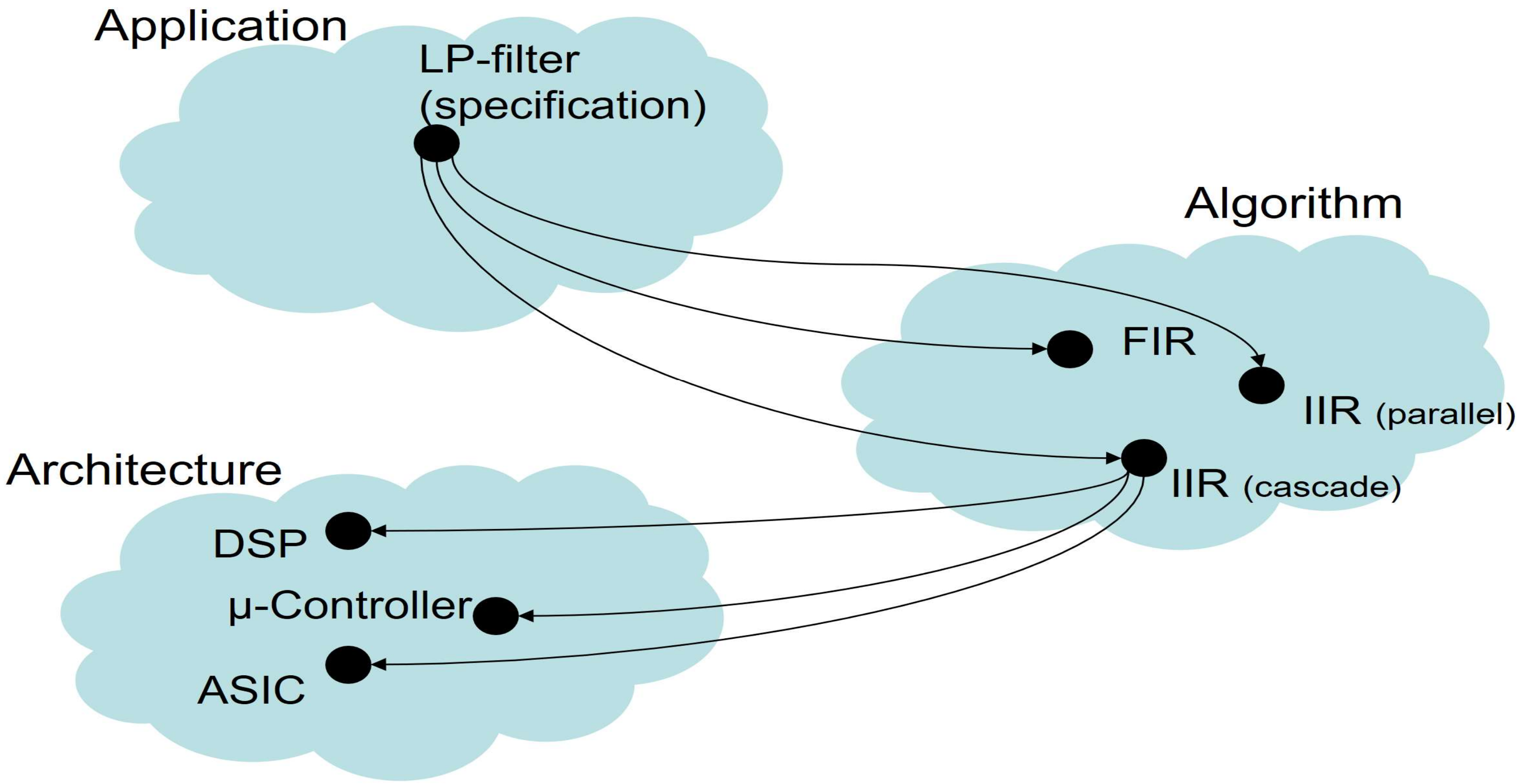
IEE 1711: Applied Signal Processing

Professor Muhammad Mahtab Alam (muhammad.alam@taltech.ee)

Lab Instructor: Julia Berdnikova



Outline

- Lecture 9: Digital Communication – Optimal Receiver
  - Brief Followup

- Lecture 10: Kalman Filter

  - Summary

Digital equalizer types

- Viterbi equalizer: Finds the maximum likelihood (ML) optimal solution to the equalization problem. Its goal is to minimize the probability of making an error over the entire sequence.
- **Linear equalizer:** processes the incoming signal with a linear filter
  - MMSE equalizer: designs the filter to minimize $E[|e|^2]$, where e is the error signal, which is the filter output minus the transmitted signal.
  - Zero forcing equalizer: approximates the inverse of the channel with a linear filter.
- Decision feedback equalizer: augments a linear equalizer by adding a filtered version of previous symbol estimates to the original filter output.
- Blind equalizer: estimates the transmitted signal without knowledge of the channel statistics, using only knowledge of the transmitted signal's statistics.
- Adaptive equalizer: is typically a linear equalizer or a DFE. It updates the equalizer parameters (such as the filter coefficients) as it processes the data. Typically, it uses the MSE cost function; it assumes that it makes the correct symbol decisions, and uses its estimate of the symbols to compute e, which is defined above.
- BCJR equalizer: uses the BCJR algorithm (also called the Forward-backward algorithm) to find the maximum *a posteriori* (MAP) solution. Its goal is to minimize the probability that a given bit was incorrectly estimated.
- Turbo equalizer: applies turbo decoding while treating the channel as a convolutional code.

Definition of Kalman filter

- It is an ***optimal (linear) estimator*** or ***optimal recursive data processing algorithm***.
- Belongs to the state space model (time domain) compared to frequency domain
- Components: system's dynamics model, control inputs, and **recursive** measurements (include noise)
- Parameters include indirect, inaccurate and uncertain observations.

- Kalman Filter is a linear adaptive filter which suits for dynamic situation. It compute the states recursively and its mathematical formulation is based on state space model.

Covariance

In statistics and probability theory, the **covariance matrix** is a matrix of covariances between elements of a vector. It is the natural generalization to higher dimensions of the concept of the variance of a scalar-valued random variable.

If X is a column vector with n scalar random variable components, and μ_k is the expected value of the k^{th} element of X, i.e., $\mu_k = \mathrm{E}(X_k)$, then the covariance matrix is defined as:

$$\Sigma = \mathrm{E}\left[(\mathbf{X} - \mathrm{E}[\mathbf{X}])(\mathbf{X} - \mathrm{E}[\mathbf{X}])^{\top}\right]$$

$$= \begin{bmatrix} \mathrm{E}[(X_1-\mu_1)(X_1-\mu_1)] & \mathrm{E}[(X_1-\mu_1)(X_2-\mu_2)] & \cdots & \mathrm{E}[(X_1-\mu_1)(X_n-\mu_n)] \\ \mathrm{E}[(X_2-\mu_2)(X_1-\mu_1)] & \mathrm{E}[(X_2-\mu_2)(X_2-\mu_2)] & \cdots & \mathrm{E}[(X_2-\mu_2)(X_n-\mu_n)] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}[(X_n-\mu_n)(X_1-\mu_1)] & \mathrm{E}[(X_n-\mu_n)(X_2-\mu_2)] & \cdots & \mathrm{E}[(X_n-\mu_n)(X_n-\mu_n)] \end{bmatrix}$$

The (i,j) element is the covariance between X_i and X_j.

This concept generalizes to higher dimensions the concept of variance of a scalar-valued random variable X, defined as

$$\sigma^2 = \mathrm{var}(X) = \mathrm{E}[(X-\mu)^2]$$

where $\mu = \mathrm{E}(X)$.

Kalman filter – Concept of States

- To understand the Kalman Filter implementation we first, need to understand the basis concepts regarding the motion of the object and how the new observations affects the state of the object.
- The state comprises of data based on the past behaviour which is then used to predict the future state.
- The predicted state could have a prediction error, therefore the new observations could be used to provide correction.

Kalman filter – Concept of Motion Models (1/5)

- Different motion models have different factors which influence the motion of an object. In regards to motion model, there are several motions which are possible. Some of them are given below:

  - motion in circular path
  - high dynamic, which consists of position, velocity and accelaration (PVA).
  - low dynamic, which consists of position and velocity (PV).
  - Stationary motion only includes position(P).

Kalman filter – Concept of Motion Models (2/5)

- For example, if the dynamics of the system is described by PV motion model. Then the state equation is defined as:

$$x(k) = A_{k-1}\, x_{k-1} + \psi_k$$

where x_k is a state vector

$$\mathbf{x}_k = \begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}$$

A_k is a state transition matrix

$$\mathbf{A}_k = \begin{bmatrix} 1 & 0 & \Delta(t) & 0 \\ 0 & 1 & 0 & \Delta(t) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where

$\Delta(t)$: is a sample rate.

Kalman filter – Concept of Motion Models (3/5)

- The state transition matrix is a reflection of the motion model. The first element of the first row represents the position in the X-direction, and the third element adds the displacement in X-direction.

$$\mathbf{A}_k = \begin{bmatrix} 1 & 0 & \Delta(t) & 0 \\ 0 & 1 & 0 & \Delta(t) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Similarly the second element of the second row represents the motion in Y-direction with displacement in that direction represented by the last element of second row.
- This is acheived when we multiply the state transition matrix with the state vector.

Kalman filter – Concept of Motion Models (4/5)

$x(k) = A_{k-1}\, x_{k-1} + \psi_k$

- ψ_k is a process noise which is assumed to be white noise with zero mean and covariance Q_k
- and it can be expressed as ψ_k = N(0,Qk). The elements of the ψ_k are assumed to be uncorrelated overtime, but the instantaneous mutual correlation is reflected in the covariance terms of Q.

$$E[\psi_k \psi(k+e)^T] = \begin{cases} Q, & e = 0 \\ 0, & e \neq 0 \end{cases}$$

- The process noise covariance matrix is a function of sample interval Δ(t). The process noise covariance matrix is given by,

Kalman filter – Concept of Motion Models (5/5)

$$\mathbf{Q}_k = q \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22}^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33}^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44}^2 \end{bmatrix}$$

- The q is a tuning factor which adjusted the variances and covariances terms.
- The noise covariance matrix determines, how much we believe that our motion model is correct.
- Note: our model will have biggest deviation from the real system when the real system shows non linear motion i.e., accelaration.

Kalman filter – Concept of Measurement Models (1)

- The relation between the state and the measurement can be described by measurement model, which is expressed by following equation:

$$b(k) = H_k x_k + \Phi_k$$

- Where H_k: is a matrix which maps from state to the measurement.

$$\mathbf{H}_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Kalman filter – Concept of Measurement Models (2)

$$b(k) = H_k x_k + \Phi_k$$

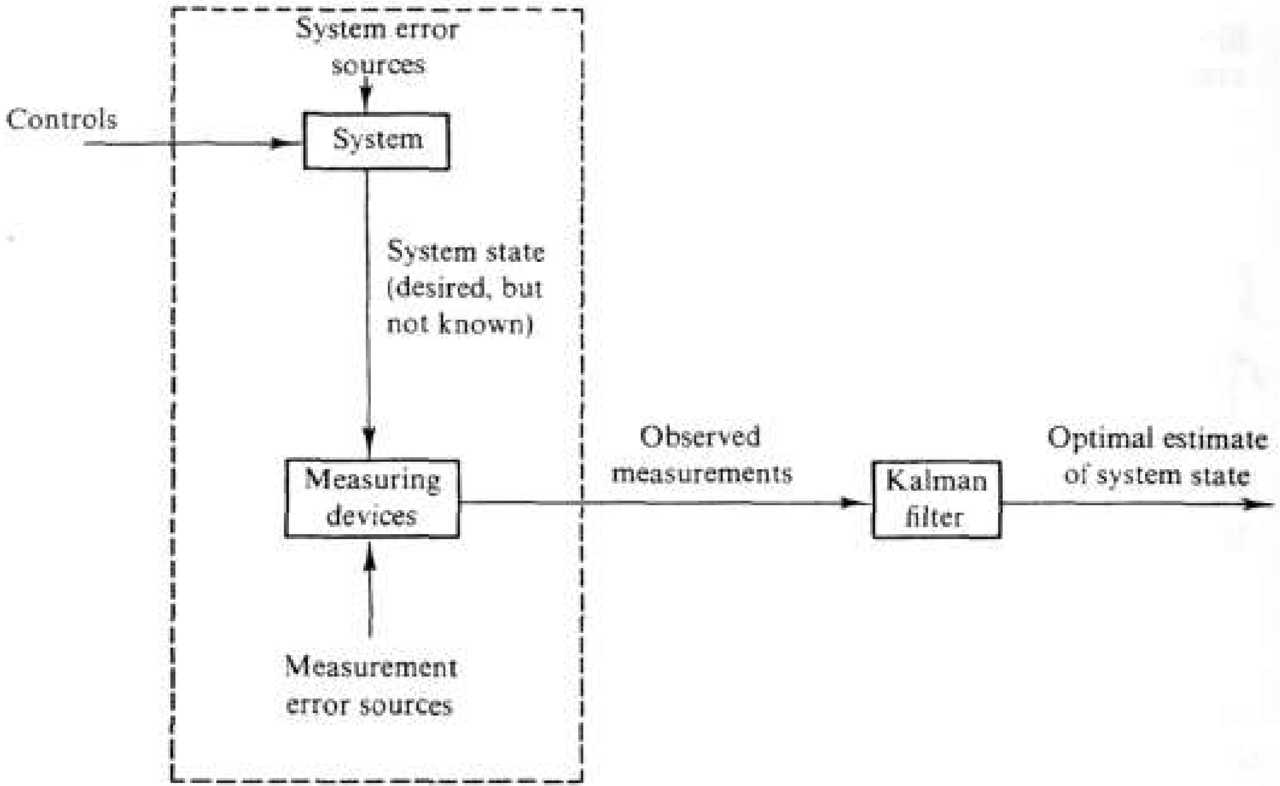
- Φk is a measurement noise (electrical, image acquision etc.,) which is also assumed to be white noise with zero mean and covariance R_k can be written as Φk = N(0, R_k). The elements of the
- measurement noise Φk are also assumed to be uncorrelated overtime, but the instantaneous non-trivial correlation is reflected in the measurement noise covariance terms of R_k, which is expressed as:

$$\mathbf{R}_k = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

Kalman filter – Concept of Measurement Models (3)

- The measurement noise is a function of particular instant in time 't'. The maximum deviation which is possible is σ, therefore the variance of measurement noise is σ^2.
- The measurement noise shows, that how much we are certain about our measurements.
  - If it is zero, then we are 100 percent sure that the measurements are correct, but in reality there is always some uncertainity which is expressed by the variance of noise of the measurement model.

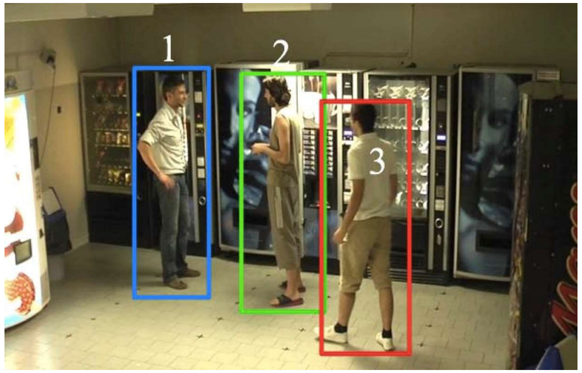
Typical Kalman filter application



Applications

http://en.wikipedia.org/wiki/Apollo_program

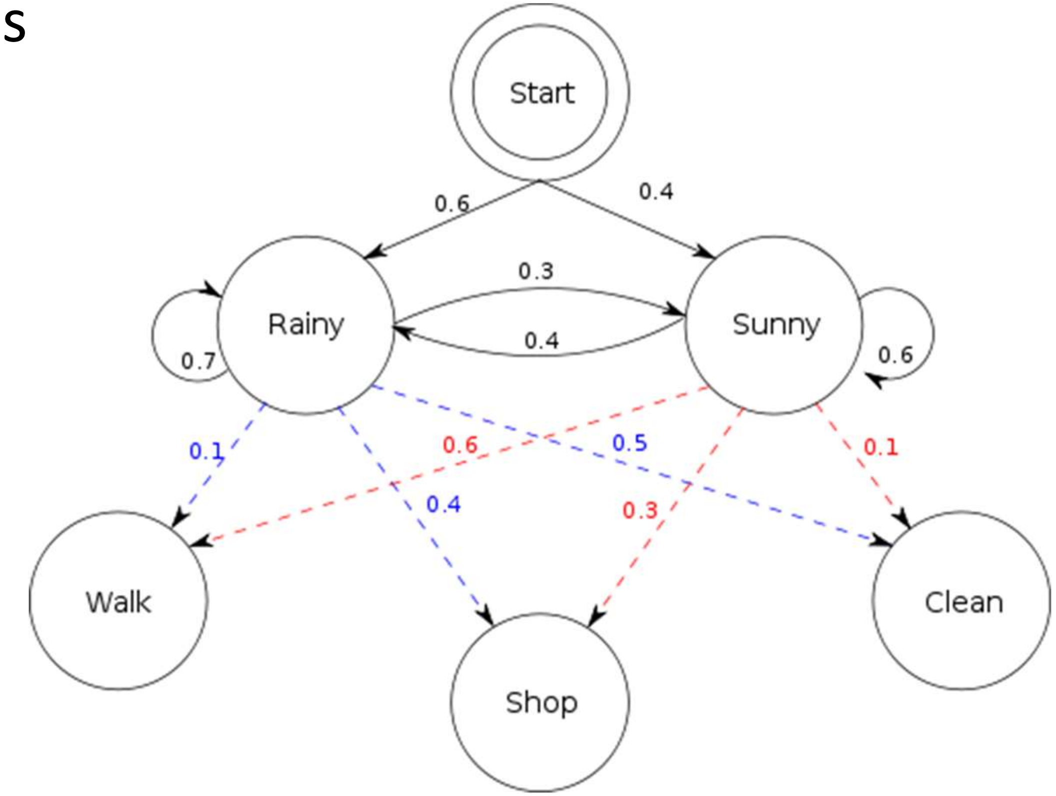
http://en.wikipedia.org/wiki/Gps



http://www.lorisbazzani.info/research-2/

Hidden Markov Model

- Markov Property :The next n+1 depends on n but not the entire past (1…n-1)
- The state is not clearly visible, but the output is visible
- The states give us information on the system.
- The task is to derive the maximum likelihood of the parameters that can predict the states



Kalman Filter

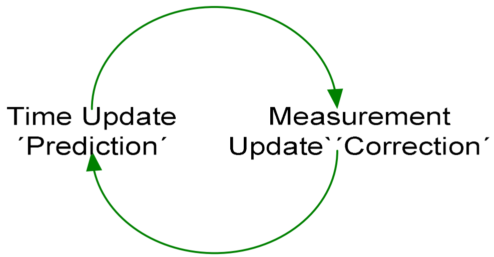
- Motion Model

  $X_k = A_{k-1} X_{k-1} + \Psi_k$

- Measurement Model

  $z_k = H_k X_k + \varphi_k$

Kalman Filter Recursive Cycle



Prediction	Correction
$x_k^- = A_{k-1}x_{k-1} + \theta_k$	$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R)^{-1}$
$P_k^- = A_k P_{k-1} A_k^T + Q_k$	$x_k = x_k^- + K_k(z_k - H_k x_k^-)$
	$P_k = (I - K_k H_k) P_k^-$

Initial Estimates
Xk-1, Pk-1

Major equation

measured value

current estimation

$$\hat{X}_k = K_k \, . \, Z_k + (1 - K_k) \, . \, \hat{X}_{k-1}$$

previous estimation

Kalman Gain

estimate error covariance

measurement variable

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1}$$

$$= \frac{P_k^- H^T}{H P_k^- H^T + R}$$

measurement error covariance

Step 1: Build a model

$$x_k = Ax_{k-1} + Bu_k + w_{k-1}$$

$$z_k = Hx_k + v_k$$

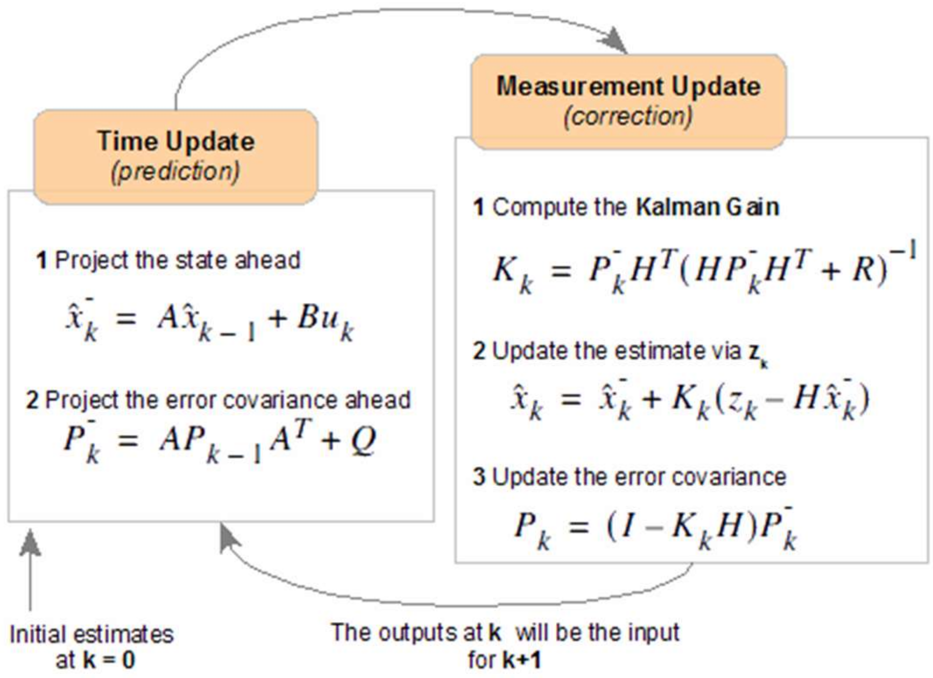
- Any $\mathbf{x_k}$ is a linear combination of its previous value plus a control signal $\mathbf{u_k}$ and a **process noise**.
- The entities **A**, **B** and **H** are in general matrices related to the states. In many cases, we can assume they are numeric value and constant.
- $\mathbf{W_{k-1}}$ is the **process noise** and $\mathbf{v_k}$ is the **measurement noise,** both are considered to be Gaussian.

Step 2: Start process

Time Update (prediction)	Measurement Update (correction)
$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$	$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$
$P_k^- = AP_{k-1}A^T + Q$	$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$
	$P_k = (I - K_k H)P_k^-$

$(z_k - H_k\widehat{x}_k^-)$ is an error between the actual measurement (z_k) and the predicted measurement $(H_k\widehat{x}_k^-)$. which is often called as forward prediction error or inovation. The kalman gain K_k gives the weight either to the actual measurement or to the predicted measurement. This blending factor K_k itself is based on the predicted error covariance P_k^-, measurement noise covariance R_k and the actual state transition matrix H_k. The actual measurement z_k is trusted more and more when the measurement noise covariance R_k approaches to zero. On the other hand as the predicted error covariance P_k approaches to zero the predicted measurements are more reliable.

Step 3: Iterate



Limiting factor

- The main problem with the Kalman filter is the estimation of the initial parameters.
- Although some parameters are updated such as the blending factor or the estimate error are updated with time, others like the process noise covariance and the measurement noise covariance need to be estimated or calculated. This is challenging since they are generally unknown

Kalman Filter

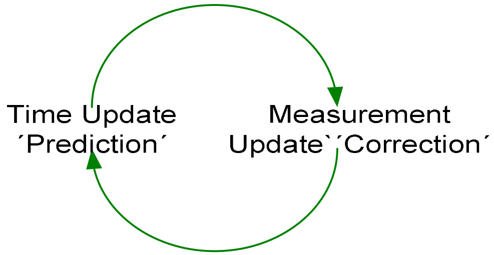
- Motion Model

  $x_k = A_{k-1} X_{k-1} + \Psi_k$

- Measurement Model

  $z_k = H_k X_k + \varphi_k$

Kalman Filter Recursive Cycle



Prediction	Correction
$x_k^- = A_{k-1} x_{k-1} + \theta_k$ $P_k^- = A_k P_{k-1} A_k^T + Q_k$	$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R)^{-1}$ $x_k = x_k^- + K_k(z_k - H_k x_k^-)$ $P_k = (I - K_k H_k) P_k^-$

Initial Estimates
Xk-1, Pk-1

Complexity Analysis of Kalman Filter (1)

The complexity of two steps i.e prediction and correction of the Kalman Filter are explaind in this section. Few complex expressions are brooken into small expressions inorder to evaluate inner complexity of each equation. Memeory Acces column have read and write operations which are mainly from external memory, intermediate expression for solving the big equation can be store in internal registers, but that is not considered during this theoritical analysis.

Expression	**Memory Access R/W**	**Mult**	**Add**	**Subtraction**	**Matrix Inversion**
$\widehat{x}_k - = A_{k-1}\widehat{x}_{k-1}$	20/4	16	12	0	0
$\mathbf{\Pi_1} = P_{k-1}\mathbf{A}_k^T$	32/16	64	48	0	0
$P_k^- = (\mathbf{A}_k\mathbf{\Pi_1}(n) + Q)$	48/16	64	64	0	0
Total	100/36	144	124	0	0

Complexity Analysis of Kalman Filter (2)

- The multiplication, additions and memory acces defining the critical path for the prediction step.
- The correction step have more equations and therefore more complex than the prediction step.
- For evaluating the Kalman Gain, inversion of a matrix of size $2 * 2$ is required whose computations are not mentioned in the table
- Simple cofactor based method for finding the inversion requires 16/13 read/write, 6 multiplications, 3 subtractions and 1 addition. The critical path for the correction step is multiplications, additions along with memory acces.
- The complexity of Kalman filter lies in the two main expressions:

Complexity Analysis of Kalman Filter (3)

- The predicted error covariance P−k in the prediction step.
- The updated error covariance Pk in the prediction step
  - The predicted error covariance P−k requires 416 operations in the prediction step, whereas updated error covariance Pk needs 260 operations dominates the overall complexity of the Kalman Filter.

Expression	Memory Access R/W	Mult	Add	Subtraction	Matrix Inversion
$\mathbf{\Pi_2} = P_k^- \mathbf{H}_k^T$	24/8	32	24	0	0
$Y1 = (\mathbf{H}_k \mathbf{\Pi_2} + R)$	16/4	16	16	0	0
$\mathbf{K}_k = \mathbf{\Pi_2} Y1^{-1}$	12/8	0	0	0	1
$\xi(n) = z_k - H_k \hat{x}_k^-$	12/2	8	6	2	0
$\hat{x_k} = \hat{x}_k^- + \mathbf{K}k\xi(n)$	14/4	8	8	0	0
$Y2 = \mathbf{K}_k \mathbf{H}_k$	16/4	32	16	0	0
$P_k = (I - Y2)P_k^-$	48/16	64	48	16	0
Total	142/50	160	118	18	1

Example 1 (1/4)

- Estimate a random constant:” voltage” reading from a source.
- It has a constant value of **a**V (volts), so there is no control signal $\mathbf{u}_k$. Standard deviation of the measurement noise is 0.1 V.
- It is a 1 dimensional signal problem: **A** and **H** are constant 1.
- Assume the error covariance $\mathbf{P}_0$ is initially 1 and initial state $\mathbf{X}_0$ is 0.

Example 1 (2/4)

$$x_k = Ax_{k-1} + Bu_k + w_k$$
$$= x_{k-1} + w_k$$

$$z_k = Hx_k + v_k$$
$$= x_k + v_k$$

Time	1	2	3	4	5	6	7	8	9	10
Value	0.39	0.50	0.48	0.29	0.25	0.32	0.34	0.48	0.41	0.45

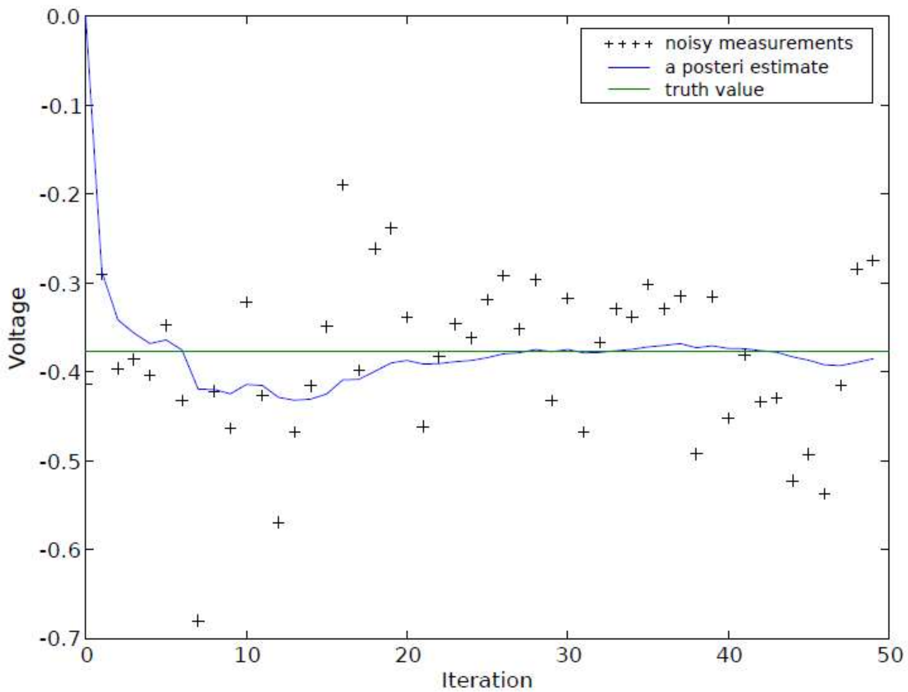
Example 1 (3/4)

Time Update *(prediction)*	**Measurement Update** *(correction)*
$\hat{x}_k^- = \hat{x}_{k-1}$ $P_k^- = P_{k-1}$	$K_k = \dfrac{P_k^-}{P_k^- + R}$ $\hat{x}_k = \hat{x}_k^- + K_k(z_k - \hat{x}_k^-)$ $P_k = (1 - K_k)P_k^-$

Example 1 (4/4)

k	1	2	3	4	5	6	7	8	9	10
z_k	0.390	0.500	0.480	0.290	0.250	0.320	0.340	0.480	0.410	0.450
$\hat{x}_{k-1}$	0	0.355	0.424	0.442	0.405	0.375	0.365	0.362	0.377	0.380
P_k^-	1	0.091	0.048	0.032	0.024	0.020	0.016	0.014	0.012	0.011
Time Update	$\hat{x}_k^- = \hat{x}_{k-1} = 0$ $P_k^- = P_{k-1} = 1$	$\hat{x}_k^- = 0.355$ $P_k^- = 0.091$								
Measurement Update	$K_k = 1 / (1\ 0.1)$ $= 0.909$ $\hat{x}_k = 0\ 0.909\ (0.390 - 0)$ $= 0.35$ $P_k = (1 - 0.909)\ .\ 1$ $= 0.091$	$K_k = 0.091 / (0.091\ 0.1)$ $= 0.476$ $\hat{x}_k = 0.355\ 0.476\ (0.500 - 0.355)$ $= 0.424$ $P_k = (1 - 0.476)\ .\ 0.091$ $= 0.048$								
$\hat{x}_k$	**0.355**	**0.424**	**0.442**	**0.405**	**0.375**	**0.365**	**0.362**	**0.377**	**0.380**	**0.387**
P_k	0.091	0.048	0.032	0.024	0.020	0.016	0.014	0.012	0.011	0.010

Result of the example 1



The Extended Kalman filter

- In simple cases, such as the linear dynamical system just, exact inference is tractable; however, in general, exact inference is infeasible, and approximate methods must be used, such as the extended Kalman filter.
- Unlike its linear counterpart, the extended Kalman filter in general is not an optimal estimator

$$x_k = f(x_{k-1}, u_k, w_{k-1})$$

Properties and conclusion

- If all noise is Gaussian, the Kalman filter minimizes the mean square error of the estimated parameters
- Convenient for online real time processing.
- Easy to formulate and implement given a basic understanding.
- To enable the convergence in fewer steps:
  - Model the system more precisely
  - Estimate the noise more precisely

Reference

- *Lindsay Kleeman, **Understanding and Applying Kalman Filtering**,* Department of Electrical and Computer Systems Engineering, Monash University, Clayton
- *Peter Maybeck, **Stochastic Models, Estimation, and Control**, Volume 1*
- Greg Welch, Gary Bishop, "**An Introduction to the Kalman Filter**", *University of North Carolina at Chapel Hill Department of Computer Science, 2001*

Thank you