Author:

Magdalena Szymkowiak

magdalena.szymkowiak@put.poznan.pl

Kalman filter

1 Motivation

1.1 Introduction

[**Rudolf Emil Kalman** (1930–2016) a Hungarian-born, American electrical engineer and mathematician]
[**Richard S. Bucy** (born 1935) an American mathematician]

References

- **Kalman, R.E.** *A New Approach to Linear Filtering and Prediction Problems*, Transaction of the ASME (American Society of Mechanical Engineers), Journal of Basic Engineering, pp. 35–45, 1960
- **Kalman, R.E.; Bucy, R.S.** *New Results in Linear Filtering and Prediction Theory*, Transaction of the ASME, Journal of Basic Engineering, pp. 95–108, 1961
- P. S. Maybeck, *Stochastic models, estimation and control. Volume 1*, Department of Electrical and Computer Engineering, Air Force Institute of Technology, Wright-Patterson Air Force Base Ohio, 1979
- R. Negenborn, *Robot Localization and Kalman Filters. On finding your position in a noisy world*, Institute of Information and Computing Sciences in partial fulfilment of the requirements for the degree of Master of Science, specialized in Intelligent Systems, 2003
- G. Welch, G. Bishop, *An Introduction to the Kalman Filter*, University of North Carolina at Chapel Hill Department of Computer Science Chapel Hill, NC 27599-3175, 2006

Goal

Kalman filter KF – a **recursive algorithm** solving
the **data filtering problem**

- **integrates**
  - **all available measured data** (received from sensors)
  - **prior knowledge** about a state transition and measuring models
- **produces an estimate of required states** (variables)
- **minimizes the mean of the squared errors**

Application

- **radar tracking** of airborne targets
- **control systems** (with full state measurement)
- **tracing moving objects** (on the basis of limited observations)
- **localization** (data fusion) and **navigation** (autopilot)
- **mapping the environment** (on the basis of noisy data)

1.2 Assumption

Basic assumption

- $\{\mathbf{X}(t)\}$ is an n-dimensional **Gauss-Markov sequence**
  - all **states distributions are Gaussian** ($\mathbb{S} = \mathbb{R}$ is uncountable)
  - sequence $\{\mathbf{X}(t)\}$ has **Markov properties**
- **KF** is a **Bayes filter** with given
  - an initial state prior distribution $\mathbf{X}(t_0)$
  - state transition model $\mathbf{X}(t_k)/\mathbf{X}(t_{k-1})$
  - measurement model $\mathbf{Z}(t_k)/\mathbf{X}(t_k)$
- state transition model and measurement model are **linear**

Normal (Gaussian) distribution $X \sim N(\mu, \sigma)$

Parameters

$$\mu \in \mathbb{R}, \sigma \in \mathbb{R}_+$$

Probability density function *pdf*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad \text{for} \quad x \in (-\infty, +\infty)$$

Expectation

$$EX = \mu$$

For $N(\mu,\ \sigma)$, $EX = \mu$ is the **expected value (expected state)**.

Variance, standard deviation

$$D^2X = \sigma^2, \qquad DX = \sigma$$

Variance σ^2 and standard deviation σ are measures of uncertainty !
Smaller σ^2 (so σ), smaller uncertainty of the expected state!

We assume that

- $\{\mathbf{X}(t)\}$ is an n-dimensional Gauss-Markov sequence

  so for each time-step $t_k \in \mathbb{T}$

  $\mathbf{X}(t_k) = [X_1(t_k), X_2(t_k), ..., X_n(t_k)]^T \sim N_n(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ where

  $\forall_{i=1}^n \quad X_i(t_k) \sim N(\mu_{k_i}, \sigma_{k_i})$

so $\quad \boldsymbol{\mu}_k = [\mu_{k_1}, \mu_{k_2}, ..., \mu_{k_n}]^T \quad$ is the **expected value (expected state)**

and $\quad \boldsymbol{\Sigma}_k = \begin{bmatrix} \sigma_{k_1}^2 & \sigma_{k_{12}} & ... & \sigma_{k_{1n}} \\ \sigma_{k_{21}} & \sigma_{k_2}^2 & ... & \sigma_{k_{2n}} \\ ... & ... & ... & ... \\ \sigma_{k_{n1}} & \sigma_{k_{n2}} & ... & \sigma_{k_n}^2 \end{bmatrix}$ **is a measure of uncertainty !**

Remark 1. Smaller diagonal elements of $\Sigma_{\mathbf{X}}$, smaller uncertainty of the expected state!

For 1-dimensional Normal variable

- $X \sim N_1(\mu,\ \Sigma)$ denotes $X \sim N_1(\mu,\ \sigma^2)$
- $X \sim N_1(\mu,\ \Sigma) \Leftrightarrow X \sim N(\mu,\ \sigma)$
- *pdf*
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

For 2-dimensional normal variable $\mathbf{X} = (X_1, X_2) \sim N_2(\boldsymbol{\mu}_{\mathbf{X}},\ \boldsymbol{\Sigma}_{\mathbf{X}})$

- $X_1 \sim N(\mu_1,\ \sigma_1)$, $X_2 \sim N(\mu_2,\ \sigma_2)$ and $\rho = \frac{Cov(X_1\ X_2)}{\sigma_1 \sigma_2}$
- $\boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, $\boldsymbol{\Sigma}_{\mathbf{X}} = \begin{bmatrix} \sigma_1^2 & Cov(X_1\ X_2) \\ Cov(X_2\ X_1) & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$
- *pdf*
$$f_{\mathbf{X}}(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

Properties of Expectation and (Co)Variance

- for 1-dimmensional random variables X, Y

  e1 $E(aX + b) = aE(X) + b$

  d1 $D^2(aX + b) = a^2 D^2(X) \quad \Rightarrow \quad D^2(X + b) = D^2(X)$

  e3 $E(X + Y) = EX + EY$

- for INDEPENDENT 1-dimmensional random variables X, Y

  i2 $D^2(X+Y) = D^2X + D^2Y$

- for random vector $\mathbf{X}$ with $\boldsymbol{\mu}_\mathbf{X} = E(\mathbf{X})$ and $\boldsymbol{\Sigma}_\mathbf{X} = Cov(\mathbf{X}\ \mathbf{X}^T)$

  n1 $E(\boldsymbol{A}\mathbf{X} + \mathbf{b}) = \boldsymbol{A}E(\mathbf{X}) + \mathbf{b} = \boldsymbol{A}\boldsymbol{\mu}_\mathbf{X} + \mathbf{b}$

  n2 $\boldsymbol{\Sigma}_{\boldsymbol{A}\mathbf{X}+\mathbf{b}} = \boldsymbol{A}\boldsymbol{\Sigma}_\mathbf{X}\boldsymbol{A}^T \quad \Rightarrow \quad \boldsymbol{\Sigma}_{\mathbf{X}+\mathbf{b}} = \boldsymbol{\Sigma}_\mathbf{X}$

  n3 for n-dimensional Gaussian random vector $\mathbf{X} \sim N_n(\boldsymbol{\mu}_\mathbf{X},\ \boldsymbol{\Sigma}_\mathbf{X})$ $\boldsymbol{A}\mathbf{X} + \mathbf{b} \sim N_n(\boldsymbol{A}\boldsymbol{\mu}_\mathbf{X} + \mathbf{b},\ \boldsymbol{A}\boldsymbol{\Sigma}_\mathbf{X}\boldsymbol{A}^T)$

- for two random vectors $\mathbf{X}$ and $\mathbf{Y}$

  n4 $E(\mathbf{X}+\mathbf{Y}) = E(\mathbf{X}) + E(\mathbf{Y})$

- for two INDEPENDENT random vectors $\mathbf{X}$ and $\mathbf{Y}$

  n5 $\boldsymbol{\Sigma}_{\mathbf{X}+\mathbf{Y}} = \boldsymbol{\Sigma}_\mathbf{X} + \boldsymbol{\Sigma}_\mathbf{Y}$

- for two INDEPENDENT n-dimensional Gaussian random vectors $\mathbf{X} \sim N_n(\boldsymbol{\mu}_\mathbf{X},\ \boldsymbol{\Sigma}_\mathbf{X})$, $\mathbf{Y} \sim N_n(\boldsymbol{\mu}_\mathbf{Y},\ \boldsymbol{\Sigma}_\mathbf{Y})$

  n6 $\mathbf{X}+\mathbf{Y} \sim N_n(\boldsymbol{\mu}_\mathbf{X} + \boldsymbol{\mu}_\mathbf{Y},\ \boldsymbol{\Sigma}_\mathbf{X} + \boldsymbol{\Sigma}_\mathbf{Y})$

Moreover, we assume that

- $\{\mathbf{Z}(t)\}$ is a set of m-dimensional noisy measurements
- $\{\mathbf{U}(t)\}$ is an optional set of l-dimensional control inputs

Methodology

We try to **estimate distribution** of $\mathbf{X}(t_k)/\mathbf{X}(t_{k-1})$

using, determined at each time-step t_k

- **linear** state transition recursive equation

$$\mathbf{x}_k = \boldsymbol{A}\ \mathbf{x}_{k-1} + \boldsymbol{B}\ \mathbf{u}_{k-1} + \mathbf{w}_{k-1}$$

- **linear** measurement equation $$\mathbf{z}_k = \boldsymbol{C}\ \mathbf{x}_k + \mathbf{v}_k$$

2 Kalman filter KF

Linear equation – deterministic part

There is a **deterministic part** in

- linear state transition recursive equation
$$\mathbf{x}_k = \boldsymbol{A}\ \mathbf{x}_{k-1} + \boldsymbol{B}\ \mathbf{u}_{k-1} + \mathbf{w}_{k-1}$$
- linear measurement equation
$$\mathbf{z}_k = \boldsymbol{C}\ \mathbf{x}_k + \mathbf{v}_k$$
- $\boldsymbol{A}$ is $n \times n$-dimensional **transition matrix** relating
  - prior state $\mathbf{x}_{k-1}$
  - to current state $\mathbf{x}_k$
- $\boldsymbol{B}$ is an **optional** $n \times l$-dimensional **control input matrix** relating
  - previous optional control input $\mathbf{u}_{k-1}$
  - to current state $\mathbf{x}_k$
- $\boldsymbol{C}$ is $m \times n$-dimensional **measurement matrix** relating
  - current state $\mathbf{x}_k$
  - to current measurement $\mathbf{z}_k$
- in practice matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ can change with each time-step t_k, but here are assumed to be **constant**

Linear equation – random part

There is a **random part** in

- linear state transition recursive equation
$$\mathbf{x}_k = \boldsymbol{A}\ \mathbf{x}_{k-1} + \boldsymbol{B}\ \mathbf{u}_{k-1} + \mathbf{w}_{k-1}$$
- linear measurement equation
$$\mathbf{z}_k = \boldsymbol{C}\ \mathbf{x}_k + \mathbf{v}_k$$
- $\mathbf{w}_{k-1}$ is a **previous state** of n-dimensional **process noise** $\mathbf{W}(t_{k-1}) \sim N_n(\mathbf{0}, \boldsymbol{Q})$ representing n-dimensional **zero-mean white Gaussian noise** $\{\mathbf{W}(t)\}$
- $\mathbf{v}_k$ is a **current state** of m-dimensional **measurement noise** $\mathbf{V}(t_k) \sim N_m(\mathbf{0}, \boldsymbol{R})$ representing m-dimensional **zero-mean white Gaussian noise** $\{\mathbf{V}(t)\}$
- in practice matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ can change with each time-step t_k, but here are assumed to be **constant**

Stationary zero-mean white Gaussian noise

So $\{\mathbf{W}(t)\}$ and $\{\mathbf{V}(t)\}$ are

- **independent (of each other)**
- **stationary**
- respectively, n-dimensional and m-dimensional, **zero-mean white Gaussian noises**

- $\mathbf{W}(t_k) \sim N_n(\mathbf{0}, \boldsymbol{Q})$ for each time-step $t_k \in \mathbb{T}$ and
- for any different time-steps $t_i, t_j \in \mathbb{T}$,
  $\mathbf{W}(t_i), \mathbf{W}(t_j)$ are **pairwise independent**

- $\mathbf{V}(t_k) \sim N_m(\mathbf{0}, \boldsymbol{R})$ for each time-step $t_k \in \mathbb{T}$ and
- for any different time-steps $t_i, t_j \in \mathbb{T}$,
  $\mathbf{V}(t_i), \mathbf{V}(t_j)$ are **pairwise independent**

Bayes filter cycle

Operations of **Bayes estimation filter**

- INITIALIZATION – a given initial prior distribution $\mathbf{X}(t_0)$ of state $\mathbf{x}_0$
- PREDICTION – time-step update
  - having old prior distribution $\mathbf{X}(t_{k-1})/\boldsymbol{Z}(t_{1:k-1})$ of state $\mathbf{x}_{k-1}$ given measurements $\boldsymbol{z}_{1:k-1}$ (and optionally, distribution $\boldsymbol{U}(t_{1:k-1})$ of control inputs $\boldsymbol{u}_{1:k-1}$)
  - using state transition model
  - we obtain predicted distribution $\mathbf{X}(t_k)/\boldsymbol{Z}(t_{1:k-1})$ of predicted state $\mathbf{x}_k$ given measurements $\boldsymbol{z}_{1:k-1}$
- CORRECTION – measurement update
  - given predicted distribution $\mathbf{X}(t_k)/\boldsymbol{Z}(t_{1:k-1})$
  - using a measurement model
  - we obtain corrected posterior distribution $\mathbf{X}(t_k)/\boldsymbol{Z}(t_{1:k})$ of state $\mathbf{x}_k$ given measurements $\boldsymbol{z}_{1:k}$

Operations of **KF estimation**

- INITIALIZATION – a given initial prior distribution $\mathbf{X}(t_0) \sim N_n(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ of state $\mathbf{x}_0$
- PREDICTION – time-step update
  - having prior distribution $\mathbf{X}(t_{k-1})/\boldsymbol{Z}(t_{1:k-1}) \sim N_n(\boldsymbol{\mu}_{k-1}, \boldsymbol{\Sigma}_{k-1})$ of state $\mathbf{x}_{k-1}$ given measurements $\boldsymbol{z}_{1:k-1}$
    (and optionally, distribution $\boldsymbol{U}(t_{1:k-1})$ of control inputs $\boldsymbol{u}_{1:k-1}$)
  - using linear state transition recursive equation $\mathbf{x}_k = \boldsymbol{A}\ \mathbf{x}_{k-1} + \boldsymbol{B}\ \mathbf{u}_{k-1} + \mathbf{w}_{k-1}$
  - we obtain predicted distribution $\mathbf{X}(t_k)/\boldsymbol{Z}(t_{1:k-1}) \sim N_n(\boldsymbol{\mu}_k^-, \boldsymbol{\Sigma}_k^-)$ of predicted state $\mathbf{x}_k$ given measurements $\boldsymbol{z}_{1:k-1}$
- CORRECTION – measurement update
  - given predicted distribution $\mathbf{X}(t_k)/\boldsymbol{Z}(t_{1:k-1})$
  - using linear measurement equation $\mathbf{z}_k = \boldsymbol{C}\ \mathbf{x}_k + \mathbf{v}_k$
  - we obtain posterior distribution $\mathbf{X}(t_k)/\boldsymbol{Z}(t_{1:k}) \sim N_n(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ of state $\mathbf{x}_k$ given measurements $\boldsymbol{z}_{1:k}$

INITIALIZATION, PREDICTION, CORRECTION

INITIALIZATION – prior distribution

- $\mathbf{X}(t_{k-1})/\boldsymbol{Z}(t_{1:k-1}) \sim N_n(\boldsymbol{\mu}_{k-1}, \boldsymbol{\Sigma}_{k-1})$
- $\boldsymbol{\mu}_{k-1} = \hat{\mathbf{x}}_{k-1}$
- $\boldsymbol{\Sigma}_{k-1} = \boldsymbol{P}_{k-1}$

at t_0, given initial distribution $\mathbf{X}(t_0) \sim N_n(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ of state $\mathbf{x}_0$

PREDICTION – linear state transition recursive equation

- $\mathbf{x}_k = \boldsymbol{A}\ \mathbf{x}_{k-1} + \boldsymbol{B}\ \mathbf{u}_{k-1} + \mathbf{w}_{k-1}$, where $\mathbf{W}(t_{k-1}) \sim N_n(\mathbf{0}, \boldsymbol{Q})$

predicted distribution

- $\mathbf{X}(t_k)/\boldsymbol{Z}(t_{1:k-1}) \sim N_n(\boldsymbol{\mu}_k^-, \boldsymbol{\Sigma}_k^-)$
- $\boldsymbol{\mu}_k^- = \hat{\mathbf{x}}_k^- = \boldsymbol{A}\hat{\mathbf{x}}_{k-1} + \boldsymbol{B}\mathbf{u}_{k-1}$
- $\boldsymbol{\Sigma}_k^- = \boldsymbol{P}_k^- = \boldsymbol{A}\boldsymbol{P}_{k-1}\boldsymbol{A}^T + \boldsymbol{Q}$

CORRECTION – linear measurement equation

- $\mathbf{z}_k = \boldsymbol{C}\ \mathbf{x}_k + \mathbf{v}_k$, where $\mathbf{V}(t_k) \sim N_m(\mathbf{0}, \boldsymbol{R})$
- measurement likelihood $\quad N_m(\mathbf{z}_k,\ \boldsymbol{R})$

posterior distribution

- $\mathbf{X}(t_k)/\boldsymbol{Z}(t_{1:k}) \sim N_n(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$

- **Kalman gain** – an $n \times m$-dimensional matrix
  $\boldsymbol{K}_k = \boldsymbol{P}_k^- \boldsymbol{C}^T \left(\boldsymbol{C} \boldsymbol{P}_k^- \boldsymbol{C}^T + \boldsymbol{R} \right)^{-1}$ allows us to determine
- $\boldsymbol{\mu}_k = \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \boldsymbol{K}_k \left(\mathbf{z}_k - \boldsymbol{C} \hat{\mathbf{x}}_k^- \right)$
- $\boldsymbol{\Sigma}_k = \boldsymbol{P}_k = \left(\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{C} \right) \boldsymbol{P}_k^-$

$\boldsymbol{K}_k$ is a factor that minimizes posterior error covariance $\boldsymbol{\Sigma}_k$
(**minimizes uncertainty of the posterior state!**).

Posterior state estimate:

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \boldsymbol{K}_k (\mathbf{z}_k - \boldsymbol{C} \hat{\mathbf{x}}_k^-)$$

- $\mathbf{z}_k - \boldsymbol{C} \hat{\mathbf{x}}_k^-$ – the measurement **innovation** (the **residual**)
  – the difference between the **actual measurement** and the **predicted measurement**
- **Kalman gain** $\boldsymbol{K}_k = \boldsymbol{P}_k^- \boldsymbol{C}^T (\boldsymbol{C} \boldsymbol{P}_k^- \boldsymbol{C}^T + \boldsymbol{R})^{-1} = \frac{\boldsymbol{P}_k^- \boldsymbol{C}^T}{\boldsymbol{C} \boldsymbol{P}_k^- \boldsymbol{C}^T + \boldsymbol{R}}$

  $\boldsymbol{K}_k$ is a **weight of FILTERING influence** on state estimate

- if **covariance of measurement noise** $\boldsymbol{R} \to \mathbf{0}$ (**measurements are reliable**) then $\boldsymbol{K}_k \to \boldsymbol{C}^{-1}$
  Kalman gain $\boldsymbol{K}_k$ has **stronger influence** on the residual
  - the **actual measurement** $\mathbf{z}_k$ **is trusted more and more**
  - the predicted measurement $\boldsymbol{C} \hat{\mathbf{x}}_k^-$ is trusted less and less
- if **covariance of prediction errors** $\boldsymbol{P}_k^- = \boldsymbol{\Sigma}_k^- \to \mathbf{0}$
  (correction is reliable) then $\boldsymbol{K}_k \to \mathbf{0}$
  Kalman gain $\boldsymbol{K}_k$ has **weaker influence** on the residual
  - the actual measurement $\mathbf{z}_k$ is trusted less and less
  - the **predicted measurement** $\boldsymbol{C} \hat{\mathbf{x}}_k^-$ **is trusted more and more**

KF Convergence

The **Kalman filter** dynamics results from the **consecutive cycles** of

- prediction
- correction (filtering)

If the **system dynamics is time-invariant** ($\boldsymbol{A}, \boldsymbol{B}$ and $\boldsymbol{C}$ are constant) and the measurement and process noises are **stationary** ($\boldsymbol{Q}$ and $\boldsymbol{R}$ are constant) then

- **KF** dynamics **converges to a steady-state**
- Posterior covariance matrix $\boldsymbol{P}_k$ **converges to a symmetric positive definite matrix** $\overline{\boldsymbol{P}}$ (it stabilizes)

- **Kalman gain $\mathbf{K}_k$ converges to a steady-state gain $\overline{\mathbf{K}}$**
  (it stabilizes)

It means that the system model is **completely observable and controllable**.

Estimate errors

State at time-step t_k

- $\mathbf{x}_k = [x_{k_1}, x_{k_2}, \cdots, x_{k_n}]^T \in \mathbb{S} \subset \mathbb{R}^n$ is n-dimensional **state vector** – a value of random variable $\mathbf{X}(t_k)$ (we try to estimate)
- $\hat{\mathbf{x}}_k^-$ – n-dimensional **state estimate after prediction**
- $\hat{\mathbf{x}}_k$ – n-dimensional posterior **state estimate after correction**

Estimate errors

- $\mathbf{e}_k^- = \mathbf{X}(t_k) - \hat{\mathbf{x}}_k^-$ – **estimate error after prediction**
- $\mathbf{e}_k = \mathbf{X}(t_k) - \hat{\mathbf{x}}_k$ – **posterior estimate error** after correction

Expectation of estimate errors

- **after prediction**
  $E[\mathbf{X}(t_k)] = \boldsymbol{\mu}_k^- = \hat{\mathbf{x}}_k^-$
  so $E[\mathbf{e}_k^-] = E[\mathbf{X}(t_k) - \hat{\mathbf{x}}_k^-] = E[\mathbf{X}(t_k)] - E[\hat{\mathbf{x}}_k^-] = 0$
- **after correction**
  $E[\mathbf{X}(t_k)] = \boldsymbol{\mu}_k = \hat{\mathbf{x}}_k$
  so $E[\mathbf{e}_k] = E[\mathbf{X}(t_k) - \hat{\mathbf{x}}_k] = E[\mathbf{X}(t_k)] - E[\hat{\mathbf{x}}_k] = 0$

Because for n-dimensional random vector $\mathbf{X}$,

$\boldsymbol{\Sigma}_{\mathbf{X}+\mathbf{b}} = \boldsymbol{\Sigma}_{\mathbf{X}}$ then **covariance of estimate errors**

- **after prediction**
  $\boldsymbol{\Sigma}_k^- = \boldsymbol{P}_k^- = \boldsymbol{\Sigma}_{\mathbf{X}(t_k)} = \boldsymbol{\Sigma}_{\mathbf{e}_k^- + \hat{\mathbf{x}}_k} = \boldsymbol{\Sigma}_{\mathbf{e}_k^-} = Cov(\mathbf{e}_k^- \ \mathbf{e}_k^{-T}) = E[\mathbf{e}_k^- \ \mathbf{e}_k^{-T}] - E[\mathbf{e}_k^-] \ E[\mathbf{e}_k^{-T}] = E[\mathbf{e}_k^- \ \mathbf{e}_k^{-T}]$
- after correction (posterior error covariance)
  $\boldsymbol{\Sigma}_k = \boldsymbol{P}_k = \boldsymbol{\Sigma}_{\mathbf{X}(t_k)} = \boldsymbol{\Sigma}_{\mathbf{e}_k + \hat{\mathbf{x}}_k} = \boldsymbol{\Sigma}_{\mathbf{e}_k} = Cov(\mathbf{e}_k \ \mathbf{e}_k^T) = E[\mathbf{e}_k \ \mathbf{e}_k^T] - E[\mathbf{e}_k] \ E[\mathbf{e}_k^T] = E[\mathbf{e}_k \ \mathbf{e}_k^T]$

Covariance of estimate errors

Covariance of estimate error after prediction

$$\boldsymbol{\Sigma}_k^- = \boldsymbol{P}_k^- = Cov(\mathbf{e}_k^- \ \mathbf{e}_k^{-T}) = \boldsymbol{\Sigma}_{\mathbf{e}_k^-} = \boldsymbol{\Sigma}_{\mathbf{x}_k - \hat{\mathbf{x}}_k^-}$$
$$= \boldsymbol{\Sigma}_{\boldsymbol{A} \ \mathbf{x}_{k-1} + \boldsymbol{B} \ \mathbf{u}_{k-1} + \mathbf{w}_{k-1} - (\boldsymbol{A} \ \hat{\mathbf{x}}_{k-1} + \boldsymbol{B} \ \mathbf{u}_{k-1})} = \boldsymbol{\Sigma}_{\boldsymbol{A}(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}) + \mathbf{w}_{k-1}} = (*)$$

variables $\mathbf{e}_{k-1}^- = \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}$ and $\mathbf{w}_{k-1}$ are independent, so

$$(*) = \boldsymbol{A}\boldsymbol{\Sigma}_{\mathbf{e}_{k-1}^-}\boldsymbol{A}^T + \boldsymbol{\Sigma}_{\mathbf{w}_{k-1}} = \boldsymbol{A}\boldsymbol{P}_{k-1}^-\boldsymbol{A}^T + \boldsymbol{Q}$$

Covariance of estimate error after correction (posterior error covariance)

$$\boldsymbol{\Sigma}_k = \boldsymbol{P}_k = Cov(\mathbf{e}_k\ \mathbf{e}_k^T) = \boldsymbol{\Sigma}_{\mathbf{e}_k} = \boldsymbol{\Sigma}_{\mathbf{x}_k - \hat{\mathbf{x}}_k}$$
$$= \boldsymbol{\Sigma}_{\mathbf{x}_k - [\hat{\mathbf{x}}_k^- + \boldsymbol{K}_k(\mathbf{z}_k - \boldsymbol{C}\hat{\mathbf{x}}_k^-)]} = \boldsymbol{\Sigma}_{\mathbf{x}_k - \hat{\mathbf{x}}_k^- - \boldsymbol{K}_k(\boldsymbol{C}\mathbf{x}_k + \mathbf{v}_k - \boldsymbol{C}\hat{\mathbf{x}}_k^-)}$$
$$= \boldsymbol{\Sigma}_{(\mathbf{x}_k - \hat{\mathbf{x}}_k^-) - \boldsymbol{K}_k\boldsymbol{C}(\mathbf{x}_k - \hat{\mathbf{x}}_k^-) - \boldsymbol{K}_k\mathbf{v}_k} = \boldsymbol{\Sigma}_{(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}) - \boldsymbol{K}_k\mathbf{v}_k} = (**)$$

variables $\mathbf{e}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k$ and $\mathbf{v}_k$ are independent, so

$$(**) = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C})\boldsymbol{\Sigma}_{\mathbf{e}_k^-}(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C})^T + \boldsymbol{\Sigma}_{\mathbf{v}_k} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C})\boldsymbol{P}_k^-(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C})^T + \boldsymbol{R}$$

Klaman gain $\boldsymbol{K}_k = \boldsymbol{P}_k^-\boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{P}_k^-\boldsymbol{C}^T + \boldsymbol{R})^{-1}$
is chosen to be a factor that minimizes posterior error covariance $\boldsymbol{\Sigma}_k$
(**minimizes uncertainty !** of the posterior state).

KF parameters and tuning

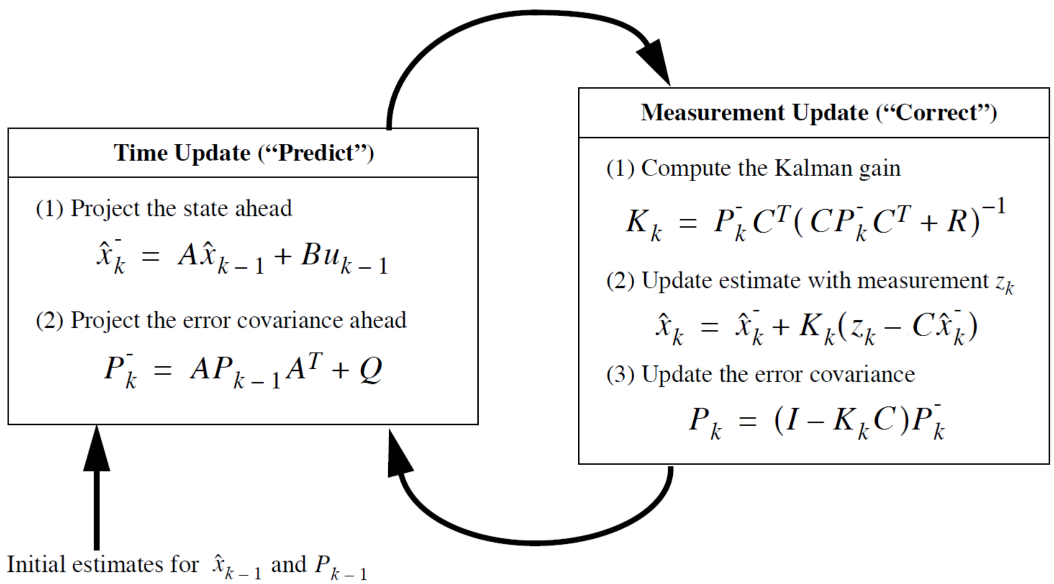
Initial conditions choice

- $\hat{\mathbf{x}}_0$ - on the basis of a priori knowledge or $\hat{\mathbf{x}}_0 = E[X(0)] = \boldsymbol{\mu}_0$
- $\boldsymbol{P}_0$ - a positive-definite matrix or $\boldsymbol{P}_0 = Cov(X(0)\ X(0)^T) = \boldsymbol{\Sigma}_0$
  (if P is chosen to small - filter converge slowly)

Tuning – process noise covariance matrix $\boldsymbol{Q}$ and measurement noise covariance matrix $\boldsymbol{R}$ should be considered as **tunable parameters**, as they are usually unknown in practice

- $\boldsymbol{Q}$ is chosen as a positive-definite matrix (usually diagonal)
  – in practice $\boldsymbol{Q}^{-1}$ represents the **confidence in the trusted** model
- $\boldsymbol{R}$ is chosen as an empirical covariance of noise (based on measurements prior to operation of the filter) or as a positive-definite matrix
  – in practice $\boldsymbol{R}^{-1}$ represents the **confidence in the measurements**
- if $\boldsymbol{Q}$ and $\boldsymbol{R}$ are constant - both $\boldsymbol{P}_k$ and $\boldsymbol{K}_k$ will stabilize quickly and remain constant (however, often $\boldsymbol{R}$ is not constant)
- elements $\boldsymbol{P}_k^-$, $\boldsymbol{K}_k$, $\boldsymbol{P}_k$ do not depend on measurements, therefore they can be calculated off-line

Kalman filter
Basic formulas



Remark 2. *For* 1*-dimensional Normal variable*

- $X \sim N_1(\mu,\ \Sigma)$ *denotes* $X \sim N_1(\mu,\ \sigma^2)$
- $X \sim N_1(\mu,\ \Sigma) \Leftrightarrow X \sim N(\mu,\ \sigma)$
- *pdf*
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

Remark 3. *For* 2*-dimensional normal variable* $\mathbf{X} = (X_1, X_2) \sim N_2(\boldsymbol{\mu}_{\mathbf{X}},\ \boldsymbol{\Sigma}_{\mathbf{X}})$

- $X_1 \sim N(\mu_1,\ \sigma_1)$, $X_2 \sim N(\mu_2,\ \sigma_2)$ *and* $\rho = \frac{Cov(X_1\ X_2)}{\sigma_1\sigma_2}$
- $\boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, $\boldsymbol{\Sigma}_{\mathbf{X}} = \begin{bmatrix} \sigma_1^2 & Cov(X_1\ X_2) \\ Cov(X_2\ X_1) & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$
- *pdf*
$$f_{\mathbf{X}}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$