in \Re^{2\times1}; \quad \mathbf{y}_j, \mathbf{v}_j \in \Re; \quad \mathbf{A}, \mathbf{Q}, \mathbf{R}, \mathbf{P}_j \in \Re^{2\times2}; \quad \mathbf{B}, \mathbf{u}_j, \mathbf{K}_j \in \Re^{2\times1}; \quad \mathbf{H} \in \Re^{1\times2}$$



Remark 4

Note the green-colored estimate tracks the red-colored exact signal quite well even with the blue-colored noisy measurement.

7. Summary

Kalman filter minimizes the impact of noise/uncertainty on the system

$$\begin{cases} \mathbf{x}_j = \mathbf{A}\mathbf{x}_{j-1} + \mathbf{B}\mathbf{u}_j + \mathbf{w}_j, & E\left(\mathbf{w}_j \mathbf{w}_j^T\right) = \mathbf{Q} \in \Re^{n \times n} \\ \mathbf{y}_j = \mathbf{H}\mathbf{x}_j + \mathbf{v}_j, & E\left(\mathbf{v}_j \mathbf{v}_j^T\right) = \mathbf{R} \in \Re^{m \times m} \end{cases}$$

where

Q is variance of noise/uncertainty of system model, **system uncertainty**

R is variance of noise/uncertainty of output measurement, **measurement tolerance**

Kalman filter is proved mathematically that there exist **Q** & **R** such that the impact of noise/uncertainty on the system is minimal. But there's no analytical calculation of such **Q** & **R**, somehow we must find some values for them to start with, then do fine-tuning using "trial-and error" approach.

8. References

[1] http://www.tramhungchau.com/CTL/dig_ctl.pdf

[2] Candy, J V, *Signal Processing: the model based approach.* McGraw-Hill, 1986