Syndrome

Definition Syndrom

Sei $C \subseteq \mathbb{F}_2^n$ ein Code mit Parity Check Matrix P und $\mathbf{x} \in \mathbb{F}_2^n$. Das Syndrom von $\mathbf{x}$ ist definiert als $S(\mathbf{x}) = \mathbf{x}P^t$.

Satz Standardarrays und Syndrome

Sei C ein linearer Code mit Standardarray A und Parity Check Matrix P. Die Elemente $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$ sind in derselben Zeile von A gdw $S(\mathbf{x}) = S(\mathbf{y})$.

Beweis:

- Sei $\mathbf{x} = \mathbf{f_i} + \mathbf{c_j}$ und $\mathbf{y} = \mathbf{f_k} + \mathbf{c}_\ell$.
- Es gilt $S(\mathbf{x}) = S(\mathbf{f_i} + \mathbf{c_j}) = S(\mathbf{f_i}) + S(\mathbf{c_j}) = S(\mathbf{f_i})$.
- Analog folgt $S(\mathbf{y}) = S(\mathbf{f_k})$. D.h.

$$\begin{aligned} S(\mathbf{y}) = S(\mathbf{x}) &\Leftrightarrow S(\mathbf{f_i}) = S(\mathbf{f_k}) \\ &\Leftrightarrow S(\mathbf{f_i} - \mathbf{f_k}) = \mathbf{0} \Leftrightarrow \mathbf{f_i} - \mathbf{f_k} \in C \Leftrightarrow i = k. \end{aligned}$$

Syndromdekodierung mittels Syndromtabelle

- Dekodierung mittels Standardarray: $\mathbf{x} = \mathbf{f_i} + \mathbf{c_j}$ mit Fehlervektor $\mathbf{f_i}$.
- Paarweise verschiedene Fehlervektoren bilden die erste Spalte eines Standardarrays.
- Berechne die folgende Syndromtabelle für C

Fehlervektor	Syndrom
$\mathbf{0}$	$\mathbf{0}$
$\mathbf{f_2}$	$S(\mathbf{f_2})$
$\mathbf{f_3}$	$S(\mathbf{f_3})$
$\vdots$	$\vdots$
$\mathbf{f}_\ell$	$S(\mathbf{f}_\ell)$

Algorithmus Syndromdekodierung

EINGABE: $\mathbf{x} \in \mathbb{F}_2^n$

1. Berechne $S(\mathbf{x})$ und vergleiche mit der Syndromspalte.
2. Falls $S(\mathbf{x}) = S(\mathbf{f_i})$, Ausgabe $\mathbf{c} = \mathbf{x} - \mathbf{f_i}$.

Äquivalente lineare Codes

Definition Äquivalenz von linearen Codes

Sei C ein linearer Code mit Generatormatrix G. Ein Code C' mit Generatormatrix G' heißt zu C *äquivalenter Code*, falls G' eine Transformation aus G mittels folgender Operationen ist.

1. Vertauschen von zwei Zeilenvektoren
2. Vertauschen von zwei Spaltenvektoren
3. Addition eines Zeilenvektors zu einem anderen Zeilenvektor

Fakt Systematische Codes

Sei C ein linearer $[n,k]$-Code mit Generatormatrix G. Dann gibt es einen zu C äquivalenten Code C' mit Generatormatrix in linker Standardform $G' = [I_k | M_{k,n-k}]$. C' nennt man *systematischen Code*.

- Für systematische C': $(x_1, \ldots, x_k)G' = (x_1, \ldots, x_k, y_1, \ldots, y_{n-k})$.
- $y_1, \ldots, y_{n-k}$ nennt man die Redundanz der Nachricht.

Umwandlung Generatormatrix in Parity Check Matrix

Satz Konversion von Generatormatrix in Parity Check Matrix

Sei C ein linearer $[n,k]$-Code mit Generatormatrix $G = [I_k|A]$. Dann ist

$$P = [A^t|I_{n-k}]$$

eine Parity Check Matrix für C.

Beweis: Sei C' der Code mit Parity Check Matrix P:

1. Zeigen: $C \subseteq C'$.
   - Für alle Zeilen $\mathbf{g_i}$ von G gilt $P\mathbf{g_i}^t = \mathbf{0}^t$, denn j-ter Eintrag von $P\mathbf{g_i}^t$:
   $$(a_{1j} \dots a_{kj} 0 \dots 1 \dots 0) \cdot (0 \dots 1 \dots 0 a_{i1} \dots a_{in-k}) = a_{ij} + a_{ij} = 0$$
   - Aus $P\mathbf{g_i}^t = \mathbf{0}^t$ folgt $C \subseteq C'$.
2. Zeigen: $\dim(C) = \dim(C')$
   - P besitzt $n-k$ linear unabhängige Zeilen.
   - D.h. Dualcode $(C')^\perp$ hat Generatormatrix P und Dimension $n-k$.
   $$\dim(C') = n - \dim((C')^\perp) = n - (n-k) = k = \dim(C).$$

Hamming-Matrix $H(h)$ und Hammingcode $\mathcal{H}(h)$

- Parametrisiert über die Zeilenanzahl h.
- Spaltenvektoren sind Binärdarstellung von $1, 2, \ldots, 2^h - 1$.
- Bsp :

$$H(3) = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

- Hammingcode $\mathcal{H}(h)$ besitzt die Parity Check Matrix $H(h)$.
- Hammingcodes unabhängig entdeckt von Golay (1949) und Hamming (1950).

Satz Hammingcode

Der Hammingcode $\mathcal{H}(h)$ mit Parity Check Matrix $H(h)$ ist ein linearer $[n, k, d]$-Code mit den Parametern

$$n = 2^h - 1, k = n - h \text{ und } d = 3.$$

k und d bei Hammingcodes

Beweis:

- $H(h)$ enthält die h Einheits-Spaltenvektoren $\mathbf{e_1}, \ldots, \mathbf{e_h}$.
- Daraus folgt, die Zeilenvektoren von $H(h)$ sind linear unabhängig.
- D.h. $H(h)$ ist eine Generatormatrix des dualen Codes $\mathcal{H}(h)^{\perp}$.
- Damit ist $\dim(\mathcal{H}(h)^{\perp}) = h$ und $k = \dim(\mathcal{H}(h)) = n - h$.

- Je zwei Spalten in $H(h)$ sind paarweise verschieden.
- Die minimale Anzahl von linear abhängigen Spalten ist mindestens 3, d.h. $d(\mathcal{H}(h)) \geq 3$.
- Die ersten drei Spalten sind stets linear abhängig, d.h. $d(\mathcal{H}(h)) = 3$.

Dekodierung mit Hammingcodes

Satz Korrigieren eines Fehlers

Sei $\mathbf{c} \in \mathcal{H}(h)$ und $\mathbf{x} = \mathbf{c} + \mathbf{e_i}$ für einen Einheitsvektor $\mathbf{e_i} \in \mathbb{F}_2^{2^h-1}$. Dann entspricht das Syndrom $S(\mathbf{x})$ der Binärdarstellung von i.

Beweis:

- Es gilt $S(\mathbf{x}) = S(\mathbf{e_i}) = \mathbf{e_i} H(h)^t = (H(h)\mathbf{e_i}^t)^t$.
- D.h. $S(\mathbf{x})$ entspricht der i-ten Spalte von $H(h)$, die wiederum die Binärkodierung von i ist.

Bsp:

- Verwenden $\mathcal{H}(3)$ und erhalten $\mathbf{x} = 1000001$.

$$S(\mathbf{x}) = (1000001)H(3)^t = (110).$$

- Da 110 die Binärkodierung von 6 ist, kodieren wir zum nächsten Nachbarn 1000011.

Simplex Code: Dualcode des Hammingcodes

Satz Simplex Code

Der Dualcode des Hammingcodes $\mathcal{H}(h)$ wird als Simplex Code $\mathcal{S}(h)$ bezeichnet. $\mathcal{S}(h)$ ist ein $[2^h - 1, h, 2^{h-1}]$-Code, bei dem für *alle verschiedenen* $\mathbf{c}, \mathbf{c}' \in \mathcal{S}(h)$ gilt, dass $d(\mathbf{c}, \mathbf{c}') = 2^{h-1}$.

Beweis:

- Hamming-Matrix $H(h)$ ist Generatormatrix von $\mathcal{S}(h) = \mathcal{H}(h)^{\perp}$.
- Da $\dim(\mathcal{S}(h)) = n - \dim(\mathcal{H}(h))$, ist $S(h)$ ein $[2^h - 1, h]$-Code.
- Rekursive Definition: Es gilt

$$H(h+1) = \left(\begin{array}{ccc|c|ccc} 0 & \dots & 0 & 1 & 1 & \dots & 1 \\ \hline & & & 0 & & & \\ & H(h) & & \vdots & & H(h) & \\ & & & 0 & & & \end{array}\right).$$

- Sei $\bar{\mathbf{c}}$ das Komplement von $\mathbf{c}$ ist. Dann gilt

$$\mathcal{S}(h+1) = \{\mathbf{c}0\mathbf{c} | \mathbf{c} \in \mathcal{S}(h)\} \cup \{\mathbf{c}1\bar{\mathbf{c}} | \mathbf{c} \in \mathcal{S}(h)\}.$$

Distanz 2^{h-1} zwischen zwei Worten im Simplex Code

Beweis von $d(\mathbf{c}, \mathbf{c}') = 2^{h-1}$ per Induktion über h.

IV $h = 1$:

- $H(1) = (1)$, d.h. $\mathcal{S} = \{0, 1\}$ und damit $d(0, 1) = 1 = 2^0$.

IS $h \to h + 1$:

- Fall 1: $d(\mathbf{c}0\mathbf{c}, \mathbf{c}'0\mathbf{c}') = 2 \cdot d(\mathbf{c}, \mathbf{c}') = 2 \cdot 2^{h-1} = 2^h$.
- Fall 2: $d(\mathbf{c}1\bar{\mathbf{c}}, \mathbf{c}'1\bar{\mathbf{c}}') = d(\mathbf{c}, \mathbf{c}') + d(\bar{\mathbf{c}}, \bar{\mathbf{c}}') = 2 \cdot d(\mathbf{c}, \mathbf{c}') = 2^h$.
- Fall 3:

$$\begin{aligned} d(\mathbf{c}0\mathbf{c}, \mathbf{c}'1\bar{\mathbf{c}}') &= d(\mathbf{c}, \mathbf{c}') + 1 + d(\mathbf{c}, \bar{\mathbf{c}}') \\ &= d(\mathbf{c}, \mathbf{c}') + 1 + (2^h - 1 - d(\mathbf{c}, \mathbf{c}')) = 2^h. \end{aligned}$$

Reed-Muller Codes

- Reed-Muller Code $\mathcal{R}(r, m)$ ist definiert für $m \in \mathbb{N}$, $0 \leq r \leq m$.
- Betrachten nur Reed-Muller Codes 1. Ordnung $\mathcal{R}(1, m) = \mathcal{R}(m)$.

Definition Rekursive Darstellung von Reed-Muller Codes

1. $\mathcal{R}(1) = \mathbb{F}_2^2 = \{00, 01, 10, 11\}$.
2. Für $m \geq 1$: $\mathcal{R}(m+1) = \{\mathbf{cc} \mid \mathbf{c} \in \mathcal{R}(m)\} \cup \{\mathbf{c}\bar{\mathbf{c}} \mid \mathbf{c} \in \mathcal{R}(m)\}$.

- $R_1 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ ist eine Generatormatrix für $\mathcal{R}(1)$.
- $\mathcal{R}(2) = \{0000, 0011, 0101, 0110, 1010, 1001, 1111, 1100\}$ mit Generatormatrix

$$R_2 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Parameter der Reed-Muller Codes

Satz Reed-Muller Parameter

$\mathcal{R}(m)$ ist ein linearer $(2^m, 2^{m+1}, 2^{m-1})$-Code. Für alle $\mathbf{c} \in \mathcal{R}(m) \setminus \{\mathbf{0}, \mathbf{1}\}$ gilt $w(\mathbf{c}) = 2^{m-1}$.

IA: $m = 1$

- $\mathcal{R}(1)$ ist ein linearer $(2^1, 2^2, 2^0)$-Code. $01, 10$ besitzen Gewicht 2^0.

IS: $m \to m + 1$

- $n = 2 \cdot 2^m = 2^{m+1}$.
- $\{\mathbf{cc} \mid \mathbf{c} \in \mathcal{R}(m)\}$ und $\{\mathbf{c\bar{c}} \mid \mathbf{c} \in \mathcal{R}(m)\}$ sind disjunkt, d.h.
$$k = 2 \cdot 2^{m+1} = 2^{m+2}.$$
- Sei $\mathbf{c} \in \mathcal{R}(m) \setminus \{\mathbf{0}, \mathbf{1}\}$.
  - Für $\mathbf{cc}$ gilt $w(\mathbf{cc}) = 2w(\mathbf{c}) = 2 \cdot 2^{m-1} = 2^m$.
  - Für $\mathbf{c\bar{c}}$ gilt $w(\mathbf{c\bar{c}}) = w(\mathbf{c}) + w(\bar{\mathbf{c}}) = 2^{m-1} + (2^m - 2^{m-1}) = 2^m$.
- Für $\mathbf{c} = \mathbf{0}$ gilt $\mathbf{c\bar{c}} = \mathbf{01}$ mit $w(\mathbf{01}) = 2^m$.
- Für $\mathbf{c} = \mathbf{1}$ gilt $\mathbf{c\bar{c}} = \mathbf{10}$ mit $w(\mathbf{10}) = 2^m$.

Reed-Muller Generatormatrizen

Satz Generatormatrix für $\mathcal{R}(m)$

Sei R_m eine Generatormatrix für $\mathcal{R}(m)$. Dann ist

$$R_{m+1} = \left(\begin{array}{ccc|ccc} 0 & \dots & 0 & 1 & \dots & 1 \\ \hline & R_m & & & R_m & \end{array}\right)$$

eine Generatormatrix für $\mathcal{R}(m+1)$.

Beweis:

- **Ann.**: $\exists$ nicht-triviale Linearkombination, die **0** liefert.
- Linearkombination kann nicht nur die erste Zeile enthalten.
- D.h. es gibt eine nicht-triviale Linearkombination der Zeilen $2 \dots m+2$, die den Nullvektor auf der ersten Hälfte liefert. (Widerspruch: R_m ist Generatormatrix für $\mathcal{R}(m)$.)
- Sei C der Code mit Generatormatrix R_{m+1}.
- Für $\mathbf{c} \in \mathcal{R}(m)$ gilt: $\mathbf{cc} \in C$ und $\mathbf{c}\bar{\mathbf{c}} \in C$. D.h. $\mathcal{R}(m+1) \subseteq C$.
- $\dim(C) = m+1 = \dim(\mathcal{R}(m+1))$ und damit $C = \mathcal{R}(m+1)$.

Charakterisierung der Generatormatrizen

Bsp:

$$R_3 = \left(\begin{array}{cccc|cccc} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right)$$

Streiche Einserzeile aus R_m. Dann

- besitzen die Spaltenvektoren Länge m und
- bestehen aus Binärkodierungen von $0, 1, \ldots, 2^m - 1$.
- D.h. Streichung der Einserzeile von R_m liefert die Hamming-Matrix $H(m)$ mit einer zusätzlichen Nullspalte.

Vergleich von Hamming, Simplex und Reed-Muller Codes

	$\mathcal{H}(m)$	$\mathcal{S}(m)$	$\mathcal{R}(m)$
Codewortlänge	$2^m - 1$	$2^m - 1$	2^m
Anzahl Codeworte	2^{2^m-1-m}	2^m	2^{m+1}
Distanz	3	2^{m-1}	2^{m-1}

Dekodierung von Reed-Muller Codes

- $\mathcal{R}(m)$ kann $\left\lfloor \frac{2^{m-1}-1}{2} \right\rfloor = 2^{m-2} - 1$ Fehler korrigieren.
- Syndrom-Tabelle besitzt $\frac{2^n}{M} = \frac{2^{2^m}}{2^{m+1}} = 2^{2^m-m-1}$ Zeilen.

Bsp: $\mathcal{R}(3)$ ist 1-fehlerkorrigierend.

$$R_3 = \begin{pmatrix} \mathbf{r_1} \\ \mathbf{r_2} \\ \mathbf{r_3} \\ \mathbf{r_4} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Sei $\mathbf{c} = \alpha_1\mathbf{r_1} + \alpha_2\mathbf{r_2} + \alpha_3\mathbf{r_3} + \alpha_4\mathbf{r_4}$. Es gilt

- $c_1+c_5 = \alpha_1(r_{11}+r_{15})+\alpha_2(r_{21}+r_{25})+\alpha_3(r_{31}+r_{35})+\alpha_4(r_{41}+r_{45}) = \alpha_1$
- $c_2+c_6 = \alpha_1(r_{12}+r_{16})+\alpha_2(r_{22}+r_{26})+\alpha_3(r_{32}+r_{36})+\alpha_4(r_{42}+r_{46}) = \alpha_1$
- Ebenso $\alpha_1 = c_3 + c_7 = c_4 + c_8$.

Mehrheitsdekodierung

- Suche für jede Zeile i Spaltenpaar (u, v), so dass sich die Spalten u, v nur in der i-ten Zeile unterscheiden. Liefert Gleichung für α_i.
- Für Zeile 1: $(1,5), (2,6), (3,7), (4,8)$, d.h. im Abstand 4.
- Für Zeile 2: $(1,3), (2,4), (5,7), (6,8)$, d.h. im Abstand 2.
- Für Zeile 3: $(1,2), (3,4), (5,6), (7,8)$, d.h. im Abstand 1.
- Für Zeile 4: nicht möglich.

- Erhalten für $\alpha_1, \alpha_2, \alpha_3$ jeweils 4 Gleichungen in verschiedenen c_i.
- Falls $\mathbf{x} = \mathbf{c} + \mathbf{e_i}$, ist genau 1 von 4 Gleichungen inkorrekt.

Algorithmus Mehrheitsdekodierung Reed-Muller Code $\mathcal{R}(m)$

1. Bestimme $\alpha_1, \ldots, \alpha_m$ per Mehrheitsentscheid.
2. Berechne $\mathbf{e} = \mathbf{x} - \sum_{i=1}^{m} \alpha_i \mathbf{r_i}$.
3. Falls $w(\mathbf{e}) \leq 2^{m-2} - 1$, dekodiere $\mathbf{c} = \mathbf{x} + \mathbf{e}$. (d.h. $\alpha_{m+1} = 0$)
4. Falls $w(\bar{\mathbf{e}}) \leq 2^{m-2} - 1$, dekodiere $\mathbf{c} = \mathbf{x} + \bar{\mathbf{e}}$. (d.h. $\alpha_{m+1} = 1$)

Beispiel Mehrheitsdekodierung

Bsp:

- Verwenden $\mathcal{R}(3)$ und erhalten $\mathbf{x} = 11011100$.
  - $\alpha_1 = x_1 + x_5 = 0$
  - $\alpha_1 = x_2 + x_6 = 0$
  - $\alpha_1 = x_3 + x_7 = 0$
  - $\alpha_1 = x_4 + x_8 = 1$
- Mehrheitsentscheid liefert $\alpha_1 = 0$.
  - $\alpha_2 = x_1 + x_3 = 1$
  - $\alpha_2 = x_2 + x_4 = 0$
  - $\alpha_2 = x_5 + x_7 = 1$
  - $\alpha_2 = x_6 + x_8 = 1$
- Mehrheitsentscheid liefert $\alpha_2 = 1$ und analog $\alpha_3 = 0$.
- $\mathbf{e} = \mathbf{x} - 0 \cdot \mathbf{r_1} - 1 \cdot \mathbf{r_2} - 0 \cdot \mathbf{r_3} = 11011100 - 00110011 = 11101111$.
- $w(\bar{\mathbf{e}}) \leq 1$, d.h. $\mathbf{c} = \mathbf{x} + \bar{\mathbf{e}} = 11001100$.

McEliece Verfahren (1978)

- **Dekodieren eines zufälligen linearen Codes ist NP-hart**.
- Verwende linearen Code C mit effizientem Dekodierverfahren (z.B. sogenannten Goppa-Code).
- Generatormatrix von C bildet den geheimen Schlüssel.
- C wird in äquivalenten linearen Code C' transformiert.

Algorithmus Schlüsselgenerierung McEliece

1. Wähle linearen $[n, k, d]$-Code C mit Generatormatrix G.
2. Wähle zufällige binäre $(k \times k)$-Matrix S mit $\det(S) = 1$.
3. Wähle zufällige binäre $(n \times n)$-Permutationsmatrix P.
4. $G' \leftarrow SGP$

öffentlicher Schlüssel: G', geheimer Schlüssel S, G, P.

McEliece Verschlüsselung

Algorithmus McEliece Verschlüsselung

EINGABE: Plaintext $\mathbf{m} \in \mathbb{F}_2^k$

1. Wähle zufälligen Fehlervektor $\mathbf{e} \in \mathbb{F}_2^n$ mit $w(\mathbf{e}) = \lfloor \frac{d-1}{2} \rfloor$.
2. $\mathbf{c} \leftarrow \mathbf{m}G' + \mathbf{e}$.

AUSGABE: Ciphertext $\mathbf{c} \in \mathbb{F}_2^n$

Vorgeschlagene Parameter:

- $[1024, 512, 101]$-Goppacode C.
- Plaintextlänge: 512 Bit, Chiffretextlänge: 1024 Bit.
- Größe des öffentlichen Schlüssels: 512×1024 Bit.

McEliece Entschlüsselung

Algorithmus McEliece Entschlüsselung

EINGABE: Ciphertext $\mathbf{c} \in \mathbb{F}_2^n$

1. $\mathbf{x} \leftarrow \mathbf{c}P^{-1}$.
2. Dekodiere $\mathbf{x}$ mittels Dekodieralgorithmus für C zu $\mathbf{m}'$.
3. $\mathbf{m} \leftarrow \mathbf{m}'S^{-1}$

AUSGABE: Plaintext $\mathbf{m} \in \mathbb{F}_2^k$

- **Korrektheit**:

$$\mathbf{x} = \mathbf{c}P^{-1} = (\mathbf{m}G' + \mathbf{e}) \cdot P^{-1} = (\mathbf{m}SGP + \mathbf{e}) \cdot P^{-1} = (\mathbf{m}S)G + \mathbf{e} \cdot P^{-1}.$$

- $\mathbf{e} \cdot P^{-1}$ besitzt Gewicht $w(\mathbf{e}P^{-1}) = w(\mathbf{e}) = \lfloor \frac{d-1}{2} \rfloor$.
- Dekodierung liefert $\mathbf{m}' = \mathbf{m}S$, d.h. $\mathbf{m} = \mathbf{m}'S^{-1}$.