Théorie de l'information et codage **2010/2011**

Cours 8 — 5 avril 2011

Enseignant: Marc Lelarge *Scribe: Charles Fougeron*

Pour information

– Page web du cours
`http://www.di.ens.fr/~lelarge/info11.html`

8.1 Le théorème de codage source-canal

Un tel système est caractérisé par :
- une v.a. $\underline{U}$ modélisant la source ;
- une fonction d'encodage modélisé par $p(\underline{X}|\underline{U})$;
- une probabilité de transition pour le canal $p(\underline{Y}|\underline{X})$;
- une fonction de décodage $p(\underline{V}|\underline{Y})$.

$$\begin{aligned} \textbf{Coût moyen} : \ \overline{\beta} &= \frac{1}{n}E[b(\underline{X})] \\ \textbf{Distorsion moyenne} : \ \overline{\delta} &= \frac{1}{k}E[d(\underline{U},\underline{V})] \\ \textbf{Taux de transmission} : \ \overline{r} &= \frac{k}{n} \end{aligned}$$

Le but est d'avoir des $\overline{\beta}$ et $\overline{\delta}$ aussi petits que possible tandis que $\overline{r}$ est aussi grand que possible.

Théorème 8.1.1 *Pour une source et un canal donné :*

(a) Les paramètres $\overline{\beta}$, $\overline{\delta}$, $\overline{r}$ doivent satisfaire $\overline{r} \leq \frac{C(\overline{\beta})}{R(\overline{\delta})}$

(b) Inversement étant donné $\beta > \beta_{min}$, $\delta > \delta_{min}$ et $r < \frac{C(\beta)}{R(\delta)}$, il est possible de construire une système tel que décrit ci-dessus avec $\overline{\beta} \leq \beta$, $\overline{\delta} \leq \delta$ et $\overline{r} \geq r$

Démonstration.

(a) La suite $(\underline{U}, \underline{X}, \underline{Y}, \underline{V})$ constitue une chaîne de Markov, donc $I(\underline{U};\underline{V}) \leq I(\underline{X};\underline{Y})$. De plus $E[b(\underline{X})] = n\overline{\beta}$ implique que $I(\underline{X};\underline{Y})) \leq C_n(\overline{\beta})$ et comme $C_n(\overline{\beta}) \leq nC(\overline{\beta})$, on a au final : $I(\underline{X};\underline{Y}) \leq nC(\overline{\beta})$. De plus $E[d(\underline{U},\underline{V})] = k\overline{\delta}$ implies that $I(\underline{U};\underline{(V)}) \geq kR(\overline{\delta})$ et donc le point (a) est vérifié.

(b) On vérifie que l'on peut choisir :

$$\begin{aligned} \beta_{min} \leq \beta_0 < \beta, \\ \delta_{min} \leq \delta_0 < \delta_1 < \delta \\ C' < C(\beta_0), \\ R' > R(\delta_0), \\ r < \frac{C'}{R'}. \end{aligned}$$

Théorème de codage de source

Pour k_0 suffisamment grand, il existe un code source C de longueur k_0 avec M_1 mots code tel que $M_1 \leq 2^{\lfloor R' k_0 \rfloor}$ et $d(C) = \frac{1}{k_0} E[d_{min}(\underline{U})] < \delta_1$, avec $d_{min}(\underline{U}) = min\{d(\underline{U}, \underline{V_i}), \underline{V_i} \in C\}$. Pour m défini plus tard, on note $k = k_0 m$. L'encodeur de source partitionne $\underline{U} = (U_1, \ldots, U_k)$ en m blocs de longueur k_0 et émet m mots code de source correspondant à ces blocs, $\underline{W} = (W_1, \ldots, W_k)$ est une suite de m mots code de C, le nombre de valeurs distinctes possibles pour $\underline{W}$ est inférieur ou égal à $M_1^m \leq 2^{k_0 m R'}$.

Encodeur de canal

On définit la distorsion pire cas du code C pour $\underline{u} \in \mathcal{U}^{k_0}$ par $d_{max}(\underline{u}) = max\{ d(\underline{u}, \underline{v_i}), \underline{v_i} \in C \}$, et la distorsion pire cas du code par $D(C) = \frac{1}{k_0} E[d_{max}(\underline{U})]$ où l'espérance est par rapport a la statistique de la source.
On note $\epsilon = \frac{\delta - \delta_1}{D(C)}$ et pour chaque $m = 1, 2, \ldots$ soit $n_m = \lceil m k_0 R' / c' \rceil$.

Théorème de codage de canal

Pour m suffisament grand, il existe un code $\{\underline{x}_1, \ldots, \underline{x}_{M_2}\}$ de longueur n_m et une règle de décodage tel que :

$$\begin{aligned} b(\underline{X}_i) &\leq n_m \beta \\ M_2 &\geq 2^{\lceil C' n_m \rceil} \geq 2^{m k_0 R'} \\ P_E^{(i)} &< \epsilon. \end{aligned}$$

On suppose de plus que m est suffisament grand pour que $\overline{r} = \frac{k}{n_m} = \frac{k_0 m}{\lceil m k_0 \frac{R'}{C'} \rceil} \geq r$, ce qui est possiblie puisque $r < \frac{C'}{R'}$. On a également $n\overline{\beta} = E[b(\underline{X})] \leq n\beta$.
Il reste donc à prouver que $\overline{\delta} \leq \delta$.
Le décodeur de canal est celui du théorème de codage de canal, le décodeur de source est celui dont la sortie $\underline{V} = (V_1, \ldots, V_k)$ de m mots code de source si possible, sinon

erreur.

$$B = \begin{cases} 0 & \text{si } \underline{Z} = \underline{X} \\ 1 & \text{sinon.} \end{cases}$$

$$E[d(\underline{U}, \underline{V})] = E[d(\underline{U}, \underline{V}) \mid B = 0\ ].P(B = 0) + E[d(\underline{U}, \underline{V}) \mid B = 1].P(B = 1)$$

$$\text{Si } B = 0,\ d(\underline{U}, \underline{V}) = \sum_{l=0}^{m} d_{min}(\underline{U}^l) \text{ avec } \underline{U} = (U^1, \ldots, U^m) \text{ blocs de taille } k_0,$$

$$\text{alors } E[d(\underline{U}, \underline{V}) \mid B = 0] = m.E[d_{min}(\underline{U})] < k_0 \delta_1 m$$

$$\text{Si } B = 1,\ E[d(\underline{U}, \underline{V}) \mid B = 1] \leq m.E[d_{max}(\underline{U}) \mid B = 1]$$

$$E[d_{max}(\underline{U}) \mid B = 1] = \sum_{i=1}^{M_2} E[d_{max}(\underline{U}) \mid B = 1, \underline{X} = \underline{X}_i].P[\underline{X} = \underline{X}_i \mid B = 1]$$

$$= \sum_{i=1}^{M_2} E[d_{max}(\underline{U}) \mid \underline{X} = \underline{X}_i].P_E^{(i)}.\frac{P(\underline{X} = \underline{X}_i)}{P(B = 1)}$$

$$E[d_{max}(U)|B = 1].P(B = 1) \leq m\epsilon.E[d_{max}(\underline{U})] = k_0 m(\delta - \delta_1) = k(\delta - \delta_1)$$

Au final, on a $\overline{\delta} = k^{-1} E\left[d(\underline{U}; \underline{V})\right] < \delta$.

□

8.2 Codes linéaires

8.2.1 Introduction

$\mathcal{X} = F_q = \{0, 1, \ldots, q-1\}$ avec q un nombre premier

Définition 8.2.1 *Un (n, k) code linéaire sur F_q est un sous-espace de dimention k de F_q^n, n est la longueur du code, k = dimension du code, taux = k/n.*

Un code est décrit par k mots code linéairement indépendant $\underline{x}_1, \ldots, \underline{x}_k$.
Chaque mot code et l'un des q^k combinaison linéaire

$$\sum_{i=1}^{k} \alpha_i \underline{x}_i, \quad \alpha_i \in F_q$$

Définition 8.2.2 *Soit C un (n, k) code linéaire sur F_q. Une matrice G dont l'espace engendré par les lignes est C est une matrice générant C.*

Exemple 8.2.1:
(5, 1) code C_1, $G = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}$

Exemple 8.2.2:
(7, 4) code

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

est le code de Hamming (7, 4)

Encodage (n, k) code linéaire a q^k mots code message $\underline{u} = (u_1, \ldots, u_k) \in F_q^k$ mot code $\underline{x} = \underline{u}G$.

Matrice de parité / parity check matrix
équation de parité $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ qui est satisfaite pour tout $\underline{x} \in C$ l'espace des vecteurs $\underline{a} = (a_1, \ldots, a_n)$ est l'espace $C^\perp$ appelé code dual de C.
$C^\perp$ a pour dimension $n - k$, $C^\perp$ est un (n, n-k) code. Une matrice de parité pour C est une matrice générant $C^\perp$

Définition 8.2.3 *C un (n,k) code linéaire sur F_q, une matrice H avec la propriété, $H\underline{x}^t = 0 \Leftrightarrow \underline{x}$ inC est appellée matrice de parité de C*

8.2.2 Géométrie de Hamming et performance du code

La distance de Hamming est $d_H(\underline{x}, \underline{y}) = w_H(\underline{x} - \underline{y})$
Soit $C = \{\underline{x}_1, \ldots, \underline{x}_M\}$ un code de longueur n.
On veut que C soit capable de corriger les erreurs de poids $\leq e$.
On envoie $\underline{x}_i$, et $\underline{y} = \underline{x}_i + \underline{z}$ est reçu ; si $w_H(\underline{z}) \leq e$ alors $\hat{\underline{x}}_i = \underline{x}$. Le code est capable de corriger les erreurs de poids $\leq e$ ssi la distance entre chaque pair de mots code est supérieur ou égale à $2e + 1$, i.e. $d_H(\underline{x}_i, \underline{x}_j) \geq 2e + 1 \ \forall i, j$
Soit $d_{min}(C) = min(d_H(\underline{x}, \underline{x}', x \neq x' \ \underline{x}, \underline{x}' \in C))$ la distance minimale du code C.

Théorème 8.2.1 *Un code $C = \{\underline{x}_1, \ldots, \underline{x}_M\}$ est capable de corriger toutes les erreurs de poids $\leq e$ ssi $d_{min}(C) \geq 2e + 1$*

Pour un code linéaire, $d_H(\underline{x}, \underline{x}') = w_H(\underline{x} - \underline{x}')$ et $\underline{x} - \underline{x}' \in C$
donc $d_{min}(C) = min\{w_H(\underline{x}), \underline{x} \in C, \underline{x} \neq 0\}$

Théorème 8.2.2 *Si C est un (n,k) code linéaire sur F_q, de matrice de parité H, $d_{min}(C)$ = le plus petite nombre de colonnes de H qui sont linéairement dépendantes.*

Démonstration.
$H\underline{x}^t = 0 = x_1\underline{c_1} + x_2\underline{c_2} + \cdots + x_n\underline{c_n}$ $\mathcal{Y} = F_q$. □

$\underline{x}$ est transmis, $\underline{y}$ est reçu, et on note $\underline{z} = \underline{y} - \underline{x}$ le motif d'erreur

Si $z_i \neq 0$ il y a erreur sur le $i^{\text{ème}}$ bit
$\underline{s} = H\underline{y}^t = H\underline{z}^t$ est appelé le syndrome et ne dépend que du motif d'erreur.
L'ensemble des solution en $\underline{z}$ de $H\underline{z}^t = \underline{s}$ forme un coset du code C
Notons $C + \underline{z}_0 = \{\underline{x} + \underline{z_0}, \underline{x} \in C\}$
q^{n-k} est le coset de C correspondants aux q^{n-k} syndromes possibles. Chaque coset contient alors q^k elements