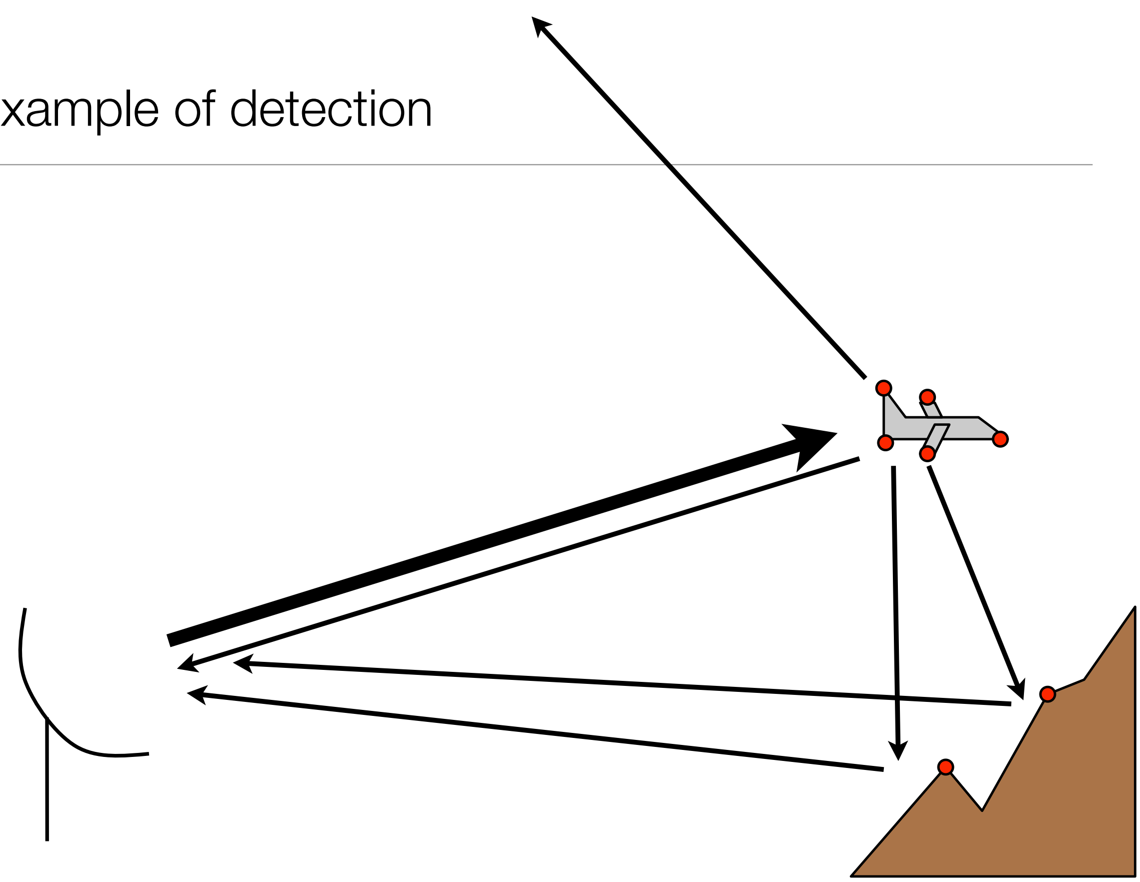
ECE 531: Detection and Estimation Theory

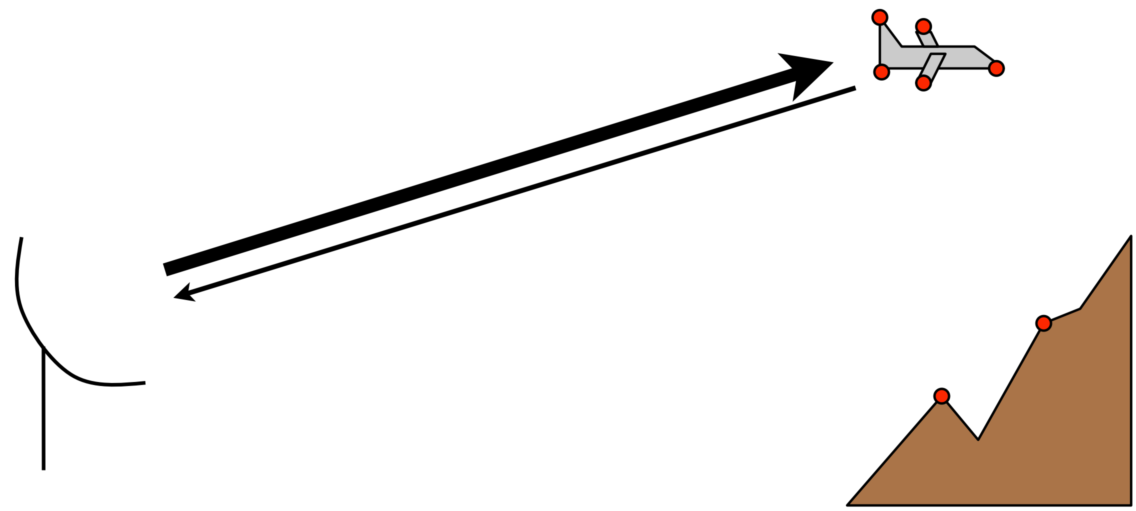
Natasha Devroye
devroye@ece.uic.edu
http://www.ece.uic.edu/~devroye

Spring 2011

Example of detection



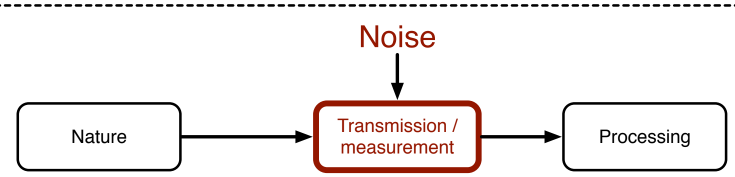
Example of estimation



Goals

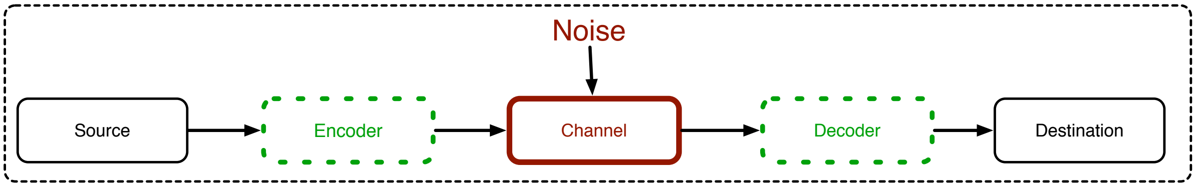
- infer value of unknown state of nature based on **noisy** observations

Mathematically, optimally



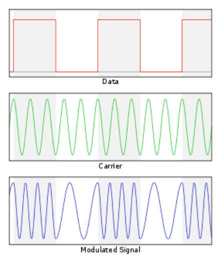
phenomenon experiment the ``truth''	model of observation or transmission process	decision rule estimation function
model of hypothesis H	$P_{Y\|H}(y\|h)$	*mapping from observation space* $\mathcal{Y}$ *to decisions/ estimates* $\hat{H}$

Detection example 1: digital communications



10001010100010

$$0 \leftrightarrow s_0(t) = \sin(\omega_0 t)$$
$$1 \leftrightarrow s_1(t) = \sin(\omega_1 t)$$



$$r(t) = \begin{cases} s_0(t) + n(t) & \text{if '0' sent} \\ s_1(t) + n(t) & \text{if '1' sent} \end{cases}$$

Detect?

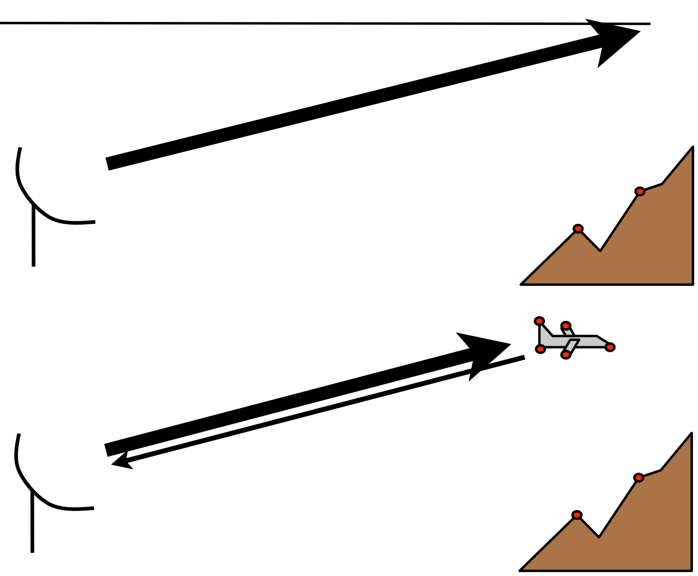
Detection example 2: Radar communication

Send $s(t) = \sin(\omega_c t), 0 \le t \le T$

Receive

Hypothesis $\mathcal{H}_0$

$$r(t) = n(t),\ 0 \le t \le T$$



Detect?

Hypothesis $\mathcal{H}_1$

$$r(t) = V_r \sin((\omega_c + \omega_d)(t - \tau) + \theta_r) + n(t),\ \tau \le t \le t + \tau$$

Further examples

- Sonar: enemy submarine
- Image processing: detect an aircraft from infrared images
- Biomedicine: cardiac arryhthmia from heartbeat sound wave
- Control: detect occurrence of abrupt change in system to be controlled
- Seismology: detect presence of oil deposit

Difference between detection and estimation?

- Detection:

Discrete set of hypotheses

Right or wrong

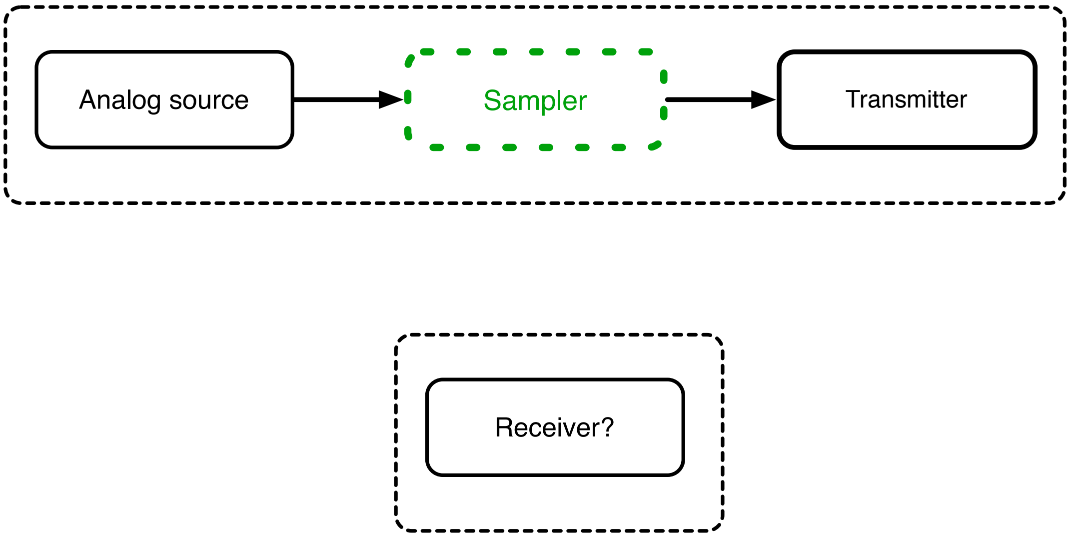
- Estimation:

Continuos set of hypotheses

Almost always wrong - minimize error instead

Estimation example 1: communications

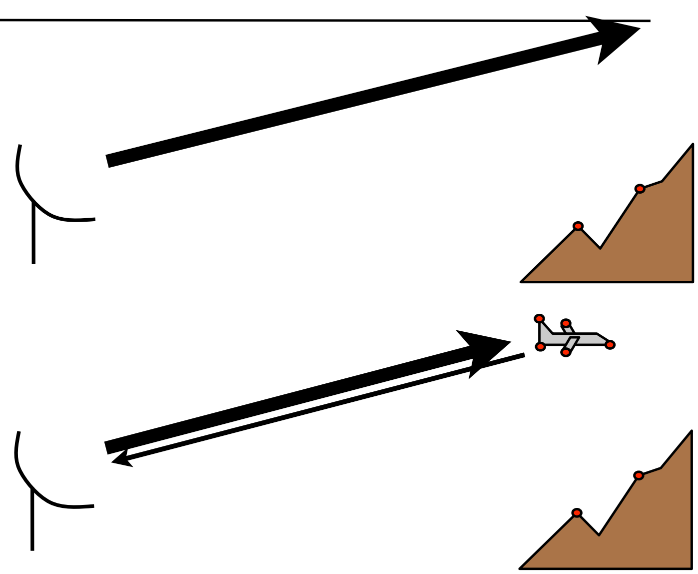
- Pulse amplitude modulation (PAM)



Estimation example 2: Radar

Send $s(t) = \sin(\omega_c t), 0 \leq t \leq T$

Receive



Hypothesis $\mathcal{H}_0$

$r(t) = n(t), \; 0 \leq t \leq T$

Estimate?

Hypothesis $\mathcal{H}_1$

$r(t) = V_r \sin((\omega_c + \omega_d)(t - \tau) + \theta_r) + n(t), \; \tau \leq t \leq t + \tau$

Our methods

- Will treat everything generally, with a unified mathematical representation
- Bias towards Gaussian noise and linear observation - parameter models
- Examples mainly drawn from communications / radar

Aside: “Classical” vs. “Bayesian”

Classical

- Hypotheses/parameters are **fixed, non-random**

Bayesian

- Hypotheses/parameters are **treated as random variables with assumed priors (or a priori distributions)**

Course outline

Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory, by Steven M. Kay, Prentice Hall, 1993

General Minimum Variance Unbiased Estimation, Ch.2, 5
Cramer-Rao Lower Bound, Ch.3
Linear Models+Unbiased Estimators, Ch.4, 6
Maximum Likelihood Estimation, Ch.7
Least squares estimation, Ch.8
Bayesian Estimation, Ch.10-12
Kalman filtering

Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory, by Steven M. Kay, Prentice Hall 1998.

Statistical Detection Theory, Ch.3
Deterministic Signals, Ch.4
Random Signals, Ch.5
Statistical Detection Theory 2, Ch.6
Non-parametric and robust detection

Estimation: General Minimum Variance Unbiased Estimation

- Bias: (expected value of estimator - true value of data)

$$\text{Bias}(\hat{\theta}) = E[\hat{\theta}] - \theta = E[\hat{\theta} - \theta]$$

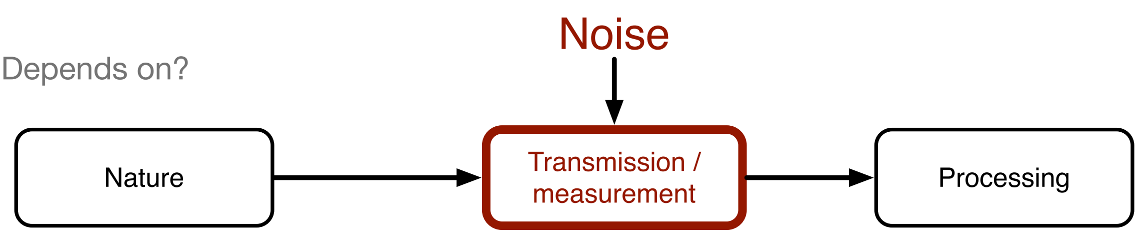
- MVUE:

Unbiased estimator of minimum variance

Always exist?

Estimation: Cramer-Rao lower bound

- Lower bound on **variance** of ANY unbiased estimator!
- Usage:
    - assert whether an estimator is MVUE
    - benchmark against which to measure the performance of an unbiased estimator
    - feasibility studies
- Depends on?



Estimation: linear models

- What's a linear model and why is it useful?

$$\mathbf{x} = \mathbf{H}\theta + \mathbf{w},$$

where

$$\begin{aligned} \theta &= \text{vector parameter to be estimated} \\ \mathbf{x} &= \text{received signal from which to estimate } \theta \\ \mathbf{H} &= \text{(known) observation matrix} \\ \mathbf{w} &= \text{noise of statistical characterization } \mathcal{N}(0, \sigma^2\mathbf{I}) \end{aligned}$$

- What can be said?
- Best Linear Unbiased Estimators (BLUE)

Estimation: Maximum Likelihood Estimation

- Alternative to MVUE which is hard to find in general
- Easy to compute - very widely used and practical
- What is the MLE?

If θ is the parameter to be estimated and $\mathbf{x}$ is the observation, then the MLE estimator $\hat{\theta_{MLE}}$ is:

$$\hat{\theta_{MLE}} = \arg\max p(\mathbf{x};\theta) \text{ for fixed (given) } \mathbf{x}$$

- Properties?

Estimation: Least Squares

- Alternative estimator with no general optimality properties, but nice and intuitive and no probabilistic assumptions on data are made - only need a signal model

The least squares estimator $\hat{\theta_{LS}}$ is equal to the value of θ that minimizes

$$J(\theta) = \sum_{n=0}^{N-1} (x[n] - s[n,\theta])^2,$$

where $s[n,\theta]$ is the sent signal (or nature) for given parameter θ.

- Advantages?
- Disadvantages?

Estimation: Bayesian Estimation

- Parameter to be estimated is assumed to be random, according to some prior distribution which models our knowledge of it

- Bayesian Minimum Mean Squared Error (MMSE):

Select the estimator $\hat{A}$ to minimize $BMSE(\hat{A}) = \int\int (A - \hat{A})^2 p(\mathbf{x}, A)\, d\mathbf{x} dA$

Obtain the famous mean of the posterior pdf, i.e.

$$\hat{A} = E[A|\mathbf{x}]$$

- Applications to Gaussian noise / linear model

Estimation: Bayesian Estimation

- General risk functions - arbitrary “cost” functions

$$\mathcal{R} = \int\int \mathcal{C}(\theta - \hat{\theta}) p(\mathbf{x}, \theta)\, d\mathbf{x} d\theta$$

- Maximum a posteriori (MAP) estimation

$$\hat{\theta} = \arg\max_{\theta} p(\theta|\mathbf{x})$$

- Linear MMSE: constrain estimator to be linear - very practical

$$\hat{\theta} = \sum_{n=0}^{N} a_n x[n] + a_N$$

where we choose the weighting coefficients a_n to minimize the Bayesian MSE

$$BMSE(\hat{\theta}) = E[(\theta - \hat{\theta})^2]$$

Estimation: Kalman filtering

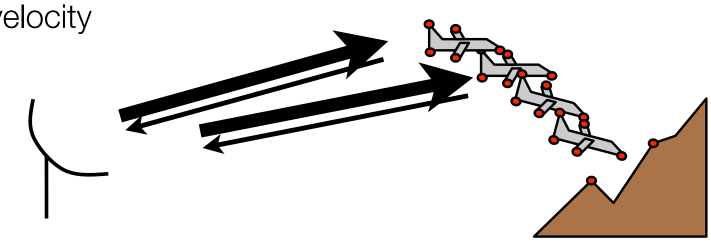
- recursive filter for estimating internal state of linear dynamical system from a series of noisy measurements

- example: tracking a moving target using radar measurements

noisy measurements

linear dynamical system

estimate the position and/or velocity



Estimation: Kalman filtering

- recursive filter for estimating internal state of linear dynamical system from a series of noisy measurements

$$
\begin{aligned}
\mathbf{X}_k &= \mathbf{F}_k \mathbf{X}_k + \mathbf{W}_k \\
\mathbf{Y}_k &= \mathbf{H}\mathbf{X}_k + \mathbf{V}_k \\
\mathbf{W}_k &\sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k), \quad E[\mathbf{W}_k \mathbf{W}_j^T] = \mathbf{Q}_k \delta_{k-j} \\
\mathbf{V}_k &\sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k, \quad E[\mathbf{V}_k \mathbf{V}_j^T] = \mathbf{R}_k \delta_{k-j},
\end{aligned}
$$

- than can recursively estimate/predict and update the state covariances as:

$$
\begin{aligned}
\mathbf{P}_k^- &= \mathbf{F}_{k-1} \mathbf{P}_{k-1}^+ \mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1} \\
\mathbf{K}_k &= \mathbf{P}_k^- \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1} = \mathbf{P}_k^+ \mathbf{H}_k^T \mathbf{R}_k^{-1} \\
\mathbf{x}_k^- &= \mathbf{F}_{k-1} \mathbf{x}_{k-1}^+ \\
\mathbf{x}_k^+ &= \mathbf{x}_k^- + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^-) \\
\mathbf{P}_k^+ &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^-
\end{aligned}
$$

Course outline

Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory, by Steven M. Kay, Prentice Hall, 1993

General Minimum Variance Unbiased Estimation, Ch.2, 5
Cramer-Rao Lower Bound, Ch.3
Linear Models+Unbiased Estimators, Ch.4, 6
Maximum Likelihood Estimation, Ch.7
Least squares estimation, Ch.8
Bayesian Estimation, Ch.10-12

Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory, by Steven M. Kay, Prentice Hall 1998.

Statistical Detection Theory, Ch.3
Deterministic Signals, Ch.4
Random Signals, Ch.5
Statistical Detection Theory 2, Ch.6
Non-parametric and robust detection

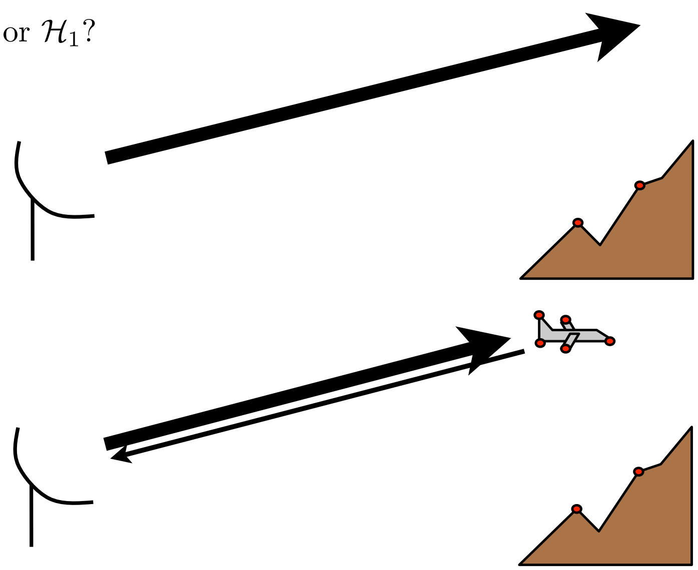
Detection: Statistical Detection Theory

- Binary hypothesis testing

$\mathcal{H}_0$ or $\mathcal{H}_1$?

Hypothesis $\mathcal{H}_0$

$$r(t) = n(t),\ 0 \le t \le T$$



Hypothesis $\mathcal{H}_1$

$$r(t) = V_r \sin((\omega_c + \omega_d)(t - \tau) + \theta_r) + n(t),\ \tau \le t \le t + \tau$$

Detection: Statistical Detection Theory

- Binary hypothesis testing

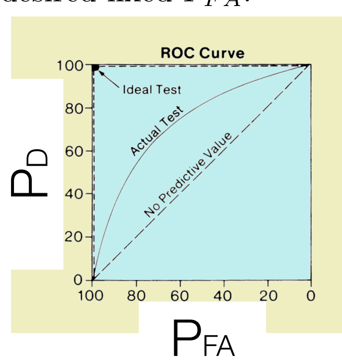
$$\mathcal{H}_0 \text{ or } \mathcal{H}_1?$$

- $P(\mathcal{H}_0;\mathcal{H}_0)$ = prob(decide $\mathcal{H}_0$ when $\mathcal{H}_0$ is true) = prob of correct non-detection
- $P(\mathcal{H}_0;\mathcal{H}_1)$ = prob(decide $\mathcal{H}_0$ when $\mathcal{H}_1$ is true) = prob of missed detection $:= P_M$
- $P(\mathcal{H}_1;\mathcal{H}_0)$ = prob(decide $\mathcal{H}_1$ when $\mathcal{H}_0$ is true) = prob of false alarm $:= P_{FA}$
- $P(\mathcal{H}_1;\mathcal{H}_1)$ = prob(decide $\mathcal{H}_1$ when $\mathcal{H}_1$ is true) = prob of detection $:= P_D$

Detection: Statistical Detection Theory

Neyman-Pearson (NP): maximize P_D subject to a desired fixed P_{FA}.

Receiver Operating Characteristics (ROC) curves



Generalized Bayesian risk which includes as special cases

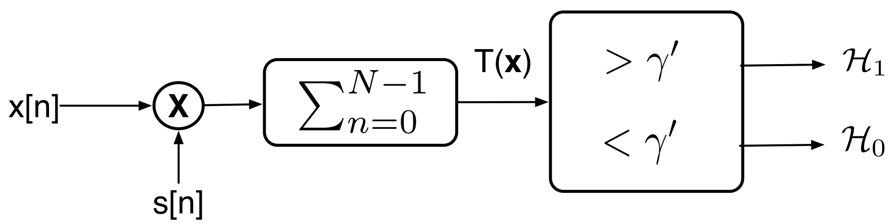
- Minimum probability of error ($\min P_E$) or maximum a posteriori (MAP): $C_{ii} = 0, C_{ij} = 1$ for $i \neq j$.
- Maximum likelihood (ML): $C_{ij} = 0, C_{ij} = 1$ for $i \neq j$ AND all priors are equal, i.e. $P(\mathcal{H}_i) = P(\mathcal{H}_j)$, $\forall i, j$.

Detection: Deterministic Signals

- How to detect known signals in noise?

$$\mathcal{H}_0 : x[n] = w[n]$$
$$\mathcal{H}_1 : x[n] = s[n] + w[n],$$

- The famous matched filter!



- Generalized matched filter
- > 2 hypotheses

Detection: Random Signals

- What if s[n] is random?

$$\mathcal{H}_0 : x[n] = w[n]$$
$$\mathcal{H}_1 : x[n] = s[n] + w[n],$$

- Key idea behind *estimator-correlator:*

Estimate the signal first, then matched-filter the estimate

- Linear model simplifies things again...

Detection: Statistical Decision Theory II

- model for the pdfs under 2 hypotheses are unknown

$$\mathcal{H}_0 : x[n] = w[n]$$
$$\mathcal{H}_1 : x[n] = s[n] + w[n],$$

- Uniformly most powerful test
- Generalized likelihood ratio test
- Bayesian approach
- Wald test
- Rao test

Course structure

http://www.ece.uic.edu/~devroye/courses/ECE531/