Aalto University

LINEAR DYNAMIC SYSTEMS WITH RANDOM INPUTS

- Gaussian pdf, mean and covariance
- Stochastic sequences, Markov property
- Discrete-time linear stochastic dynamic systems
  Prediction, propagation of mean and covariance
- Continuous-time linear stochastic dynamic systems
  Propagation of mean and covariance

Gaussian pdf, mean and covariance

- Scalar Gaussian or Normal pdf

$$p(x) = \mathcal{N}(x; \bar{x}, \sigma^2) \triangleq \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

$$\boxed{x \sim \mathcal{N}(\bar{x}, \sigma^2)}$$

- Vector Gaussian pdf

$$\boxed{\mathcal{N}(x; \bar{x}, P) \triangleq |2\pi P|^{-1/2} e^{-\frac{1}{2}(x-\bar{x})' P^{-1} (x-\bar{x})}}$$

- Mean or expected value

$$\bar{x} = E[x]$$

- Covariance P

$$P = E[(x - \bar{x})(x - \bar{x})']$$

Stochastic sequences, Markov property

- Markov process and property

Markov processes are defined by the following *Markov property*

$$p[x(t)|x(\tau), \tau \leq t_1] = p[x(t)|x(t_1)] \qquad \forall t > t_1 \qquad (1.4.21\text{-}1)$$

that is, the past up to any t_1 is *fully characterized* by the value of the process at t_1.

- Random Sequences, Markov Sequences

$$X^k = \{x(j)\}_{j=1}^{k} \qquad k = 1, 2, \ldots \qquad (1.4.22\text{-}1)$$

$$p[x(k)|X^j] = p[x(k)|x(j)] \qquad \forall k > j \qquad (1.4.22\text{-}2)$$

Markov property continues

discrete-time white noise (a white sequence) scalar

$$E[v(k)v(j)] = q(k)\delta_{kj} \qquad (1.4.22\text{-}3)$$

Kronecker delta function

$$\delta_{kj} = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases} \qquad (1.4.22\text{-}4)$$

The state of linear dynamic equation with white noise process

$$\boxed{E[v(k)] = 0} \qquad (4.3.1\text{-}6)$$

$$\boxed{E[v(k)v(j)'] = Q(k)\delta_{kj}} \qquad (4.3.1\text{-}7)$$

$$\boxed{x(k+1) = F(k)x(k) + G(k)u(k) + v(k)} \qquad (4.3.1\text{-}9)$$

has the following analytical solution, next slide

The State as a Markov Process

$$x(k) = [\prod_{j=0}^{k-l-1} F(k-1-j)]x(l) + \sum_{i=l}^{k-1} [\prod_{j=0}^{k-i-2} F(k-1-j)][G(i)u(i) + v(i)] \tag{4.3.3-1}$$

Thus, since $v(i)$, $i = l, \ldots, k-1$, are independent of

$$X^l \triangleq \{x(j)\}_{j=0}^{l} \tag{4.3.3-2}$$

which depend only on $v(i)$, $i = 0, \ldots, l-1$, one has

$$p[x(k)|X^l, U^{k-1}] = p[x(k)|x(l), U_l^{k-1}] \qquad \forall k > l \tag{4.3.3-3}$$

$$U_l^{k-1} \triangleq \{u(j)\}_{j=l}^{k-1} \tag{4.3.3-4}$$

Thus, **the state vector is** a Markov process, or, more correctly, a **Markov sequence**.

State of a stochastic system described by a Markov process — **summarizes probabilistically its past**.

Discrete-time linear stochastic dynamic systems - prediction

$$x(k+1) = F(k)x(k) + G(k)u(k) + \Gamma(k)v(k) \quad (4.3.4\text{-}1)$$

with the known input u(k) and the process noise v(k) white, but for the sake of generality, nonstationary with nonzero mean:

$$E[v(k)] = \bar{v}(k) \quad (4.3.4\text{-}2)$$

$$\text{cov}[v(k), v(j)] = E[[v(k) - \bar{v}(k)][v(j) - \bar{v}(j)]'] = Q(k)\delta_{kj} \quad (4.3.4\text{-}3)$$

Then the expected value of the state

$$\bar{x}(k) \triangleq E[x(k)] \quad (4.3.4\text{-}4)$$

evolves according to the difference equation

$$\boxed{\bar{x}(k+1) = F(k)\bar{x}(k) + G(k)u(k) + \Gamma(k)\bar{v}(k)} \quad (4.3.4\text{-}5)$$

The above, which is the **propagation equation of the mean**, follows immediately by applying the expectation operator to (4.3.4-1)

Propagation of covariance

The covariance of the state

$$P_{xx}(k) \triangleq E[[x(k) - \bar{x}(k)][x(k) - \bar{x}(k)]'] \tag{4.3.4-6}$$

evolves according to the difference equation — **the covariance propagation Equation**

$$\boxed{P_{xx}(k+1) = F(k)P_{xx}(k)F(k)' + \Gamma(k)Q(k)\Gamma(k)'} \tag{4.3.4-7}$$

This follows by subtracting (4.3.4-5) from (4.3.4-1), which yields

$$x(k+1) - \bar{x}(k+1) = F(k)[x(k) - \bar{x}(k)] + \Gamma(k)[v(k) - \bar{v}(k)] \tag{4.3.4-8}$$

It can be easily shown that multiplying (4.3.4-8) with its transpose and taking the expectation yields (4.3.4-7).

Continuous-time linear stochastic dynamic systems – for comparison

The Continuous-Time State-Space Model

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + D(t)\tilde{v}(t)$$

$\tilde{v}$ is the (continuous-time) input disturbance or ***process noise***, also called ***plant noise***, a vector of dimension n_v,

$$z(t) = C(t)x(t) + \tilde{w}(t)$$

$\tilde{w}$ is the (unknown) output disturbance or ***measurement noise***

In the stochastic case, the noises are usually assumed to be

1. zero-mean,
2. white, and
3. mutually independent

stochastic processes.

Propagation of the State's Mean and Covariance

- The state of a dynamic system driven by white noise is a **Markov process**

Let's assume known input u(t) and nonstationary white process noise with nonzero mean

$$E[\tilde{v}(t)] = \bar{v}(t)$$

and autocovariance function

$$E[[\tilde{v}(t) - \bar{v}(t)][\tilde{v}(\tau) - \bar{v}(\tau)]'] = V(t)\delta(t-\tau)$$

The expected value of the state $\bar{x}(t) \triangleq E[x(t)]$

The propagation equation of the mean follows from taking the expected value of continuous state space equation.

$$\dot{\bar{x}}(t) = A(t)\bar{x}(t) + B(t)u(t) + D(t)\bar{v}(t)$$

The covariance of the state $P_{xx}(t) \triangleq E[[x(t) - \bar{x}(t)][x(t) - \bar{x}(t)]']$

evolves according to the differential equation, known as the Lyapunov equation

$$\dot{P}_{xx}(t) = A(t)P_{xx}(t) + P_{xx}(t)A(t)' + D(t)V(t)D(t)'$$

This is the propagation equation of the covariance

Aalto University

STATE ESTIMATION IN DISCRETE TIME LINEAR DYNAMIC SYSTEMS

The estimation of the state vector of a stochastic linear dynamic system is considered.

The state estimator for discrete-time linear dynamic systems driven by white noise — the (discrete-time) Kalman filter — is introduced.

Estimation of Gaussian random vectors.

Kalman Filter

For linear systems, white noise Gaussian processes

Linear equations used for state prediction, prediction of the measurement and for measurement update.

Exact propagation and measurement update equations for a priori and a posteriori covariances

All pdfs stay exactly Gaussian, no need for approximations

Reasoning of Kalman filter

Prediction, propagation of covariance, with linear dynamic models with white noise

Measurement update, direct application of fundamental equations of linear estimation

Estimation of Gaussian random vectors

x and z jointly Gaussian

z is the measurement

x random variable to be estimated

$$y \triangleq \begin{bmatrix} x \\ z \end{bmatrix} \qquad \bar{y} = \begin{bmatrix} \bar{x} \\ \bar{z} \end{bmatrix} \qquad P_{yy} = \begin{bmatrix} P_{xx} & P_{xz} \\ P_{zx} & P_{zz} \end{bmatrix}$$

$$y \sim \mathcal{N}[\bar{y}, P_{yy}] \qquad P_{xx} = E[(x-\bar{x})(x-\bar{x})'] \qquad P_{xz} = E[(x-\bar{x})(z-\bar{z})'] = P'_{zx}$$

The MMSE Minimum Mean Square Error -estimator, conditional mean of x given z, for linear Gaussian case also Maximum a Posteriori MAP -estimator

$$\boxed{\hat{x} \triangleq E[x|z] = \bar{x} + P_{xz}P_{zz}^{-1}(z-\bar{z})}$$

$$\boxed{P_{xx|z} \triangleq E[(x-\hat{x})(x-\hat{x})'|z] = P_{xx} - P_{xz}P_{zz}^{-1}P_{zx}}$$

Fundamental equations of linear estimation - Interpretations

- *A priori* estimate is updated/corrected on the basis of measurement information in calculation of *a posteriori* estimate
- Correction gain depends directly on P_{xz} the crosscovariance between x and measurement z.
- Correction effect depends inversely proportional on P_{zz} . The better measurements, the ’smaller’ covariance, the bigger the correction gain.

The Dynamic Estimation Problem

$$x(k+1) = F(k)x(k) + G(k)u(k) + v(k) \qquad k = 0, 1, \ldots$$

v(k), k = 0, 1, . . . , is the sequence of zero-mean white Gaussian process noise with covariance

$$E[v(k)v(k)'] = Q(k)$$

$$z(k) = H(k)x(k) + w(k) \qquad k = 1, \ldots$$

with w(k) the sequence of zero-mean white Gaussian measurement noise with covariance

$$E[w(k)w(k)'] = R(k)$$

the **linear Gaussian (LG)** assumption. The linearity leads to the preservation of the Gaussian property of the state and measurements — this is **a Gauss-Markov process**

The **conditional mean**

$$\hat{x}(j|k) \triangleq E[x(j)|Z^k] \qquad Z^k \triangleq \{z(i), i \leq k\}$$

Estimate of the state if *j* = *k* (also called filtered value)

Smoothed value of the state if *j* < *k*

Predicted value of the state if *j* > *k*

The estimation error is defined as

$$\tilde{x}(j|k) \triangleq x(j) - \hat{x}(j|k)$$

The **conditional covariance** matrix of *x(j)* given the data or the covariance associated with the estimate is

$$P(j|k) \triangleq E[[x(j) - \hat{x}(j|k)][x(j) - \hat{x}(j|k)]'|Z^k] = E[\tilde{x}(j|k)\tilde{x}(j|k)'|Z^k]$$

MMSE criterion for estimation leads to the conditional mean as the optimal estimate

The Estimation Algorithm

The estimation algorithm starts with the initial estimate and the associated initial covariance *P(0|0)*, assumed to be available. One cycle of the dynamic estimation algorithm — the Kalman filter (KF) — will thus consist of mapping the estimate

$$\hat{x}(k|k) \triangleq E[x(k)|Z^k]$$

and the associated covariance matrix

$$P(k|k) = E[[x(k) - \hat{x}(k|k)][x(k) - \hat{x}(k|k)]'|Z^k]$$

into the corresponding variables at the next stage

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

Gaussian random variable is fully characterized by its first two moments

Dynamic Estimation as a Recursive Static Estimation

the state estimate at k + 1 and its covariance can be obtained from the static estimation equations

$$\hat{x} \triangleq E[x|z] = \bar{x} + P_{xz}P_{zz}^{-1}(z - \bar{z})$$

$$P_{xx|z} \triangleq E[(x - \hat{x})(x - \hat{x})'|z] = P_{xx} - P_{xz}P_{zz}^{-1}P_{zx}$$

The **prior** (unconditional) expectations from the static case become **prior to the availability of the measurement at time k + 1** in the dynamic case, that is, given the data up to and including k.

The **posterior** (conditional) expectations become **posterior to obtaining the measurement at time k + 1**, that is, given the data up to and including k + 1.

Measurement update, direct application of fundamental equations of linear estimation

- The *prior* (unconditional) expectations from the static case become *prior to the availability of the measurement at time k + 1* in the dynamic case, that is, given the data up to and including k.
- The *posterior* (conditional) expectations become *posterior to obtaining the measurement at time k + 1*, that is, given the data up to and including k + 1.
- The variable to be estimated is the state at *k* + 1

$$x \rightarrow x(k+1) \tag{5.2.2-3}$$

- Its mean *prior to k + 1* — the **(one-step) predicted state** — is

$$\bar{x} \rightarrow \bar{x}(k+1) \triangleq \hat{x}(k+1|k) \triangleq E[x(k+1)|Z^k] \tag{5.2.2-4}$$

Measurement update, …

Based on the observation (measurement)

$$z \longrightarrow z(k+1) \tag{5.2.2-5}$$

with prior mean — the predicted measurement

$$\bar{z} \longrightarrow \bar{z}(k+1) \triangleq \hat{z}(k+1|k) \triangleq E[z(k+1)|Z^k] \tag{5.2.2-6}$$

one computes the estimate posterior to *k+1* — the updated state estimate

$$\hat{x} \longrightarrow \hat{x}(k+1) \triangleq \hat{x}(k+1|k+1) \triangleq E[x(k+1)|Z^{k+1}] \tag{5.2.2-7}$$

The Covariances

- The prior covariance matrix of the state variable *x(k + 1)* to be estimated — the state prediction covariance or predicted state covariance — is

$$P_{xx} \rightarrow \bar{P}(k+1) \triangleq P(k+1|k) \triangleq \mathrm{cov}[x(k+1)|Z^k] = \mathrm{cov}[\tilde{x}(k+1|k)|Z^k] \tag{5.2.2-8}$$

Measurement update, covariances

- The (prior) covariance of the observation z(k + 1) — the measurement prediction covariance — is

$$P_{zz} \rightarrow S(k+1) \triangleq \text{cov}[z(k+1)|Z^k] = \text{cov}[\tilde{z}(k+1|k)|Z^k] \qquad (5.2.2\text{-}9)$$

- The covariance between the variable to be estimated x(k + 1) and the observation z(k + 1) is

$$P_{xz} \rightarrow \text{cov}[x(k+1), z(k+1)|Z^k] = \text{cov}[\tilde{x}(k+1|k), \tilde{z}(k+1|k)|Z^k] \quad (5.2.2\text{-}10)$$

- The posterior covariance of the state x(k + 1) — the updated state covariance — is

$$\begin{aligned} P_{xx|z} \rightarrow P(k+1) &\triangleq P(k+1|k+1) = \text{cov}[x(k+1)|Z^{k+1}] \\ &= \text{cov}[\tilde{x}(k+1|k+1)|Z^{k+1}] \end{aligned} \qquad (5.2.2\text{-}11)$$

In Kalman filter: a priori estimate by prediction

- mean *prior to k + 1* — the **(one-step) predicted state** — is

$$\bar{x} \rightarrow \bar{x}(k+1) \triangleq \hat{x}(k+1|k) \triangleq E[x(k+1)|Z^k] \qquad (5.2.2\text{-}4)$$

$$\boxed{\hat{x}(k+1|k) = F(k)\hat{x}(k|k) + G(k)u(k)} \qquad (5.2.3\text{-}2)$$

- The corresponding covariance

$$P_{xx} \rightarrow \bar{P}(k+1) \triangleq P(k+1|k) \triangleq \text{cov}[x(k+1)|Z^k] = \text{cov}[\tilde{x}(k+1|k)|Z^k] \qquad (5.2.2\text{-}8)$$

$$\boxed{P(k+1|k) = F(k)P(k|k)F(k)' + Q(k)} \qquad (5.2.3\text{-}5)$$

In Kalman filter: Measurement update

- With prior mean — the predicted measurement

$$\bar{z} \to \bar{z}(k+1) \triangleq \hat{z}(k+1|k) \triangleq E[z(k+1)|Z^k] \qquad (5.2.2\text{-}6)$$

$$\boxed{\hat{z}(k+1|k) = H(k+1)\hat{x}(k+1|k)} \qquad (5.2.3\text{-}7)$$

- The corresponding covariance

$$P_{zz} \to S(k+1) \triangleq \mathrm{cov}[z(k+1)|Z^k] = \mathrm{cov}[\tilde{z}(k+1|k)|Z^k] \qquad (5.2.2\text{-}9)$$

$$\boxed{S(k+1) = H(k+1)P(k+1|k)H(k+1)' + R(k+1)} \qquad (5.2.3\text{-}9)$$

Measurement update

- one computes the estimate posterior to *k+1* — the updated state estimate

$$\hat{x} \rightarrow \hat{x}(k+1) \triangleq \hat{x}(k+1|k+1) \triangleq E[x(k+1)|Z^{k+1}] \qquad (5.2.2\text{-}7)$$

$$\boxed{\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + W(k+1)\nu(k+1)} \qquad (5.2.3\text{-}12)$$

$$\boxed{\nu(k+1) \triangleq z(k+1) - \hat{z}(k+1|k) = \tilde{z}(k+1|k)} \qquad (5.2.3\text{-}13)$$

$$P_{xz} \rightarrow \text{cov}[x(k+1), z(k+1)|Z^k] = \text{cov}[\tilde{x}(k+1|k), \tilde{z}(k+1|k)|Z^k] \quad (5.2.2\text{-}10)$$

$$\begin{aligned} &E[\tilde{x}(k+1|k)\tilde{z}(k+1|k)'|Z^k] \\ &\qquad = E\left[\tilde{x}(k+1|k)[H(k+1)\tilde{x}(k+1|k) + w(k+1)]'|Z^k\right] \\ &\qquad = P(k+1|k)H(k+1)' \end{aligned} \qquad (5.2.3\text{-}10)$$

$$P_{xz}P_{zz}^{-1} \rightarrow W(k+1) \triangleq \text{cov}[x(k+1), z(k+1)|Z^k]S(k+1)^{-1} \qquad (5.2.2\text{-}12)$$

$$\boxed{W(k+1) \triangleq P(k+1|k)H(k+1)'S(k+1)^{-1}} \qquad (5.2.3\text{-}11)$$

In Kalman filter: Measurement update, covariances

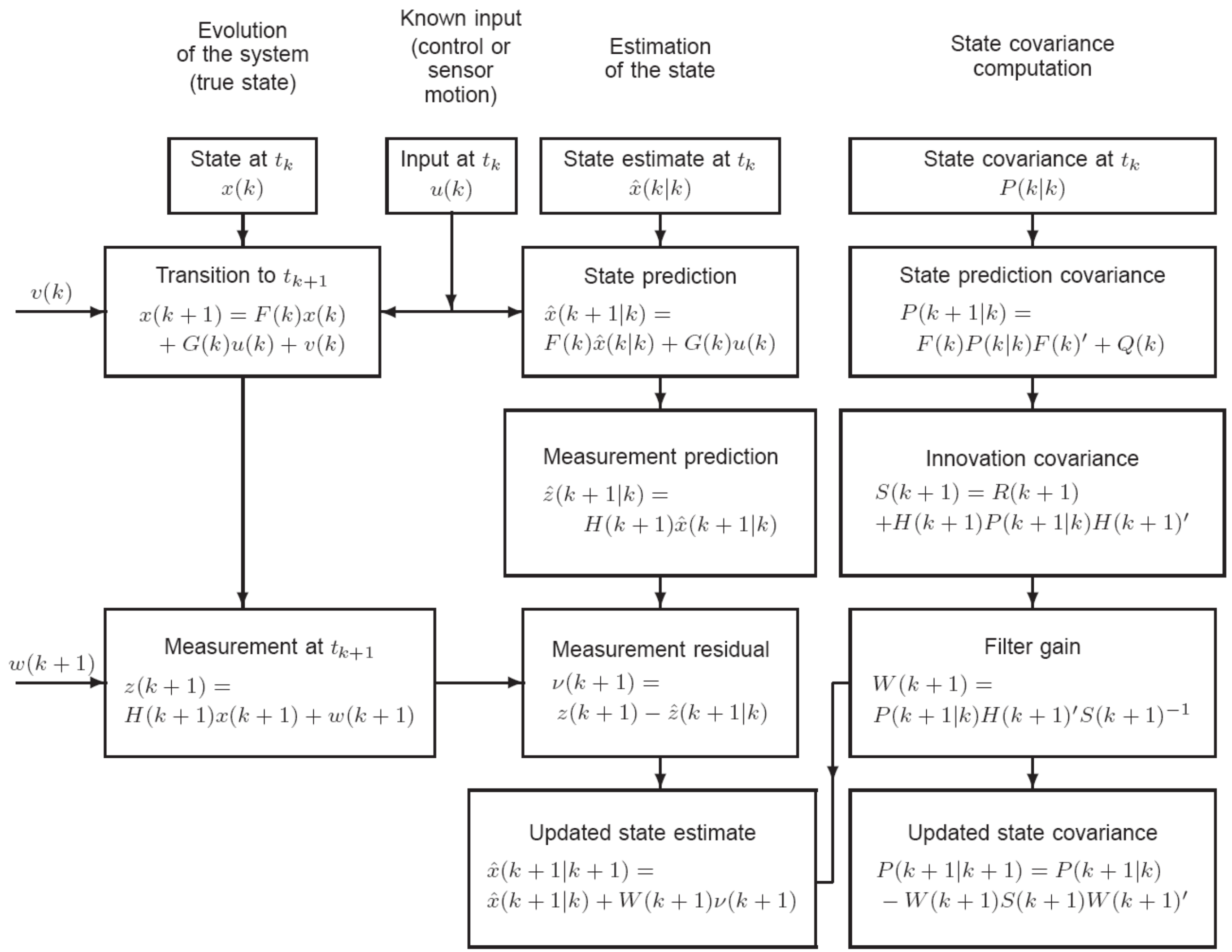
- The corresponding covariance

$$P_{xx|z} \rightarrow P(k+1) \triangleq P(k+1|k+1) = \text{cov}[x(k+1)|Z^{k+1}]$$
$$= \text{cov}[\tilde{x}(k+1|k+1)|Z^{k+1}] \qquad (5.2.2\text{-}11)$$

$$P(k+1|k+1) = P(k+1|k) - P(k+1|k) \cdot H(k+1)' S(k+1)^{-1} H(k+1) P(k+1|k)$$
$$= [I - W(k+1)H(k+1)]P(k+1|k) \qquad (5.2.3\text{-}14)$$

$$\boxed{P(k+1|k+1) = P(k+1|k) - W(k+1)S(k+1)W(k+1)'} \qquad (5.2.3\text{-}15)$$

Kalman filter



Intuitive Interpretation of the Gain

the optimal filter gain is

1. "Proportional" to the state prediction variance
2. "Inversely proportional" to the innovation variance

Thus, the gain is

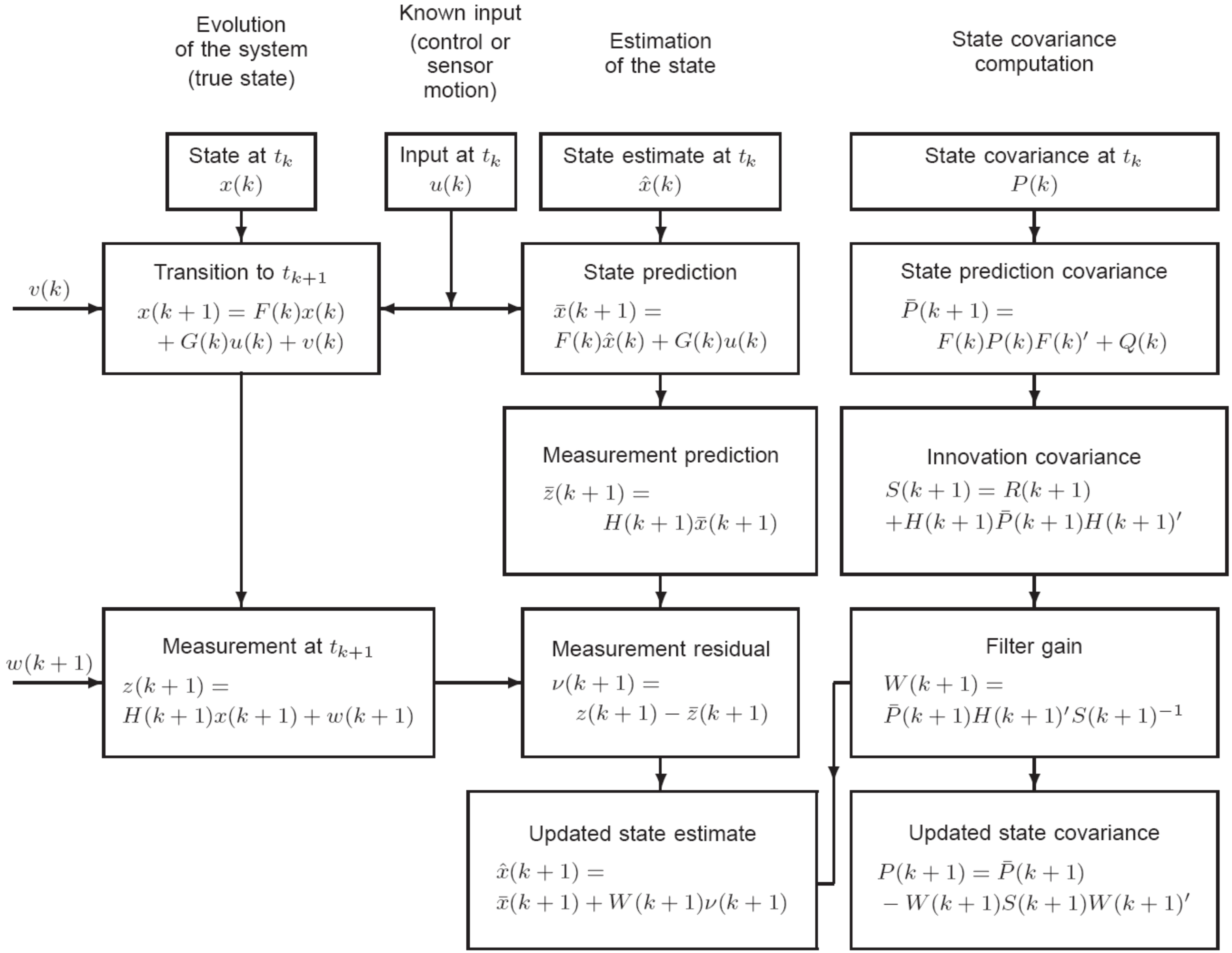
- "Large" if the state prediction is "inaccurate" (has a large variance) and the measurement is "accurate" (has a relatively small variance)
- "Small" if the state prediction is "accurate" (has a small variance) and the measurement is "inaccurate" (has a relatively large variance)

A large gain indicates a "rapid" response to the measurement in updating the state, while a small gain yields a slower response to the measurement.

Overview of the Kalman Filter Algorithm

- **Under the Gaussian assumption** for the initial state (or initial state error) and all the noises entering into the system, the Kalman filter is the **optimal MMSE state estimator**
- **If these random variables are not Gaussian** and one has only their first two moments, then, the Kalman filter algorithm is **the best linear state estimator**, that is, **the LMMSE state estimator**
- Note that at every stage (cycle) k the entire past is summarized by the sufficient statistic $\hat{x}(k|k)$ and the associated covariance P(k|k).
- The state update requires the filter gain, obtained in the course of the covariance calculations. The covariance calculations are independent of the state and measurements (and control — assumed to be known) and can, therefore, be performed offline.
- **Statistical assumptions**: The initial state has the known mean and covariance. The process and measurement noise sequences are zero mean and white with known covariance matrices. **All the above are mutually uncorrelated**

Kalman filter with single-argument notations



The Matrix Riccati Equation

The covariance equations in the static MMSE estimation problem are independent of the measurements

the following recursion can be written for the one-step prediction covariance

$$
\begin{aligned}
P(k+1|k) &= F(k)\{P(k|k-1) - P(k|k-1)H(k)' \\
&\quad \cdot [H(k)P(k|k-1)H(k)' + R(k)]^{-1}H(k)P(k|k-1)\}F(k)' + Q(k)
\end{aligned}
$$

This is the discrete-time (difference) matrix Riccati equation, or just the Riccati equation.

The solution of the above Riccati equation for a time-invariant system converges to a finite steady-state covariance if the pair {F,H} is completely observable.

The steady-state gain for the Kalman filter

The steady-state covariance matrix is the solution of the algebraic matrix Riccati equation (or just the algebraic Riccati equation)

$$P = F[P - PH'(HPH' + R)^{-1}HP]F' + Q$$

and this yields the steady-state gain for the Kalman filter.

Stability: The convergence of the covariance to a finite steady state — that is, the error becoming a stationary process in the MS sense — is equivalent to filter stability in the bounded input bounded output sense.

CRLB: The lower bound on the minimum achievable covariance in state estimation is given by the (posterior) CRLB. In the linear Gaussian case it can be shown that this is given by the solution of the Riccati equation.

The Innovations — a Zero-Mean White Sequence

An important property of the innovation sequence is that it is an orthogonal sequence, that is

$$\boxed{E[\nu(k)\nu(j)'] = S(k)\delta_{kj}}$$

the **innovation sequence is zero mean and white**.

In fault sitations, the dynamics of the real systems does not anymore behave as described with the dynamic model and the measurement model in the Kalman filter, **the innovation sequence becomes then non-white**.

Fault situations can be detected by testing the whiteness of the innovation sequence of the Kalman filter.