6. Detection

Statistical Signal Processing
Version 1.7

Prof. Peter Schreier, Ph.D.
Signal and System Theory Group
Faculty of Electrical Engineering, Computer Science, and Mathematics
Universität Paderborn

6.1 Neyman-Pearson detectors

Binary hypothesis testing

- A hypothesis is a statement about the distribution of a measurement $\mathbf{y} \in \mathcal{Y}$, coded by a parameter $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ that indexes the pdf $f_{\boldsymbol{\theta}}(\mathbf{y})$. The value of $\boldsymbol{\theta}$ **modulates** the distribution of the measurement (e.g., through the mean or covariance matrix).

- In **binary hypothesis testing** there are two competing hypotheses, H_0 (the **null hypothesis**) and H_1 (the **alternative**).

- Each of the hypotheses may be
    - **simple**, in which case the pdf of the measurement is $f_{\theta_i}(\mathbf{y})$ under hypothesis H_i
    - **composite**, in which case the distribution of the measurement is $f_{\boldsymbol{\theta}}(\mathbf{y}), \boldsymbol{\theta} \in \boldsymbol{\Theta}_i$, under hypothesis H_i

- There are two main lines of development for detection theory:
    - **Neyman-Pearson**: frequentist theory that assigns no prior probabilities to competing models (we will study only Neyman-Pearson detectors)
    - **Bayes**: assigns prior probabilities to competing models; assigns costs to incorrect decisions, and then minimizes average cost

Test statistic

A binary test of H_1 vs. H_0 is a **test statistic**, or **threshold detector** $\phi : \mathcal{Y} \to \{0, 1\}$ of the form

$$\phi(\mathbf{y}) = \begin{cases} 1, & \mathbf{y} \in \mathcal{Y}_1 \\ 0, & \mathbf{y} \in \mathcal{Y}_0 \end{cases}$$

Two types of errors:

- **Missed detection:** Classify a measurement drawn from the pdf $f_{\boldsymbol{\theta}}(\mathbf{y}), \boldsymbol{\theta} \in \boldsymbol{\Theta}_1$, as a measurement drawn from the pdf $f_{\boldsymbol{\theta}}(\mathbf{y}), \boldsymbol{\theta} \in \boldsymbol{\Theta}_0$
- **False alarm:** Classify a measurement drawn from the pdf $f_{\boldsymbol{\theta}}(\mathbf{y}), \boldsymbol{\theta} \in \boldsymbol{\Theta}_0$, as a measurement drawn from the pdf $f_{\boldsymbol{\theta}}(\mathbf{y}), \boldsymbol{\theta} \in \boldsymbol{\Theta}_1$

For simple hypothesis and alternative the **detection probability** (power) P_D and the **false alarm probability** (size) P_F of a test are defined as

$$P_D = P_{\boldsymbol{\theta}_1}(\phi = 1) = E_{\boldsymbol{\theta}_1}[\phi]$$
$$P_F = P_{\boldsymbol{\theta}_0}(\phi = 1) = E_{\boldsymbol{\theta}_0}[\phi]$$

The subscripts $\boldsymbol{\theta}_i$ denote probability and expectation under pdf $f_{\boldsymbol{\theta}_i}(\mathbf{y})$.

Neyman-Pearson lemma

Neyman-Pearson (NP) detector

Among all competing binary tests ϕ' of a simple alternative vs. a simple hypothesis, with probability of false alarm less than or equal to α, none has larger detection probability than the **likelihood ratio test**

$$\phi(\mathbf{y}) = \begin{cases} 1, & \ell(\mathbf{y}) > \eta \\ \gamma, & \ell(\mathbf{y}) = \eta \\ 0, & \ell(\mathbf{y}) < \eta. \end{cases}$$

In this equation, $0 \leq \ell(\mathbf{y}) < \infty$ is the *likelihood ratio*

$$\ell(\mathbf{y}) = \frac{f_{\boldsymbol{\theta}_1}(\mathbf{y})}{f_{\boldsymbol{\theta}_0}(\mathbf{y})}$$

and the threshold parameters (η, γ) are chosen to meet the constraint on false alarm probability $P_{\boldsymbol{\theta}_0}[\ell(\mathbf{y}) > \eta] + \gamma P_{\boldsymbol{\theta}_0}[\ell(\mathbf{y}) = \eta] = \alpha$. When $\phi(\mathbf{y}) = \gamma$, then select H_1 with probability γ, and H_0 with probability $1 - \gamma$.

Sufficient statistics

The test function ϕ, given its likelihood ratio ℓ, is 0 if $\ell < \eta$, γ if $\ell = \eta$, and 1 if $\ell > \eta$, independently of $\boldsymbol{\theta}$. Thus, the likelihood ratio statistic is a **sufficient statistic** for testing H_1 vs. H_0.

Fisher-Neyman factorization theorem

We call a statistic $\mathbf{t}(\mathbf{y})$ **sufficient** if the pdf $f_{\boldsymbol{\theta}}(\mathbf{y})$ may be factored as $f_{\boldsymbol{\theta}}(\mathbf{y}) = a(\mathbf{y})q_{\boldsymbol{\theta}}(\mathbf{t})$.

In this case the likelihood ratio equals the likelihood ratio for $\mathbf{t}$:

$$\ell(\mathbf{y}) = \frac{f_{\boldsymbol{\theta}_1}(\mathbf{y})}{f_{\boldsymbol{\theta}_0}(\mathbf{y})} = \frac{q_{\boldsymbol{\theta}_1}(\mathbf{t})}{q_{\boldsymbol{\theta}_0}(\mathbf{t})} = \ell'(\mathbf{t})$$

We may thus rewrite the NP Lemma in terms of the sufficient statistic:

$$\phi(\mathbf{t}) = \begin{cases} 1, & \ell'(\mathbf{t}) > \eta \\ \gamma, & \ell'(\mathbf{t}) = \eta \\ 0, & \ell'(\mathbf{t}) < \eta \end{cases}$$

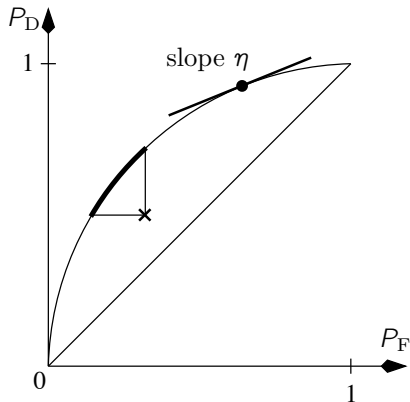where (η, γ) are chosen to meet the P_{F} constraint.

Minimal statistic

A sufficient statistic is **minimal**, meaning it is most memory efficient, if it is a function of every other sufficient statistic.

Invariant statistic

A sufficient statistic is an **invariant statistic** with respect to the transformation group G if $\phi(g(\mathbf{t})) = \phi(\mathbf{t})$ for all $g \in G$. The composition rule for any two transformations within G is $g_1 \circ g_2(\mathbf{t}) = g_1(g_2(\mathbf{t}))$.

Receiver operating characteristics (ROC)

For $t = \ell(\mathbf{y})$, the pdfs under the two hypotheses are connected as

$$f_{\boldsymbol{\theta}_1}(\ell) = \ell f_{\boldsymbol{\theta}_0}(\ell)$$

The probability of detection and probability of false alarm may be written

$$1 - P_{\mathrm{D}} = \int_0^\eta f_{\boldsymbol{\theta}_1}(\ell)d\ell = \int_0^\eta \ell f_{\boldsymbol{\theta}_0}(\ell)d\ell$$

$$1 - P_{\mathrm{F}} = \int_0^\eta f_{\boldsymbol{\theta}_0}(\ell)d\ell$$

Differentiate each of these formulas with respect to η to find

$$\frac{\partial}{\partial \eta}P_{\mathrm{D}} = -\eta f_{\boldsymbol{\theta}_0}(\eta) \quad \text{and} \quad \frac{\partial}{\partial \eta}P_{\mathrm{F}} = -f_{\boldsymbol{\theta}_0}(\eta)$$

and therefore $\dfrac{dP_{\mathrm{D}}}{dP_{\mathrm{F}}} = \eta$.

6.2 Simple Gaussian hypothesis testing

Uncommon means and common covariance

- Test the alternative H_1 that a Gaussian measurement **y** has mean value $\boldsymbol{\mu}_1$ vs. the hypothesis H_0 that it has mean value $\boldsymbol{\mu}_0$.

- Under both hypotheses **y** has covariance matrix
  $\mathbf{R} = E_{\boldsymbol{\theta}_0}[(\mathbf{y} - \boldsymbol{\mu}_0)(\mathbf{y} - \boldsymbol{\mu}_0)^H] = E_{\boldsymbol{\theta}_1}[(\mathbf{y} - \boldsymbol{\mu}_1)(\mathbf{y} - \boldsymbol{\mu}_1)^H]$.

- The pdf of **y** under hypothesis H_i is

$$f_{\boldsymbol{\theta}_i}(\mathbf{y}) = \frac{1}{\pi^n \det \mathbf{R}} \exp\left\{-(\mathbf{y} - \boldsymbol{\mu}_i)^H \mathbf{R}^{-1}(\mathbf{y} - \boldsymbol{\mu}_i)\right\}$$

- After some algebra, the log-likelihood ratio may be expressed as a function of **y**:

$$L = \log \frac{f_{\boldsymbol{\theta}_1}(\mathbf{y})}{f_{\boldsymbol{\theta}_0}(\mathbf{y})} = -(\mathbf{y} - \boldsymbol{\mu}_1)^H \mathbf{R}^{-1}(\mathbf{y} - \boldsymbol{\mu}_1) + (\mathbf{y} - \boldsymbol{\mu}_0)^H \mathbf{R}^{-1}(\mathbf{y} - \boldsymbol{\mu}_0)$$

$$= 2\mathrm{Re}\left\{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^H \mathbf{R}^{-1}(\mathbf{y} - \mathbf{y}_0)\right\}$$
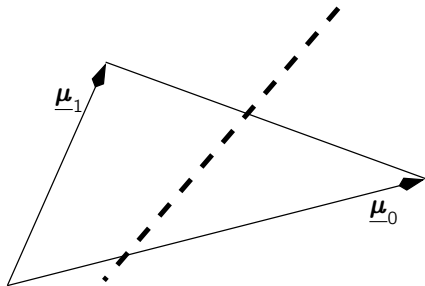
with $\mathbf{y}_0 = \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_0)$.

The log-likelihood ratio may be written as the inner product

$$L = 2\mathrm{Re}\left\{\left[\mathbf{R}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)\right]^H (\mathbf{y} - \mathbf{y}_0)\right\} = 2\mathrm{Re}\left\{\mathbf{w}^H(\mathbf{y} - \mathbf{y}_0)\right\},$$

with $\mathbf{w} = \mathbf{R}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)$. This inner product can be interpreted as **matched filtering**.

The centered measurement $\mathbf{y} - \mathbf{y}_0$ is resolved onto the line $\langle \mathbf{R}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)\rangle$ to produce the **real** log-likelihood ratio L, which is compared to a threshold.

In the figure, we choose lines parallel to the thick dashed line to get the desired probability of false alarm P_F.

Uncommon means and common covariance

But how do we determine the threshold to get the desired P_F? We will see that the deflection, or output signal-to-noise ratio, plays a key role.

The **deflection** for a likelihood ratio test with log-likelihood ratio L is

$$d = \frac{[E_{\boldsymbol{\theta}_1}(L) - E_{\boldsymbol{\theta}_0}(L)]^2}{\text{var}_{\boldsymbol{\theta}_0}(L)},$$

where $\text{var}_{\boldsymbol{\theta}_0}(L)$ denotes the variance of L under H_0.
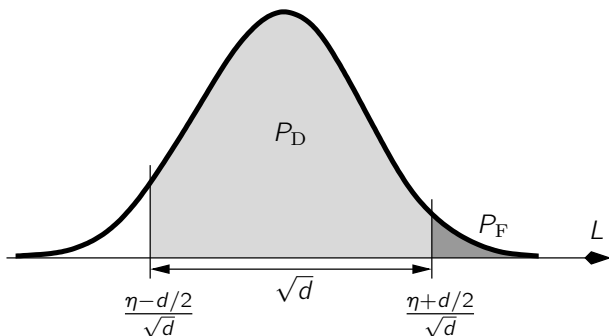
The likelihood ratio L derived on the previous slide results in

$$d = 2(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^H \mathbf{R}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0).$$

Proof: The mean of L is $d/2$ under H_1 and $-d/2$ under H_0, and its variance is d under both hypotheses.

Thus, the log-likelihood ratio statistic L is distributed as a real Gaussian random variable with mean value $\pm d/2$ and variance d.

Now P_F and P_D for the detector ϕ that compares L to a threshold η are

$$P_\mathrm{F} = 1 - \int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi d}} \exp\left\{-\frac{1}{2d}(L+d/2)^2\right\} dL = 1 - \Phi\left(\frac{\eta+d/2}{\sqrt{d}}\right)$$

$$P_\mathrm{D} = 1 - \int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi d}} \exp\left\{-\frac{1}{2d}(L-d/2)^2\right\} dL = 1 - \Phi\left(\frac{\eta-d/2}{\sqrt{d}}\right)$$

where Φ is the cumulative distribution function of a zero-mean, variance one, real Gaussian random variable.

Common mean and uncommon covariances

- Assume w.l.o.g. that the mean is zero. Then the likelihood for the measurement $\mathbf{y}$ under hypothesis H_i is

$$f_{\boldsymbol{\theta}_i}(\mathbf{y}) = \frac{1}{\pi^n \det \mathbf{R}_i} \exp\left\{-\mathbf{y}^H \mathbf{R}_i^{-1} \mathbf{y}\right\}$$

where $\mathbf{R}_i$ is the covariance matrix under hypothesis H_i.

- The log-likelihood ratio for comparing H_1 to H_0 is

$$L = \mathbf{y}^H (\mathbf{R}_0^{-1} - \mathbf{R}_1^{-1})\mathbf{y} = \mathbf{y}^H \mathbf{R}_0^{-H/2}(\mathbf{I} - \mathbf{S}^{-1})\mathbf{R}_0^{-1/2}\mathbf{y}$$

where $\mathbf{S} = \mathbf{R}_0^{-1/2}\mathbf{R}_1\mathbf{R}_0^{-H/2}$ is the **signal-to-noise ratio matrix**.

- The transformed measurement $\mathbf{R}_0^{-1/2}\mathbf{y}$ has covariance matrix $\mathbf{I}$ under H_0 and covariance matrix $\mathbf{S}$ under H_1.

- Thus L is the log-likelihood ratio for testing that the linearly transformed measurement $\mathbf{R}_0^{-1/2}\mathbf{y}$ is **white** vs. the alternative that it has covariance $\mathbf{S}$.