Multiple-Model Adaptive Estimation (MMAE)

MICHAEL ATHANS

MIT & ISR/IST

Last Revised: October 17, 2001

Ref. No. KF # 15

Theme

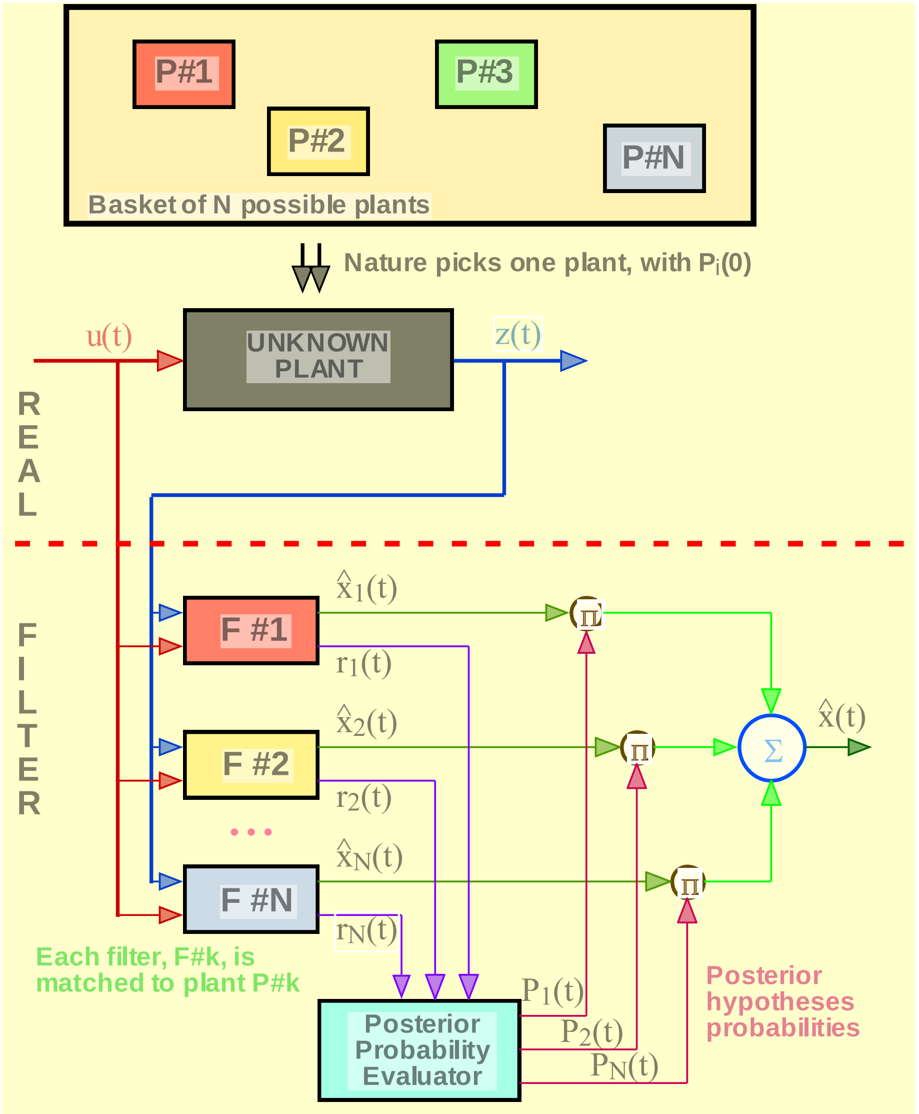
- We discuss a very powerful method, the so-called "Multiple-Model Adaptive Estimation (MMAE)" algorithm, for combined state-estimation and system-identification problems, [1]-[4]
  - the MMAE combines hypotheses-testing and state-estimation
- The detailed results will be presented for linear time-varying (LTV) discrete-time systems with Gaussian uncertainties
  - we shall take full advantage of the classical discrete-time Kalman filter theory for linear-gaussian problems
  - we shall provide complete proofs, through extensive use of Gaussian conditional density functions and repeated applications of Bayes rule, following the development in [1]
- The MMAE framework can be used for suboptimal estimation in nonlinear nongaussian situations

Typical MMAE Applications

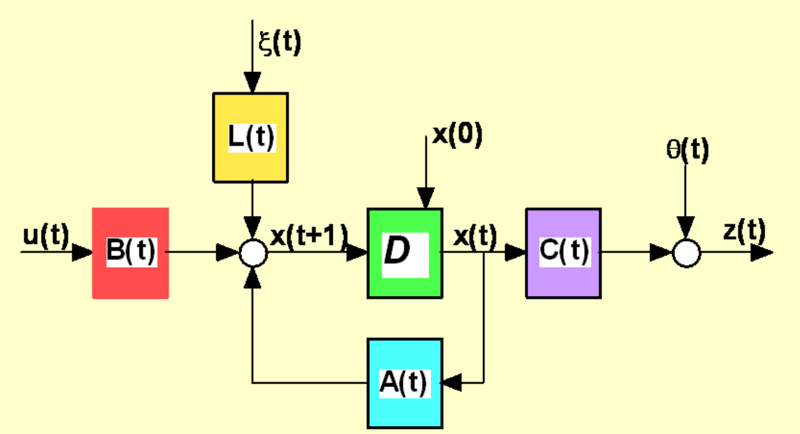
- Tracking maneuvering vehicles in aerospace, ground, and marine applications, taking into account that applied maneuvers are not known to the observer a-priori
- Doing accurate state-estimation for plants with wide uncertainty in the parameters of the dynamic system and sensors
- Doing accurate state-estimation for nonlinear systems with large uncertainties, by employing a family of distinct linearized models
- Initial framework for dealing with multi-target multi-sensor surveillance and tracking systems

The MMAE Structure

- The (unknown) plant, that generates the data, is one (or close to one) of the N possible models
- The MMAE algorithm uses a bank of N parallel (Kalman) filters, each matched to one of the N models
- Each (Kalman) filter generates its own state-estimate and residual
- Posterior probabilities are generated on-line and weigh local state-estimates



LTV Discrete-Time Models



MODEL OF K - TH PLANT

- Time index: $t = 0, 1, 2, \ldots.$
- Model index: $k = 1, 2, \ldots, N$

$$x(t+1) = A_k(t)x(t) + B_k(t)u(t) + L_k(t)\xi(t)$$

$$z(t+1) = C_k(t+1)x(t+1) + \theta(t+1)$$

PROBABILISTIC INFORMATION

- Initial state: $x(0) \sim N(\bar{x}_{0k}, \Sigma_{0k})$
- Plant disturbance: $\xi(t) \sim N(0, \Xi_k(t)\delta_{t\tau})$
- Sensor noise: $\theta(t) \sim N(0, \Theta_k(t)\delta_{t\tau})$
- $x(0), \xi(t), \theta(\tau)$ independent for all t, τ

NOTES

- For each model, indexed by $k = 1, 2, \ldots, N,$ some or all plant and sensor parameter matrices can be different
- For each model, indexed by $k = 1, 2, \ldots, N,$ the statistics of the initial state and / or plant disturbance and / or sensor noise can be different

MMAE: Problem Formulation

- GIVEN: prior probabilities, $P_k(0) \quad k = 1, 2, \ldots, N,$
  that nature selects the k-th model to generate data, with

  $$P_k(0) \geq 0, \quad \sum_{k=0}^{N} P_k(0) = 1$$

- GIVEN: the set of past controls

  $$u(0), u(1), u(2), \ldots, u(t-1)$$

  and the set of past measurements, including the one at the "present" time t

  $$z(1), z(2), \ldots, z(t-1), z(t)$$

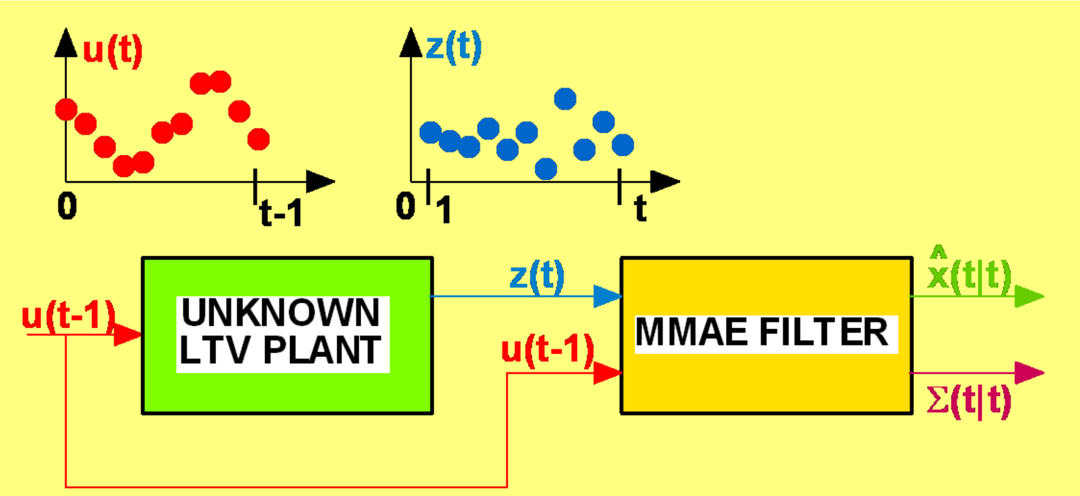
- DETERMINE: (1) the true conditional mean of the present state vector, $x(t)$, i.e.

  $$\hat{x}(t \mid t) \equiv E\left\{ x(t) \mid \underbrace{u(0), u(1), u(2), \ldots, u(t-1); z(1), z(2), \ldots, z(t-1), z(t)}_{Z(t)} \right\}$$

  and (2) the true conditional covariance matrix of the present state vector, $x(t)$, i.e.

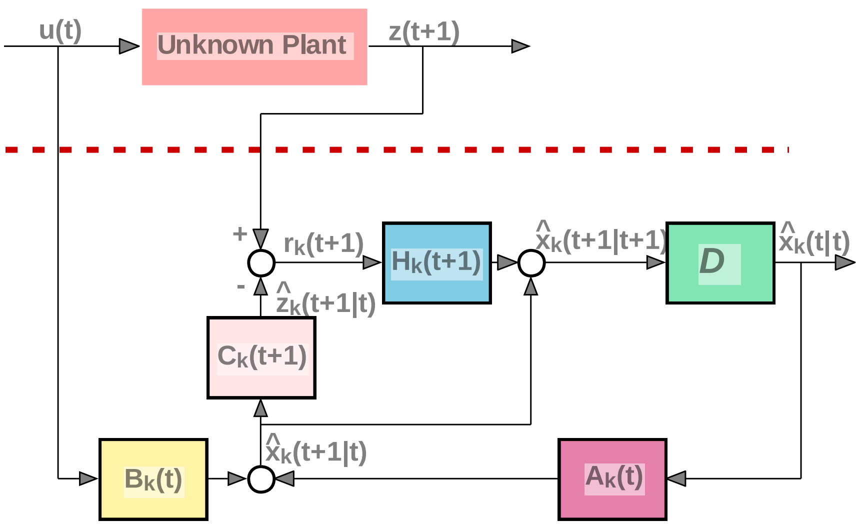
  $$\Sigma(t \mid t) \equiv E\left[\left(x(t) - \hat{x}(t \mid t)\right)\left(x(t) - \hat{x}(t \mid t)\right)' \mid Z(t) \right]$$

MMAE: Problem Visualization



- The MMAE filter is driven by the sequence of past controls and noisy sensor measurements
- The MMAE filter generates both a state-estimate vector and a corresponding error-covariance matrix
- The MMAE is a recursive algorithm: it updates the state-estimate and covariance every time a new control is applied and a new sensor measurement is obtained

Review: The Discrete-Time Kalman Filter



Kth KALMAN FILTER EQUATIONS

- Predict Cycle:

$$\hat{x}_k(t+1 \mid t) = A_k(t)\hat{x}_k(t \mid t) + B_k(t)u(t)$$

$$\hat{z}_k(t+1 \mid t) = C_k(t+1)\hat{x}_k(t+1 \mid t)$$

- Residual: $r_k(t+1) \equiv z(t+1) - \hat{z}_k(t+1 \mid t)$
- Update Cycle:

$$\hat{x}_k(t+1 \mid t+1) = \hat{x}_k(t+1 \mid t) + H_k(t+1)r_k(t+1)$$

- KF Gain Matrix:

$$H_k(t+1) = \Sigma_k(t+1 \mid t+1)C_k'(t+1)\Theta_k^{-1}(t+1)$$

Kth KF COVARIANCE EQUATIONS

- Initialization: $\Sigma_k(0 \mid 0) = \Sigma_{0k}$
- Predict Cycle:

$$\Sigma_k(t+1 \mid t) = A_k(t)\Sigma_k(t \mid t)A_k'(t) + L_k(t)\Xi_k(t)L_k'(t)$$

- Update Cycle:

$$\Sigma_k(t+1 \mid t+1) = \Sigma_k(t+1 \mid t) + \Sigma_k(t+1 \mid t)C_k'(t+1) \bullet$$

$$\bullet \left[C_k(t+1)\Sigma_k(t+1 \mid t)C_k'(t+1) + \Theta_k(t+1)\right]^{-1} \bullet$$

$$\bullet\, C_k(t+1)\Sigma_k(t+1 \mid t)$$

Kth RESIDUAL INFORMATION

- Residual definition:

$$r_k(t+1) \equiv z(t+1) - \hat{z}_k(t+1 \mid t)$$

- Residual covariance matrix:

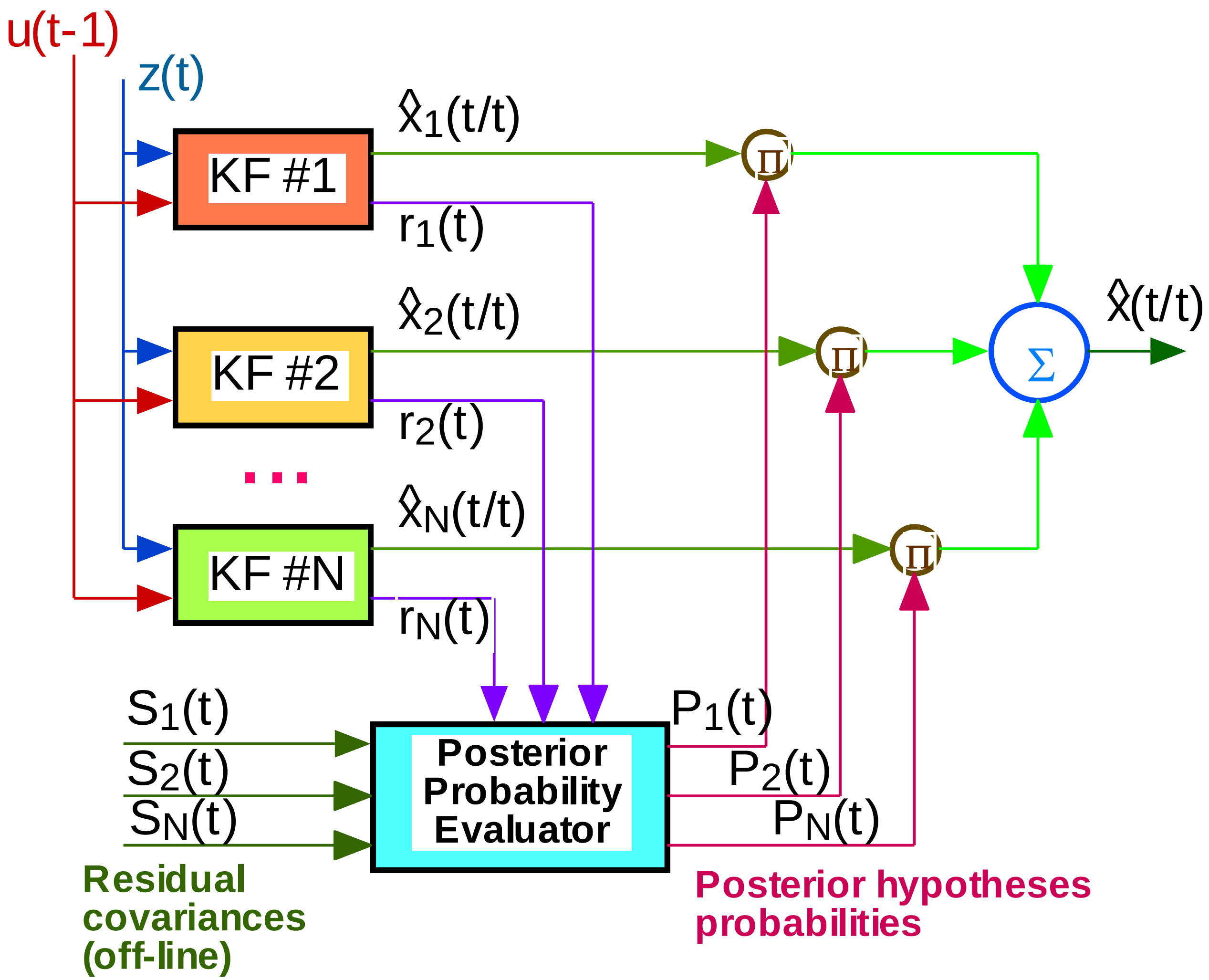
$$S_k(t+1) \equiv cov\left[r_k(t+1), r_k(t+1)\right] =$$

$$= C_k(t+1)\Sigma_k(t+1 \mid t)C_k'(t+1) + \Theta(t+1)$$

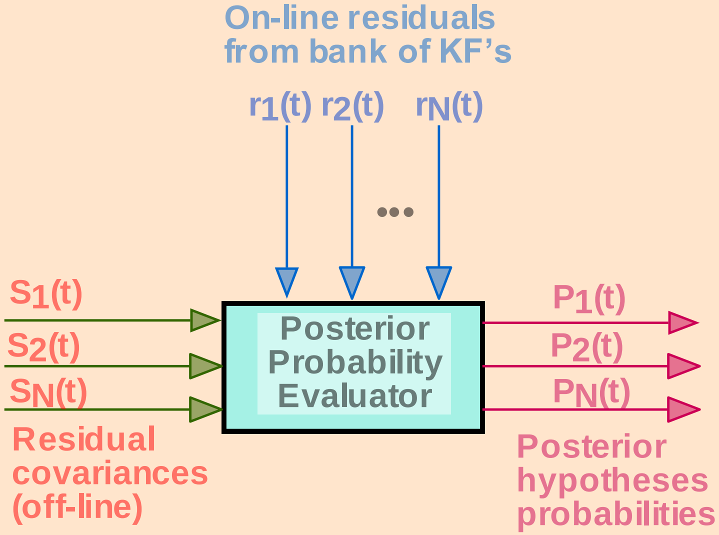
Basic Idea of the MMAE Method

- Construct a bank of N discrete-time Kalman filters, each KF "matched" to each of the N possible models
- Each KF generates (in real-time) a local state-estimate vector and a residual vector
- All of the N available KF residual vectors are used to compute (on-line) the posterior probability $P_k(t)$, $k=1,2, \ldots, N$, that the *kth* model is indeed the true one (I.e. the one that generates the data)
- The overall MMAE state-estimate is formed by weighting the local state-estimates by the corresponding posterior probability
- The overall MMAE state-covariance matrix is formed by weighting the local state-covariance matrices by the corresponding posterior probability, including a correction that involves the global conditional mean

The MMAE Filter



Posterior Probability Evaluator



POSTERIOR PROBABILITIES

- Posterior probability:

$$P_k(t), \quad t = 1, 2, \ldots, ; \quad k = 1, 2, \ldots, N$$

$$P_k(t) = \text{Prob}\{k^{th} \text{ model is true } \mid Z(t)\}$$

- Prior probabilities $P_k(0)$ are known

DEFINITIONS

- k^{th} KF residual: $r_k(t); \quad r_k(t) \in R^m$
- k^{th} KF residual covariance: $S_k(t);\ m \times m$ matrix
- Define the scalar quantities

$$\beta_k(t+1) \equiv \frac{1}{(2\pi)^{m/2}\sqrt{\det S_k(t+1)}}$$

$$w_k(t+1) \equiv r_k'(t+1) S^{-1}{}_k(t+1) r_k(t+1)$$

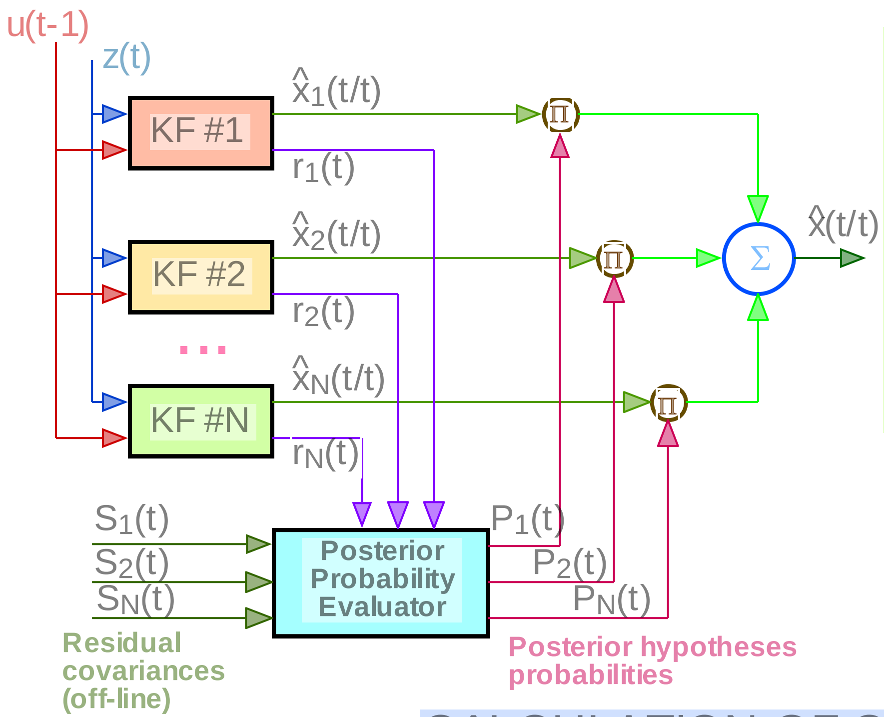
DYNAMIC PROBABILITY EVALUATION

- For $t = 0, 1, 2, \ldots$ and $k = 1, 2, \ldots, N$

$$P_k(t+1) = \left(\frac{\beta_k(t+1) e^{-\frac{1}{2} w_k(t+1)}}{\sum_{j=1}^{N} \beta_j(t+1) e^{-\frac{1}{2} w_j(t+1)} P_j(t)} \right) \bullet P_k(t)$$

with the initial probabilities $P_k(0) = \text{given}$

Result Summary: Global State Estimate



CALCULATION OF STATE-ESTIMATE

- Sum individual KF state-estimates multiplied by the associated posterior probabilities

$$\hat{x}(t \mid t) = \sum_{k=1}^{N} P_k(t)\hat{x}_k(t \mid t)$$

CALCULATION OF COVARIANCE MATRIX

$$\Sigma(t \mid t) = \sum_{k=1}^{N} P_k(t)\left[\Sigma_k(t \mid t) + \left(\hat{x}_k(t \mid t) - \hat{x}(t \mid t)\right)\left(\hat{x}_k(t \mid t) - \hat{x}(t \mid t)\right)'\right]$$

- Note that the covariance matrix $\Sigma(t \mid t)$ must be computed on-line

Important Remark

- Both the MMAE state-estimate and state-covariance matrix represent true conditional expectation and conditional covariance
- This fact will be proven in the sequel

DEFINITIONS AND NOTATION

- Data set at time t: $Z(t) = \{u(0), u(1), \ldots, u(t-1); z(1), z(2), \ldots, z(t)\}$
- Then, $\hat{x}(t \mid t) = E\{x(t) \mid Z(t)\}$, and

  $\Sigma(t \mid t) = cov[x(t); x(t) \mid Z(t)]$
- To prove these assertions, we must explicitly calculate the conditional probability density function $p(x(t) \mid Z(t))$
- We shall show that the desired pdf, $p(x(t) \mid Z(t))$, turns out to be a weighted sum of gaussian densities, where the weights are found from the posterior probability evaluator

Elements of Proof

- The problem is a combination of a hypothesis-testing problem and a state-estimation problem
- The fact that one of the *N* models is the true one is modeled by hypothesis random variable that must belong to a discrete set of hypotheses $H_1, H_2, \ldots, H_N$
- The focal point is to calculate the conditional probability density function, $p(x(t)|Z(t))$, of the state at time *t*, given measurements up to time *t*. Then,
  - the conditional expectation of the state, $E\{x(t)|Z(t)\}$, provides the global state estimate
  - the conditional covariance of the state, $cov[x(t);x(t)|Z(t)]$, provides the measure of uncertainty
- It also turns out that on-line generation of the posterior conditional probabilities determines which hypothesis is true

Proof: Hypotheses

- Hypothesis random variable (scalar) is H
- H can attain only one of N possible values,

  (1) $H \in \{H_1, H_2, \ldots, H_N\}$
- The event $H = H_k$ means that the k - th system is the true one, i.e. the one that is generating the data inside the black box
- Prior probability: $P_k(0) \equiv Prob(H = H_k)$ at initial time $t = 0$

  (2) $P_k(0) \geq 0, \quad \sum_{k=1}^{N} P_k(0) = 1$
- Data set at time $t, t = 0, 1, 2, \ldots,$

  (3) $Z(t) = \{u(0), u(1), \ldots, u(t-1); z(1), z(2), \ldots, z(t)\}$

  consists of the set of past applied controls and observed sensor measurements, including the latest one at time $t, z(t)$
- Posterior probability: $P_k(t) = Prob(H = H_k \mid Z(t))$

  (4) $P_k(t) \geq 0, \quad \sum_{k=1}^{N} P_k(t) = 1$

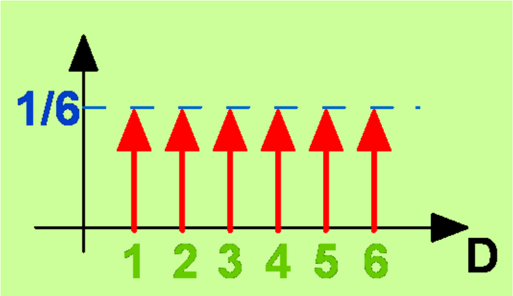
Digression: Discrete Random Variables, I

OUTCOME OF PERFECT DIE

- Let D be the number; $D \in \{1, 2, 3, 4, 5, 6\}$
- The probability of each outcome is $\frac{1}{6}$
- The probability density function is

$$p(D) = \frac{1}{6}\delta(D-1) + \frac{1}{6}\delta(D-2) + \frac{1}{6}\delta(D-3) +$$

$$+ \frac{1}{6}\delta(D-4) + \frac{1}{6}\delta(D-5) + \frac{1}{6}\delta(D-6) = \sum_{k=1}^{6} \frac{1}{6}\delta(D-k)$$



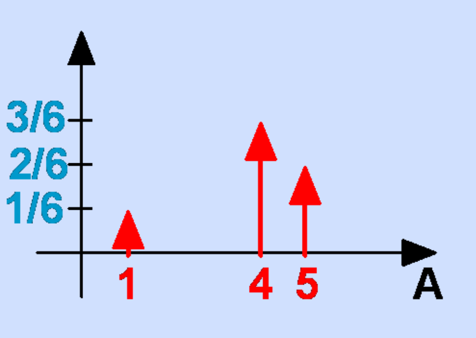
3 - VALUED RANDOM VARIABLE

- Suppose that a discrete random variable A can only attain three values, 1, 4, and 5, $A \in \{1, 4, 5\}$
- Assume that the probabilities of each outcome are

$$Prob(A=1) = \frac{1}{6},\ Prob(A=4) = \frac{3}{6},\ Prob(A=5) = \frac{2}{6}$$

- Then, the PDF of A is given by

$$p(A) = \frac{1}{6}\delta(A-1) + \frac{3}{6}\delta(A-4) + \frac{2}{6}\delta(A-5)$$



Digression: Discrete Random Variables, II

GENERAL CASE OF DISCRETE RANDOM VARIABLES

- Let x be a discrete-valued scalar random variable

  $$x \in \{X_1, X_2, \ldots X_M\}$$

- Suppose that the probability that x attains a particular value is given by

  $$Prob(x = X_k) = P_k \,; \quad k = 1, 2, \ldots, M$$

  $$P_k \geq 0 \,; \quad \sum_{k=1}^{M} P_k = 1$$

- Then the probability density function, $p(x)$, of the RV x is

  $$p(x) = \sum_{k=1}^{M} P_k \delta(x - X_k) \,; \quad \delta(x - X_k) = \text{unit impulse at } x = X_k$$

- Note that the area under the pdf is unity, since the area of each unit impulse is unity, i.e.

  $$\int p(x)dx = \int \sum_{k=1}^{M} P_k \delta(x - X_k)dx = \sum_{k=1}^{M} P_k \underbrace{\int \delta(x - X_k)dx}_{1} = 1$$

Proof: Density Relations

- Key quantity of interest is the conditional density function, $p\left(x(t) \mid Z(t)\right)$
- Consider the joint density function, $p\left(x(t), H \mid Z(t)\right)$
- Using the marginal density function we have

(5) $$p\left(x(t) \mid Z(t)\right) = \int p\left(x(t), H \mid Z(t)\right) dH$$

- From Bayes rule we have

(6) $$p\left(x(t), H \mid Z(t)\right) = p\left(x(t) \mid H, Z(t)\right) p\left(H \mid Z(t)\right)$$

- Substitute (6) into (5) and use (4) to obtain

(7) $$p\left(x(t) \mid Z(t)\right) = \int p\left(x(t) \mid H, Z(t)\right) p\left(H \mid Z(t)\right) dH =$$

$$= \int p\left(x(t) \mid H, Z(t)\right) \sum_{k=1}^{N} P_k(t) \delta\left(H - H_k\right) dH =$$

$$= \sum_{k=1}^{N} P_k(t) \int p\left(x(t) \mid H, Z(t)\right) \delta\left(H - H_k\right) dH =$$

$$= \sum_{k=1}^{N} P_k(t) p\left(x(t) \mid H_k, Z(t)\right)$$

Proof: Conditional Density Relations

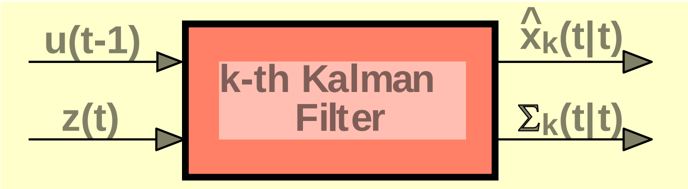
- Key equation (7) repeated: $p(x(t) \mid Z(t)) = \sum_{k=1}^{N} P_k(t) p(x(t) \mid H_k, Z(t))$
- But the conditional density $p(x(t) \mid H_k, Z(t))$ is precisely the conditional density of the k-th Kalman filter which assumes that $H = H_k$, i.e. that the true system is the k-th model
- Thus, we know that

(8) $$p(x(t) \mid H_k, Z(t)) \sim N(\hat{x}_k(t \mid t), \Sigma_k(t \mid t))$$

- Therefore, we see that the desired conditional density $p(x(t) \mid Z(t))$ is a probabilistically weighted sum of N gaussian densities, $p(x(t) \mid H_k, Z(t))$, each of which is generated by the bank of the N Kalman filters
- All that remains is to calculate the posterior probabilities

(9) $$P_k(t) = Prob(H = H_k \mid Z(t))$$

Proof: The Conditional Mean



- k-th Kalman filter assumes $H = H_k$
- Then it generates

$$(10) \qquad \hat{x}_k(t \mid t) = E\left\{x(t) \mid H_k, Z(t)\right\} = $$

$$= \int x(t) p\left(x(t) \mid H_k, Z(t)\right) dx(t)$$

$$(11) \qquad \Sigma_k(t \mid t) = cov\left[x(t); x(t) \mid H_k, Z(t)\right] =$$

$$= \int \left(x(t) - \hat{x}_k(t \mid t)\right)\left(x(t) - \hat{x}_k(t \mid t)\right)' p\left(x(t) \mid H_k, Z(t)\right) dx(t)$$

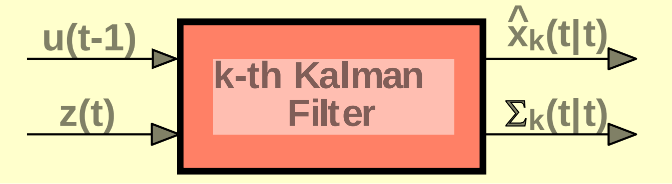
- Key equation (7) repeated: $p\left(x(t) \mid Z(t)\right) = \sum_{k=1}^{N} P_k(t) p\left(x(t) \mid H_k, Z(t)\right)$
- Global conditional mean: $\hat{x}(t \mid t)$

$$(13) \qquad \hat{x}(t \mid t) \equiv E\left(x(t) \mid Z(t)\right) =$$

$$= \int x(t) p\left(x(t) \mid Z(t)\right) dx(t) == \int \sum_{k=1}^{N} P_k(t) x(t) p\left(x(t) \mid H_k, Z(t)\right) dx(t) =$$

$$= \sum_{k=1}^{N} P_k(t) \underbrace{\int x(t) p\left(x(t) \mid H_k, Z(t)\right) dx(t)}_{\hat{x}_k(t \mid t)} = \sum_{k=1}^{N} P_k(t) \hat{x}_k(t \mid t)$$

Proof: The Conditional Covariance, I



- k-th Kalman filter assumes $H = H_k$
- Then it generates

$$(10) \quad \hat{x}_k(t \mid t) = E\{x(t) \mid H_k, Z(t)\} = \int x(t) p(x(t) \mid H_k, Z(t)) dx(t)$$

$$(11) \quad \Sigma_k(t \mid t) = cov[x(t); x(t) \mid H_k, Z(t)] =$$

$$= \int (x(t) - \hat{x}_k(t \mid t))(x(t) - \hat{x}_k(t \mid t))' p(x(t) \mid H_k, Z(t)) dx(t)$$

- Key equation (7) repeated: $p(x(t) \mid Z(t)) = \sum_{k=1}^{N} P_k(t) p(x(t) \mid H_k, Z(t))$
- Global conditional covariance: $\Sigma(t \mid t)$

$$(14) \quad \Sigma(t \mid t) = E\{x(t) - \hat{x}(t \mid t))(x(t) - \hat{x}(t \mid t))' \mid Z(t)\} =$$

$$= \int (x(t) - \hat{x}(t \mid t))(x(t) - \hat{x}(t \mid t))' p(x(t) \mid Z(t)) dx(t) =$$

$$= \sum_{k=1}^{N} P_k(t) \int (x(t) - \hat{x}(t \mid t))(x(t) - \hat{x}(t \mid t))' p(x(t) \mid H_k, Z(t)) dx(t)$$

Proof: The Conditional Covariance, II

- Eq. (14) repeated

$$(14)\qquad \Sigma(t\mid t)=\sum_{k=1}^{N}P_k(t)\int\big(x(t)-\hat{x}(t\mid t)\big)\big(x(t)-\hat{x}(t\mid t)\big)'\,p\big(x(t)\mid H_k,Z(t)\big)dx(t)$$

- Add and subtract $\hat{x}_k(t\mid t)$ in $\big(x(t)-\hat{x}(t\mid t)\big)$

$$\big(x(t)-\hat{x}_k(t\mid t)+\hat{x}_k(t\mid t)-\hat{x}(t\mid t)\big)\big(x(t)-\hat{x}_k(t\mid t)+\hat{x}_k(t\mid t)-\hat{x}(t\mid t)\big)'=$$

$$=\big(x(t)-\hat{x}_k(t\mid t)\big)\big(x(t)-\hat{x}_k(t\mid t)\big)'+\big(\hat{x}_k(t\mid t)-\hat{x}(t\mid t)\big)\big(\hat{x}_k(t\mid t)-\hat{x}(t\mid t)\big)'+$$

$$+\big(x(t)-\hat{x}_k(t\mid t)\big)\big(\hat{x}_k(t\mid t)-\hat{x}(t\mid t)\big)'+\big(\hat{x}_k(t\mid t)-\hat{x}(t\mid t)\big)\big(x(t)-\hat{x}_k(t\mid t)\big)'$$

$$\therefore\ (15)\qquad \int\big(x(t)-\hat{x}(t\mid t)\big)\big(x(t)-\hat{x}(t\mid t)\big)'p\big(x(t)\mid H_k,Z(t)\big)dx(t)=$$

$$=\underbrace{\int\big(x(t)-\hat{x}_k(t\mid t)\big)\big(x(t)-\hat{x}_k(t\mid t)\big)'p\big(x(t)\mid H_k,Z(t)\big)dx(t)}_{\Sigma_k(t\mid t)}+$$

$$+\big(\hat{x}_k(t\mid t)-\hat{x}(t\mid t)\big)\big(\hat{x}_k(t\mid t)-\hat{x}(t\mid t)\big)'\underbrace{\int p\big(x(t)\mid H_k,Z(t)\big)dx(t)}_{1}+$$

$$+\underbrace{\int\big(x(t)-\hat{x}_k(t\mid t)\big)p\big(x(t)\mid H_k,Z(t)\big)dx(t)}_{0}\cdot\big(\hat{x}_k(t\mid t)-\hat{x}(t\mid t)\big)'+$$

$$+\big(\hat{x}_k(t\mid t)-\hat{x}(t\mid t)\big)\cdot\underbrace{\int\big(x(t)-\hat{x}_k(t\mid t)\big)'p\big(x(t)\mid H_k,Z(t)\big)dx(t)}_{0'}$$

Proof: The Conditional Covariance, III

- From eqs. (14) and (15) we deduce that:

(16) $$\Sigma(t \mid t) = \sum_{k=1}^{N} P_k(t) \left[\Sigma_k(t \mid t) + \left(\hat{x}_k(t \mid t) - \hat{x}(t \mid t) \right) \left(\hat{x}_k(t \mid t) - \hat{x}(t \mid t) \right)' \right]$$

- The global covariance matrix $\Sigma(t \mid t)$ cannot precomputed off-line, even though the (local) KF covariances $\Sigma_k(t \mid t)$ are computed off-line
- The posterior probabilities $P_k(t) = Prob\left(H = H_k \mid Z(t)\right)$ must be also computed on-line
- The mean correction terms $\left(\hat{x}_k(t \mid t) - \hat{x}(t \mid t)\right)$ must also be computed on-line
- All that remains is to derive the recursive relation that generates the posterior probabilities, $P_k(t) = Prob\left(H = H_k \mid Z(t)\right)$

Proof: Model Probabilities

- Data at time t: $Z(t) \equiv \{u(0),\dots,u(t-1); z(1),\dots,z(t)\}$
- Data at time $t+1$: $Z(t+1) \equiv \{u(0),\dots,u(t-1), u(t); z(1),\dots,z(t), z(t+1)\}$
- Note: $Z(t+1) = \{u(t), z(t+1), Z(t)\}$; $u(t)$ deterministic
- Recall: $P_k(t) \equiv Prob[H = H_k \mid Z(t)]$; $P_k(t+1) \equiv Prob[H = H_k \mid Z(t+1)]$
- Recall that since H is a discrete random variable, the associated probability density function is a weighted sum of impulses, i.e.

$$p(H \mid Z(t)) = \sum_{k=1}^{N} P_k(t)\delta(H - H_k) \tag{17}$$

$$p(H \mid Z(t+1)) = \sum_{k=1}^{N} P_k(t+1)\delta(H - H_k) \tag{18}$$

- PROBLEM: Relate $P_k(t+1)$ to $P_k(t)$

Proof: Probability Relations

- Use Bayes rule to obtain:

(19) $$p(H \mid Z(t+1)) = p(H \mid u(t), z(t+1), Z(t)) = \frac{p(H, z(t+1) \mid u(t), Z(t))}{p(z(t+1) \mid u(t), Z(t))} =$$

$$= \frac{p(z(t+1) \mid u(t), H, Z(t)) \cdot p(H \mid u(t), Z(t))}{p(z(t+1) \mid u(t), Z(t))} =$$

$$= \frac{p(z(t+1) \mid u(t), H, Z(t)) \cdot p(H \mid Z(t))}{p(z(t+1) \mid u(t), Z(t))}$$

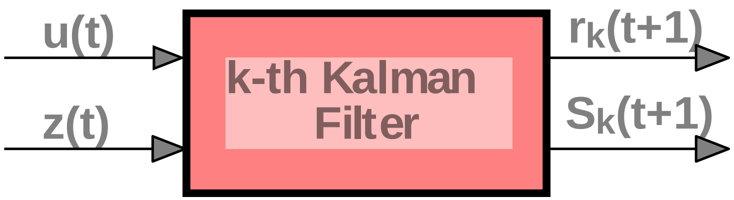
- Substitute eqs. (17) and (18) into eq. (19) to obtain

(20) $$\sum_{k=1}^{N} P_k(t+1)\delta(H - H_k) = \sum_{k=1}^{N} \frac{p(z(t+1) \mid u(t), H, Z(t))}{p(z(t+1) \mid u(t), Z(t))} \cdot P_k(t)\delta(H - H_k)$$

- Equate coefficients of delta functions (impulses) to obtain

(21) $$P_k(t+1) = \frac{p(z(t+1) \mid u(t), H_k, Z(t))}{p(z(t+1) \mid u(t), Z(t))} \cdot P_k(t)$$

Proof: Probability Calculations, I



k - TH KF RESIDUAL

$$r_k(t+1) \equiv z(t+1) - C_k(t+1)\hat{x}(t+1 \mid t)$$

$$E\{r_k(t+1) \mid u(t), H_k, Z(t)\} = 0$$

$$S_k(t+1) \equiv cov[r_k(t+1); r_k(t+1) \mid u(t), H_k, Z(t)] =$$

$$= C_k(t+1)\Sigma_k(t+1 \mid t)C_k'(t+1) + \Theta_k(t+1)$$

- In eq. (21) we need to evaluate $p(z(t+1) \mid u(t), H_k, Z(t))$
- But, for the k - th model,

(22) $z(t+1) = C_k(t+1)x(t+1) + \theta(t+1) = r_k(t+1) + C_k(t+1)\hat{x}_k(t+1 \mid t)$

(23) $E\{z(t+1) \mid u(t), H_k, Z(t)\} = C_k(t+1)\hat{x}_k(t+1 \mid t)$

(24) $cov[z(t+1); z(t+1) \mid u(t), H_k, Z(t)] =$

$$= C_k(t+1)\Sigma_k(t+1 \mid t)C_k'(t+1) + \Theta_k(t+1) = S_k(t+1)$$

$p(z(t+1) \mid u(t), H_k, Z(t))$ is gaussian with mean (23) and covariance (24)

(25) $p(z(t+1) \mid u(t), H_k, Z(t)) = \dfrac{1}{(2\pi)^{m/2}\sqrt{\det S_k(t+1)}} e^{-\frac{1}{2}\cdot r_k'(t+1)S_k^{-1}(t+1)r_k(t+1)}$

Proof: Probability Calculations, II

- In eq. (21) we also need to evaluate $p\big(z(t+1) \mid u(t), Z(t)\big)$
- Using the marginal density and Bayes rule we deduce that

(26) $$p\big(z(t+1) \mid u(t), Z(t)\big) = \int p\big(z(t+1), H \mid u(t), Z(t)\big) dH =$$

$$= \int p\big(z(t+1) \mid H, u(t), Z(t)\big) p\big(H \mid u(t), Z(t)\big) dH =$$

$$= \int p\big(z(t+1) \mid H, u(t), Z(t)\big) p\big(H \mid Z(t)\big) dH =$$

$$= \int p\big(z(t+1) \mid H, u(t), Z(t)\big) \sum_{j=1}^{N} P_j(t) \delta(H - H_j) dH =$$

$$= \sum_{j=1}^{N} P_j(t) p\big(z(t+1) \mid H_j, u(t), Z(t)\big)$$

- We have already calculated $p\big(z(t+1) \mid H_j, u(t), Z(t)\big)$ in eq. (25)!

Proof: Probability Calculations, III

- From eqs. (21) and (26) we obtain the general recursion

(27) $$P_k(t+1) = \frac{p\big(z(t+1) \mid H_k, u(t), Z(t)\big)}{\sum_{j=1}^{N} P_j(t) \cdot p\big(z(t+1) \mid H_j, u(t), Z(t)\big)} \cdot P_k(t)$$

- For notational simplicity define

(28) $$\beta_i(t+1) \equiv \frac{1}{(2\pi)^{m/2}\sqrt{\det S_i(t+1)}}$$

(29) $$w_i(t+1) \equiv r_i'(t+1) S_i^{-1}(t+1) r_i(t+1)$$

- Then, from eqs. (27), (28), (29) and (25), we deduce that

(30) $$P_k(t+1) = \frac{\beta_k(t+1) e^{-(1/2) w_k(t+1)}}{\sum_{j=1}^{N} \beta_j(t+1) e^{-(1/2) w_j(t+1)} \cdot P_j(t)} \cdot P_k(t); \quad P_k(0) = \text{prior model probs.}$$

QED

Observations

- Clearly, as in the single-model case, the past control sequence *{u(0), u(1), …, u(t-1)}* influences the conditional state-estimate and residuals at time t
- Unlike the single-model case, in the MMAE algorithm, the past control sequence *{u(0), u(1), …, u(t-1)}* also influences the conditional covariance matrix, $\Sigma(t|t)$, and hence the accuracy of the state-estimate
- This implies that some control sequences are "better" for improving the accuracy of the state-estimates and model identification probabilities

Model Identification

DYNAMIC PROBABILITY EVALUATION

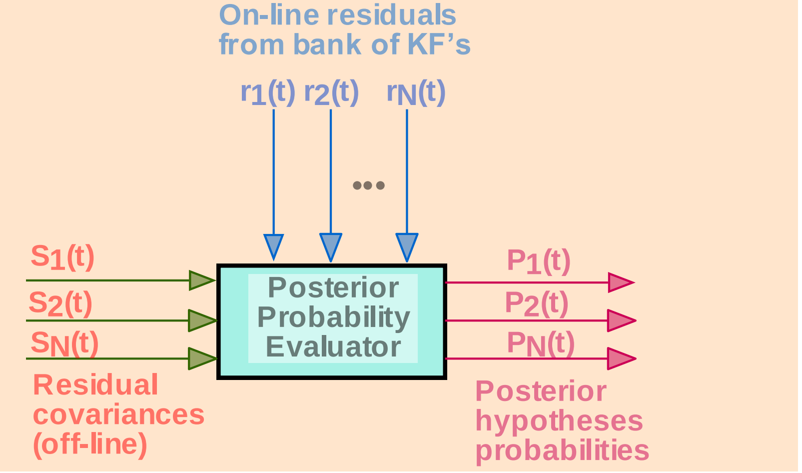
- For $t = 0, 1, 2, \ldots$ and $k = 1, 2, \ldots, N$

$$P_k(t+1) = \left(\frac{\beta_k(t+1) e^{-\frac{1}{2} w_k(t+1)}}{\sum_{j=1}^{N} \beta_j(t+1) e^{-\frac{1}{2} w_j(t+1)} P_j(t)} \right) \bullet P_k(t)$$

with the initial probabilities $P_k(0) =$ given, and

$$\beta_k(t+1) \equiv \frac{1}{(2\pi)^{m/2} \sqrt{\det S_k(t+1)}}$$

$$w_k(t+1) \equiv r_k'(t+1) S^{-1}{}_k(t+1) r_k(t+1)$$



- It can be shown that, if $H = H_i$ is true, i.e. the i-th model is the true one, then

$$\lim_{t\to\infty} P_i(t) = 1, \quad \lim_{t\to\infty} P_j(t) = 0 \quad \forall j \neq i$$

which means that the true model is identified with probability ONE.

- If none of the models is the true one, then the model "nearest" to the true one (in a probabilistic sense) will be identified. See [5], pp. 270-277 for mathematical details on convergence

Concluding Remarks

- The basic concept of the MMAE is extremely powerful, since it combines system identification and state estimation
- Results are also available for the continuous-time case, but care must be exercised for their proper implementation
- In nonlinear problems, one can obtain suboptimal solutions by using extended Kalman filters (EKFs), rather than linear KFs
- The structure of the MMAE algorithm is very similar to that used in the so-called “Sum of Gaussian Methods” for nonlinear estimation
- Advances in computational speed and digital parallel computer hardware and software make MMAE-type of implementations more and more practical for very complex identification and state estimation problems

Extensions

DYNAMIC PROBABILITY EVALUATION

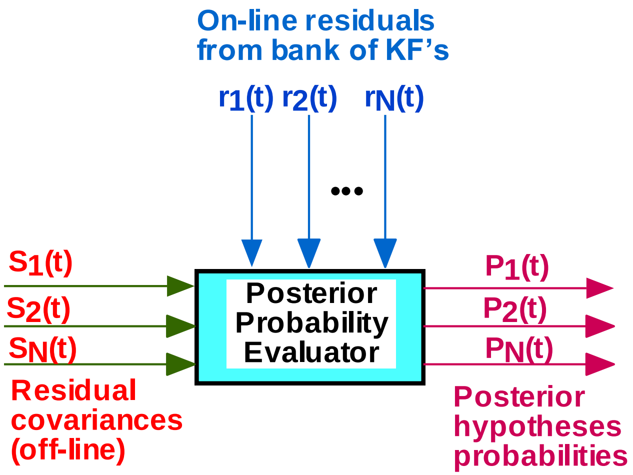
- For $t = 0, 1, 2, \ldots$ and $k = 1, 2, \ldots, N$

$$P_k(t+1) = \left(\frac{\beta_k(t+1) e^{-\frac{1}{2} w_k(t+1)}}{\sum_{j=1}^{N} \beta_j(t+1) e^{-\frac{1}{2} w_j(t+1)} P_j(t)} \right) \bullet P_k(t)$$

with the initial probabilities $P_k(0) =$ given, and

$$\beta_k(t+1) \equiv \frac{1}{(2\pi)^{m/2} \sqrt{\det S_k(t+1)}}$$

$$w_k(t+1) \equiv r_k'(t+1) S^{-1}{}_k(t+1) r_k(t+1)$$

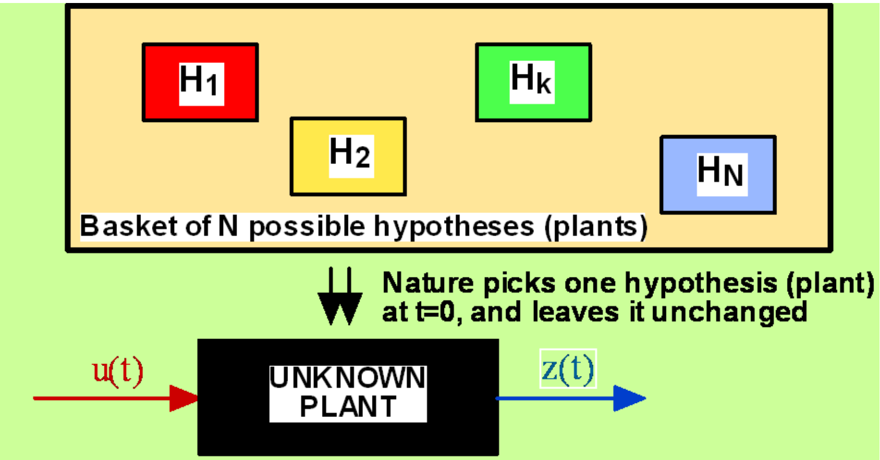


- ADVICE: During the on-line calculation of the posterior probabilities $P_k(t)$, put a lower bound ε, say $\varepsilon = 10^{-2}$, on each probability, i.e. $P_k(t) \geq \varepsilon$ for all k, t
- This will enable the MMAE to respond if the true model, say H_j, changes to another, say H_i

Time-Dependent Hypotheses

BASIC MMAE ASSUMPTION

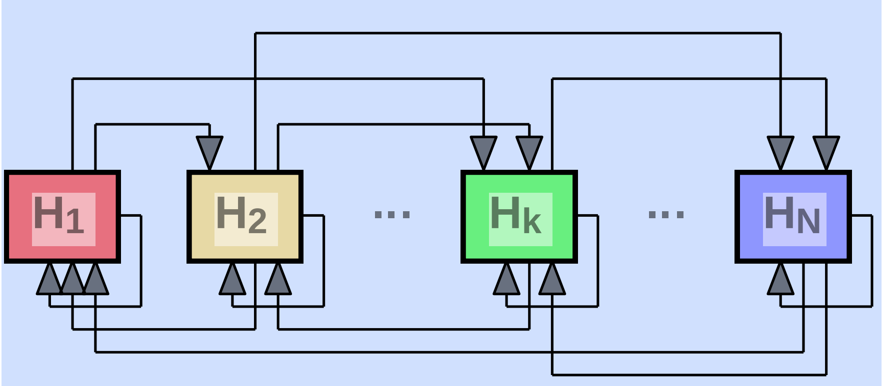
- The true hypothesis, $H = H_k$, does not change during the experiment
- Then we guarantee that $P_k(t) \to 1$



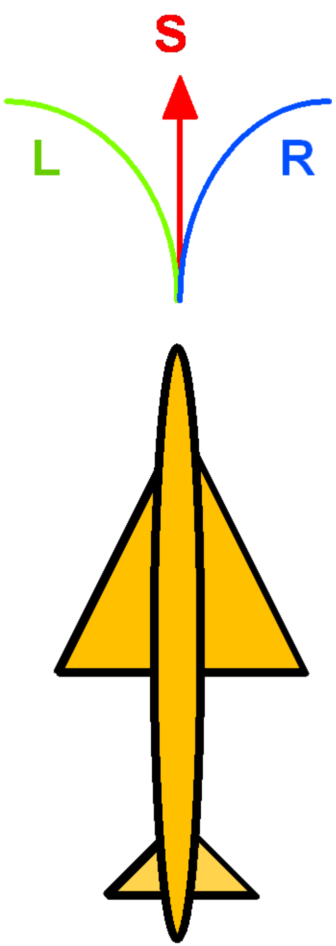
DYNAMIC HYPOTHESES

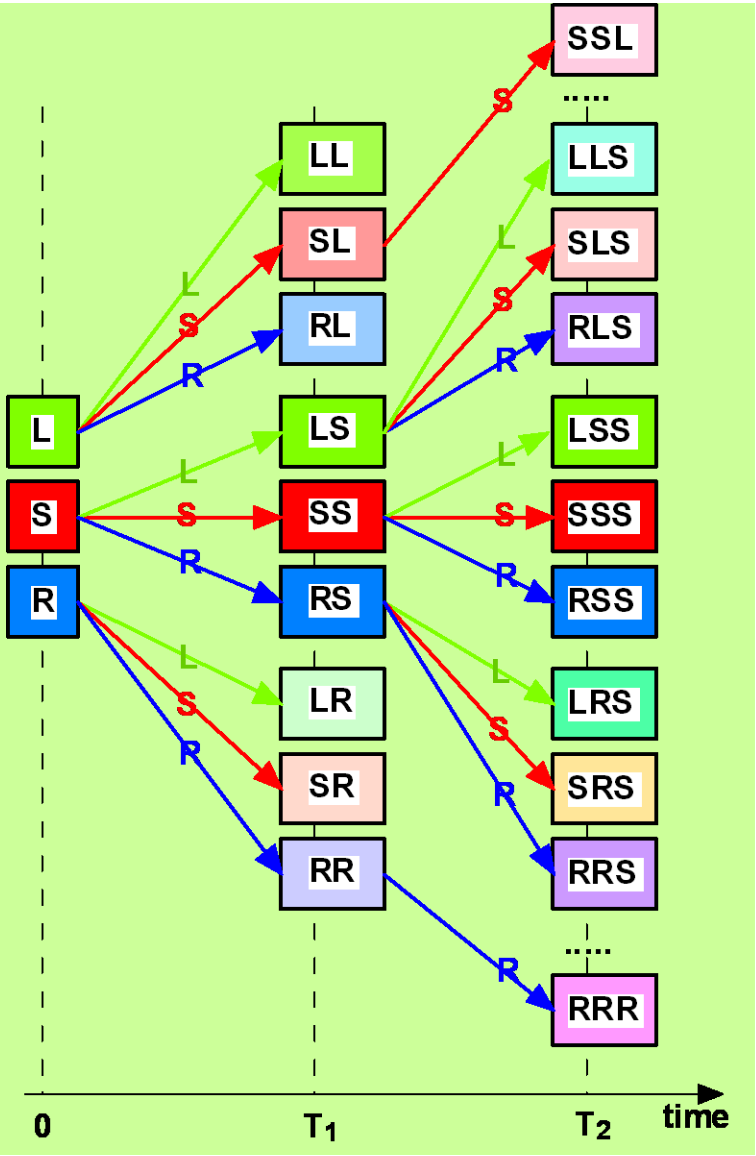
- Suppose that the true hypothesis changes, one or more times, during the experiment
- For example,

$$H_{true}(t) = \begin{cases} H_1 & \text{for } 0 \le t < t_1 \\ H_2 & \text{for } t_1 \le t < t_2 \\ & \cdots \\ H_k & \text{for } t_{k-1} \le t < t_k \end{cases}$$

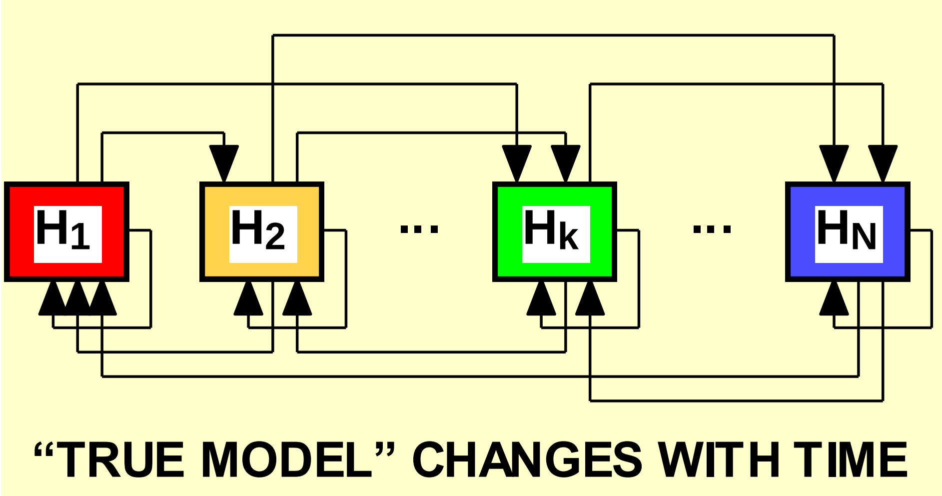


Dynamic Hypotheses Example





Estimation With Dynamic Hypotheses



- The truly optimal extension of the MMAE algorithm to handle the dynamic hypotheses problem requires to take into account the exploding hypotheses-tree, and at each time construct each Kalman filter so that takes into account the postulated past history of the hypotheses
- The optimal algorithm requires a growing number of parallel Kalman filters
- The combinatorial explosion prohibits the implementation of the truly optimal algorithm

Suboptimal MMAE Estimation

- Assumption: the time-interval between changes in the true hypothesis is “larger” than the convergence time of the standard MMAE algorithm
  - unfortunately, estimates of the convergence rate are not known
- As long as we enforce a lower-bound on each posterior probability, we can expect the standard MMAE algorithm to recover when the hypothesis changes
  - it will only “work well” only if the changes are infrequent
- Allowing even a small number of dynamic hypotheses trees help the MMAE to respond more rapidly
- These are precisely the issues in complex surveillance problems where several sensors track several maneuvering targets, including “newborn” targets and “dying” targets

References

[1]. M. Athans and C.B. Chang, “Adaptive Estimation and Parameter Identification using Multiple Model Estimation Algorithm,” Technical Note 1976-28, MIT Lincoln Lab., Lexington, Mass., 23 June 1976 (copy included in handouts)

[2]. D.T. Magill, “Optimal Adaptive Estimation of Sampled Stochastic Processes,” *IEEE Trans. Auto. Control*, Vol. AC-10, 1965, pp. 434-439

[3]. D.G. Lainiotis, “Optimal Adaptive Estimation: Structure and Parameter Adaptation,” *IEEE Trans. Auto. Control*, Vol. AC-15, 1971, pp. 160-170

[4]. B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Prentice-Hall, 1979, Chapter 10