Introduction to Kalman Filtering

A set of two lectures

Maria Isabel Ribeiro

Associate Professor

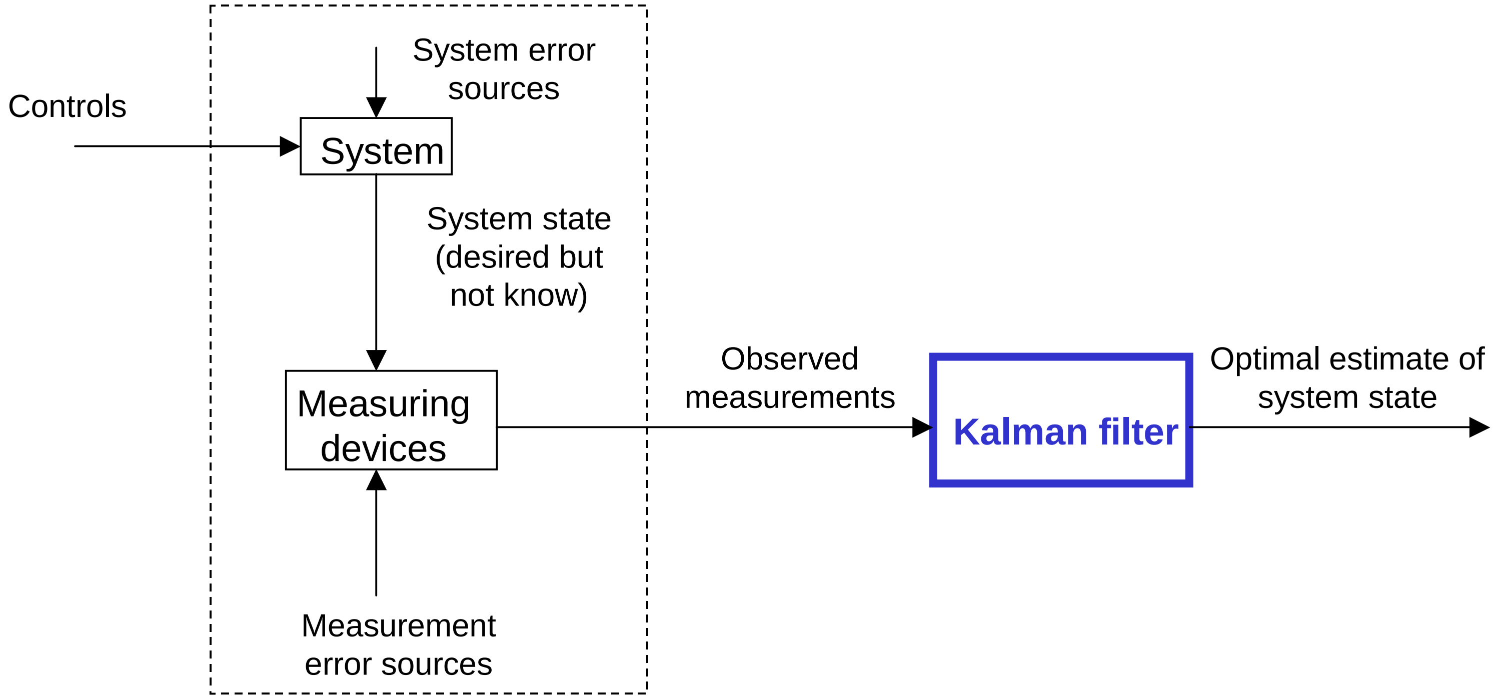
Instituto Superior Técnico / Instituto de Sistemas e Robótica

June 2000

INTRODUCTION TO KALMAN FILTERING

WHAT IS A KALMAN FILTER?

- Optimal Recursive Data Processing Algorithm
- Typical Kalman filter application



WHAT IS A KALMAN FILTER?
Introduction to the Concept

- **Optimal** Recursive Data Processing Algorithm
  - Dependent upon the criteria chosen to evaluate performance
  - Under certain assumptions, KF is optimal with respect to virtually any criteria that makes sense.
  - KF incorporates all available information
    - knowledge of the system and measurement device dynamics
    - statistical description of the system noises, measurement errors, and uncertainty in the dynamics models
    - any available information about initial conditions of the variables of interest

WHAT IS A KALMAN FILTER?
Introduction to the concept

- **Optimal** Recursive Data Processing Algorithm

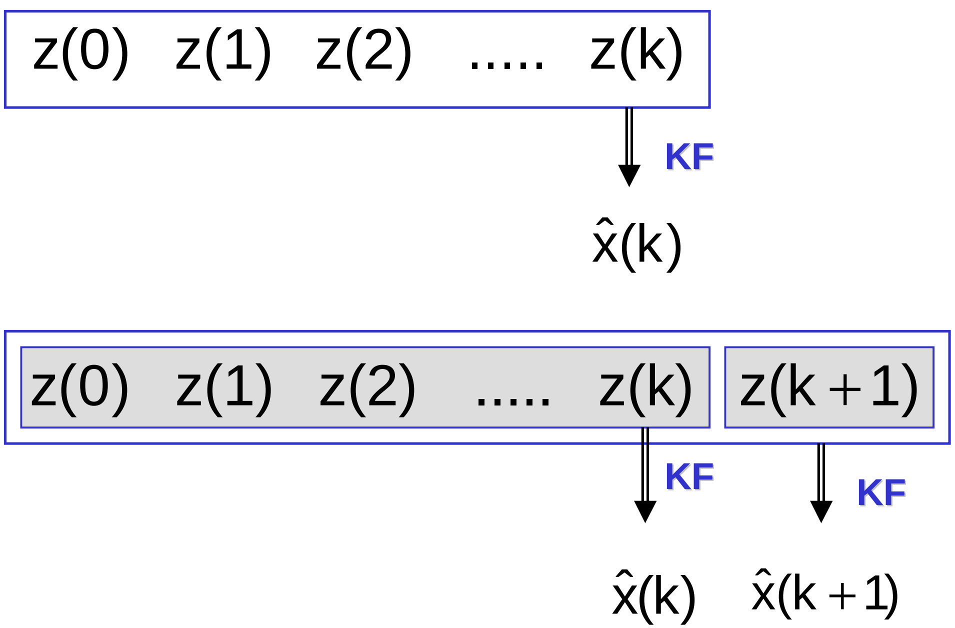
$$x(k+1) = f(x(k), u(k), w(k))$$
$$z(k+1) = h(x(k+1), v(k+1))$$

- x - state
- f - system dynamics
- h - measurement function
- u - controls
- w - system error sources
- v - measurement error sources
- z - observed measurements

- **Given**
  - f, h, noise characterization, initial conditons
  - z(0), z(1), z(2), …, z(k)
- **Obtain**
  - the **“best”** estimate of x(k)

WHAT IS A KALMAN FILTER?
Introduction to the concept

- Optimal **Recursive** Data Processing Algorithm
  - the KF does not require all previous data to be kept in storage and reprocessed every time a new measurement is taken.



To evaluate $\hat{x}(k+1)$ the KF only requires $\hat{x}(k)$ and z(k+1)

WHAT IS A KALMAN FILTER?
Introduction to the concept

- Optimal Recursive **Data Processing** Algorithm
  - The KF is a data processing algorithm
  - The KF is a computer program runing in a central processor

WHAT IS THE KALMAN FILTER ?
Which is the best estimate?

- Any type of filter tries to obtain an optimal estimate of desired quantities from data provided by a noisy environment.
- Best = minimizing errors in some respect.
- Bayesian viewpoint - the filter propagates the conditional probability density of the desired quantities, conditioned on the knowledge of the actual data coming from measuring devices

- Why base the state estimation on the conditional probability density function ?

WHAT IS A KALMAN FILTER?
Which is the best estimate?

Example

- x(i) one dimensional position of a vehicle at time instant i
- z(j) two dimensional vector describing the measurements of position at time j by two separate radars

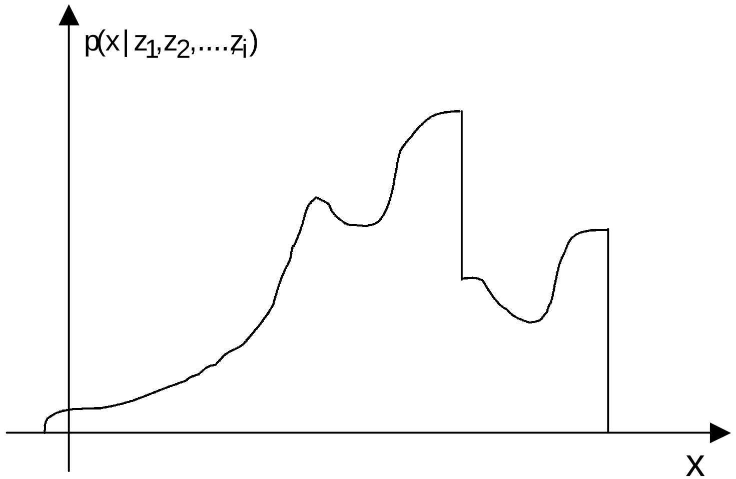
- If $z(1)=z_1$, $z(2)=z_2$, …., $z(j)=z_j$

$$p_{x(i)|z(1),z(2),\ldots,z(i)}(x\,|\,z_1,z_2,\ldots,z_i)$$

  - represents all the information we have on x(i) based (conditioned) on the measurements acquired up to time i
  - given the value of all measurements taken up time i, this conditional pdf indicates what the probability would be of x(i) assuming any particular value or range of values.

WHAT IS A KALMAN FILTER? Which is the best estimate?

- The shape of $p_{x(i)|z(1),z(2),...,z(i)}(x \mid z_1, z_2, ..., z_i)$ conveys the amount of certainty we have in the knowledge of the value x.



- Based on this conditional pdf, the estimate can be:
  - the **mean** - the center of probability mass (MMSE)
  - the **mode** - the value of x that has the highest probability (MAP)
  - the **median** - the value of x such that half the probability weight lies to the left and half to the right of it.

WHAT IS THE KALMAN FILTER ?
Basic Assumptions

- The Kalman Filter performs the conditional probability density propagation
  - for systems that can be described through a LINEAR model
  - in which system and measurement noises are WHITE and GAUSSIAN
- Under these assumptions,
  - the conditional pdf is Gaussian
  - mean=mode=median
  - there is a unique best estimate of the state
  - the KF is the best filter among all the possible filter types

- What happens if these assumptions are relaxed?
- Is the KF still an optimal filter? In which class of filters?

DISCRETE KALMAN FILTER
Problem Formulation

MOTIVATION

- Given a discrete-time, linear, time-varying plant
  - with random initial state
  - driven by white plant noise
- Given noisy measurements of linear combinations of the plant state variables
- Determine the best estimate of the system state variable

STATE DYNAMICS AND MEASUREMENT EQUATION

$$x_{k+1} = A_k x_k + B_k u_k + G_k w_k, \quad k \geq 0$$

$$z_k = C_k x_k + v_k$$

DISCRETE KALMAN FILTER
Problem Formulation

VARIABLE DEFINITIONS

$x_k \in R^n$ state vector (stochastic non - white process)

$u_k \in R^m$ deterministic input sequence

$w_k \in R^n$ white Gaussian system noise (assumed with zero mean)

$v_k \in R^r$ white Gaussian measurement noise (assumed with zero mean)

$z_k \in R^r$ measurement vector (stochastic non - white sequence)

DISCRETE KALMAN FILTER
Problem Formulation

INITIAL CONDITIONS

- x_0 is a Gaussian random vector, with
  - mean $E[x_0] = \bar{x}_0$
  - covariance matrix $E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = P_0 = P_0^T \geq 0$

STATE AND MEASUREMENT NOISE

- zero mean $E[w_k]=E[v_k]=0$
- $\{w_k\}$, $\{v_k\}$ - white Gaussian sequences

$$E\left[\begin{pmatrix} w_k \\ v_k \end{pmatrix}\begin{pmatrix} w_k^T & v_k^T \end{pmatrix}\right] = \begin{bmatrix} Q_k & 0 \\ 0 & R_k \end{bmatrix}$$

$x(0)$, w_k and v_j are independent for all k and j

DISCRETE KALMAN FILTER
Problem Formulation

DEFINITION OF FILTERING PROBLEM

- Let k denote present value of time
- Given the sequence of past inputs

$$U_0^{k-1} = \{u_0, u_1, \ldots u_{k-1}\}$$

- Given the sequence of past measurements

$$Z_1^k = \{z_1, z_2, \ldots z_k\}$$

- Evaluate the best estimate of the state x(k)

DISCRETE KALMAN FILTER
Problem Formulation

$$x_{k+1} = A_k x_k + B_k u_k + G_k w_k, \quad k \geq 0$$
$$z_k = C_k x_k + v_k$$

- **Given x_0**
  - "Nature" apply w_0
  - We apply u_0
  - The system moves to state x_1
  - We make a measurement z_1

Question: which is the best estimate of x_1?

Answer: obtained from $p(x_1 \mid Z_1^1, U_0^0)$

  - "Nature" apply w_1
  - We apply u_1
  - The system moves to state x_2
  - We make a measurement z_2

Question: which is the best estimate of x_2?

Answer: obtained from $p(x_2 \mid Z_1^2, U_0^1)$

.
.
.

DISCRETE KALMAN FILTER
Problem Formulation

.
.
.

Question: **which is the best estimate of x_{k-1}?**

Answer: **obtained from** $p(x_{k-1} \mid Z_1^{k-1}, U_0^{k-2})$

- "Nature" apply w_{k-1}
- We apply u_{k-1}
- The system moves to state x_k
- We make a measurement z_k

Question: **which is the best estimate of x_k?**

Answer: **obtained from** $p(x_k \mid Z_1^k, U_0^{k-1})$

.
.
.

DISCRETE KALMAN FILTER
Towards the Solution

- The filter has to propagate the conditional probability density functions

$p(x_0)$

$p(x_1 \mid Z_1^1, U_0^0) \longrightarrow \hat{x}(1 \mid 1)$

$p(x_2 \mid Z_1^2, U_0^1) \longrightarrow \hat{x}(2 \mid 2)$

$\vdots \qquad \vdots$

$p(x_{k-1} \mid Z_1^{k-1}, U_0^{k-2}) \longrightarrow \hat{x}(k-1 \mid k-1)$

$p(x_k \mid Z_1^k, U_0^{k-1}) \longrightarrow \hat{x}(k \mid k)$

$\vdots \qquad \vdots$

DISCRETE KALMAN FILTER
From the Assumptions to the Problem Solution

- The LINEARITY of
  - the system state equation
  - the system observation equation
- The GAUSSIAN nature of
  - the initial state, x_0
  - the system white noise, w_k
  - the measurement white noise, v_k

$p(x_k \mid Z_1^k, U_0^{k-1})$

is Gaussian

Uniquely characterized by

- the conditional mean $\hat{x}(k \mid k) = E[x_k \mid Z_1^k, U_0^{k-1}]$
- the conditional covariance $P(k \mid k) = \mathrm{cov}[x_k; x_k \mid Z_1^k, U_1^{k-1}]$

$$p(x_k \mid Z_0^k, U_0^{k-1}) \sim \mathcal{N}(\hat{x}(k \mid k), P(k \mid k))$$

DISCRETE KALMAN FILTER
Towards the Solution

- As the conditional probability density functions are Gaussian, the Kalman filter only propagates the first two moments

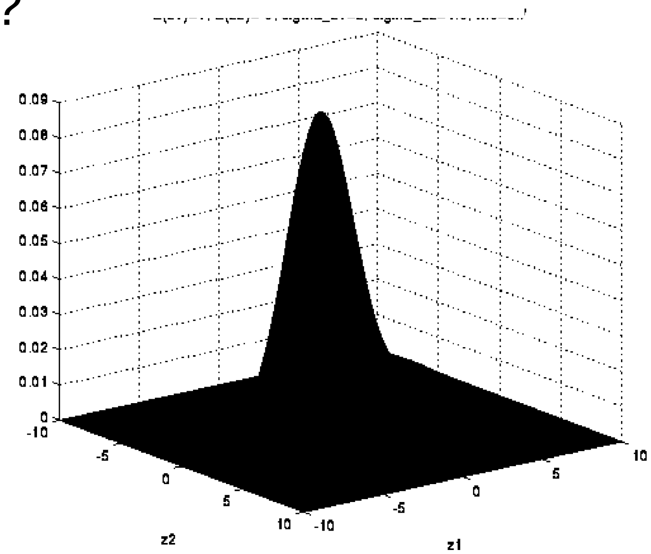
$p(x_0)$	$p(x_0)$	
~~$p(x_1 \mid Z_1^1, U_0^0)$~~	$E[x_1 \mid Z_1^1, U_0^0] = \hat{x}(1 \mid 1)$	$P(1 \mid 1)$
~~$p(x_2 \mid Z_1^2, U_0^1)$~~	$E[x_2 \mid Z_1^2, U_0^1] = \hat{x}(2 \mid 2)$	$P(2 \mid 2)$
⋮	⋮	⋮
~~$p(x_{k-1} \mid Z_1^{k-1}, U_0^{k-2})$~~	$E[x_{k-1} \mid Z_1^{k-1}, U_0^{k-2}] = \hat{x}(k-1 \mid k-1)$	$P(k-1 \mid k-1)$
~~$p(x_k \mid Z_1^k, U_0^{k-1})$~~	$E[x_k \mid Z_1^k, U_0^{k-1}] = \hat{x}(k \mid k)$	$P(k \mid k)$
⋮	⋮	⋮

DISCRETE KALMAN FILTER
Towards the Solution

- We stated that the state estimate equals the conditional mean

$$\hat{x}(k \mid k) = E[x_k \mid Z_1^k, U_0^{k-1}]$$

- Why?
- Why not the mode of $p(x_k \mid Z_1^k, U_0^{k-1})$?
- Why not the median of $p(x_k \mid Z_1^k, U_0^{k-1})$?



- As $p(x_k \mid Z_1^k, U_0^{k-1})$ is Gaussian
  - mean = mode = median

DISCRETE KALMAN FILTER
Filter dynamics

- KF dynamics is recursive

$Z_1^k = \{z_1, z_2, \ldots, z_k\}$

$U_0^{k-1} = \{u_0, u_1, \ldots, u_{k-1}\}$

$Z_1^{k+1} = \{Z_1^k, z_{k+1}\}$

$U_0^k = \{U_0^{k-1}, u_k\}$

$p(x_k \mid Z_1^k, U_0^{k-1})$ ⟶ $p(x_{k+1} \mid Z_1^{k+1}, U_0^k)$

Prediction cycle

Filtering cycle

$p(x_{k+1} \mid Z_1^k, U_0^k)$

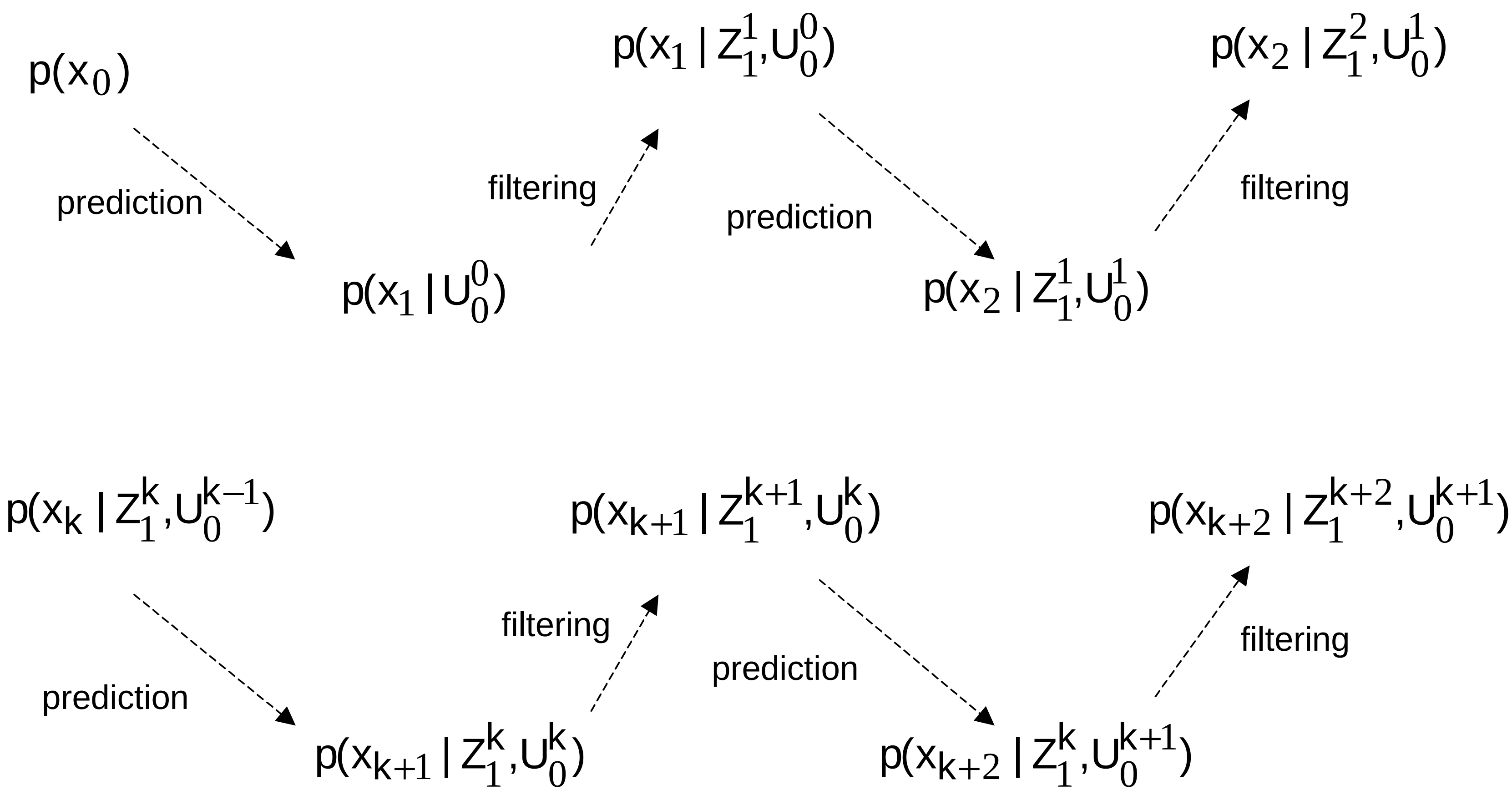
What can you say about x_{k+1} before we make the measurement z_{k+1}

$Z_1^k = \{z_1, z_2, \ldots, z_k\}$

$U_0^k = \{U_0^{k-1}, u_k\}$

How can we improve our information on x_{k+1} after we make the measurement z_{k+1}

DISCRETE KALMAN FILTER
Filter dynamics



DISCRETE KALMAN FILTER
Filter dynamics - Prediction cycle

- Prediction cycle

$$p(x_k \mid Z_1^k, U_0^{k-1}) \dashrightarrow p(x_{k+1} \mid Z_1^k, U_0^k)$$

$$p(x_k \mid Z_1^k, U_0^{k-1}) \sim \mathcal{N}(\hat{x}(k \mid k), P(k \mid k))$$

assumed known

?

- Is Gaussian

$$\hat{x}(k+1 \mid k) = E(x_{k+1} \mid Z_1^k, U_0^k) \;\; ?$$

$$P(k+1 \mid k) = \mathrm{cov}[x_{k+1}; x_{k+1} \mid Z_1^k, U_0^k] \;\; ?$$

$$x_{k+1} = A_k x_k + B_k u_k + G_k w_k$$

$$E[x_{k+1} \mid Z_1^k, U_0^k] = A_k E[x_k \mid Z_1^k, U_0^k] + B_k E[u_k \mid Z_1^k, U_0^k] + G_k E[w_k \mid Z_1^k, U_0^k]$$

$$\hat{x}(k+1 \mid k) = A_k \hat{x}(k \mid k) + B_k u_k$$

DISCRETE KALMAN FILTER
Filter dynamics - Prediction cycle

- Prediction cycle

$$P(k+1|k) = \text{cov}[x_{k+1}; x_{k+1} \mid Z_1^k, U_0^k]$$

$$\tilde{x}(k+1|k) = x_{k+1} - \hat{x}(k+1|k)$$

→ prediction error

$$x(k+1) - \hat{x}(k+1|k) = A_k x_k + B_k u_k + G_k w_k - (A_k \hat{x}(k|k) + B_k u_k)$$

$$\tilde{x}(k+1|k) = A_k \tilde{x}(k|k) + G_k w_k$$

$$\text{cov}[y;y] = E[(y-\bar{y})(y-\bar{y})^T]$$

$$P(k+1|k) = E[\tilde{x}(k+1|k)\tilde{x}(k+1|k)^T \mid Z_1^k, U_0^k]$$

$$P(k+1|k) = A_k P(k|k) A_k^T + G_k Q_k G_k^T$$

DISCRETE KALMAN FILTER
Filter dynamics - Filtering cycle

- Filtering cycle

$$p(x_{k+1} \mid Z_1^k, U_0^k) \quad + \quad z_{k+1} \quad \dashrightarrow \quad p(x_{k+1} \mid Z_1^{k+1}, U_0^k) \text{ ?}$$

$$N(\hat{x}(k+1|k), P(k+1|k))$$

1º Passo	Measurement prediction

What can you say about z_{k+1} before we make the measurement z_{k+1}

$$p(z_{k+1} \mid Z_1^k, U_0^k)$$

$$p(C_{k+1}x_{k+1} + v_{k+1} \mid Z_1^k, U_0^k)$$

$$E[z_{k+1} \mid Z_1^k, U_0^k] = \hat{z}(k+1|k) = C_{k+1}\hat{x}(k+1|k)$$

$$\text{cov}[z_{k+1}; z_{k+1} \mid Z_1^k, U_0^k] = P_z(k+1|k) = C_{k+1}P(k+1|k)C_{k+1}^T + R_{k+1}$$

DISCRETE KALMAN FILTER
Filter dynamics - Filtering cycle

- Filtering cycle

2º Passo $p(x_{k+1} \mid Z_1^k, U_0^k)$

$$E[x_{k+1} \mid Z_1^{k+1}, U_0^k] = E[x_{k+1} \mid Z_1^k, z_{k+1}, U_0^k]$$

Z_1^{k+1} e $\{Z_1^k, \tilde{z}(k+1 \mid k)\}$ São equivalentes do ponto de vista de infirmação contida

$$E[x_{k+1} \mid Z_1^{k+1}, U_0^k] = E[x_{k+1} \mid Z_1^k, \tilde{z}(k+1 \mid k), U_0^k]$$

Required result

If x, y and z are jointly Gaussian and y and z are statistically independent

$$E[x \mid y, z] = E[x \mid y] + E[x \mid z] - m_x$$

DISCRETE KALMAN FILTER
Filter dynamics - Filtering cycle

- Filtering cycle

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + \underbrace{P(k+1|k)C_{k+1}^T\left[C_{k+1}P(k+1|k)C_{k+1}^T + R_{k+1}\right]^{-1}}_{K(k+1)}\underbrace{(z_{k+1} - \underbrace{C_{k+1}\hat{x}(k+1|k)}_{\hat{z}(k+1|k)})}_{\tilde{z}(k+1|k)}$$

$K(k+1)$

Kalman Gain

$\hat{z}(k+1|k)$ measurement prediction

$\tilde{z}(k+1|k)$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)(z_{k+1} - C_{k+1}\hat{x}(k+1|k))$$

$$P(k+1|k+1) = P(k+1|k) - P(k+1|k)C_{k+1}^T\left[C_{k+1}P(k+1|k)C_{k+1}^T + R_{k+1}\right]^{-1}C_{k+1}P(k+1|k)$$

DISCRETE KALMAN FILTER
Dynamics

- Linear System $$x_{k+1} = A_k x_k + B_k u_k + G_k w_k, \quad k \geq 0$$ $$z_k = C_k x_k + v_k$$

- Discrete Kalman Filter

prediction

$$\hat{x}(k+1|k) = A_k \hat{x}(k|k) + B_k u_k$$

$$P(k+1|k) = A_k P(k|k) A_k^T + G_k Q_k G_k^T$$

filtering

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)(z_{k+1} - C_{k+1}\hat{x}(k+1|k))$$

$$P(k+1|k+1) = P(k+1|k) - P(k+1|k)C_{k+1}^T\left[C_{k+1}P(k+1|k)C_{k+1}^T + R_{k+1}\right]^{-1} C_{k+1}P(k+1|k)$$

$$K(k+1) = P(k+1|k)C_{k+1}^T\left[C_{k+1}P(k+1|k)C_{k+1}^T + R_{k+1}\right]^{-1}$$

Initial conditions

$$\hat{x}(0|0) = \bar{x}_0 \qquad P(0|0) = P_0$$

DISCRETE KALMAN FILTER
Properties

- The Discrete KF is a time-varying linear system

$$\hat{x}_{k+1|k+1} = (I - K_{k+1}C_{k+1})A_k\hat{x}_{k|k} + K_{k+1}z_{k+1} + B_k u_k$$

  – even when the system is time-invariant and has stationary noise

$$\hat{x}_{k+1|k+1} = (I - K_{k+1}C)A\hat{x}_{k|k} + K_{k+1}z_{k+1} + Bu_k$$

    • the Kalman gain is not constant

- Does the Kalman gain matrix converges to a constant matrix? In which conditions?

DISCRETE KALMAN FILTER
Properties

- The state estimate is a linear function of the measurements

KF dyamics in terms of the filtering estimate

$$\hat{x}_{k+1|k+1} = \underbrace{(I - K_{k+1}C_{k+1})A_k}_{\Phi_k}\hat{x}_{k|k} + K_{k+1}z_{k+1} + B_k u_k$$

Assuming null inputs for the sake of simplicity

$$\hat{x}_{0|0} = \bar{x}_0$$

$$\hat{x}_{1|1} = \Phi_0\hat{x}_{0|0} + K_1 z_1$$

$$\hat{x}_{2|2} = \Phi_1\Phi_0\hat{x}_{0|0} + \Phi_1 K_1 z_1 + K_2 z_2$$

$$\hat{x}_{3|3} = \Phi_2\Phi_1\Phi_0\hat{x}_{0|0} + \Phi_2\Phi_1 K_1 z_1 + \Phi_2 K_2 z_2 + K_3 z_3$$

DISCRETE KALMAN FILTER
Properties

- Innovation process

$$r_{k+1} = z_{k+1} - C_{k+1}\hat{x}(k+1\,|\,k)$$

$$\hat{x}(k+1\,|\,k) = E(x_{k+1}\,|\,Z_1^k, U_0^k)\ ?$$

  – z(k+1) carries information on x(k+1) that was not available on Z_1^k
  – this new information is represented by r(k+1) - innovation process

- Properties of the innovation process
  – the innovations r(k) are orthogonal to z(i)

$$E[r(k)z^T(i)] = 0, \quad i = 1,2,\ldots,k-1$$

  – the innovations are uncorrelated/white noise

$$E[r(k)r^T(i)] = 0, \quad i \neq k$$

    - this test can be used to acess if the filter is operating correctly

DISCRETE KALMAN FILTER
Properties

- Covariance matrix of the innovation process

$$S(k+1) = C_{k+1}P(K+1\,|\,K)C_{k+1}^{T} + R_{k+1}$$

$$\Downarrow$$

$$K(k+1) = P(k+1\,|\,k)C_{k+1}^{T}\left[C_{k+1}P(k+1\,|\,k)C_{k+1}^{T} + R_{k+1}\right]^{-1}$$

$$K(k+1) = P(k+1\,|\,k)C_{k+1}^{T}S_{k+1}^{-1}$$

DISCRETE KALMAN FILTER
Properties

- The Discrete KF provides an unbiased estimate of the state

  – $\hat{x}_{k+1|k+1}$ is an unbiased estimate of the state x(k+1), providing that the initial conditions are $\hat{x}(0 | 0) = \bar{x}_0$ $P(0 | 0) = P_0$

  – Is this still true if the filter initial conditions are not the specified ?

DISCRETE KALMAN FILTER
Steady state Kalman Filter

- Time invariant system and stationay white system and observation noise

$$x_{k+1} = Ax_k + Gw_k, \quad k \geq 0 \qquad E[w_k w_k^T] = Q$$

$$z_k = Cx_k + v_k \qquad E[v_k v_k^T] = R$$

- Filter dynamics

$$\hat{x}(k+1 \mid k+1) = A\hat{x}(k+1 \mid k) + K(k+1)(z_{k+1} - C\hat{x}(k+1 \mid k))$$

$$P(k+1|k) = AP(k|k\text{-}1)A^T - AP(k-1|k)C^T[CP(k|k-1)C^T + R]^{-1}CP(k|k-1)A^T + GQG^T$$

Discrete Riccati Equation

K(k)

DISCRETE KALMAN FILTER
Steady state Kalman Filter

- If Q is positive definite, $(A,G\sqrt{Q})$ is controllable, and (A,C) is observable, then
  - the steady state Kalman filter exists
  - the limit exists $\lim_{k \to \infty} P(k+1|k) = P_\infty^-$
  - P_∞^- is the unique, finite positive-semidefinite solution to the algebraic equation

$$P_\infty^- = AP_\infty^- A^T - AP_\infty^- C^T [CP_\infty^- C^T + R]^{-1} CP_\infty^- A^T + GQG^T$$

  - P_∞^- is independent of P_0 provided that $P_0 \geq 0$
  - the steady-state Kalman filter is assymptotically unbiased

$$K_\infty = P_\infty^- C^T [CP_\infty^- C^T + R]^{-1}$$

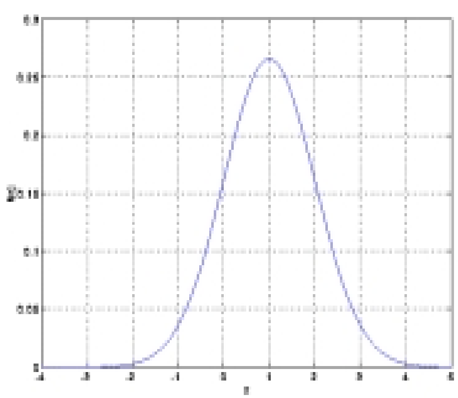
MEANING OF THE COVARIANCE MATRIX
Generals on Gaussian pdf

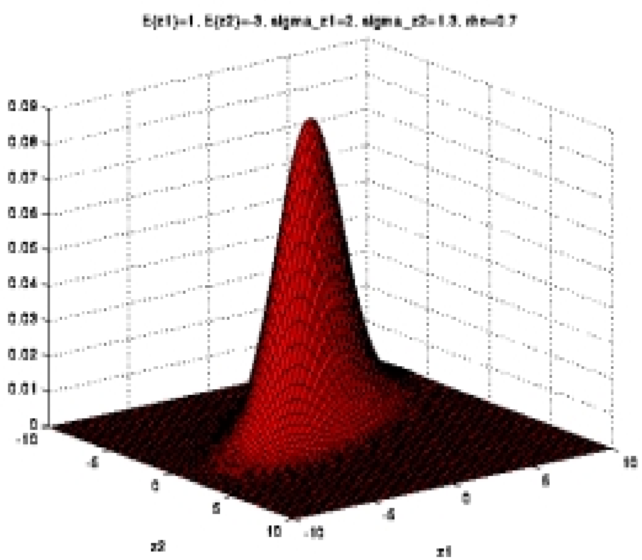
- Let z be a Gaussian random vector of dimension n

$$E[z] = m, \quad E\left[(z-m)(z-m)^T\right] = P$$

- P - covariance matrix - symetric, positive defined
- Probability density function $p(z) = \frac{1}{\sqrt{(2\pi)^n \det P}} \exp\left[-\frac{1}{2}(z-m)^T P^{-1}(z-m)\right]$

n=1





n=2

MEANING OF THE COVARIANCE MATRIX
Generals on Gaussian pdf

- Locus of points where the fdp is greater or equal than a given threshold

$$(z-m)^T P^{-1}(z-m) \leq K$$

n=1 line segment

n=2 ellipse and inner points

n=3 3D ellipsoid and inner points

n>3 hiperellipsoid and inner points

- If $P = \text{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2)$
  - the ellipsoid axis are aligned with the axis of the referencial where the vector z is defined

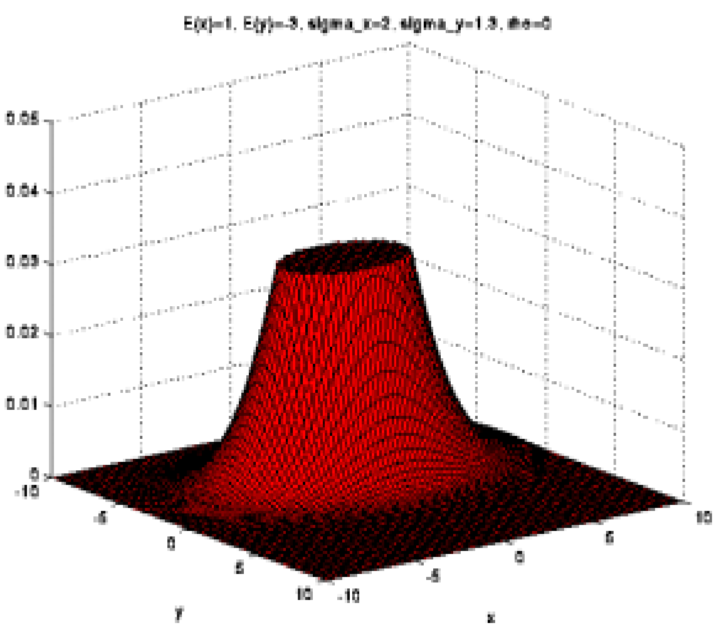
$$(z-m)^T P^{-1}(z-m) \leq K \Leftrightarrow \sum_{i=1}^{n} \frac{(z_i - m_i)^2}{\sigma_i^2 K} \leq 1$$

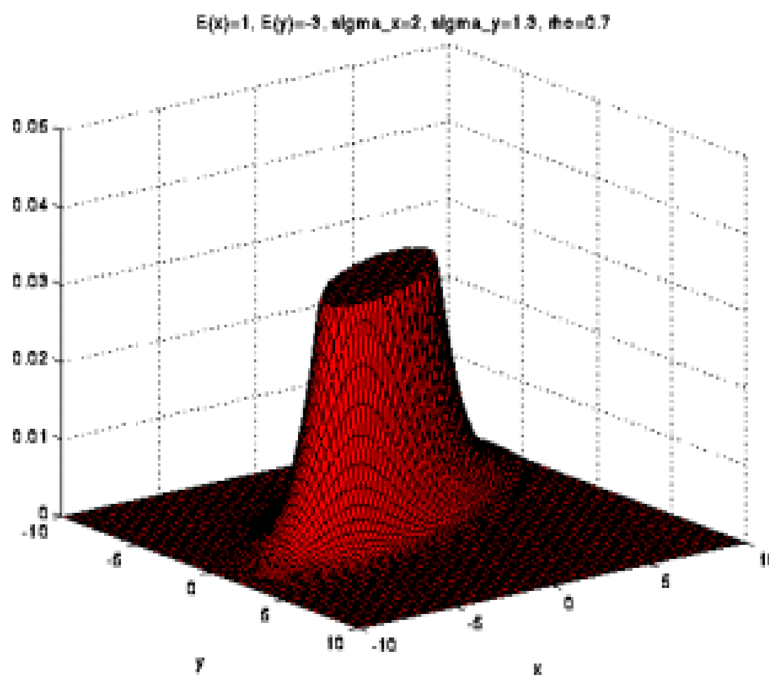
  - length of the ellipse semi-axis = $\sigma_i \sqrt{K}$

MEANING OF THE COVARIANCE MATRIX
Generals on Gaussian pdf - Error elipsoid

Example
n=2

$$P = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \qquad P = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$





MEANING OF THE COVARIANCE MATRIX
Generals on Gaussian pdf -Error ellipsoid and axis orientation

- Error ellipsoid $(z-m_z)^T P^{-1}(z-m_z) \leq K$
- $P=P^T$ - to distinct eigenvalues correspond orthogonal eigenvectors
- Assuming that P is diagonalizable

$$P = TDT^{-1} \qquad \text{with} \qquad D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

$$TT^T = I$$

- Error ellipoid (after coordinate transformation)

$$w = T^T z \qquad (z-m_z)^T T D^{-1} T^T (z-m_z) \leq K$$

$$(w-m_w)^T D^{-1}(w-m_w) \leq K$$

- At the new coordinate system, the ellipsoid axis are aligned with the axis of the new referencial

MEANING OF THE COVARIANCE MATRIX
Generals on Gaussian pdf -Error elipsis and referencial axis

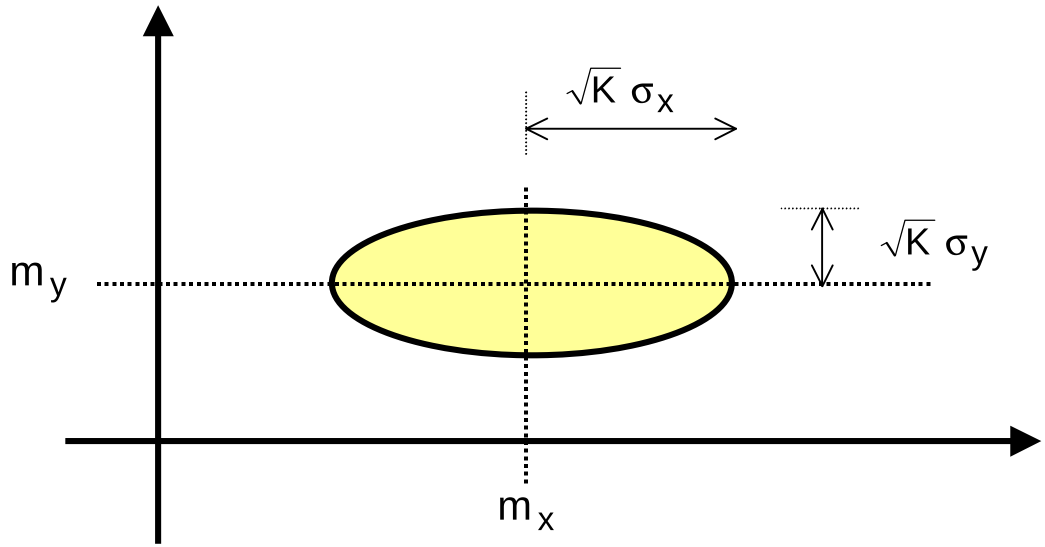
- n=2

$$z = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\begin{bmatrix} x - m_x & y - m_y \end{bmatrix} \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}^{-1} \begin{bmatrix} x - m_x \\ y - m_y \end{bmatrix} \leq K$$

ellipse

$$\frac{(x - m_x)^2}{K\sigma_x^2} + \frac{(y - m_y)^2}{K\sigma_y^2} \leq 1$$



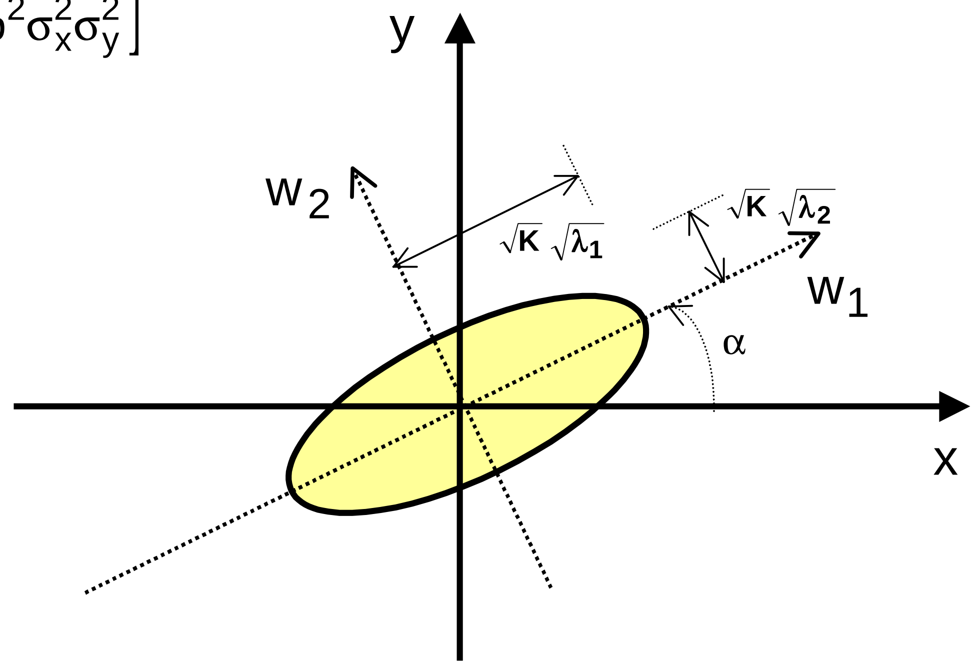
MEANING OF THE COVARIANCE MATRIX
Generals on Gaussian pdf -Error ellipse and referencial axis

- n=2 $z = \begin{bmatrix} x \\ y \end{bmatrix}$ $P = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$ $[x \quad y] \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \leq K$

$$\lambda_1 = \frac{1}{2}\left[\sigma_x^2 + \sigma_y^2 + \sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4\rho^2\sigma_x^2\sigma_y^2}\right]$$

$$\lambda_2 = \frac{1}{2}\left[\sigma_x^2 + \sigma_y^2 - \sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4\rho^2\sigma_x^2\sigma_y^2}\right]$$

$$\frac{w_1^2}{K\lambda_1} + \frac{w_2^2}{K\lambda_2} \leq 1$$



$$\alpha = \frac{1}{2}\tan^{-1}\left(\frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2}\right),$$

$$-\frac{\pi}{4} \leq \alpha \leq \frac{\pi}{4}, \quad \sigma_x^2 \neq \sigma_y^2$$

DISCRETE KALMAN FILTER
Probabilistic interpretation of the error ellipsoid

$$p(x_k \mid Z_0^k, U_0^{k-1}) \sim \mathcal{N}(\hat{x}(k \mid k), P(k \mid k))$$

- Given $\hat{x}(k \mid k)$ and $P(k \mid k)$ it is possible to define the locus where, with a **given probability**, the values of the random vector x(k) ly.

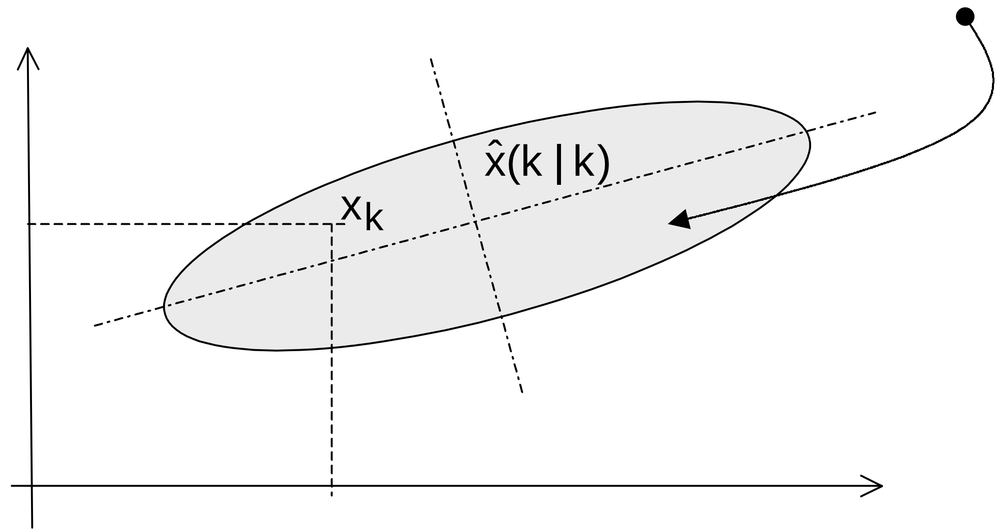
Hiperellipsoid with center in $\hat{x}(k \mid k)$ and with semi-axis proportional to the eigenvalues of $P(k \mid k)$

DISCRETE KALMAN FILTER
Probabilistic interpretation of the error ellipsoid

$$p(x_k \mid Z_0^k, U_0^{k-1}) \sim \mathcal{N}(\hat{x}(k|k), P(k|k))$$

- Example for n=2

$$M = \{x_k : \quad [x_k - \hat{x}(k|k)]^T P(k|k)^{-1} [x_k - \hat{x}(k|k)] \leq K\}$$



$$\Pr\{x_k \in M\}$$

- is a function of K
- a pre-specified values of this probability can be obtained by an apropriate choice of K

DISCRETE KALMAN FILTER
Probabilistic interpretation of the error ellipsoid

$$p(x_k \mid Z_0^k, U_0^{k-1}) \sim \mathcal{N}(\hat{x}(k \mid k), P(k \mid k))$$

$x_k \in R^n$

$$\underbrace{[x_k - \hat{x}(k \mid k)]^T P(k \mid k)^{-1} [x_k - \hat{x}(k \mid k)]}_{} \leq K$$

(Scalar) random variable with a χ^2 distribution with n degrees of reedom

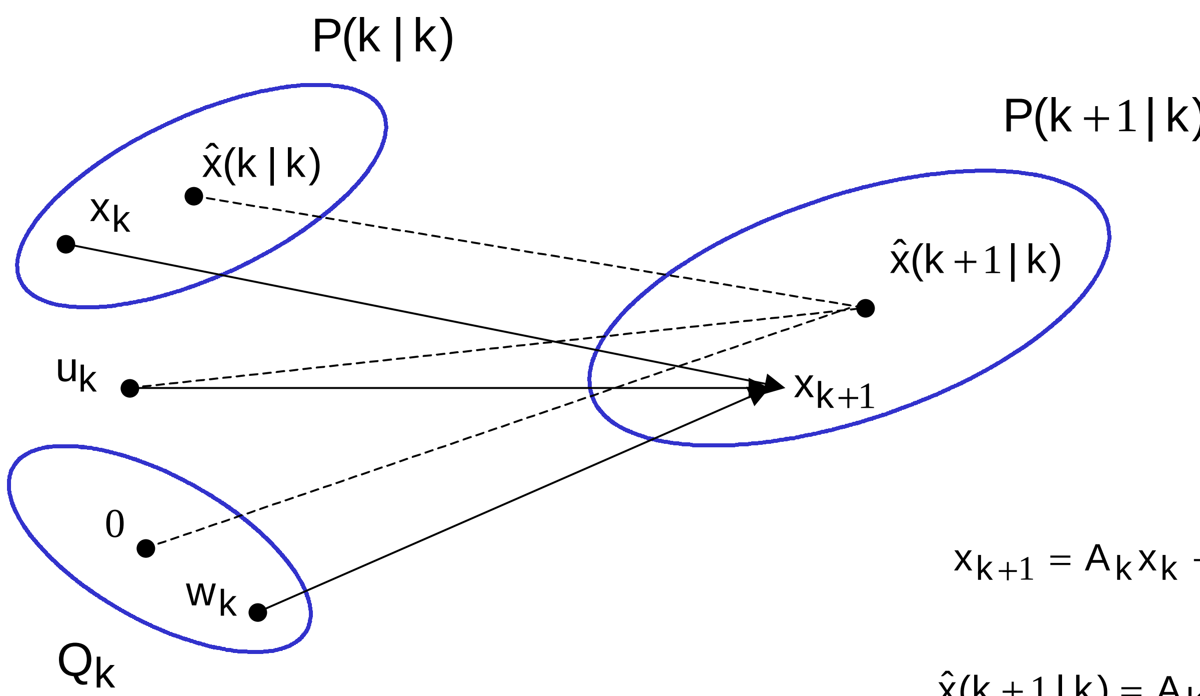
<u>Probability = 90%</u>
n=1 K=2.71
n=2 K=4.61

- How to chose K for a desired probability?
  - Just consult a Chi square distribution table

<u>Probability = 95%</u>
n=1 K=3.84
n=2 K=5.99

DISCRETE KALMAN FILTER
The error ellipsoid and the filter dynamics

- Prediction cycle



$$x_{k+1} = A_k x_k + B_k u_k + G_k w_k$$

$$\hat{x}(k+1|k) = A_k \hat{x}(k|k) + B_k u_k$$

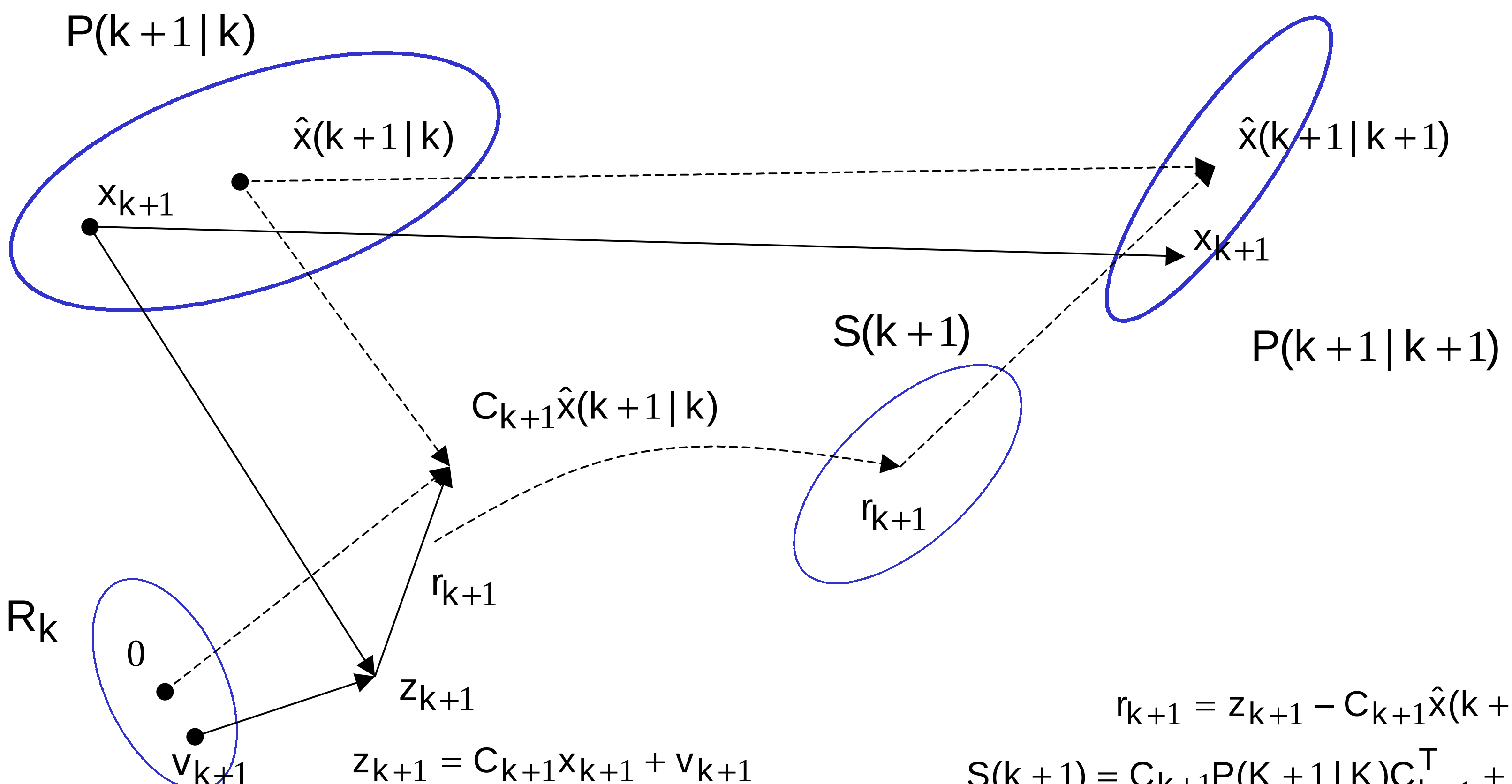
$$P(k+1|k) = A_k P(k|k) A_k^T + G_k Q_k G_k^T$$

DISCRETE KALMAN FILTER
The error ellipsoid and the filter dynamics

- Filtering cycle

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)r(k+1)$$

$$P(k+1|k+1) = P(k+1|k) - K(k+1)C_{k+1}P(k+1|k)$$



$$z_{k+1} = C_{k+1}x_{k+1} + v_{k+1}$$

$$r_{k+1} = z_{k+1} - C_{k+1}\hat{x}(k+1|k)$$

$$S(k+1) = C_{k+1}P(K+1|K)C_{k+1}^T + R_{k+1}$$

Extended Kalman Filter

- **Non linear** dynamics
- **White Gaussian** system and observation noise

$$x_{k+1} = f_k(x_k, u_k) + w_k$$
$$z_k = h_k(x_k) + v_k$$

$$x_0 \sim N(\bar{x}_0, P_0)$$
$$E[w_k w_j^T] = Q_k \delta_{kj}$$
$$E[v_k v_j^T] = R_k \delta_{kj}$$

- QUESTION: Which is the MMSE (minimum mean-square error) estimate of x(k+1)?
  - Conditional mean $\hat{x}(k+1 \mid k) = E(x_{k+1} \mid Z_1^k, U_0^k)$?
  - Due to the non-linearity of the system,

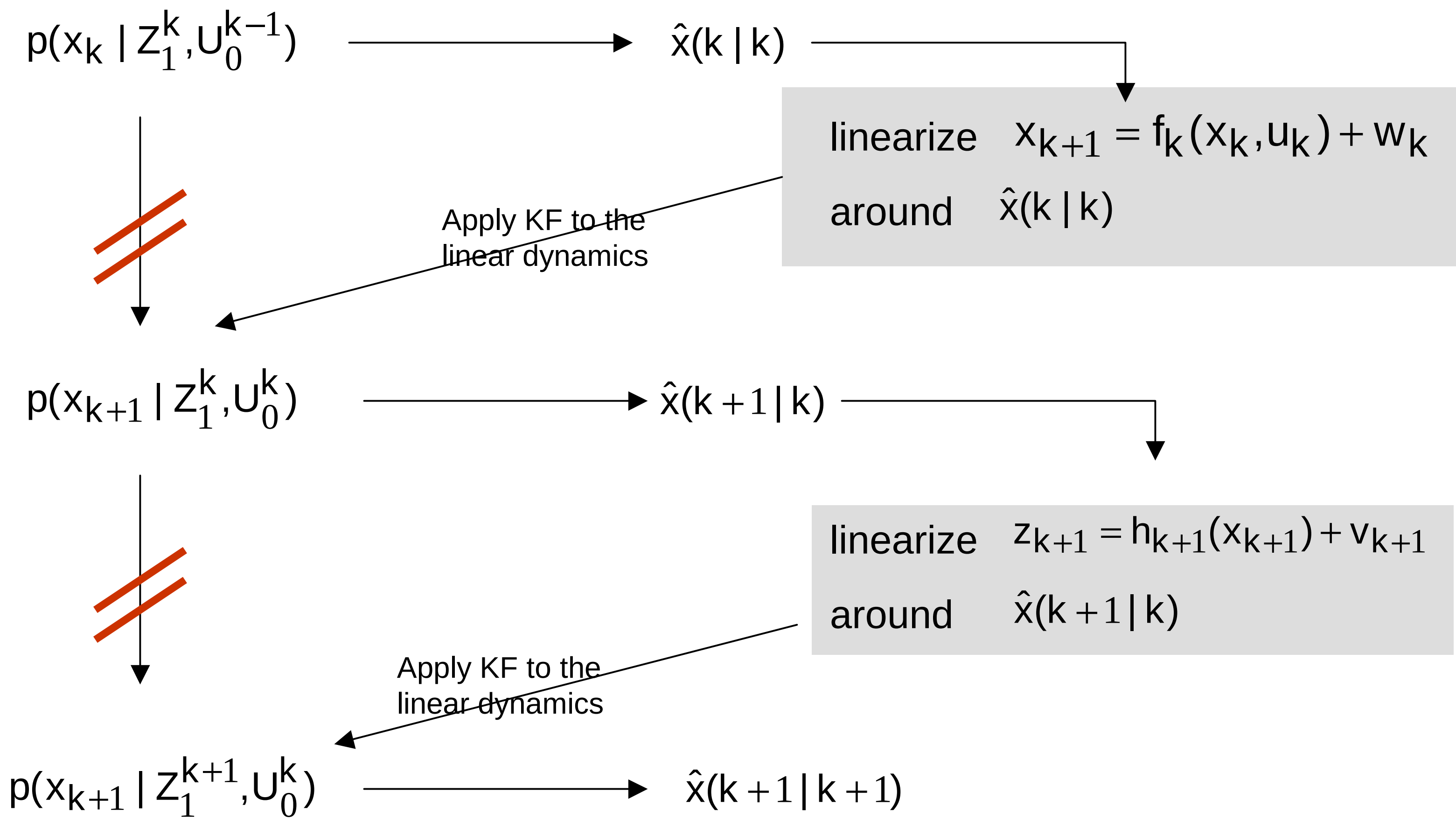
$$p(x_k \mid Z_1^k, U_0^{k-1}) \qquad p(x_{k+1} \mid Z_1^k, U_0^k)$$

are non Gaussian

Extended Kalman Filter

- **(Optimal)** ANSWER: The MMSE estimate is given by a non-linear filter, that propagates the conditonal pdf.
- The **EKF** gives an approximation of the optimal estimate
  - The non-linearities are approximated by a linearized version of the non-linear model around the last state estimate.
  - For this approximation to be valid, this linearization should be a good approximation of the non-linear model in all the unceratinty domain associated with the state estimate.

Extended Kalman Filter



Extended Kalman Filter

linearize $x_{k+1} = f_k(x_k, u_k) + w_k$

around $\hat{x}(k \mid k)$

$$f_k(x_k, u_k) \cong f_k(\hat{x}_{k|k}, u_k) + \nabla f_k (x_k - \hat{x}_{k|k}) + \ldots.$$

$$x_{k+1} \cong \nabla f_k x_k + w_k + \underbrace{(f_k(\hat{x}_{k|k}, u_k) - \nabla f_k \hat{x}_{k|k})}_{\text{known input}}$$

Prediction cycle of KF

$$\hat{x}_{k+1|k} = \nabla f_k \hat{x}_{k|k} + (f_k(\hat{x}_{k|k}, u_k) - \nabla f_k \hat{x}_{k|k})$$

$$P(k+1 \mid k) = \nabla f_k P(k \mid k) \nabla f_k{}^T + Q_k$$

Extended Kalman Filter

linearize $z_{k+1} = h_{k+1}(x_{k+1}) + v_{k+1}$

around $\hat{x}(k+1 \mid k)$

$$h_{k+1}(x_{k+1}) \cong h_{k+1}(\hat{x}_{k+1|k}) + \nabla h_{k+1}(x_{k+1} - \hat{x}_{k+1|k}) + \ldots.$$

$$z_{k+1} \cong \nabla h_{k+1} x_{k+1} + v_k + \underbrace{(h_{k+1}(\hat{x}_{k+1|k}) - \nabla h_{k+1}\hat{x}_{k+1|k})}_{\text{known input}}$$

Update cycle of KF

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + P(k+1 \mid k)\nabla h^T_{k+1}(\nabla h_{k+1} P(k+1 \mid k)\nabla h^T_{k+1} + R_{k+1})^{-1}[z_{k+1} - h_{k+1}(\hat{x}_{k+1|k})]$$

$$P(k+1 \mid k+1) = P(k+1 \mid k) - P(k+1 \mid k)\nabla h^T_{k+1}[\nabla h_{k+1} P(k+1 \mid k)\nabla h^T_{k+1} + R_{k+1}]^{-1}\nabla h_{k+1} P(k+1 \mid k)$$

References

- Anderson, Moore, “Optimal Filtering”, Prentice-Hall, 1979.
- M. Athans, “Dynamic Stochastic Estimation, Prediction and Smoothing,” Series of Lectures, Spring 1999.
- E. W. Kamen, J. K. Su, “Introduction to Optimal Estimation”, Springer, 1999.
- Peter S. Maybeck, “The Kalman Filter: an Introduction to Concepts”
- Jerry M. Mendel, “Lessons in Digital Estimation Theory”, Prentice-Hall, 1987.
- M.Isabel Ribeiro, “Notas Dispersas sobre Filtragem de Kalman” CAPS, IST, June 1989 (http://www.isr.ist.utl.pt/~mir)