SF2943: Time Series Analysis
Kalman Filtering

Timo Koski

09.05.2013

KTH Matematik

Contents

Introduction: Optimal Estimation of a Random Variable

In the preceding lectures of sf2943 we have been dealing with the various of instances of the general problem of estimating X with a linear combination $a_1 Y_1 + \ldots + a_N Y_N$ selecting the parameters $a_1, \ldots, a_N$ so that

$$E\left[\left(X - (a_1 Y_1 + \ldots + a_N Y_N)\right)^2\right] \tag{1}$$

is minimized (Minimal Mean Squared Error). The following has been shown.

Introduction: Optimal Solution

Suppose $Y_1, \ldots, Y_N$ and X are random variables, all with zero means, and

$$\begin{aligned} \gamma_{mk} &= E\left[Y_m Y_k\right], m = 1, \ldots, N; k = 1, \ldots, N \\ \gamma_{om} &= E\left[Y_m X\right], m = 1, \ldots, N. \end{aligned} \tag{2}$$

then

$$E\left[\left(X - \left(a_1 Y_1 + \ldots + a_N Y_N\right)\right)^2\right] \tag{3}$$

is minimized if the coefficients $a_1, \ldots, a_N$ satisfy the Wiener-Hopf equations requiring the inversion of an $N \times N$ -matrix,

$$\sum_{k=1}^{N} a_k \gamma_{mk} = \gamma_{om}; m = 1, \ldots, N. \tag{4}$$

Introduction: Wiener-Hopf Equations

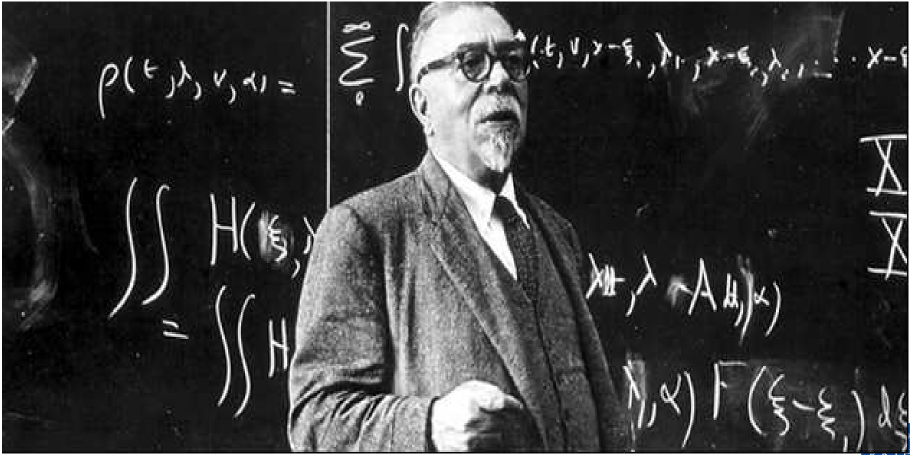
The equations

$$\sum_{k=1}^{N} a_k \gamma_{mk} = \gamma_{om}; m = 1, \ldots, N. \tag{5}$$

are often called the **Wiener-Hopf Equations**.

KTH Matematik

Norbert Wiener 1894 -1964 ; Professor of Mathematics at MIT.



Rudolf E. Kalman

The Kalman filter has brought a fundamental reformation in the classical theory of time series prediction originated by N. Wiener. The recursive algorithm (to be derived) was invented by Rudolf E. Kalman[1]. His original work is found in [3].

[1] b. 1930 in Budapest, but studied and graduated in electrical engineering in USA, Professor Emeritus in Mathematics at ETH, the Swiss Federal Institute of Technology in Zürich.

U.S. President Barack Obama (R) presents a 2008 National Medal of Science to Rudolf Kalman (L) An East Room ceremony October 7, 2009 at the White House in Washington.

KTH Matematik

Introduction: Additional Samples

We augment the notations by a dependence on the number of data points, N, as

$$\widehat{Z}_N = a_1(N)Y_1 + \ldots + a_N(N)Y_N.$$

Suppose now that we obtain one more measurement, Y_{N+1}. Then we need to find

$$\widehat{Z}_{N+1} = a_1(N+1)Y_1 + \ldots + a_{N+1}(N+1)Y_{N+1}.$$

In principle we can by the above find $a_1(N+1), \ldots, a_{N+1}(N+1)$ by solving the Wiener-Hopf equations requiring the inversion of an $N+1 \times N+1$ -matrix,

$$\sum_{k=1}^{N+1} a_k(N+1) r_{mk} = r_{om}; m = 1, \ldots, N+1. \tag{6}$$

but if new observations $Y_{N+1}, Y_{N+2}, \ldots$ are gathered sequentially in real time, this will soon become practically unfeasible.

Kalman Filtering

Kalman filtering is a technique by which we calculate $\widehat{Z}_{N+1}$ recursively using $\widehat{Z}_N$, and the latest sample Y_{N+1}. This requires a dynamic **state space representation** for the observed time series $Y \mapsto Y_n$ with $X \mapsto X_n$ as the state process. We consider the simplest special case. The Kalman Recursions are usually established for multivariate time series applying matrix equations, see, e.g., pp. $137 - 142$ in [5]. However, some of the basic principles can be made intelligible by a simpler approach involving only scalar time series[2]. The presentation in this lecture is to a large degree based on the treatment in [2] .

[2] R.M. du Plessis: Poor Man's Explanation of Kalman Filtering. North American Rockwell Electronics Group, June 1967

The State Model (1): AR(1)

- The *state model* is AR(1), i.e.,

$$X_n = \phi X_{n-1} + Z_{n-1}, n = 0, 1, 2, \ldots, \tag{1}$$

where $\{Z_n\}$ is a white noise with expectation zero, and $|\phi| \leq 1$ (so that we may regard Z_{n-1} as non correlated with $X_{n-1}, X_{n-2}, \ldots$). We have

$$R_Z(k) = E\left[Z_n Z_{n+k}\right] = \sigma^2 \cdot \delta_{0,k} = \begin{cases} \sigma^2 & k = 0 \\ 0 & k \neq 0 \end{cases} \tag{2}$$

We assume that $E\left[X_0\right] = 0$ and $\mathrm{Var}(X_0) = \sigma_0^2$.

The Observavtion Equation (2): State plus Noise

- The true state X_n is hidden from us, we see X_n with added white measurement noise V_n, or, as Y_n in

$$Y_n = cX_n + V_n, n = 0, 1, 2, \ldots, \tag{3}$$

where

$$R_V(k) = E\left[V_n V_{n+k}\right] = \sigma_V^2 \cdot \delta_{0,k} = \begin{cases} \sigma_V^2 & k = 0 \\ 0 & k \neq 0 \end{cases} \tag{4}$$

- The state and measurement noises $\{Z_n\}$ and $\{V_n\}$, respectively, are independent.

KTH Matematik

A Special Case: Poor Man's Kalman Filter (PMKF)

Assume now in (1) that $\phi = 1$ and $\sigma^2 = 0$, i.e.,

$$X_n = X_0, \quad n = 0, 1, \ldots, \tag{5}$$

We take that $c = 1$, so that

$$Y_n = X_0 + V_n, \quad n = 0, 1, 2, \ldots, \tag{6}$$

This is the statistical model of several measurements of one random variable. We shall obtain the Kalman recursions for estimating X_0 using sequentially $Y_0, Y_1, \ldots, Y_n, \ldots$, as a special case of the Kalman predictor of X_n in (1) (to be derived).

KTH Matematik

Recursive Prediction

We want to estimate X_n in (1) using $Y_0, \ldots, Y_{n-1}$ accrued according to (3), so that

$$E\left[\left(X_n - \left(a_1(n)Y_{n-1} + \ldots + a_n(n)Y_0\right)\right)^2\right] \tag{7}$$

is minimized. Next, we obtain Y_n and want to estimate X_{n+1} using $Y_0, \ldots, Y_n$ so that

$$E\left[\left(X_{n+1} - \left(a_1(n+1)Y_n + \ldots + a_{n+1}(n+1)Y_0\right)\right)^2\right] \tag{8}$$

is minimized.

Recursive Prediction

Let us set

$$\widehat{X}_{n+1} \stackrel{\text{def}}{=} a_1(n+1)Y_n + \ldots + a_{n+1}(n+1)Y_0$$

or

$$\widehat{X}_{n+1} = \sum_{k=1}^{n+1} a_k(n+1)Y_{n+1-k}, \tag{9}$$

and

$$\widehat{X}_n \stackrel{\text{def}}{=} a_1(n)Y_{n-1} + \ldots + a_n(n)Y_0$$

or

$$\widehat{X}_n = \sum_{k=1}^{n} a_k(n)Y_{n-k}. \tag{10}$$

As stated above, $a_1(n+1), \ldots, a_{n+1}(n+1)$ satisfy the Wiener-Hopf equations

$$\sum_{k=1}^{n+1} a_k(n+1)E\left[Y_m Y_{n+1-k}\right] = E\left[Y_m X_{n+1}\right]; m = 0, \ldots, n. \tag{11}$$

Recursive Prediction: the road map

The road map is as follows:

- We express $a_k(n+1)$ for $k = 2, \ldots, n+1$ recursively as functions of $a_k(n)$s.
- We show that

$$\widehat{X}_{n+1} = \phi \widehat{X}_n + a_1(n+1)\left(Y_n - c\widehat{X}_n\right),$$

- We find a recursion for

$$e_{n+1} = X_{n+1} - \widehat{X}_{n+1},$$

- We determine finally $a_1(n+1)$ by minimizing $E\left[e_{n+1}^2\right]$.

Two Auxiliary Formulas

Lemma

$$E\left[Y_m X_{n+1}\right] = \phi E\left[Y_m X_n\right], m = 0, \ldots, n-1. \tag{12}$$

and for $n \geq 1$

$$E\left[Y_m X_n\right] = \frac{E\left[Y_m Y_n\right]}{c}; m = 0, \ldots, n-1. \tag{13}$$

Proof: From the state equation (1) we get

$$E\left[Y_m X_{n+1}\right] = E\left[Y_m\left(\phi X_n + Z_n\right)\right]$$
$$= \phi E\left[Y_m X_n\right] + E\left[Y_m Z_n\right].$$

Here $E\left[Y_m Z_n\right] = E\left[Y_m\right] E\left[Z_n\right] = 0$, since Z_n is a white noise independent of V_m and X_m for $m \leq n$. Next, from (3)

$$E\left[Y_m X_n\right] = E\left[Y_m\left(Y_n - V_n\right)/c\right]$$
$$= E\left[Y_m Y_n\right]/c,$$

as claimed.

KTH Matematik

The crucial step

We start from (11), and use (12) and (13) together with the fact that $a_k(n)$:s satisfy the Wiener-Hopf equations

$$\sum_{k=1}^{n} a_k(n) E\left[Y_m Y_{n-k}\right] = E\left[Y_m X_n\right]; m = 0, \ldots, n-1. \tag{14}$$

This lemma still leaves us with one free parameter, i.e., $a_1(n+1)$, which will be determined in the sequel.

KTH Matematik

The crucial step

Lemma

$$a_{k+1}(n+1) = a_k(n)\left(\phi - a_1(n+1)c\right); k = 1, \ldots, n. \tag{15}$$

KTH Matematik

The crucial step: the proof

Proof: Let $m < n$. From (9) and (11) we get

$$E\left[Y_m X_{n+1}\right] = \sum_{k=1}^{n+1} a_k(n+1) E\left[Y_m Y_{n+1-k}\right] \tag{16}$$

$$= a_1(n+1) E\left[Y_m Y_n\right] + \sum_{k=2}^{n+1} a_k(n+1) E\left[Y_m Y_{n+1-k}\right] \tag{17}$$

We change the index of summation from k to $l = k-1$. Then

$$\sum_{k=2}^{n+1} a_k(n+1) E\left[Y_m Y_{n+1-k}\right] = \sum_{l=1}^{n} a_{l+1}(n+1) E\left[Y_m Y_{n-l}\right]. \tag{18}$$

The crucial step: the proof

On the other hand, (12) yields in the left hand side of (16) that

$$E\left[Y_m X_{n+1}\right] = \phi E\left[Y_m X_n\right]$$

and from (13) that

$$a_1(n+1)E\left[Y_m Y_n\right] = a_1(n+1)cE\left[Y_m X_n\right].$$

KTH Matematik

The crucial step: the proof

Thus we can write (16), (17) by means of (18) as

$$\left(\phi - a_1(n+1)c\right) E\left[Y_m X_n\right] = \sum_{l=1}^{n} a_{l+1}(n+1) E\left[Y_m Y_{n-l}\right]. \tag{19}$$

Now we compare with the system of equations in (14), i.e.,

$$\sum_{k=1}^{n} a_k(n) E\left[Y_m Y_{n-k}\right] = E\left[Y_m X_n\right]; m = 0, \ldots, n-1.$$

It must hold that

$$a_k(n) = \frac{a_{k+1}(n+1)}{\left(\phi - a_1(n+1)c\right)}$$

or

$$a_{k+1}(n+1) = a_k(n)\left(\phi - a_1(n+1)c\right).$$

We can keep $a_1(n+1)$ undetermined and still get the following result.

Lemma

$$\widehat{X}_{n+1} = \phi \widehat{X}_n + a_1(n+1)\left(Y_n - c\widehat{X}_n\right) \tag{20}$$

KTH Matematik

Proof: By (9) and the same trick of changing the index of summation as above we obtain

$$\widehat{X}_{n+1} = \sum_{k=1}^{n+1} a_k(n+1) Y_{n+1-k} = a_1(n+1) Y_n + \sum_{k=1}^{n} a_{k+1}(n+1) Y_{n-k}$$

so from (15) in the preceding lemma

$$= a_1(n+1) Y_n + \sum_{k=1}^{n} a_k(n) \left[\phi - a_1(n+1)c\right] Y_{n-k}$$

$$= a_1(n+1) Y_n + \phi \sum_{k=1}^{n} a_k(n) Y_{n-k} - c a_1(n+1) \sum_{k=1}^{n} a_k(n) Y_{n-k}.$$

From (10) we obviously find in the right hand side

$$= \phi \widehat{X}_n + a_1(n+1) \left(Y_n - c\widehat{X}_n \right),$$

as was to be proved.

We choose the value for $a_1(n+1)$ which minimizes the second moment of **the state prediction error/innovation** defined as

$$e_{n+1} = X_{n+1} - \widehat{X}_{n+1}.$$

First we find recursions for e_{n+1} and $E\left[e_{n+1}^2\right]$.

Lemma

$$e_{n+1} = \left[\phi - a_1(n+1)c\right] e_n + Z_n - a_1(n+1)V_n, \tag{21}$$

and

$$E\left[e_{n+1}^2\right] = \left[\phi - a_1(n+1)c\right]^2 E\left[e_n^2\right] + \sigma^2 + a_1^2(n+1)\sigma_V^2. \tag{22}$$

A Recursion with $a_1(n+1)$ undetermined: proof

Proof : We have from the equations above that

$$e_{n+1} = X_{n+1} - \widehat{X}_{n+1} = \phi X_n + Z_n - \phi \widehat{X}_n - a_1(n+1)\left(cX_n + V_n - c\widehat{X}_n\right)$$

$$= \phi\left(X_n - \widehat{X}_n\right) - a_1(n+1)c\left(X_n - \widehat{X}_n\right) + Z_n - a_1(n+1)V_n,$$

and with $e_n = X_n - \widehat{X}_n$ the result is (21).
If we square both sides of (21) we get

$$e_{n+1}^2 = \left[\phi - a_1(n+1)c\right]^2 e_n^2 + \left(Z_n - a_1(n+1)V_n\right)^2 + 2\tau,$$

where

$$\tau = \left[\phi - a_1(n+1)c\right] e_n \left(Z_n - \left(a_1(n+1)\right) V_n\right)$$

By the properties above

$$E\left[e_n Z_n\right] = E\left[Z_n V_n\right] = E\left[e_n V_n\right] = 0$$

(Note that $\widehat{X}_n$ uses $X_{n-1}, X_{n-2}, \ldots$ and is thus uncorrelated with V_n.)
Thus we have $E\left[e_{n+1}^2\right]$ as asserted.

A Recursion with $a_1(n+1)$ determined

We can apply orthogonality to find the expression for $a_1(n+1)$, but the differentiation to be invoked in the next lemma gives a faster argument.

Lemma

$$a_1(n+1) = \frac{\phi c E\left[e_n^2\right]}{\sigma_V^2 + c^2 E\left[e_n^2\right]}. \tag{23}$$

minimizes $E\left[e_{n+1}^2\right]$.

A Recursion with $a_1(n+1)$ determined: proof

Proof: We differentiate (22) w.r.t. $a_1(n+1)$ and set the derivative equal to zero. This entails

$$-2c\left[\phi - a_1(n+1)c\right] E\left[e_n^2\right] + 2a_1(n+1)\sigma_V^2 = 0$$

or

$$a_1(n+1)\left(c^2 E\left[e_n^2\right] + \sigma_V^2\right) = \phi c E\left[e_n^2\right],$$

which yields (23) as claimed. $\square$

We change notation for $a_1(n+1)$ determined in lemma 5, (23). The traditional way (in the engineering literature) is to write this quantity as the *predictor gain*, also known as the *Kalman gain* $K(n)$ or

$$K(n) \stackrel{\text{def}}{=} \frac{\phi c E\left[e_n^2\right]}{\sigma_V^2 + c^2 E\left[e_n^2\right]}. \tag{24}$$

The initial conditions are

$$\widehat{X}_0 = 0, E\left[e_0^2\right] = \sigma_0^2. \tag{25}$$

Then we have obtained above that

$$E\left[e_{n+1}^2\right] = \left[\phi - K(n)c\right]^2 E\left[e_n^2\right] + \sigma^2 + K^2(n)\sigma_V^2, \tag{26}$$

and

$$\widehat{X}_{n+1} = K(n)\left[Y_n - c\widehat{X}_n\right] + \phi\widehat{X}_n. \tag{27}$$

The Kalman Prediction Filter

Proposition

$$K(n) \stackrel{\text{def}}{=} \frac{\phi c E\left[e_n^2\right]}{\sigma_V^2 + c^2 E\left[e_n^2\right]}. \tag{28}$$

$$E\left[e_{n+1}^2\right] = \left[\phi - K(n)c\right]^2 E\left[e_n^2\right] + \sigma^2 + K^2(n)\sigma_V^2, \tag{29}$$

and

$$\widehat{X}_{n+1} = K(n)\left[Y_n - c\widehat{X}_n\right] + \phi\widehat{X}_n. \tag{30}$$

The equations (28) -(30) are an algorithm for recursive computation of $\widehat{X}_{n+1}$ for all $n \geq 0$. This is a simple case of an important statistical processor of time series data known as the **Kalman filter**.

The Kalman Prediction Filter

A Kalman filter is in view of (30) *predicting* X_{n+1} using $\phi\widehat{X}_n$ and additively *correcting* the prediction by the measured *innovations*

$$\varepsilon_n = \left[Y_n - c\widehat{X}_n \right].$$

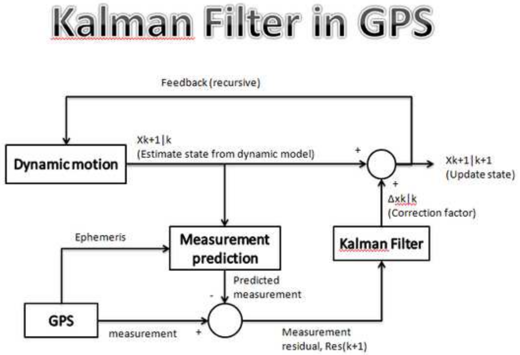
Here ε_n is the part of Y_n which is not exhausted by $c\widehat{X}_n$. The innovations are modulated by the filter gains $K(n)$ that depend on the error variances $E\left[e_n^2\right]$.

KTH Matematik

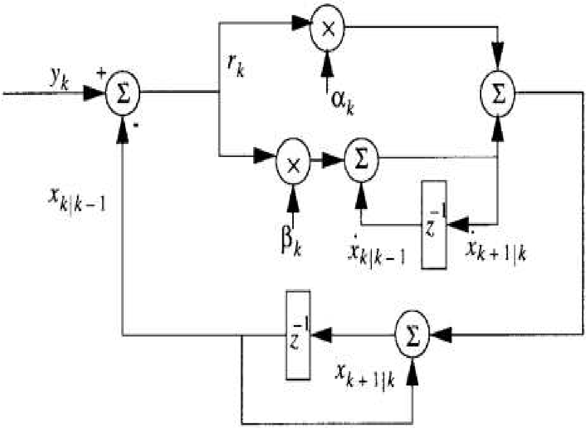
The Kalman Prediction Filter

The recursive nature of the Kalman filter cannot be overemphasized: *the filter processes one measurement Y_n at a time, instead of all measurements*. The data preceding Y_n, i.e., $Y_0, \ldots, Y_{n-1}$ are summarized in $\widehat{X}_n$, no past data need be stored. Each estimation is identical in procedure to those that took place before it, but each has a new weighing factor computed to take into account the sum total effect of all the previous estimates.

Block-Diagrams



Block-Diagrams



The Kalman Prediction Filter: A Special Case

Assume $\sigma_V = 0$ and $c = 1$. Then (28) becomes

$$K(n) = \phi \tag{31}$$

and (29) boils down to

$$E\left[e_{n+1}^2\right] = \sigma^2, \tag{32}$$

and (30) yields, since $c = 1$,

$$\widehat{X}_{n+1} = \phi Y_n = \phi X_n. \tag{33}$$

But this verifies the familiar formula about one-step MSE-prediction of an AR(1) process. $\square$

We can write the filter also with **an innovations representation**

$$\widehat{X}_{n+1} = \phi \widehat{X}_n + K(n)\varepsilon_n \tag{34}$$

$$Y_n = c\widehat{X}_n + \varepsilon_n, \tag{35}$$

which by comparison with (1) and (3) shows that the Kalman filter follows equations similar to the original ones, but is driven by the innovations ε_n as noise.

Riccati Recursion for the Variance of the Prediction Error

We find a further recursion for $E\left[e_{n+1}^2\right]$. We start with a general identity for Minimal Mean Squared Error estimation. We have

$$E\left[e_{n+1}^2\right] = E\left[X_{n+1}^2\right] - E\left[\widehat{X}_{n+1}^2\right]. \tag{36}$$

To see this, let us note that

$$\begin{aligned} E\left[e_{n+1}^2\right] &= E\left[\left(X_{n+1} - \widehat{X}_{n+1}\right)^2\right] \\ &= E\left[X_{n+1}^2\right] - 2E\left[X_{n+1}\widehat{X}_{n+1}\right] + E\left[\widehat{X}_{n+1}^2\right]. \end{aligned}$$

Here

$$E\left[X_{n+1}\widehat{X}_{n+1}\right] = E\left[\left(\widehat{X}_{n+1} + e_{n+1}\right)\widehat{X}_{n+1}\right] = E\left[\widehat{X}_{n+1}^2\right] + E\left[e_{n+1}\widehat{X}_{n+1}\right].$$

But by the **orthogonality principle** of Minimal Mean Squared Error estimation we have

$$E\left[e_{n+1}\widehat{X}_{n+1}\right] = 0,$$

and this proves (36).

KTH Matematik

Riccati Recursion

We note next writing

$$\varepsilon_n = Y_n - c\widehat{X}_n = cX_n + V_n - c\widehat{X}_n$$

$$= ce_n + V_n$$

that

$$E\left[\varepsilon_n^2\right] = \sigma_V^2 + c^2 E\left[e_n^2\right]. \tag{37}$$

KTH Matematik

When we use this formula we obtain from (30) or $\widehat{X}_{n+1} = K(n)\varepsilon_n + \phi \widehat{X}_n$ that

$$E\left[\widehat{X}_{n+1}^2\right] = \phi^2 E\left[\widehat{X}_n^2\right] + K(n)^2\left(\sigma_V^2 + c^2 E\left[e_n^2\right]\right). \tag{38}$$

But in view of our definition of the Kalman gain in (28) we have

$$K(n)^2\left(\sigma_V^2 + c^2 E\left[e_n^2\right]\right) = \frac{\phi^2 c^2 E\left[e_n^2\right]^2}{\sigma_V^2 + c^2 E\left[e_n^2\right]}$$

As in [1, p.273] we set

$$\theta_n = \phi c E\left[e_n^2\right], \tag{39}$$

and

$$\nabla_n = \sigma_V^2 + c^2 E\left[e_n^2\right], \tag{40}$$

KTH Matematik

Riccati Recursion for the Variance of the Prediction Error

Next we use the state equation (1) to get

$$E\left[X_{n+1}^{2}\right]=\phi^{2}E\left[X_{n}^{2}\right]+\sigma^{2}. \tag{41}$$

Hence we have by (36), (41) and (38)

$$E\left[e_{n+1}^{2}\right]=\phi^{2}E\left[X_{n}^{2}\right]+\sigma^{2}-\phi^{2}E\left[\widehat{X}_{n}^{2}\right]-\frac{\theta_{n}^{2}}{\nabla_{n}}$$

or, again using (36),

$$E\left[e_{n+1}^{2}\right]=\phi^{2}E\left[e_{n}^{2}\right]+\sigma^{2}-\frac{\theta_{n}^{2}}{\nabla_{n}}. \tag{42}$$

KTH Matematik

Hence we have shown the following proposition (c.f., [1, p.273]) about Kalman prediction.

Proposition

$$\widehat{X}_{n+1} = \phi \widehat{X}_n + \frac{\theta_n}{\nabla_n} \varepsilon_n. \tag{43}$$

and if

$$e_{n+1} = X_{n+1} - \widehat{X}_{n+1}$$

then

$$E\left[e_{n+1}^2\right] = \phi^2 E\left[e_n^2\right] + \sigma^2 - \frac{\theta_n^2}{\nabla_n}, \tag{44}$$

where

$$\theta_n = \phi c E\left[e_n^2\right], \tag{45}$$

and

$$\nabla_n = \sigma_V^2 + c^2 E\left[e_n^2\right]. \tag{46}$$

PMKF

In PMKF we have $\phi = 1, c = 1$ and $\sigma = 0$, we are estimating sequentially a random variable without time dynamics. We need to rewrite the preceding a bit, see [5]. First we have

$$\frac{1}{E\left[e_{n+1}^2\right]} = \frac{1}{E\left[e_n^2\right]} + \frac{1}{\sigma_V^2}, \tag{47}$$

since in this case (44) becomes

$$E\left[e_{n+1}^2\right] = E\left[e_n^2\right] - \frac{E\left[e_n^2\right]^2}{\sigma_V^2 + E\left[e_n^2\right]},$$

and by some elementary algebra

$$E\left[e_{n+1}^2\right] = \frac{\sigma_V^2 E\left[e_n^2\right]^2}{\sigma_V^2 + E\left[e_n^2\right]}.$$

From this

$$\frac{1}{E\left[e_{n+1}^2\right]} = \frac{\sigma_V^2 + E\left[e_n^2\right]}{\sigma_V^2 E\left[e_n^2\right]^2} = \frac{1}{E\left[e_n^2\right]} + \frac{1}{\sigma_V^2}.$$

PMKF

Then iteration gives that

$$\frac{1}{E\left[e_n^2\right]} = \frac{1}{\sigma_0^2} + \frac{n}{\sigma_V^2}. \tag{48}$$

We need in view of (43) to compute the Kalman gain for the case at hand, or

$$\frac{\theta_n}{\nabla_n} = \frac{E\left[e_n^2\right]}{\sigma_V^2 + E\left[e_n^2\right]} = \frac{1}{\frac{\sigma_V^2}{E\left[e_n^2\right]} + 1},$$

where we now use (48) to get

$$= \frac{1}{\frac{\sigma_V^2}{\sigma_0^2} + n + 1} = \frac{\sigma_0^2}{\sigma_V^2 + \sigma_0^2\,(n+1)},$$

KTH Matematik

PMKF

Thus we have found the Poor Man's Kalman Filter

$$\widehat{X}_{n+1} = \widehat{X}_n + \frac{\sigma_0^2}{(n+1)\sigma_0^2 + \sigma_V^2}\left(Y_n - \widehat{X}_n\right). \tag{49}$$

The poor man's cycle of computation: You have computed $\widehat{X}_n$.

- You receive Y_n.
- Update the gain to $\frac{\sigma_0^2}{(n+1)\sigma_0^2+\sigma_V^2}$
- Compute $\widehat{X}_{n+1}$ by adding the correction to $\widehat{X}_n$.
- $\widehat{X}_{n+1} \mapsto \widehat{X}_n$.

Recall (44), or,

$$E\left[e_{n+1}^2\right] = \phi^2 E\left[e_n^2\right] + \sigma^2 - \frac{\theta_n^2}{\nabla_n}, \tag{50}$$

The recursion in (50) is a first order nonlinear difference equation known as the *Riccati equation*[3] for the prediction error variance.

[3] named after Jacopo Francesco Riccati, 1676 - 1754, who was a mathematician born in Venice, who wrote on philosophy, physics and differential equations. http://www-groups.dcs.st-and.ac.uk/~history/Biographies/Riccati.html

The Stationary Kalman Filter

We say that the prediction filter has reached a *steady state*, if $E\left[e_n^2\right]$ is a constant, say $P = E\left[e_n^2\right]$, that does not depend on n. Then the Riccati equation in (44) becomes a quadratic algebraic Riccati equation

$$P = \phi^2 P + \sigma^2 - \frac{\phi^2 c^2 P^2}{\sigma_V^2 + c^2 P}, \tag{51}$$

or

$$P = \sigma^2 + \frac{\sigma_V^2 \phi^2 P}{\sigma_V^2 + c^2 P}.$$

KTH Matematik

and further by some algebra

$$c^2 P^2 + \left(\sigma_V^2 - \sigma_V^2 \phi^2 - \sigma^2 c^2\right) P - \sigma^2 \sigma_V^2 = 0. \tag{52}$$

This algebraic second order equation is solvable by the ancient formula of Indian mathematics. We take only the non negative root into account. Given the stationary P we have the stationary Kalman gain as

$$K \stackrel{\text{def}}{=} \frac{\phi c P}{\sigma_V^2 + c^2 P}. \tag{53}$$

Example of Kalman Prediction: Random Walk Observed in Noise

Consider a discrete time Brownian motion (Z_n is a Gaussian white noise) or a random walk

$$X_{n+1} = X_n + Z_n, n = 0, 1, 2, \ldots,$$

observed in noise

$$Y_n = X_n + V_n, n = 0, 1, 2, \ldots,$$

KTH Matematik

Examples of Kalman Prediction: Random Walk Observed in Noise

Then

$$\widehat{X}_{n+1} = \widehat{X}_n + \frac{E\left[e_n^2\right]}{\sigma_V^2 + E\left[e_n^2\right]}\left(Y_n - \widehat{X}_n\right).$$

and

$$\widehat{X}_{n+1} = \left(1 - \frac{E\left[e_n^2\right]}{\sigma_V^2 + E\left[e_n^2\right]}\right)\widehat{X}_n + \frac{E\left[e_n^2\right]}{\sigma_V^2 + E\left[e_n^2\right]} Y_n.$$

Example of Kalman Prediction: Random Walk Observed in Noise

It can be shown that there is convergence to the stationary filter

$$\widehat{X}_{n+1} = \left(1 - \frac{P}{\sigma_V^2 + P}\right)\widehat{X}_n + \frac{P}{\sigma_V^2 + P} Y_n,$$

where P is found by solving (52).
We have in this example $\phi = 1, c = 1$. We select $\sigma^2 = 1$, $\sigma_V^2 = 1$. Then (52) becomes

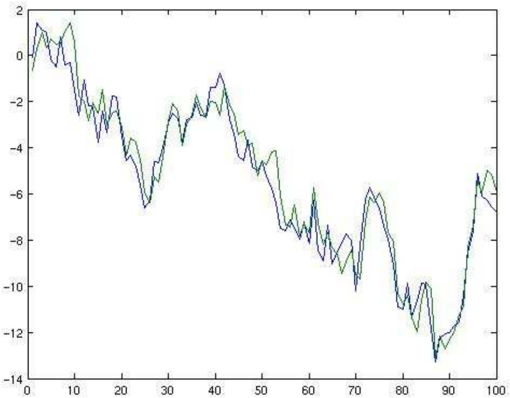
$$P^2 - P - 1 = 0 \Leftrightarrow P = \frac{1}{2} + \sqrt{\frac{1}{4} + 1} = 1.618,$$

and the stationary Kalman gain is $K = 0.618$.

KTH Matematik

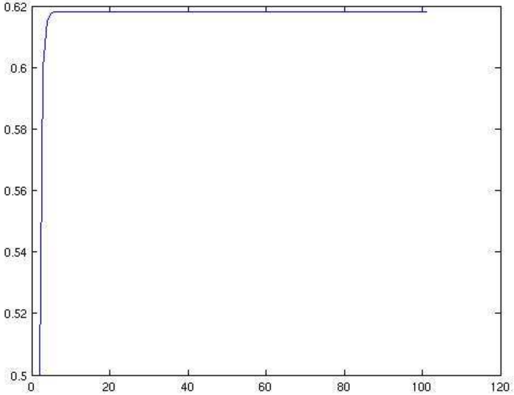
Examples of Kalman Prediction: Random Walk Observed in Noise

In the first figure there is depicted a simulation of the state process and the computed trajectory of the Kalman predictor.



Example of Kalman Prediction: Random Walk Observed in Noise

In the next figure we see the fast convergence of the Kalman gain $K(n)$ to



$K = 0.618.$

Example of Kalman Prediction: Exponential Decay Observed in Noise

Consider an exponential decay, $0 < \phi < 1$,

$$X_{n+1} = \phi X_n, n = 0.1, 2, \ldots, \Leftrightarrow X_n = \phi^n X_0$$

observed in noise

$$Y_n = X_n + V_n, n = 0, 1, 2, \ldots.$$

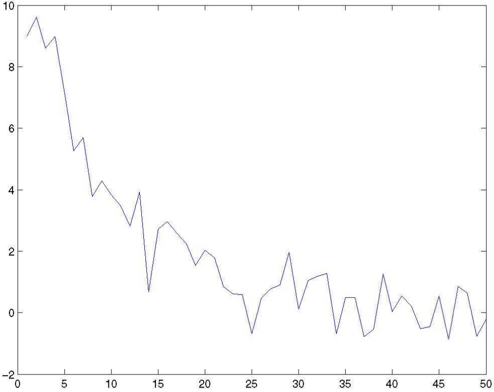
Then there is convergence to the stationary filter

$$\widehat{X}_{n+1} = \phi \left(1 - \frac{P}{\sigma_V^2 + P} \right) \widehat{X}_n + \frac{\phi P}{\sigma_V^2 + P} Y_n.$$

KTH Matematik

Example of Kalman Prediction: Exponential Decay Observed in Noise

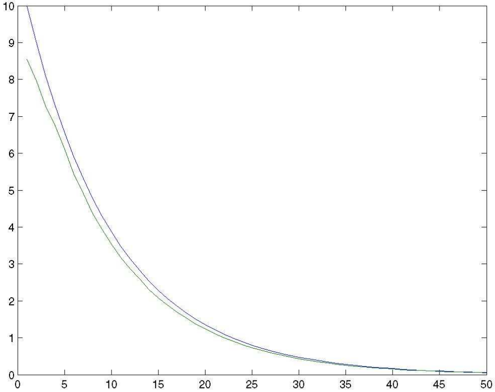
We see the corresponding noisy measurements of the exponential decay with $c = 1$, $\phi = 0.9$ $\sigma^2 = 0$, and $\sigma_V^2 = 1$.



KTH Matematik

Example of Kalman Prediction: Exponential Decay Observed in Noise

In the figures there is first depicted the state process and the computed trajectory of the Kalman predictor.



Example of Kalman Prediction: Exponential Decay Observed in Noise

With $c = 1$, $\phi = 0.9$, $\sigma^2 = 0$, and $\sigma_V^2 = 1$ (52) becomes

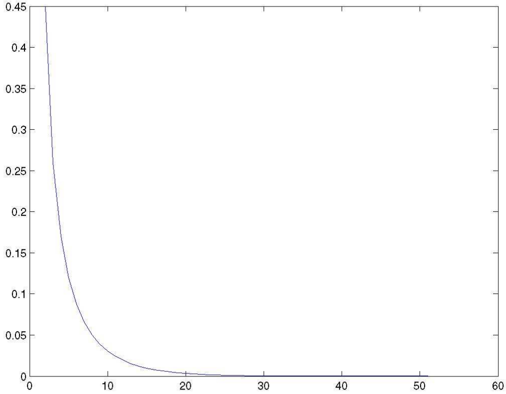
$$P^2 + \left(1 - 0.9^2\right) P = 0 \Leftrightarrow P = 0,$$

(where we neglect the negative root) and the stationary Kalman gain is $K = 0$.

KTH Matematik

Example of Kalman Prediction: Exponential Decay Observed in Noise

In the next figure we see the fast convergence of the Kalman gain $K(n)$ to



$K = 0$.

Example of Kalman Prediction

The state process

$$X_n = 0.6X_{n-1} + Z_{n-1}, n = 0, 1, 2, \ldots,$$

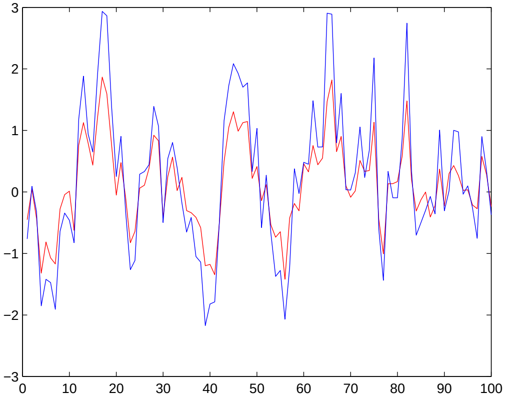
and the observation equation

$$Y_n = X_n + V_n, n = 0, 1, 2, \ldots,$$

$\sigma^2 = 1$, $\sigma_V^2 = 0.1$,$\sigma_0^2 = 1$.

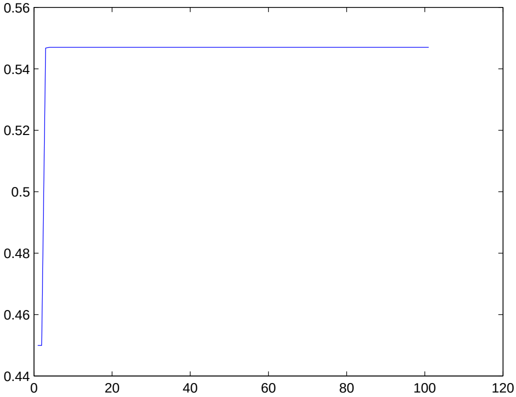
Example of Kalman Prediction

The time series in red is $\widehat{X}_n$, the time series in blue is X_n,



$n = 1, 2, \ldots, 100.$

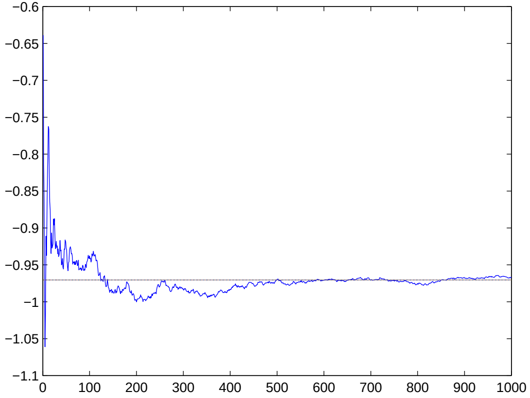
Example of Kalman Prediction



The Kalman gain:

PMKF

PMKF, where $X_0 = -0.9705$. Here $\widehat{X}_{1000} = -0.9667$.



Recall the state-space model:

$$\mathbf{X}_{t+1} = F_t \mathbf{X}_t + \mathbf{V}_t, \quad \{\mathbf{V}_t\} \sim \mathrm{WN}(\mathbf{0}, \{Q_t\}),$$

$$\mathbf{Y}_t = G_t \mathbf{X}_t + \mathbf{W}_t, \quad \{\mathbf{W}_t\} \sim \mathrm{WN}(\mathbf{0}, \{R_t\}).$$

The Kalman recursions for Prediction: the General Case

Linear estimation of $\mathbf{X}_t$ in terms of

- $\mathbf{Y}_0, \ldots, \mathbf{Y}_{t-1}$ defines the *prediction problem*;
- $\mathbf{Y}_0, \ldots, \mathbf{Y}_t$ defines the *filtering problem*;
- $\mathbf{Y}_0, \ldots, \mathbf{Y}_n, \;\; n > t,$ defines the *smoothing problem*.

The Kalman recursions: the General Case

The predictors $\widehat{\mathbf{X}}_t \stackrel{\text{def}}{=} P_{t-1}(\mathbf{X}_t)$ and the error covariance matrices

$$\Omega_t \stackrel{\text{def}}{=} E[(\mathbf{X}_t - \widehat{\mathbf{X}}_t)(\mathbf{X}_t - \widehat{\mathbf{X}}_t)']$$

are uniquely determined by the initial conditions

$$\widehat{\mathbf{X}}_1 = P(\mathbf{X}_1 \mid \mathbf{Y}_0), \quad \Omega_1 \stackrel{\text{def}}{=} E[(\mathbf{X}_1 - \widehat{\mathbf{X}}_1)(\mathbf{X}_1 - \widehat{\mathbf{X}}_1)']$$

KTH Matematik

and the recursions, for $t = 1, \ldots,$

$$\widehat{\mathbf{X}}_{t+1} = F_t \widehat{\mathbf{X}}_t + \Theta_t \Delta_t^{-1} (\mathbf{Y}_t - G_t \widehat{\mathbf{X}}_t) \tag{54}$$

$$\Omega_{t+1} = F_t \Omega_t F_t' + Q_t - \Theta_t \Delta_t^{-1} \Theta_t', \tag{55}$$

where

$$\Delta_t = G_t \Omega_t G_t' + R_t,$$
$$\Theta_t = F_t \Omega_t G_t'.$$

The matrix $\Theta_t \Delta_t^{-1}$ is called the *Kalman gain*.

KTH Matematik

Notes and Literature on the Kalman Filter

The Kalman filter was first applied to the problem of trajectory estimation for the Apollo space program of the NASA (in the 1960s), and incorporated in the Apollo space navigation computer. Perhaps the most commonly used type of Kalman filter is nowadays the phase-locked loop found everywhere in radios, computers, and nearly any other type of video or communications equipment. New applications of the Kalman Filter (and of its extensions like particle filtering) continue to be discovered, including for radar, global positioning systems (GPS), hydrological modelling, atmospheric observations e.t.c..

Wikipedia on the Kalman Filter

Kalman filters have been vital in the implementation of the navigation systems of U.S. Navy nuclear ballistic missile submarines, and in the guidance and navigation systems of cruise missiles such as the U.S. Navy's Tomahawk missile and the U.S. Air Force's Air Launched Cruise Missile. It is also used in the guidance and navigation systems of the NASA Space Shuttle and the attitude control and navigation systems of the International Space Station.

Notes and Literature on the Kalman Filter

A Kalman filter webpage with lots of links to literature, software, and extensions (like particle filtering) is found at
http://www.cs.unc.edu/~welch/kalman/

KTH Matematik

It has been understood only recently that the Danish mathematician and statistician Thorvald N. Thiele[4] discovered the principle (and a special case) of the Kalman filter in his book published in Copenhagen in 1889 (!): *Forelæsningar over Almindelig Iagttagelseslære: Sandsynlighedsregning og mindste Kvadraters Methode*. A translation of the book and an exposition of Thiele′s work is found in [4].

[4] 1838−1910, a short biography is in `http://www-groups.dcs.st-ac.uk./∼history/Biographies/Thiele.html`

Notes and Literature on the Kalman Filter

1. P.J. Brockwell and R.A. Davies: *Introduction to Time Series and Forecasting. Second Edition* Springer New York, 2002.
2. R.M. Gray and L.D. Davisson: *An introduction to statistical signal processing.* Cambridge University Press, Cambridge, U.K., 2004.
3. R.E. Kalman: A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME – Journal of Basic Engineering, 1960, 82, (Series D): 35–45.*
4. S.L. Lauritzen: *Thiele: pioneer in statistics.* Oxford University Press, 2002.
5. T. Söderström: *Discrete Time Stochastic Systems*, Springer 2002.

KTH VETENSKAP OCH KONST

KTH Matematik