M-ary Detection & the MAP Test

Let H be a $\{0,\dots,M-1\}$-valued RV with

$$p_i := P(H=i).$$

Let Y be an $\mathbb{R}^d$-valued random vector with conditional density $f_i(y)$ so that

$$P(Y\in B|H=i) = \int_B f_i(y)\,dy.$$

Having observed Y, we would like to decide what the most likely value of H is. In other words, we want to find a function $\psi(y)$ taking values in $\{0,\dots,M-1\}$ so that

$$P(\psi(Y)\neq H)$$

is minimized.

Since $\psi(y)$ takes only the values $0,\dots,M-1$, ψ must have the form

$$\psi(y) = \sum_{i=0}^{M-1} i\, I_{D_i}(y)$$

where the sets D_i are disjoint and $\bigcup_{i=0}^{M-1} D_i = \mathbb{R}^d$. So really, our goal is to find the <u>decision regions</u> $D_0,\dots,D_{M-1}$. To find the best ones, we proceed as follows:

Use the law of total probability to write

$$P(\psi(Y)=H) = \sum_{i=0}^{M-1} P(\psi(Y)=H \mid H=i)\,p_i$$

$$= \sum_{i=0}^{M-1} P(\psi(Y)=i \mid H=i)\,p_i, \quad \text{by substitution}$$

$$= \sum_{i=0}^{M-1} P(Y \in D_i \mid H=i)\,p_i$$

$$= \sum_{i=0}^{M-1} \left(\int_{D_i} f_i(y)\,dy \right) p_i$$

$$= \sum_{i=0}^{M-1} \left(\int I_{D_i}(y)\, f_i(y)\,dy \right) p_i$$

$$= \int \left[\sum_{i=0}^{M-1} I_{D_i}(y)\, f_i(y)\, p_i \right] dy.$$

For fixed y, only one term in the sum is nonzero since the D_i are disjoint. To maximize $P(\psi(Y)=H)$, we should put $y \in D_i \Leftrightarrow f_i(y)p_i \geq f_j(y)p_j$ for all j. This is the <u>maximum</u> <u>a</u> <u>posteriori</u> <u>probability</u> rule (MAP rule). If the p_i are all the same; i.e., H is uniformly distributed with $p_i = 1/M$, then the rule of putting $y \in D_i \Leftrightarrow f_i(y) \geq f_j(y)$ for all j is called the maximum likelihood (ML) rule.

<u>Example.</u> Suppose $f_i(y) \sim N(m_i, C)$ and that $p_i = 1/M$. Then

$$f_i(y) \geq f_j(y)$$

if and only if

$$\frac{\exp[-(y-m_i)'C^{-1}(y-m_i)/2]}{(2\pi)^{d/2}\sqrt{\det C}} \geq \frac{\exp[-\frac{1}{2}(y-m_j)'C^{-1}(y-m_j)]}{(2\pi)^{d/2}\sqrt{\det C}}$$

$\Leftrightarrow$

$$-\tfrac{1}{2}(y-m_i)'C^{-1}(y-m_i) \geq -\tfrac{1}{2}(y-m_j)'C^{-1}(y-m_j) \qquad (*)$$

$$\Leftrightarrow \quad y'C^{-1}y - 2m_i'C^{-1}y + m_i'C^{-1}m_i \leq y'C^{-1}y - 2m_j'C^{-1}y + m_j'C^{-1}m_j \qquad (**)$$

$$\Leftrightarrow \quad (m_j-m_i)'C^{-1}y \leq \tfrac{1}{2}\{m_j'C^{-1}m_j - m_i'C^{-1}m_i\}.$$

If $C=\sigma^2 I$, then $(*)$ is equivalent to

$$\|y-m_i\| \leq \|y-m_j\|, \qquad (\#)$$

where here $\|\cdot\|$ denote Euclidean distance. Thus, we choose i if y is closer to m_i than any other m_j. If $C=\sigma^2 I$, then $(**)$ becomes

$$\|y\|^2 - 2m_i'y + \|m_i\|^2 \leq \|y\|^2 - 2m_j'y + \|m_j\|^2. \qquad (\#\#)$$

<u>Example.</u> Let us repeat the preceding example, but instead of having the observation be a <u>real-valued</u> random vector, let us use a complex circularly symmetric Gaussian random vector $Z \sim N(m_i, K)$. Then $f_i(z) \geq f_j(z) \iff$

$$\frac{e^{-(z-m_i)^H K^{-1}(z-m_i)}}{\pi^d \det K} \geq \frac{e^{-(z-m_j)' K^{-1}(z-m_j)}}{\pi^d \det K}$$

$$\iff (z-m_i)^H K^{-1}(z-m_i) \leq (z-m_j)^H K^{-1}(z-m_j) \qquad (*)$$

$$\iff z^H K^{-1} z - 2\,\mathrm{Re}\, m_i^H K^{-1} z + m_i^H K^{-1} m_i \leq z^H K^{-1} z - 2\,\mathrm{Re}\, m_j^H K^{-1} z + m_j^H K^{-1} m_j \qquad (**)$$

$$\iff \mathrm{Re}\,(m_j - m_i)^H K^{-1} z \leq \tfrac{1}{2}\{m_j^H K^{-1} m_j - m_i^H K^{-1} m_i\}$$

If $K = \sigma^2 I$, then $(*)$ is equivalent to

$$\|z - m_i\| \leq \|z - m_j\|, \qquad (\#)$$

and $(**)$ becomes

$$\|z\|^2 - 2\,\mathrm{Re}\, m_i^H z + \|m_i\|^2 \leq \|z\|^2 - 2\,\mathrm{Re}\, m_j^H z + \|m_j\|^2. \qquad (\#\#)$$

Detection of Known Waveforms in AWGN

In this problem, given $H = i$, $Y(t)$ is a random process of the form

$$Y(t) = s_i(t) + Z(t)$$

where $s_i(t)$ is a known, finite-energy waveform and $Z(t)$ is AWGN. In the complex case, $Z(t)$ is assumed circularly symmetric; i.e., any vector of the form $[Z(t_1), \ldots, Z(t_n)]'$ is circularly symmetric, and hence, so is any vector whose components have the form

$$\int g_k(t) Z_t \, dt.$$

Plan of Analysis.

1) Let $\mathcal{S} := \text{span}\{s_0, \ldots, s_{M-1}\}$. Thus, any $x \in \mathcal{S}$ has the form $x(t) = \sum_{i=0}^{M-1} x_i s_i(t)$ for suitable coefficients x_i.

2) Let $\varphi_1, \ldots, \varphi_N$ be orthonormal and such that $\text{span}\{\varphi_1, \ldots, \varphi_N\} = \mathcal{S}$. Then $N \leq M$, with equality $\Leftrightarrow$ $s_0, \ldots, s_{M-1}$ are linearly independent. The φ_k can be found by the Gram-Schmidt procedure.

Notation: For any finite-energy waveforms x and y, $\langle x, y \rangle := \int x(t) y(t)^* \, dt$. The energy of y is $\|y\|^2$, where $\|y\| := \langle y, y \rangle^{1/2}$. Thus $\|y\|^2 = \int |y(t)|^2 dt$.

We say x & y are orthogonal if $\langle x, y \rangle = 0$, and (denoted by $x \perp y$)

we say x and y are <u>orthonormal</u> (o.n.) if they are orthogonal AND $\|x\|=\|y\|=1$. (i.e., they have unit energy).

3) Every finite-energy waveform y can be written <u>uniquely</u> in the form

$$y = \hat{y} + \tilde{y},$$

where $\hat{y} \in \mathcal{S}$ and $\tilde{y}$ is orthogonal to every element of $\mathcal{S}$, including $\hat{y}$. To see this, put

$$\hat{y} := \sum_{k=1}^{N} \langle y, \varphi_k \rangle \varphi_k .$$

Then#

$$\begin{aligned}
\langle \tilde{y}, \varphi_\ell \rangle &= \langle y - \hat{y}, \varphi_\ell \rangle \\
&= \langle y, \varphi_\ell \rangle - \langle \hat{y}, \varphi_\ell \rangle \\
&= \langle y, \varphi_\ell \rangle - \Big\langle \sum_{k=1}^{N} \langle y, \varphi_k \rangle \varphi_k, \varphi_\ell \Big\rangle \\
&= \langle y, \varphi_\ell \rangle - \sum_{k=1}^{N} \langle y, \varphi_k \rangle \underbrace{\langle \varphi_k, \varphi_\ell \rangle}_{= \delta_{k\ell}} \\
&= \langle y, \varphi_\ell \rangle - \langle y, \varphi_\ell \rangle = 0. \qquad (*)
\end{aligned}$$

The decomposition $y = \hat{y} + \tilde{y}$ is unique, for if $y = \hat{y}_1 + \tilde{y}_1$ as well, then from

$$\hat{y} + \tilde{y} = \hat{y}_1 + \tilde{y}_1$$

we would have

$$\hat{y} - \hat{y}_1 = \tilde{y}_1 - \tilde{y}$$

Then the LHS $\in \mathcal{S}$, but the RHS is $\perp$ to $\mathcal{S}$. We

Eq. (*) shows that $\langle y, \varphi_k \rangle = \langle \hat{y}, \varphi_k \rangle$

could then write

$$\begin{aligned}\|\hat{y}-\hat{y}_1\|^2 &= \langle \hat{y}-\hat{y}_1, \hat{y}-\hat{y}_1 \rangle \\ &= \langle \hat{y}-\hat{y}_1, (y-\tilde{y})-(y-\tilde{y}_1) \rangle \\ &= \langle \hat{y}-\hat{y}_1, \tilde{y}_1-\tilde{y} \rangle \\ &= 0.\end{aligned}$$

$\therefore$ $\hat{y}=\hat{y}_1$, and then $\tilde{y}=\tilde{y}_1$ as well.

4) Since the above decomposition is unique, if $s\in\mathcal{S}$, then $\hat{s}=s\in\mathcal{S}$ and $\tilde{s}=0$; i.e., $s=s+0$.

5) Take the received signal $Y(t)=s_i(t)+Z(t)$ and write it as $Y=\hat{Y}+\tilde{Y}$, where $\hat{Y}=\hat{s}_i+\hat{Z}=s_i+\hat{Z}$, and $\tilde{Y}=\tilde{s}_i+\tilde{Z}=\tilde{Z}$. The first key observation is that $\tilde{Y}$ does NOT contain any information about the signal s_i. However, $\tilde{Y}=\tilde{Z}$ may contain information about the noise $\hat{Z}$. Fortunately, because the noise is Gaussian, we can show that $\hat{Z}$ and $\tilde{Z}$ are statistically independent. Because they are Gaussian, it is enough to show they are uncorrelated, which you will do in the HW. Because $\tilde{Y}$ does not contain any information about the signal s_i or the noise term $\hat{Z}$, we can make our decision based on $\hat{Y}$ instead of Y itself.

6) Our next key observation is that the waveform $\hat{Y}(t)$ and the vector $\underline{Y}=[\langle Y,\varphi_1\rangle,\ldots,\langle Y,\varphi_N\rangle]'$ are

equivalent in that each is a function of the other. If we know $\hat{Y}$, then we know $\langle \hat{Y}, \varphi_k \rangle = \langle Y, \varphi_k \rangle$ by the footnote on p. D-5, and so we can compute the vector $\underline{Y}$ from the waveform $\hat{Y}(t)$. Conversely, since

$$\hat{Y}(t) := \sum_{k=1}^{N} \langle Y, \varphi_k \rangle \varphi_k(t),$$

we can compute the waveform $\hat{Y}(t)$ from the column vector $\underline{Y}$.

7) We can write $\underline{Y} = \underline{s}_i + \underline{Z}$, where $(\underline{s}_i)_k = \langle s_i, \varphi_k \rangle$ and $Z_k = \langle Z, \varphi_k \rangle$.

Thus, given $H=i$, $\underline{Y}$ has density $f_i(\underline{y}) = f(\underline{y} - \underline{s}_i)$

where $f(\underline{z}) = \dfrac{e^{-\|\underline{z}\|^2/\sigma^2}}{\pi^N (\sigma^2)^N}$ in the complex case,

and $f(\underline{z}) = \dfrac{e^{-\frac{1}{2}\|\underline{z}\|^2/\sigma^2}}{(2\pi)^{N/2}\, \sigma^N}$ in the real case,

and where $\|\underline{z}\|^2 = \sum_{k=1}^{N} |z_k|^2$.

We have from ## on p. D-4 (p. D-3 or) that we should decide i if

$$\|\underline{Y}\|^2 - 2\,\mathrm{Re}\, \underline{s}_i^H \underline{Y} + \|\underline{s}_i\|^2 \le \|\underline{Y}\|^2 - 2\,\mathrm{Re}\, \underline{s}_j^H \underline{Y} + \|\underline{s}_j\|^2. \quad (*)$$

8) Since $s_i \in \mathcal{S}$, $s_i = \sum_{k=1}^{N} \langle s_i, \varphi_k \rangle \varphi_k$, we have :

for any waveform x,

$$\langle x, s_i \rangle = \langle x, \sum_{k=1}^{N} \langle s_i, \varphi_k \rangle \varphi_k \rangle$$

$$= \sum_{k=1}^{N} \langle s_i, \varphi_k \rangle^* \langle x, \varphi_k \rangle$$

$$= \underline{s}_i^H \begin{bmatrix} \langle x, \varphi_1 \rangle \\ \vdots \\ \langle x, \varphi_N \rangle \end{bmatrix}.$$

Thus, $\|s_i\|^2 = \langle s_i, s_i \rangle = \underline{s}_i^H \underline{s}_i = \|\underline{s}_i\|^2$.

Also,

$$\langle Y, s_i \rangle = \underline{s}_i^H \underline{Y}.$$

Thus, after replacing $\|\underline{Y}\|^2$ by $\|Y\|^2$ in ⊛ on p. D-8, and then changing $\underline{s}_i^H \underline{Y}$ to $\langle Y, s_i \rangle$ and $\|\underline{s}_i\|^2$ to $\|s_i\|^2$, etc, we obtain

$$\|Y - s_i\| \le \|Y - s_j\|.$$

Thus, the optimal decision can be expessed in terms of the received waveform $Y(t)$ and the signaling waveforms $s_i(t)$.

Probability of Error for Binary Signaling

The false alarm probability is

$$P(\text{decide } i=1 \mid H=0) = P(\ \|(s_0 + Z) - s_1\| \le \|(s_0 + Z) - s_0\| \mid H=0)$$

The required event can be rewritten as

$$\|(s_0-s_1)+z\|^2 \le \|z\|^2$$

or

$$\|s_0-s_1\|^2 + 2\,\mathrm{Re}\langle s_0-s_1, z\rangle + \|z\|^2 \le \|z\|^2$$

or

$$2\,\mathrm{Re}\,\langle s_1-s_0, z\rangle \ge \|s_1-s_0\|^2$$

or

$$\mathrm{Re}\left\langle \frac{s_1-s_0}{\|s_1-s_0\|}, z\right\rangle \ge \tfrac{1}{2}d,$$

where $d := \|s_1-s_0\|$.

Now, since $(s_1-s_0)/\|s_1-s_0\|$ has unit energy,

$$\left\langle \frac{s_1-s_0}{\|s_1-s_0\|}, z\right\rangle \sim N(0,\sigma^2)$$

and its real part is $N(0,\sigma^2/2)$. Thus, in the complex case, the false-alarm prob. is

$$P_{FA} = Q\left(\frac{d/2}{\sigma/\sqrt{2}}\right) = Q\left(\frac{d}{\sqrt{2}\,\sigma}\right),$$

where $Q(x) := \int_x^\infty \frac{e^{-\theta^2/2}}{\sqrt{2\pi}}\,d\theta$. In the real case, we would have $Q(d/(2\sigma))$.

The total prob. of error is

$$P_e = P(\text{decide } i=1 \mid H=0)\tfrac{1}{2} + P(\text{decide } 0 \mid H=1)\,1/2$$

By symmetry, $P_e = P_{FA}$.

Examples

1) Antipodal Signaling (BPSK), $s_1 = -s_0$. Then

$$d = \| s_1 - (-s_1) \| = 2\| s_1 \| = 2\sqrt{\mathcal{E}} \text{ if}$$

$$\mathcal{E} := \| s_1 \|^2 = \text{energy in } s_1$$

In the complex case,

$$P_e = Q\left(\sqrt{\frac{2\mathcal{E}}{\sigma^2}}\right).$$

In the real case

$$P_e = Q\left(\sqrt{\mathcal{E}/\sigma^2}\right).$$

2) Orthogonal Signaling. $s_1 \perp s_0$. Then

$$d^2 = \| s_1 - s_0 \|^2 = \| s_1 \|^2 - 2\,\mathrm{Re}\langle s_1, s_0 \rangle + \| s_0 \|^2$$
$$= \| s_1 \|^2 + \| s_0 \|^2.$$

For equal-energy signals,

$$d = \sqrt{2\mathcal{E}}.$$

In the complex case, $P_e = Q\left(\sqrt{\mathcal{E}/\sigma^2}\right)$.

In the real case, $P_e = Q\left(\sqrt{\frac{\mathcal{E}}{2\sigma^2}}\right)$.