from Schweppe, F.C., Uncertain Dynamic Systems Ch. 9 & 10.

Hypothesis Testing: Static Systems I

Review:

Estimation: Given observations on specified process, what is the state of the process?

Hypothesis Testing: Given observations on one of M specified processes, which process is observed?

single observation vector $\underline{z}$.

M hypotheses : $(1, 2, \ldots, m)$

$H_j : z = z_j$

z_j : uncertain vector produced by process j.

Example:

Given a disk which may be wood (5' long) or metal (6' long), which is it?

H_1 : disk is wood

H_2 : disk is metal.

or

$H_1 : z = 5$ stochastic: $H_1 : z = 5 + v$

$H_2 : z = 6$ $H_2 : z = 6 + v$

$v \sim N(0, 1)$

Reasonable decision rule:

$$\text{if } z_{obs} < 5.5 \Rightarrow H_1$$

$$\text{if } z_{obs} > 5.5 \Rightarrow H_2.$$

Two steps in hypothesis testing:

1. Decision function: map of observations into a vector space.

2. Decision rule: Rule which takes output of decision function & decides which hypothesis is correct.

$$\text{observation} \Rightarrow \boxed{\begin{array}{c}\text{Decision}\\ \text{Function}\end{array}} \xrightarrow{\xi} \boxed{\begin{array}{c}\text{Decision}\\ \text{Rule}\end{array}} \longrightarrow \text{Final Decision}$$

Ex: Pattern Recognition ($\subset$ Hypothesis testing)

given a lot of information (data)

extract certain features (decision function) called the pattern

apply decision rule to the pattern.

P.R. is largely heuristic methods for choosing decision function (feature extraction) and decision rule (learning mode of pattern classification)

Applications:

Communications:

Radar: Decide, based on data, whether or not signal is present:

H_1 : observe signal + noise

H_2 : observe noise only.

Channel: Decide which of M signals was transmitted over channel

H_j : observe $s_j(t)$ + noise
$j = 1, 2, \ldots, m$

Tracking problem:

Decide whether an observed vehicle is ballistic or maneuvering. I.e., if $\delta(t)$ = control input to vehicle,

H_1 : $\delta(t) = 0$ (ballistic)

H_2 : $\delta(t)$ = large (maneuver)

Reliability:

H_j : j^{th} component has failed

Parameter Estimation

$$z = h(\alpha) + v$$

let $\alpha \in \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ - model parameter

$$H_j : z = z_j$$
$$z_j = h(\alpha_j) + v$$

Model Verification:

If you have developed a model for a physical process, you must test to see if model accurately describes process

H_1 : observations came from model

H_2 : observations did not come from model.

Bayesian Hypothesis testing

Bayesian model $\triangleq$ one which is completely defined ~~statistically~~.

Let H_j : z a random vector with prob. den. $p_j(z)$, $j = 1, 2, \ldots, m$

Define ρ_j : a priori prob that H_j is true.

$$\sum_{j=1}^{m} \rho_j = 1$$

likelihood function $= p_j p_j(z)\Big|_{z=z_{actual}}$

$$\boxed{\begin{array}{l}\text{Max-likelihood hyp. test:} \\ \quad \text{Choose } H_K \text{ if } j=K \text{ maximizes } p_j p_j(z)\Big|_{z=z_{actual}} \\ \qquad\qquad , j=1,2,\ldots,m.\end{array}}$$

"reasonable", but can also, if desired, be shown "optimal" with proper choice of criterion.

if, for $m=2$,

$\alpha_{1,2}$ = prob. H_1 chosen | H_2 true

$\alpha_{2,1}$ = prob H_2 chosen | H_1 true

(also called type I & II error probabilities, & failure to detect & false alarm prob.)

the prob. of error is defined as

$$P_e = p_2 \alpha_{1,2} + p_1 \alpha_{2,1}$$

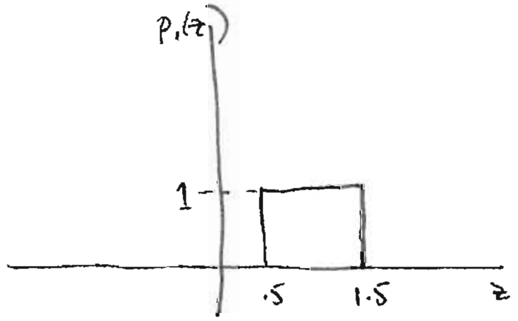
then max-likelihood hyp. test. minimizes P_e.

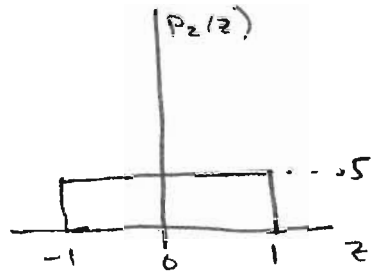
EX:

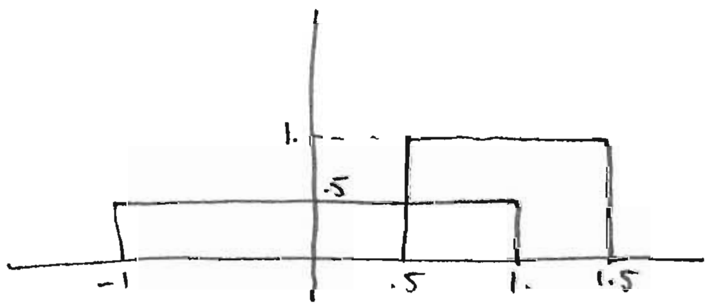
$$H_1 : z = z_1 \quad ; \quad p_1(z)$$

$$H_2 : z = z_2 \quad ; \quad p_2(z)$$

$$P_1 = P_2 = \frac{1}{2}$$







decision rule:
choose H_1 if
$z \in (.5, 1.5]$
choose H_2 if
$z \in [-1, .5]$

here

$$\alpha_{1,2} = \int_{z \in \Omega_1} p_2(z)\,dz = \frac{1}{4}$$

$$P_e = \frac{1}{8}$$

$$\alpha_{2,1} = \int_{z \in \Omega_2} p_1(z)\,dz = 0$$

where Ω_1 = region of $z \ni H_1$ is chosen

Ω_2 = region of $z \ni H_2$ is chosen.

Normally, use log-likelihood function

$$\xi_j(z) : \text{ log likelihood function}$$

$$\xi_j(z) = \ln\{p_j(z)\} \qquad j = 1, 2, \ldots, m$$

<u>Log-likelihood decision rule</u>:

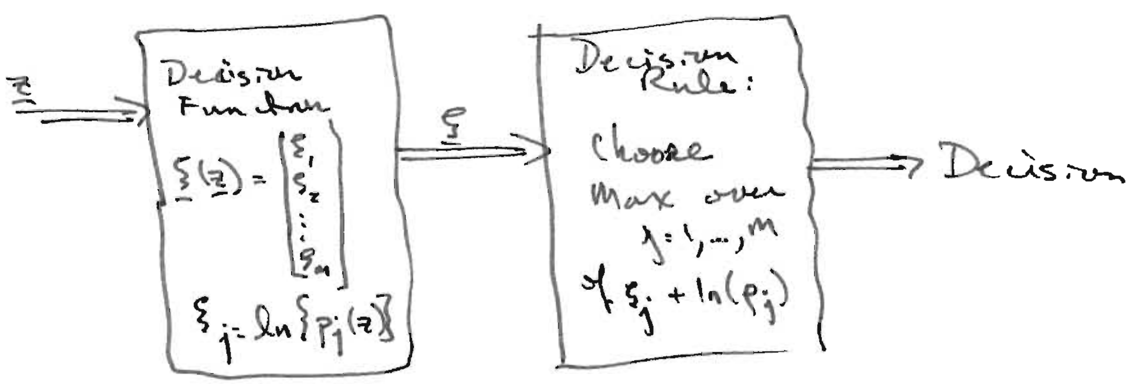
Choose H_K if $j = k$ yields $\max\left\{ \xi_j(z_{actual}) + \ln\{P_j\} \right\}$

$$j = 1, 2, \ldots, m.$$

if

$$\underline{\xi}(\underline{z}) = \begin{bmatrix} \xi_1(\underline{z}) \\ \xi_2(\underline{z}) \\ \vdots \\ \xi_m(\underline{z}) \end{bmatrix}$$

then



Note: You may wish to avoid certain types of errors; this can be done by selecting weights other than ρ_j.

Ex: If you are a man taking a woman to a dance,

H_1 : She does not expect you to bring Flowers

H_2 : She expects you to bring flowers.

choose $H_2 \mid H_1$ true   no problem

choose $H_1 \mid H_2$ true   no dance!

The Gaussian Hypothesis Testing Case

$$H_j : \underline{z} = \underline{z}_j \qquad \underline{z}_j \sim N(\underline{m}_j, \underline{\Gamma}_j) \qquad j = 1, \dots, m$$

$$z \in R^k \qquad \underline{\Gamma}_j^{-1} \text{ exists}$$

Gaussian:

$$p_j(\underline{z}) = \text{prob. den. of } \underline{z} \mid H_j$$

$$= \frac{1}{[(2\pi)^k |\underline{\Gamma}_j|]^{1/2}} \exp\left\{-\frac{1}{2}(\underline{z}-\underline{m}_j)^T \underline{\Gamma}_j^{-1} (\underline{z}-\underline{m}_j)\right\}$$

∴ log-likelihood function

$$\boxed{\begin{aligned} 2\xi_j(\underline{z}) &= 2\ln(p_j(\underline{z})) \\ &= -K\ln(2\pi) - \ln|\underline{\Gamma}_j| - [\underline{z}-\underline{m}_j]^T \underline{\Gamma}_j^{-1} [\underline{z}-\underline{m}_j] \end{aligned}}$$

if $\quad H_j : \underline{z} = \underline{H}_j \underline{x}_j + \underline{v}_j \qquad j = 1, \dots, m$

$$\left.\begin{aligned} \underline{x}_j &\sim N(\underline{0}, \underline{\Psi}_j) \\ \underline{v}_j &\sim N(\underline{0}, \underline{R}_j) \end{aligned}\right\} \underline{x}, \underline{v} \text{ independent}$$

then $\quad H_j : \underline{z} \sim N(\underline{0}, \underline{\Gamma}_j) \quad j = 1, \dots, m$

$$\underline{\Gamma}_j = \underline{H}_j \underline{\Psi}_j \underline{H}_j^T + \underline{R}_j \qquad \text{is equivalent}$$

and

$$\boxed{2\xi_j(\underline{z}) = -K\ln(2\pi) - \ln|\underline{H}_j \underline{\Psi}_j \underline{H}_j^T + \underline{R}_j| - \underline{z}^T [\underline{H}_j \underline{\Psi}_j \underline{H}_j^T + \underline{R}_j]^{-1} \underline{z}}$$

Note that for Gaussian case:

$$2\xi_j(\underline{z}) = -k\ln(2\pi) - \ln|\underline{\Gamma}_j| - [\underline{z}-\underline{m}_j]^T \underline{\Gamma}_j^{-1}[\underline{z}-\underline{m}_j]$$

~~an equivalent likelihood function is~~

$$= -k\ln(2\pi) - \underline{z}^T\underline{\Gamma}_j^{-1}\underline{z} + 2\underline{z}^T\underline{\Gamma}_j^{-1}\underline{m}_j - \underline{m}_j^T\underline{\Gamma}_j^{-1}\underline{m}_j - \ln|\underline{\Gamma}_j|$$

an equivalent likelihood function is

$$2\tilde{\xi}_j(\underline{z}) = -\underline{m}_j^T\underline{\Gamma}_j^{-1}\underline{m}_j - \ln|\underline{\Gamma}_j| + 2\underline{z}^T\underline{\Gamma}_j^{-1}\underline{m}_j$$

and, if the decision rule is modified appropriately,

$\boxed{\underset{\sim}{\xi}_j(\underline{z}) = \underline{z}^T\underline{\Gamma}_j^{-1}\underline{m}_j}$ is also equivalent, since

$$-\underline{m}_j^T\underline{\Gamma}_j^{-1}\underline{m}_j - \ln|\underline{\Gamma}_j|$$

can be incorporated in the decision rule.

Hypothesis Testing: Linear Dynamical Systems

Discrete-Time Case

Assume observe $z(n)$, $n = 1, \ldots, N$

$$H_j : \; z(n) = z_j(n) \quad ; \; n = 1, \ldots, N$$

z_j is discrete-time Gaussian process, $j = 1, 2, \ldots, M$

H_j hypothesis that j^{th} process produced all observations.

can use static results.

$\xi_j(N)$ - log-likelihood ~~ratio~~ function for $\underline{Z}_N$

$$\underline{Z}_N = \begin{bmatrix} \underline{z}(1) \\ \underline{z}(2) \\ \vdots \\ \underline{z}(N) \end{bmatrix}$$

however, if we can obtain $\xi_j(N)$ recursively, as with the KBF, less computation.

Main Result:

$\xi_j(N)$ can be computed recursively, essentially by squaring and integrating the innovations process of the KBF.

If

$$\underline{z}_N = \begin{bmatrix} z(1) \\ \vdots \\ z(N) \end{bmatrix}$$

$$H_j : \underline{z}_N = \underline{z}_{N,j} \quad \text{with } p_j(z_N) \qquad j = 1, \dots, M$$

$$\xi_j(N) = \ln\{p_j(z_N)\}$$

now, using Bayes' rule

$$p_j(z_N) = p_j(z_{N-1}, z(N)) = p_j(z_{N-1})\, p_j(z(N) | z_{N-1})$$

and

$$\xi_j(N) = \ln\{p_j(z_{N-1})\, p_j(z(N) | z_{N-1})\}$$

$$\boxed{\xi_j(N) = \xi_j(N-1) + \ln\{p_j(z(N) | z_{N-1})\} \; ; \quad \xi_j(0) = 0}$$

If, as notation, we use

$$\hat{z}_j(N|N-1) = E\{z(N) | z_{N-1}, H_j\}$$

$$\delta_j(N) = z(N) - \hat{z}_j(N|N-1)$$

$$\Sigma_{z,j}(N|N-1) = E\{\delta_j(N)\delta_j^T(N)\}$$

and, since z is gaussian

$$p_j(z(N) | z_{N-1}) = \left[(2\pi)^k \, |\Sigma_{z,j}(N|N-1)|\right]^{-1/2}$$

$$\cdot \exp\{-\tfrac{1}{2}[\delta_j^T(N)\, \Sigma_{z,j}^{-1}(N|N-1)\, \delta_j(N)]\} \qquad \text{where } z(N) \in \mathbb{R}^k$$

Then

$$2\varsigma_j(N) = 2\varsigma_j(N-1) - \ln\{|\Sigma_{z_j}(N|N-1)|\}$$
$$- K\ln(2\pi) - \delta_j^T(N)\,\Sigma_{z,j}^{-1}(N|N-1)\,\delta_j(N)$$

$$\varsigma_j(0) = 0$$

If $H_j : z(n) = y_j(n) + v_j(n) \quad n = 1,\dots,N$
$\qquad j = 1,\dots,M$

y_j : Gaussian process
$v_j(n) \sim N(0, R_j(n))$

then

$$2\varsigma_j(N) = \varsigma_{j,bias}(N) + \varsigma_{j,observation}(N)$$

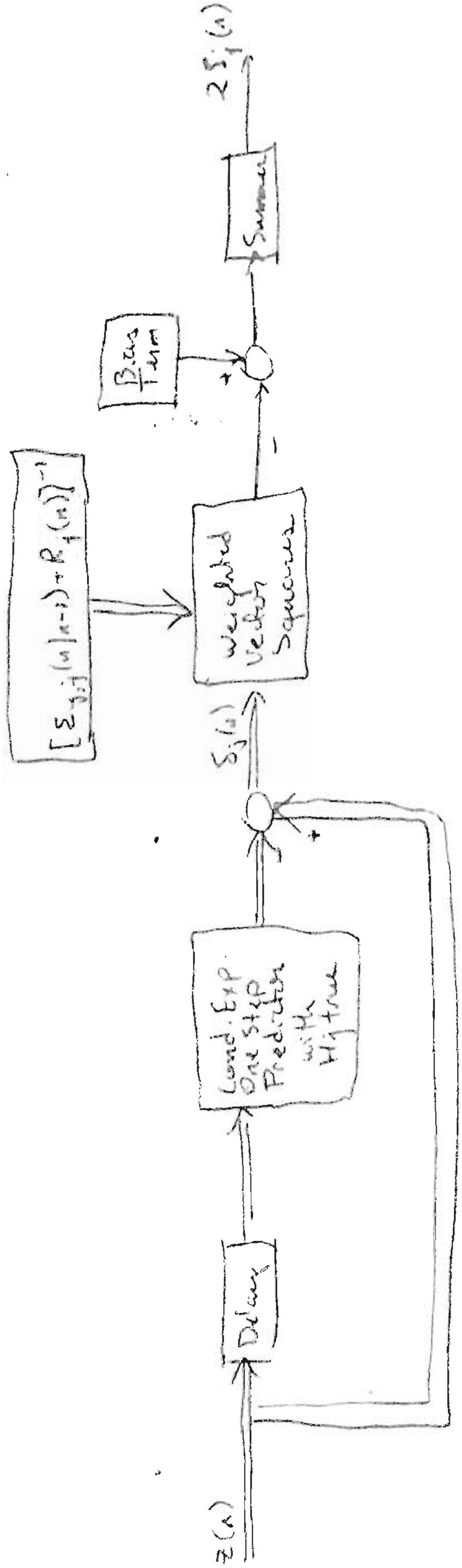
$$\varsigma_{j,bias}(N) = -\sum_{n=1}^{N}\ln\{|\Sigma_{y,j}(n|n-1) + R_j(n)|\} - NK\ln\{2\pi\}$$

$$\varsigma_{j,obs}(N) = -\sum_{n=1}^{N}\delta_j^T(n)\left[\Sigma_{y,j}(n|n-1) + R_j(n)\right]^{-1}\delta_j(n)$$

$$\hat{y}_j(n|n-1) = E\{y_j(n)|Z_{N-1}\} \quad H_j \text{ true}$$

$$\Sigma_{y,j}(N|N-1) = E\{(y_j(N) - \hat{y}_j(N|N-1))(y_j(N) - \hat{y}_j(N|N-1))^T\}$$

$$\delta_j(N) = z(N) - \hat{y}_j(N|N-1)$$



$z(n)$
Delay
Cond. Exp. One Step Predictor with Hj true
$S_j(n)$
$[\Sigma_{yj}(n|n-1) + R_1(n)]^{-1}$
Weighted Vector Squarer
Bias Term
Summer
$2S_j(n)$

If

$$y_j(n) = H_j(n)x_j(n) + v_j(n)$$

$$x_j(n+1) = \Phi_j(n)x_j(n) + G_j(n)w_j(n)$$

$$w_j(n) \sim N(0, Q_j(n))$$

$$v_j(n) \sim N(0, R_j(n))$$

$$x(0) \sim N(0, \Psi)$$

then

$$\rho_{j,obs}(N) = -\sum_{n=1}^{N} \delta_j^T(n)\left[H_j(n)\Sigma_j(n|n-1)H_j^T(n) + R_j(n)\right]^{-1}\delta_j(n)$$

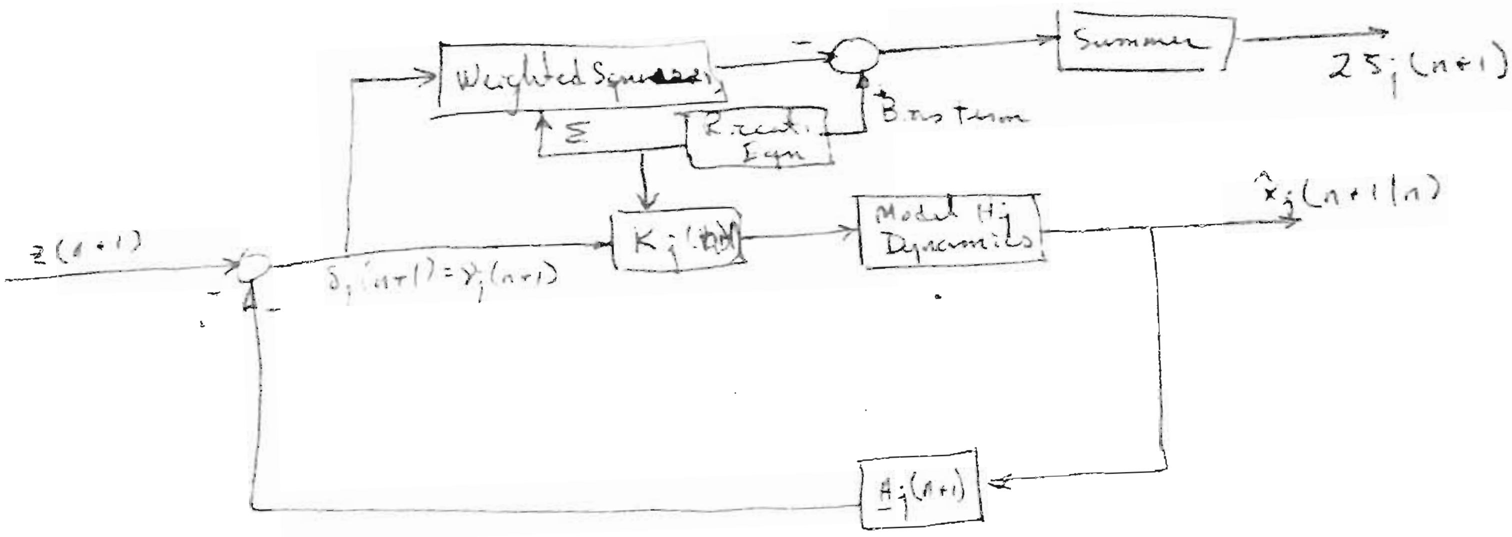
$$= \rho_{j,obs}(N-1) - \delta_j^T(N)\left[H_j(N)\Sigma_j(N|N-1)H_j^T(N) + R_j(N)\right]^{-1}\delta_j(N)$$

$$\delta_j(n+1) = z(n+1) - H_j(n+1)\hat{x}_j(n+1|n) \quad \text{(innovations)}$$

$$\rho_{j,bias}(N) = -NK\ln(2\pi) - \sum_{n=1}^{N}\ln\left|H_j(n)\Sigma_j(n|n-1)H_j^T(n) + R_j(n)\right|$$

$$= \rho_{j,bias}(N-1) - K\ln(2\pi) - \ln\left|H_j(N)\Sigma_j(N|N-1)H_j^T(N) + R_j(N)\right|$$

where $\hat{x}_j(n+1|n)$ & $\Sigma_j(N|N-1)$ are from the j^{th} KBF.



Decision Rule: Choose model with largest $\varsigma_j(n+1)$