

EE881 - Princípios de Comunicações I

Henrique Koji Miyamoto

1 Introdução e objetivos

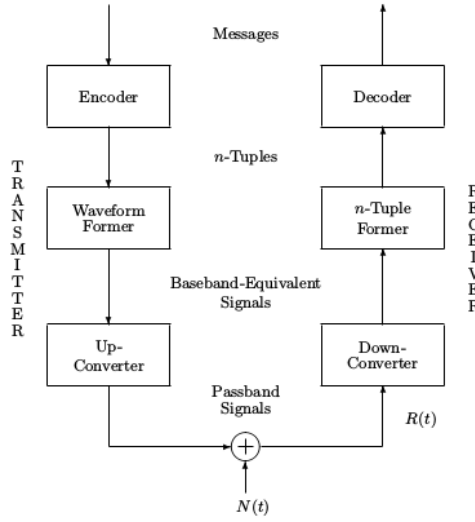


Figura 1: Visão geral do problema de comunicação.

2 Projeto do receptor para observações no tempo discreto: primeira camada

2.1 Introdução

O sistema de comunicação considerado nessa seção é composto pelas seguintes partes:

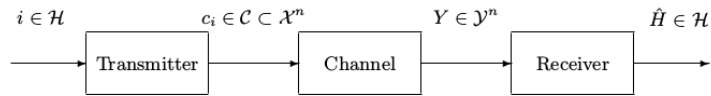


Figura 2: Sistema de comunicação considerado nessa análise.

- A **fonte** produz mensagens a serem transmitidas. É modelada como uma variável aleatória H que assume valores no conjunto de índices $\mathcal{H} = \{0, 1, \dots, m-1\}$, cada um com probabilidade $P_H(i)$.
- O **canal** é descrito por um alfabeto de entrada \mathcal{X} , um alfabeto de saída \mathcal{Y} (assumiremos $\mathcal{Y} = \mathbb{R}$) e uma descrição estatística da saída dada a entrada (verossimilhança). Se $Y \in \mathcal{Y}$ é uma variável aleatória discreta, descrevemos a distribuição de probabilidade $P_{Y|X}(\cdot|x)$; se é uma variável aleatória contínua, descrevemos a função densidade de probabilidade $f_{Y|X}(\cdot|x)$ ¹.
- O **transmissor** é modelado como uma aplicação do conjunto de mensagens $\mathcal{H} = \{0, \dots, m-1\}$ no conjunto de sinais (constelação de sinais) $\mathcal{C} = \{c_0, \dots, c_{m-1}\}$, em que $c_i \in \mathcal{X}^n$ para algum n .
- A tarefa do **receptor** é “adivinhar” a realização da hipótese H a partir da realização da saída do canal. O palpite do receptor é denotado \hat{i} e a variável aleatória associada a esse processo aleatório é $\hat{H} \in \mathcal{H}$.

¹Em geral, o canal não tem memória, i.e., a probabilidade da ocorrência de uma sequência de variáveis aleatórias na saída é o produto das probabilidades de ocorrência de cada variável.

2.2 Teste de hipóteses

Teste de hipóteses é o problema de “adivinhar” a realização da variável aleatória $H \in \mathcal{H}$ com base na realização da variável aleatória Y (observação). É o mesmo que *decodificação* ou *tomada de decisão*. Em comunicações, a hipótese H é a mensagem transmitida e a observação Y é a saída do canal.

Objetivo: Queremos estabelecer uma estratégia de decisão que maximiza a probabilidade de acerto $P_c = \Pr\{\hat{H} = H\}$, ou, equivalentemente, minimiza a probabilidade de erro $P_e = \Pr\{\hat{H} \neq H\} = 1 - P_c$.

Da regra de Bayes, temos

$$P_{H|Y}(i|y) = \frac{P_H(i)f_{Y|H}(y|i)}{f_Y(y)}$$

em que $f_Y(y) = \sum_i P_H(i)f_{Y|H}(y|i)$. Ao observar $Y = y$, a probabilidade de $H = i$ vai do *prior* (probabilidade *a priori*) $P_H(i)$ para o *posterior* (probabilidade *a posteriori*) $P_{H|Y}(i|y)$.

Regra MAP (máximo a posteriori): A decisão ótima para $\hat{H} = i$, que maximiza a probabilidade de acerto, é a que maximiza a probabilidade a posteriori $P_{H|Y}(i|y)^2$.

$$\hat{H}(y) = \arg \max_{i \in \mathcal{H}} P_{H|Y}(i|y)$$

A probabilidade de acerto da variável aleatória $P_{H|Y}(\hat{H}(Y)|Y)$ é

$$P_c = \mathcal{E}[P_{H|Y}(\hat{H}(Y)|Y)] = \int_y P_{H|Y}(\hat{H}(Y)|Y)f_Y(y)dy$$

Regra ML (máxima verossimilhança): É uma regra de decisão subótima que maximiza a verossimilhança $f_{Y|H}(y|i)$. Pode ser usada quando os *priors* são uniformes (neste caso, MAP \equiv ML) ou quando são desconhecidos.

$$\hat{H}(y) = \arg \max_{i \in \mathcal{H}} f_{Y|H}(y|i)$$

2.2.1 Teste de hipóteses binário

Trata-se do caso particular em que $\mathcal{H} = \{0, 1\}$.

Regra MAP:

$$P_{H|Y}(1|y) = \frac{f_{Y|H}(y|1)P_H(1)}{f_Y(y)} \stackrel{\hat{H}=1}{\underset{\hat{H}=0}{\geq}} \frac{f_{Y|H}(y|0)P_H(0)}{f_Y(y)} = P_{H|Y}(0|y)$$

$$\boxed{f_{Y|H}(y|1)P_H(1) \stackrel{\hat{H}=1}{\underset{\hat{H}=0}{\geq}} f_{Y|H}(y|0)P_H(0)}$$

$$\underbrace{\Lambda(y)}_{\text{razão de verossimilhança}} := \frac{f_{Y|H}(y|1)}{f_{Y|H}(y|0)} \stackrel{\hat{H}=1}{\underset{\hat{H}=0}{\geq}} \frac{P_H(0)}{P_H(1)} =: \underbrace{\eta}_{\text{limiar}}$$

Regra ML:

$$\boxed{f_{Y|H}(y|1) \stackrel{\hat{H}=1}{\underset{\hat{H}=0}{\geq}} f_{Y|H}(y|0)}$$

Uma função $\hat{H} : \mathcal{Y} \rightarrow \mathcal{H}$ é chamada *função de decisão* ou *função de decodificação*. Podemos descrevê-las em função de *regiões de decisão*:

$$\mathcal{R}_i = \{y \in \mathcal{Y} : \hat{H}(y) = i\}$$

Dessa forma, a probabilidade de erro pode ser computada:

$$P_e(0) = \Pr\{Y \in \mathcal{R}_1 | H = 0\} = \int_{\mathcal{R}_1} f_{Y|H}(y|0)dy$$

$$P_e(1) = \Pr\{Y \in \mathcal{R}_0 | H = 1\} = \int_{\mathcal{R}_0} f_{Y|H}(y|1)dy$$

$$\boxed{P_e = P_e(1)P_H(1) + P_e(0)P_H(0)}$$

²Se ocorre um empate, não faz diferença por qual hipótese escolhemos, já que elas têm a mesma probabilidade de erro.

2.2.2 Teste de hipóteses m -ário

Esse é o caso mais geral, em que $\mathcal{H} = \{0, 1, \dots, m-1\}$.

Regra MAP:

$$\hat{H}_{MAP}(y) = \arg \max_{i \in \mathcal{H}} P_{H|Y}(i|y) = \arg \max_{i \in \mathcal{H}} f_{Y|H}(y|i) P_H(i)$$

Regra ML:

$$\hat{H}_{ML}(y) = \arg \max_{i \in \mathcal{H}} f_{Y|H}(y|i)$$

Regiões de decodificação e probabilidade de erro:

$$\mathcal{R}_i := \{y : \hat{H}(y) = i\}$$

$$P_e(i) := \Pr\{\hat{H} \neq H | H = i\} = \int_{\mathcal{R}_i^c} f_{Y|H}(y|i) dy = 1 - \underbrace{\int_{\mathcal{R}_i} f_{Y|H}(y|i) dy}_{P_c(i)}$$

$$P_e := \Pr\{\hat{H} \neq H\} = \sum_{i=0}^{m-1} P_H(i) P_e(i)$$

2.3 A função Q

A função Q é definida como

$$Q(x) := \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{\xi^2}{2}} d\xi$$

Propriedade básica: Se Z é uma variável aleatória com distribuição normal de média nula e variância unitária, $Z \sim N(0, 1)$, então $\Pr\{Z \geq x\} = Q(x)$. Se Z tem média m e variância σ^2 , $Z \sim N(m, \sigma^2)$, então $\Pr\{Z \geq x\} = Q(\frac{x-m}{\sigma})$.

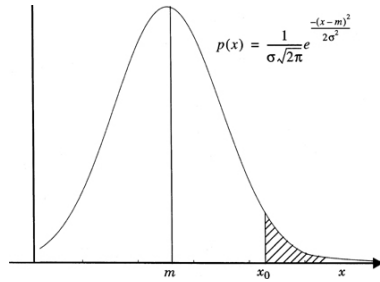


Figura 3: Distribuição normal de média m e variância σ^2 .

Outras propriedades da função Q :

- Se $Z \sim N(0, 1)$, $F_Z := \Pr\{Z \leq z\} = 1 - Q(z)$.
- $Q(0) = 1/2$, $Q(-\infty) = 1$, $Q(\infty) = 0$.
- $Q(-x) + Q(x) = 1$

2.4 Projeto de receptor para canal AWGN no tempo-discreto

Considere comunicação por um canal ruidoso no tempo-discreto. A hipótese $H \in \mathcal{H}$ representa uma mensagem gerada aleatoriamente. O transmissor mapeia $H = i$ em uma n -upla $c_i \in \mathbb{R}^n$. O canal adiciona um vetor ruído aleatório Z com média nula e componentes de distribuição gaussiana independentes e identicamente distribuídas de variância σ^2 , i.e., $Z \sim N(0, \sigma^2 I_n)$. A observação é $Y = c_i + Z$.

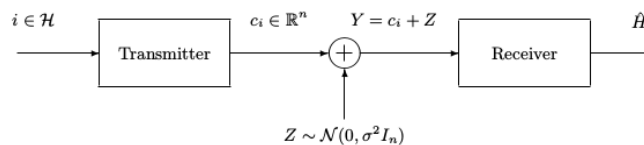


Figura 4: Diagrama de comunicação por um canal ruidoso.

2.4.1 Decisão binária para observação escalar

Seja a mensagem $H \in \{0, 1\}$ com um transmissor que mapeia $H = 0 \mapsto c_0 \in \mathbb{R}$ e $H = 1 \mapsto c_1 \in \mathbb{R}$.

$$\begin{cases} H = 0 : Y \sim N(c_0, \sigma^2) \rightarrow f_{Y|H}(y|0) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y-c_0)^2}{2\sigma^2} \right\} \\ H = 1 : Y \sim N(c_1, \sigma^2) \rightarrow f_{Y|H}(y|1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y-c_1)^2}{2\sigma^2} \right\} \end{cases}$$

Regra MAP:

$$\Lambda(y) = \frac{f_{Y|H}(y|1)}{f_{Y|H}(y|0)} \underset{\hat{H}=0}{\overset{\hat{H}=1}{\gtrless}} \eta \Rightarrow \ln \Lambda(y) = \frac{(y-c_0)^2 - (y-c_1)^2}{2\sigma^2} \underset{\hat{H}=0}{\overset{\hat{H}=1}{\gtrless}} \ln \eta \Rightarrow y \frac{c_1 - c_0}{\sigma^2} + \frac{c_0^2 - c_1^2}{2\sigma^2} \underset{\hat{H}=0}{\overset{\hat{H}=1}{\gtrless}} \ln \eta$$

Sem perda de generalidade, podemos assumir $c_1 > c_0$ e escrever:

$$\boxed{y \underset{\hat{H}=0}{\overset{\hat{H}=1}{\gtrless}} \frac{\sigma^2}{c_1 - c_0} \ln \eta + \frac{c_0 + c_1}{2} =: \theta}$$

Regra ML:

Se $P_H(0) = P_H(1) = \frac{1}{2}$, então MAP \equiv ML. A regra é de *distância mínima*:

$$\boxed{y \underset{\hat{H}=0}{\overset{\hat{H}=1}{\gtrless}} \frac{c_0 + c_1}{2}}$$

Probabilidade de erro

$$P_e = P_H(0)P_e(0) + P_H(1)P_e(1) = P_H(0)Q\left(\frac{\theta - c_0}{\sigma}\right) + P_H(1)Q\left(\frac{\theta - c_1}{\sigma}\right)$$

No caso $P_H(0) = P_H(1) = \frac{1}{2}$, temos $\frac{\theta - c_0}{\sigma} = \frac{c_1 - \theta}{\sigma}$, logo:

$$P_e = Q\left(\frac{c_1 - c_0}{2\sigma}\right) = Q\left(\frac{d}{2\sigma}\right)$$

2.4.2 Decisão binária para observação de n -uplas

Nesse caso, $H \in \{0, 1\}$ e $c_i \in \mathbb{R}^n$, $i = 0, 1$. O ruído adicionado pelo canal é $Z \sim N(0, \sigma^2 I_n)$.

$$\begin{cases} H = 0 : Y = c_0 + Z \sim N(c_0, \sigma^2 I_n) \rightarrow f_{Y|H}(y|0) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{\|y-c_0\|^2}{2\sigma^2} \right\} \\ H = 1 : Y = c_1 + Z \sim N(c_1, \sigma^2 I_n) \rightarrow f_{Y|H}(y|1) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{\|y-c_1\|^2}{2\sigma^2} \right\} \end{cases}$$

Regra MAP:

$$\Lambda(y) = \frac{f_{Y|H}(y|1)}{f_{Y|H}(y|0)} = \exp \left\{ \frac{\|y-c_0\|^2 - \|y-c_1\|^2}{2\sigma^2} \right\} \Rightarrow \ln \Lambda(y) = \langle y, \frac{c_1 - c_0}{\sigma^2} \rangle + \frac{\|c_0\|^2 - \|c_1\|^2}{2\sigma^2}$$

$$\ln \Lambda(y) \underset{\hat{H}=0}{\overset{\hat{H}=1}{\gtrless}} \ln \eta \Rightarrow \langle y, \frac{c_1 - c_0}{\sigma^2} \rangle + \frac{\|c_0\|^2 - \|c_1\|^2}{2\sigma^2} \underset{\hat{H}=0}{\overset{\hat{H}=1}{\gtrless}} \ln \eta \Rightarrow \underbrace{\langle y, \frac{c_1 - c_0}{d} \rangle}_{\psi} \underset{\hat{H}=0}{\overset{\hat{H}=1}{\gtrless}} \frac{\sigma^2}{d} \ln \eta + \frac{\|c_1\|^2 - \|c_0\|^2}{2d} =: \theta$$

As regiões de decisão \mathcal{R}_0 e \mathcal{R}_1 são delimitadas pelo plano afim $\{y \in \mathbb{R}^n : \langle y, \psi \rangle = \theta\}$.

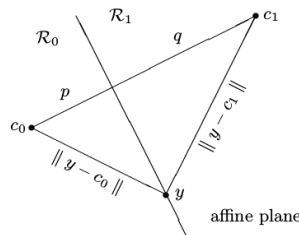


Figura 5: Visualização geométrica das regiões de decodificação.

É possível mostrar que as distâncias p e q (Figura 5) são tais que

$$p = \frac{d}{2} + \frac{\sigma^2 \ln \eta}{d} \text{ e } q = \frac{d}{2} - \frac{\sigma^2 \ln \eta}{d}$$

Regra ML: Se $P_H(0) = P_H(1) = \frac{1}{2}$, então MAP \equiv ML. A regra é de *distância mínima*:

$$\boxed{\|y - c_0\| \underset{\hat{H}=0}{\overset{\hat{H}=1}{\gtrless}} \|y - c_1\|}$$

Observações:

- Quando $P_H(0) = P_H(1) = \frac{1}{2}$, temos $p = q$ e o plano afim é equidistante de c_0 e c_1 .
- O vetor ψ não é afetado pelos priors, mas o limiar θ é, i.e., os priors afetam a posição, mas não a orientação do plano afim.
- O efeito anterior é aumentado quando a variância σ^2 aumenta.

Probabilidade de erro

Podemos deduzir as probabilidades de erro geometricamente a partir da Figura 5. Lembrando que $\langle Z, \psi \rangle \sim N(0, \sigma^2)$.

$$P_e(0) = Pr\{\langle Z, \psi \rangle \geq p\} = Q\left(\frac{p}{\sigma}\right)$$

$$P_e(1) = Pr\{\langle Z, \psi \rangle \geq q\} = Q\left(\frac{q}{\sigma}\right)$$

$$P_e = P_H(0) \left(\frac{p}{\sigma}\right) + P_H(1) \left(\frac{q}{\sigma}\right)$$

Para o caso $P_H(0) = P_H(1) = \frac{1}{2}$:

$$P_e = P_e(0) = P_e(1) = Pr\{\langle Z, \psi \rangle \geq \frac{d}{2}\} = Q\left(\frac{d}{2\sigma}\right)$$

2.5 Decisão m -ária para observação de n -uplas

Nesse caso, $H = i$, $i \in \{0, \dots, m-1\}$ e $c_i \in \mathbb{R}^n$. Assumiremos a simplificação $P_H(i) = \frac{1}{m}$.

Receptor ML:

$$\hat{H}_{ML} = \arg \max_{i \in \mathcal{H}} f_{Y|H}(y|i) = \arg \max_{i \in \mathcal{H}} \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{\|y - c_i\|^2}{2\sigma^2}\right\} = \boxed{\arg \min_{i \in \mathcal{H}} \|y - c_i\|}$$

Mais uma vez, a regra de decisão para o receptor ML é de distância mínima. A menos de empates, a região de decisão \mathcal{R}_i é a região de Voronoi³.

Referências

- [1] RIMOLDI, Bixio. *Principles of Digital Communication: A Top-Down Approach*. Cambridge: Cambridge University Press, 2016.

³A região de Voronoi do ponto $c_i \in \mathbb{R}^n$ é o conjunto $R(c_i) := \{x \in \mathbb{R}^n : \|c_i - x\| \leq \|c_i - y\| \forall y \in \mathbb{R}^n\}$.