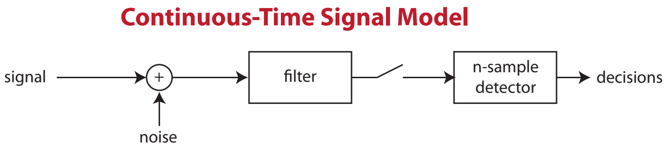
ECE531 Lecture 4b: Detection of Deterministic Discrete-Time Signals

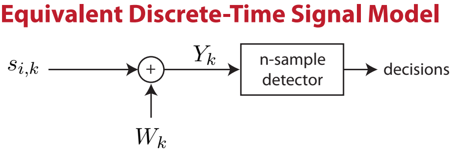
D. Richard Brown III

Worcester Polytechnic Institute

12-February-2009

Detection of Deterministic Discrete-Time Signals in Noise





In the **deterministic** model, we have a finite number of **known** signal vectors $s_i = [s_{i,0}, \dots, s_{i,n-1}]^\top$, $i = 0, \dots, M-1$, observed in additive noise $W = [W_0, \dots, W_{n-1}]^\top$. The problem is to use the n-sample observation to determine which of the M signals was sent.

Coherent Detection in Binary Communication Systems

- We have two known discrete-time signals $s_0, s_1 \in \mathbb{R}^n$ that we can transmit through a noisy communication channel.
- A signal $x \in \{s_0, s_1\}$ is transmitted and we observe a realization $y \in \mathbb{R}^n$ of the random variable $Y = x + W$ where, in this lecture, $W \sim \mathcal{N}(0, \Sigma)$ is zero mean additive Gaussian noise.
- The noise covariance matrix Σ is defined as

$$\Sigma = \mathrm{E}[WW^\top]$$

- We assume that the receiver knows the noise distribution $\mathcal{N}(0, \Sigma)$.
- Given the observation y, we must decide whether s_0 or s_1 was transmitted.
- This is called a **coherent** detection problem because the signals s_0 and s_1 are deterministic and are completely known in advance to the detector/receiver.

Conditional Distributions and Decision Rules

Conditioned on $x = s_j$, we can write density of the vector observation as

$$p_j(y) \quad = \quad \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{(y-s_j)^\top \Sigma^{-1}(y-s_j)}{2}\right)$$

This coherent detection problem is just a simple binary HT problem. We know we are going to have decision rules of the form:

$$\rho(y) \quad = \quad \begin{cases} 1 & \text{if } L(y) > v \\ \gamma & \text{if } L(y) = v \\ 0 & \text{if } L(y) < v \end{cases}$$

Hence, we are going to have to compute the likelihood ratio...

Likelihood Ratio

$$
\begin{aligned}
L(y) &= \frac{p_1(y)}{p_0(y)} \\
&= \exp\left(\frac{(y-s_0)^\top \Sigma^{-1}(y-s_0) - (y-s_1)^\top \Sigma^{-1}(y-s_1)}{2}\right)
\end{aligned}
$$

It is more convenient to work with the log-likelihood ratio here. Let

$$
\begin{aligned}
\ell(y) &:= 2\ln(L(y)) \\
&= (y-s_0)^\top \Sigma^{-1}(y-s_0) - (y-s_1)^\top \Sigma^{-1}(y-s_1)
\end{aligned}
$$

Then we can write the decision rule template as

$$
\rho(y) = \begin{cases} 1 & \text{if } \ell(y) \geq 2\ln v \\ 0 & \text{if } \ell(y) < 2\ln v \end{cases}
$$

Note that we dropped the case $\rho(y) = \gamma$. Why?

Decorrelation

Lemma

A real symmetric matrix P is positive definite if and only if there exists a nonsingular matrix S such that $P = S^\top S$.

Since Σ is positive definite, then so is Σ^{-1}. Hence, we can write

$$\Sigma^{-1} = S^\top S$$

where S is invertible. Now let

$$\begin{aligned}
\bar{y} &= Sy \\
&= S(x+w) \\
&= \begin{cases} \bar{s}_0 + \bar{w} & \text{if } x = s_0 \\ \bar{s}_1 + \bar{w} & \text{if } x = s_1 \end{cases}
\end{aligned}$$

Decorrelation

Note that S just specifies a one-to-one coordinate transformation between $\mathbb{R}^n$ and $\mathbb{R}^n$. In this new coordinate system

$$\begin{aligned}(y - s_j)^\top \Sigma^{-1} (y - s_j) &= (y - s_j)^\top S^\top S (y - s_j) \\ &= [S(y - s_j)]^\top S(y - s_j) \\ &= (\bar{y} - \bar{s}_j)^\top (\bar{y} - \bar{s}_j) \\ &= ||\bar{y} - \bar{s}_j||^2\end{aligned}$$

What happened to the noise after this coordinate transformation? It is still Gaussian, of course, with

$$\begin{aligned}\mathrm{E}[\bar{W}] &= \mathrm{E}[SW] = 0 \\ \mathrm{E}[\bar{W}\bar{W}^\top] &= \mathrm{E}[SWW^\top S^\top] = S\mathrm{E}[WW^\top]S^\top = S\Sigma S^\top = I\end{aligned}$$

Hence, the coordinate transformation has decorrelated the noise.

Correlation Form of the Decision Rule

By applying our coordinate transformation S, we can write the statistic used in our decision rule as

$$\begin{aligned}\ell(y) &= (y-s_0)^\top \Sigma^{-1}(y-s_0) - (y-s_1)^\top \Sigma^{-1}(y-s_1) \\ &= (\bar{y}-\bar{s}_0)^\top(\bar{y}-\bar{s}_0) - (\bar{y}-\bar{s}_1)^\top(\bar{y}-\bar{s}_1) \\ &= 2(\bar{s}_1-\bar{s}_0)^\top \bar{y} + ||\bar{s}_0||^2 - ||\bar{s}_1||^2\end{aligned}$$

Hence, our decision rule template can be further simplified as

$$\rho(\bar{y}) = \begin{cases} 1 & (\bar{s}_1-\bar{s}_0)^\top \bar{y} \quad \geq \ln v + \frac{1}{2}(||\bar{s}_1||^2 - ||\bar{s}_0||^2) \\ 0 & \qquad\qquad\qquad < \end{cases} \tag{1}$$

where $\bar{s}_j = Ss_j$ for $j = 0, 1$, $\bar{y} = Sy$, and $S^\top S = \Sigma^{-1}$.

Statistics of the Decision Variable $(\bar{s}_1 - \bar{s}_0)^\top \bar{y}$

Notation: Let $\bar{s} := \bar{s}_1 - \bar{s}_0$.

- Note that $\bar{s}^\top \bar{y}$ is the inner product (or deterministic correlation) between the coordinate-transformed observation $\bar{y}$ and the coordinate-transformed signal difference vector $\bar{s}$. The decision rule template (1) is sometimes called the **correlation form** decision rule.
- When $x = s_j$, the coordinate transformed observation $\bar{Y} \sim \mathcal{N}(\bar{s}_j, I)$.
- When $x = s_j$, how is $Z = \bar{s}^\top \bar{Y}$ distributed?
  - Conditioned on $x = s_j$, $Z \sim \mathcal{N}(\mu_j, \sigma^2)$ is a Gaussian random variable.

$$
\begin{aligned}
\mu_j &:= \mathrm{E}[Z \,|\, x = s_j] = \mathrm{E}[\bar{s}^\top \bar{Y} \,|\, x = s_j] = \bar{s}^\top \mathrm{E}[Y \,|\, x = s_j] = \bar{s}^\top \bar{s}_j \\
&= (s_1 - s_0)^\top \Sigma^{-1} s_j \\
\sigma^2 &:= \mathrm{E}[(Z - \bar{s}^\top \bar{s}_j)^2 \,|\, x = s_j] = \mathrm{E}[(\bar{s}^\top \bar{Y} - \bar{s}^\top \bar{s}_j)^2 \,|\, x = s_j] \\
&= \mathrm{E}[(\bar{s}^\top (\bar{Y} - \bar{s}_j))^2 \,|\, x = s_j] = \mathrm{E}[(\bar{s}^\top \bar{W})^2] = \bar{s}^\top \mathrm{E}[\bar{W}\bar{W}^\top]\bar{s} = \bar{s}^\top \bar{s} \\
&= (s_1 - s_0)^\top \Sigma^{-1} (s_1 - s_0).
\end{aligned}
$$

Types of Detectors

All of the detectors we have considered will have the following form:

$$\rho(\bar{y}) = \begin{cases} 1 & \bar{s}^\top \bar{y} \geq \ln v + \frac{1}{2}(||\bar{s}_1||^2 - ||\bar{s}_0||^2) \\ 0 & \phantom{\bar{s}^\top \bar{y}} < \end{cases}$$

The only difference will be in the choice of v.

- Bayes detector: Given a prior $\pi_j = \text{Prob}(x = s_j)$ for $j = 0, 1$ and a cost assignment C, we can find v to minimize the Bayes risk.
- Minimax detector: We already know the least favorable prior is $\pi_0 = \pi_1 = \frac{1}{2}$. Hence, for this problem, the minimax detector is simply the Bayes detector for this prior. No randomization is necessary.
- Neyman-Pearson detector: Letting $\phi := \frac{1}{2}(||\bar{s}_1||^2 - ||\bar{s}_0||^2)$, we can write the false positive probability as

$$P_{\mathsf{fp}} = \text{Prob}(Z \geq \ln v + \phi \,|\, x = s_0) = Q\left(\frac{\ln v + \phi - \mu_0}{\sigma}\right)$$

  We can find the N-P detector by finding v such that $P_{\mathsf{fp}} = \alpha$.

Bayes Detection with UCA: Probability of Error

Defining $\mathcal{H}_0 : x = s_0$, $\mathcal{H}_1 : x = s_1$ and given a prior $\pi_0 = \mathrm{Prob}(x = s_0)$ and $\pi_1 = \mathrm{Prob}(x = s_1) = 1 - \pi_0$, we can write the probability of making an incorrect detection as

$$\begin{aligned} P_e &= \pi_0 \mathrm{Prob}(\text{decide } \mathcal{H}_1 \,|\, x = s_0) + (1 - \pi_0)\mathrm{Prob}(\text{decide } \mathcal{H}_0 \,|\, x = s_1) \\ &= \pi_0 \mathrm{Prob}\left(\bar{s}^\top \bar{Y} \geq \tau \,|\, x = s_0\right) + (1 - \pi_0)\mathrm{Prob}\left(\bar{s}^\top \bar{Y} < \tau \,|\, x = s_1\right) \end{aligned}$$

where $\tau = \ln v + \frac{1}{2}(||\bar{s}_1||^2 - ||\bar{s}_0||^2)$. Hence

$$P_e = \pi_0 Q\left(\frac{\tau - \mu_0}{\sigma}\right) + (1 - \pi_0) Q\left(\frac{\mu_1 - \tau}{\sigma}\right)$$

We know from our study of Bayesian hypothesis testing (lectures 3-4) that the optimum threshold τ is

$$\tau = \frac{\mu_0 + \mu_1}{2} + \frac{\sigma^2}{\mu_1 - \mu_0} \ln \frac{\pi_0}{1 - \pi_0}.$$

Uniform Prior and Optimum Detection Threshold

Suppose that we have a uniform prior $\pi_0 = \pi_1 = \frac{1}{2}$ and that the optimum threshold for this prior, i.e. $\tau = \frac{\mu_0+\mu_1}{2}$, is used by the detector. Then the probability of error can be written as

$$\begin{aligned}
P_e &= \frac{1}{2}Q\left(\frac{\mu_1-\mu_0}{2\sigma}\right) + \frac{1}{2}Q\left(\frac{\mu_1-\mu_0}{2\sigma}\right) \\
&= Q\left(\frac{(s_1-s_0)^\top\Sigma^{-1}(s_1-s_0)}{2\sqrt{(s_1-s_0)^\top\Sigma^{-1}(s_1-s_0)}}\right) \\
&= Q\left(\frac{1}{2}\sqrt{(s_1-s_0)^\top\Sigma^{-1}(s_1-s_0)}\right) \\
&= Q\left(\frac{||\bar{s}||}{2}\right)
\end{aligned}$$

In communication problems, we often can choose our signal alphabet. Since $Q(x)$ is monotonically decreasing in x, we should choose s_0 and s_1 to maximize $||\bar{s}||$.

Optimum Signal Design for Minimum Error Probability

To get an useful result, we require the power of each signal s_0 and s_1 to be upper bounded by B, i.e.

$$\frac{1}{n}||s_j||^2 \leq B.$$

Let $\lambda_{max}(\Sigma^{-1})$ be the largest eigenvalue of the positive definite matrix Σ^{-1} and let ν be the eigenvector associated with this eigenvalue. Then we can say

$$(s_1 - s_0)^\top \Sigma^{-1}(s_1 - s_0) \leq \lambda_{max}(\Sigma^{-1})||s_1 - s_0||^2$$

with equality only if $s_1 - s_0 = \alpha\nu$ for some scalar α. Hence

$$(s_1 - s_0)^\top \Sigma^{-1}(s_1 - s_0) = ||\bar{s}||^2 \leq 2nB\lambda_{max}(\Sigma^{-1}).$$

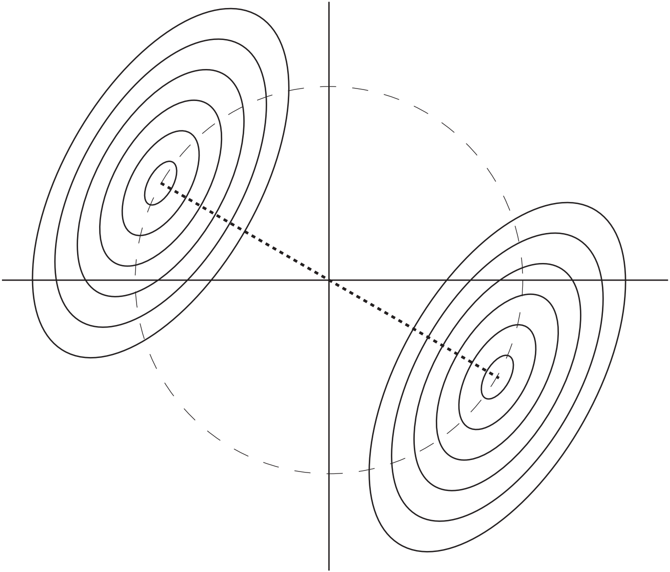
Optimum Signal Design: Interpretation

- The probability of error is minimized when $s_1 - s_0$ is aligned with the eigenvector of Σ^{-1} corresponding to the maximum eigenvalue $\lambda_{max}(\Sigma^{-1})$ and arranged antipodally on the sphere of radius $\sqrt{nB}$.
- Note that $\lambda_{max}(\Sigma^{-1}) = 1/\lambda_{min}(\Sigma)$.
- The probability of error is minimized when $s_1 - s_0$ is aligned with the eigenvector of Σ corresponding to the smallest eigenvalue $\lambda_{min}(\Sigma)$, i.e. $s_1 - s_0$ is aligned in the direction of least noise variance.
- The minimum achievable error probability with optimum signaling (and equiprobable signals) is then

$$P_e^* = Q\left(\frac{\sqrt{nB}}{\sigma_{min}}\right)$$

where $\sigma_{min}^2 = \lambda_{min}(\Sigma)$.

Optimum Signal Design: Geometric Interpretation



Special Case: Independent Gaussian Noise

When $\Sigma = \sigma^2 I$, no coordinate transformation is needed and the decision rule is simply of the form

$$\rho(y) = \begin{cases} 1 & (s_1 - s_0)^\top y \quad \geq \ln v + \frac{1}{2}(||s_1||^2 - ||s_0||^2) \\ 0 & \qquad\qquad\quad < \end{cases}$$

where v is chosen to satisfy the objective, e.g. minimum Bayes risk. Moreover, since all of the eigenvalues of Σ are the same, the error probability will be the same for all orientations of the signal difference vector $s_1 - s_0$.

Conclusions

- Optimum detection of **deterministic** signals in noise: just simple hypothesis testing.
- We discussed the detection of discrete-time deterministic signals in Gaussian noise with arbitrary covariance.
- Your textbook also discusses detection of discrete-time deterministic signals in i.i.d. non-Gaussian noise, pp. 47-50. You should read this.
- Lots of detection problems, however, have signals with unknown parameters, e.g. non-coherent communication systems, radar, etc.
  - Example: Detect the presence of a signal $s_1 = a\sin(\omega t + \phi)$ versus $s_0 = 0$ where a and ϕ are unknown.
  - How can we approach these types of detection problems?
  - Answer: Composite hypothesis testing.