Perfektní lineární kódy

Odpřednesenou látku naleznete v dodatku I skript *Abstraktní a konkrétní lineární algebra*.

Minulé přednášky

1. Detekce a oprava chyb v lineárních kódech.
2. Hammingův $(7, 4)$-kód.

Dnešní přednáška

1. Kódy perfektní pro opravu t chyb.
2. Obecné Hammingovy kódy.
3. Jako aplikaci teorie kódů vyřešíme Hat Problem (nepovinné).

Další možné doplňující informace

1. J. Adámek, *Foundations of coding*, John Wiley & Sons, 1991
2. D. J. C. MacKay, *Information theory, inference and learning algorithms*, Cambridge Univ. Press, 2003

Definice (perfektní kód pro t chyb)

Řekneme, že lineární kód W délky n nad $\mathbb{Z}_p$ je perfektní pro t chyb, pokud pro každé $\mathbf{v}$ ze $(\mathbb{Z}_p)^n$ existuje právě jedno $\mathbf{w}$ z W tak, že $d_H(\mathbf{w}, \mathbf{v}) \leq t$.

Poznámky

1. Kód W je perfektní pro t chyb právě tehdy, když platí
$$(\mathbb{Z}_p)^n = \bigcup_{\mathbf{w} \in W} \mathrm{Ball}_t(\mathbf{w})$$
kde
$$\mathrm{Ball}_t(\mathbf{w}) = \{\mathbf{v} \in (\mathbb{Z}_p)^n \mid d_H(\mathbf{w}, \mathbf{v}) \leq t\}$$
je koule v $(\mathbb{Z}_p)^n$ se středem ve $\mathbf{w}$ a poloměrem t. Toto sjednocení je navíc disjunktní.[a]
Je-li W perfektní pro t chyb, pak $\mathrm{dist}(W) = 2t + 1$. Takže W opravuje t chyb.

[a] Slogan: kód W je perfektní pro t chyb právě tehdy, když $(\mathbb{Z}_p)^n$ lze pokrýt disjunktními koulemi se středy v kódových slovech a poloměrem t.

Poznámky (pokrač.)

2. Počet prvků jedné koule

$$\begin{aligned}\mathrm{Ball}_t(\mathbf{w}) &= \{\mathbf{v} \mid d_H(\mathbf{w},\mathbf{v}) = 0\} \cup \{\mathbf{v} \mid d_H(\mathbf{w},\mathbf{v}) = 1\} \cup \ldots \\ &\quad \cdots \cup \{\mathbf{v} \mid d_H(\mathbf{w},\mathbf{v}) = t-1\} \cup \{\mathbf{v} \mid d_H(\mathbf{w},\mathbf{v}) = t\}\end{aligned}$$

je roven součtu

$$\sum_{i=0}^{t} \underbrace{\binom{n}{i}}_{\substack{\text{výběr}\\ \text{lišících}\\ \text{se posic}}} \cdot \overbrace{(p-1)^i}^{\substack{\text{výběr}\\ \text{lišících}\\ \text{se znaků}}}, \quad \text{kde } \binom{n}{i} = \frac{n!}{i! \cdot (n-i)!}.$$

Poznámky (pokrač.)

3. Kód W je perfektní pro t chyb právě tehdy, když platí rovnost

$$p^n = (\text{počet slov ve } W) \cdot \sum_{i=0}^{t} \binom{n}{i} \cdot (p-1)^i$$

Víme-li navíc, že W má dimensi k, lze tuto rovnost psát jako

$$p^{n-k} = \sum_{i=0}^{t} \binom{n}{i} \cdot (p-1)^i$$

protože takové W obsahuje přesně p^k slov.

4. Pro obecný lineární kód W dimense k opravující t chyb platí pouze nerovnost

$$p^{n-k} \geq \sum_{i=0}^{t} \binom{n}{i} \cdot (p-1)^i$$

které se říká sphere-packing bound.

Poznámky

1. Klasifikace perfektních kódů nad $\mathbb{Z}_p$ existuje, jde však o těžký a hluboký výsledek. Viz například knihu
   - W. C. Huffman a V. Pless, *Fundamentals of error-correcting codes*, Cambridge Univ. Press, 2003
2. Uvidíme příklad třídy kódů, perfektních pro 1 chybu. Jde o třídu takzvaných Hammingových kódů.
3. Vesmírný program NASA používá perfektní kódy pro přenos fotografií z družic.

   Například sondy Voyager 1 a Voyager 2 používaly pro přenos fotografií Jupitera a Saturnu perfektní kód (Golay (24,12,8) code).

Definice (Hammingův kód)

Hammingův kód je kód nad $\mathbb{Z}_2$ délky $n = 2^m - 1$ s kontrolní maticí $\mathbf{H}^T$, která má m řádků a sloupce matice $\mathbf{H}^T$ přesně odpovídají binárním zápisům všech čísel $1, \ldots, 2^m - 1$, kde $m \geq 1$ je přirozené číslo.

Příklady

1. Kontrolní matice Hammingova kódu pro $m = 1$:
$$\mathbf{H}^T = (1)$$
Délka tohoto kódu je 1, dimense 0.[a] Inf. rate $= \frac{0}{1} = 0$.
2. Kontrolní matice Hammingova kódu pro $m = 2$:
$$\mathbf{H}^T = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$
Délka tohoto kódu je 3, dimense 1. Inf. rate $= \frac{1}{3} = 0.33\ldots$

[a] Jde tedy o triviální kód $(\mathbb{Z}_2)^0 = \{\mathbf{o}\}$, který není příliš použitelný.

Příklady (pokrač.)

3. Kontrolní matice Hammingova $(7,4)$-kódu (zde je $m = 3$):

$$\mathbf{H}^T = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

   Délka tohoto kódu je 7, dimense 4. Inf. rate $= \frac{4}{7} = 0.57\ldots$

4. Kontrolní matice Hammingova kódu pro $m = 4$:

$$\mathbf{H}^T = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

   Délka tohoto kódu je 15, dimense 11. Inf. rate $= \frac{11}{15} = 0.73\ldots$

5. Pro obecné $m \geq 1$ je délka příslušného Hammingova kódu $2^m - 1$, dimense je rovna $2^m - 1 - m$ a inf. rate je roven

$$\frac{2^m - 1 - m}{2^m - 1} = 1 - \frac{m}{2^m - 1}$$

Tvrzení

Ať W je Hammingův kód délky $n = 2^m - 1$, $m \geq 2$. Potom W je perfektní pro 1 chybu.

Důkaz.

Chceme ukázat rovnost

$$p^{n-k} = \sum_{i=0}^{t} \binom{n}{i} \cdot (p-1)^i$$

kde $p = 2$, $t = 1$, $n = 2^m - 1$ a $k = 2^m - 1 - m$.
Opravdu, platí:

$$\underbrace{\binom{2^m - 1}{0}}_{=1} + \underbrace{\binom{2^m - 1}{1}}_{=2^m-1} = 1 + 2^m - 1 = 2^m = 2^{n-k}$$

The Hat Problem (Todd Ebert, 1998)

Skupina vězňů hraje následující hru o svobodu:

1. Každý vězeň dostane buď černý nebo bílý klobouk (klobouky se rozdávájí náhodně s pravděpodobností 1/2).
2. Každý vidí barvy klobouků ostatních, barvu svého klobouku nevidí nikdo.
3. Skupina hraje jako tým. Vyhrají, pokud alespoň jeden uhodne správně barvu svého klobouku a nikdo ze skupiny nehádá špatně.
4. Před začátkem hry se vězni na strategii domlouvat mohou, po začátku hry spolu komunikovat nesmí.

Simple-minded strategie: hádejme náhodně, pravděpodobnost výhry je pak 1/2.

Existuje stategie lepší?

Ano: je-li $m \geq 2$, pak pro skupinu $n = 2^m - 1$ vězňů pomůže příslušný Hammingův kód.

Optimální vyhrávací strategie pro Hat Problem

Označme jako W Hammingův kód délky $n = 2^m - 1$, kde $m \geq 2$. Slova $\mathbf{v}$ ze $(\mathbb{Z}_2)^n$ budeme považovat za distribuci klobouků:

1. 0 v i-té položce slova $\mathbf{v}$ znamená: vězeň i má černý klobouk.
2. 1 v i-té položce slova $\mathbf{v}$ znamená: vězeň i má bílý klobouk.

Pro zadanou distribuci $\mathbf{v}$ definujme slovo $\mathbf{v}_i$ jako distribuci, která má shodné položky se slovem $\mathbf{v}$, kromě i-té položky, kde je 0.[a] Platí rovnost $\mathbf{v} = \mathbf{v}_i + a_i\mathbf{e}_i$, kde $a_i \in \mathbb{Z}_2$ je pevné.

Strategie i-tého vězně:

1. Jesliže $\mathbf{v}_i + b_i\mathbf{e}_i \notin W$ pro jakékoli b_i ze $\mathbb{Z}_2$, vězeň i mlčí.
2. Jesliže $\mathbf{v}_i + b_i\mathbf{e}_i \in W$ pro nějaké b_i ze $\mathbb{Z}_2$, vězeň i prohlásí, že má klobouk barvy $1 + b_i$.

[a] Slovo $\mathbf{v}_i$ je tedy definováno jako distribuce, kterou i-tý vězeň skutečně vidí a který předpokládá, že má černý klobouk.

Optimální vyhrávací strategie pro Hat Problem (pokrač.)

Strategie je dobře definovaná: nemůže současně platit $\mathbf{v}_i + 0 \cdot \mathbf{e}_i \in W$ a $\mathbf{v}_i + 1 \cdot \mathbf{e}_i \in W$.

Kdyby platilo $\mathbf{v}_i + 0 \cdot \mathbf{e}_i \in W$ a $\mathbf{v}_i + 1 \cdot \mathbf{e}_i \in W$, pak součet $(\mathbf{v}_i + 0 \cdot \mathbf{e}_i) + (\mathbf{v}_i + 1 \cdot \mathbf{e}_i) = \mathbf{e}_i$ leží ve W. To není možné, protože syndrom $\mathbf{e}_i$ je i-tý sloupec kontrolní matice $\mathbf{H}^T$ a ten je nenulový.[a]

1. Jestliže $\mathbf{v}$ není ve W, strategie dává vítězství.

   Pokud $\mathbf{v}$ není ve W, existuje jediné j tak, že $\mathbf{v} + \mathbf{e}_j$ je ve W. Hammingův kód je totiž perfektní pro 1 chybu. Vězeň j tedy správně uhodl barvu svého klobouku. Navíc všichni ostatní vězni museli mlčet: pro $i \neq j$ by v opačném případě muselo platit $\mathbf{v}_i + b_i \mathbf{e}_i \in W$ pro nějaké b_i.

[a] Připomenutí: kontrolní matice Hammingova kódu má ve sloupcích binární zápisy nenulových čísel $1, \ldots, 2^m - 1$.

Optimální vyhrávací strategie pro Hat Problem (pokrač.)

Protože $\mathbf{v} + \mathbf{e}_j = \mathbf{v}_i + a_i\mathbf{e}_i + \mathbf{e}_j \in W$ a $\mathbf{v}_i + b_i\mathbf{e}_i \in W$, platí $(\mathbf{v}_i + a_i\mathbf{e}_i + \mathbf{e}_j) + (\mathbf{v}_i + b_i\mathbf{e}_i) = (a_i + b_i)\mathbf{e}_i + \mathbf{e}_j \in W$.

Syndrom slova $(a_i + b_i)\mathbf{e}_i + \mathbf{e}_j$ je ale nenulový, to je spor.

2. Jestliže $\mathbf{v}$ je ve W, každý hádá špatně.

   Opravdu: pro každé i platí $\mathbf{v} = \mathbf{v}_i + a_i\mathbf{e}_i \in W$ a i-tý vězeň tedy prohlásí, že má klobouk barvy $1 + a_i$, ačkoli má ve skutečnosti klobouk barvy a_i.

Optimální vyhrávací strategie pro Hat Problem (pokrač.)

To znamená, že pravděpodobnost výhry při této strategii je přesně

$$1 - \frac{\text{počet slov ve } W}{\text{počet slov v } (\mathbb{Z}_2)^n} = 1 - \frac{2^{2^m-1-m}}{2^{2^m-1}} = 1 - \frac{1}{2^m}$$

1. Například pro $m = 2$ (tedy pro skupinu $n = 2^2 - 1 = 3$ vězňů) je pravděpodobnost výhry $3/4 = 0.75$.[a]
2. Pro $m = 3$ (tj. pro $n = 2^3 - 1 = 7$ vězňů) je pravděpodobnost výhry $7/8 = 0.875$.
3. Pro $m = 4$ (tj. pro $n = 2^4 - 1 = 15$ vězňů) je pravděpodobnost výhry $15/16 = 0.9375$.
4. Pro $m = 5$ (tj. pro $n = 2^5 - 1 = 31$ vězňů) je pravděpodobnost výhry $31/32 = 0.96875$.
5. Atd.

To je vždy lepší výsledek než simple-minded strategie (která dává vždy pravděpodobnost 1/2).

[a] Je užitečným cvičením si optimální strategii pro $m = 2$ vyzkoušet.

Důležité upozornění

V teorii lineárních kódů je zvykem psát vektory z $\mathbb{F}^n$ do řádku (na rozdíl od B6B01LAG). Co tím ztrácíme a co tím získáváme?

1. Vycvičeni dosavadním průběhem této přednášky, ztrácíme okamžitý geometrický přehled o tom, co se při kódování skutečně děje.
   Pro zájemce: ve skutečnosti geometrický přehled neztrácíme; pracujeme jen s kovektory místo s vektory, viz kapitolu 3.5 skript.
   To znamená, že kódování má jasnou geometrickou interpretaci v duálním prostoru.
2. Získáváme kompatibilitu s rozsáhlou literaturou o kódování.

Protože nám šlo jen o velmi krátký úvod do lineárních kódů, psali jsme vektory nadále do sloupců. Kdo bude číst jinou literaturu z kódování, bude velmi pravděpodobně muset všechny matice a maticové rovnice z teorie kódů transponovat.

Která lineární algebra je tedy ta „správná“?

1. Psaní vektorů z $\mathbb{F}^n$ do sloupců nám umožnilo chápat součin $\mathbf{A} \cdot \mathbf{x}$ jako funkční hodnotu lineárního zobrazení $\mathbf{A}$ v bodě $\mathbf{x}$. Chápání součinu $\mathbf{A} \cdot \mathbf{x}$ jako funkční hodnoty vedlo k přirozené geometrické interpretaci maticových výpočtů.

   To je ve shodě s tím, jak značíme funkční hodnoty ve zbytku matematiky: značku $f(x)$ chápeme jako funkční hodnotu funkce f v bodě x.

2. Při psaní vektorů z $\mathbb{F}^n$ do řádku bychom museli hodnotu lineárního zobrazení $\mathbf{A}$ v bodě $\mathbf{x}$ značit $\mathbf{x} \cdot \mathbf{A}$.

   Tento způsob uvažování o maticovém součinu je ve shodě s (menšinovým) názorem, že funkční hodnotu funkce f v bodě x bychom měli značit $(x)f$. Takovému značení funkčních hodnot se říká reverse Polish notation (RPN).

Proč jsme v přednášce zvolili „sloupcovou“ lineární algebru?

1. Protože na RPN nejsme zvyklí, zvolili jsme „sloupcovou“ lineární algebru. Je totiž ve shodě s tím, jak uvažujeme ve zbytku matematiky.
2. „Řádková“ lineární algebra navíc není ve svém značení úplně důsledná. Dochází tak k absurditám:[a] například soustava rovnic
$$\left(\begin{array}{cc|c} 2 & -1 & 3 \\ 4 & 7 & 11 \end{array}\right)$$
ze „sloupcové“ lineární algebry by se v „řádkové“ lineární algebře měla správně zapisovat
$$\left(\begin{array}{cc} 2 & 4 \\ -1 & 7 \\ \hline 3 & 11 \end{array}\right)$$
ale neděje se tak. „Řádková“ lineární algebra pro soustavy rovnic přebírá zápis „sloupcové“ lineární algebry!

[a] Vzpomeňte si, kolikrát se v textech z „řádkové“ lineární algebry objevuje rčení: „*. . . jednotlivé vektory nyní napíšeme do sloupců matice. . .*“

Závěrečné poznámky

1. Poslední čtyři přednášky byly pouze nahlédnutím do teorie kódů.
2. Skutečné studium teorie kódů vyžaduje zvládnutí dalších partií matematiky:
   1. Teorie informace.
   2. Teorie pravděpodobnosti.
   3. Teorie grup.
   4. A dalších...

   Více se lze dozvědět v doporučené (tj. nepovinné) literatuře.