Lineární kódy

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

Dnešní přednáška

1. Analýza Hammingova $(7,4)$-kódu.
2. Lineární kódy nad $\mathbb{Z}_p$.
3. Oprava a detekce chyb.
4. Syndrome decoding a nearest neighbour decoding.

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
3. W. C. Huffman a V. Pless, *Fundamentals of error-correcting codes*, Cambridge Univ. Press, 2003

Připomenutí (minulé přednášky):

Teorie lineárních prostorů nad obecným tělesem $\mathbb{F}$ byla vybudována.

Příkladem těles jsou $\mathbb{Z}_p$, kde p je prvočíslo.

To znamená:

1. Umíme určit bázi a dimensi podprostorů W lineárního prostoru $(\mathbb{Z}_p)^n$.
2. Umíme pracovat s maticemi nad $\mathbb{Z}_p$ a řešit soustavy lineárních rovnic nad $\mathbb{Z}_p$.

Navíc:

3. V prostoru $(\mathbb{Z}_p)^n$ rozumíme relaci ortogonality a Hammingově vzdálenosti.
4. V prostoru $(\mathbb{Z}_p)^n$ rozumíme generujícím a kontrolním maticím lineárních podprostorů.

Příklad (Hammingův $(7,4)$-kód)

V $(\mathbb{Z}_2)^7$ zvolme lineární podprostor W dimense 4 s bází danou vektory

$$\mathbf{g}_1 = \begin{pmatrix}1\\0\\0\\0\\0\\1\\1\end{pmatrix}, \mathbf{g}_2 = \begin{pmatrix}0\\1\\0\\0\\1\\0\\1\end{pmatrix}, \mathbf{g}_3 = \begin{pmatrix}0\\0\\1\\0\\1\\1\\0\end{pmatrix}, \mathbf{g}_4 = \begin{pmatrix}0\\0\\0\\1\\1\\1\\1\end{pmatrix}$$

Generující matice $\mathbf{G}$ podprostoru W je

$$\mathbf{G} = (\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3, \mathbf{g}_4) = \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\0&1&1&1\\1&0&1&1\\1&1&0&1\end{pmatrix}$$

Příklad (Hammingův $(7,4)$-kód, pokrač.)

1. $\mathrm{rank}(\mathbf{G}) = 4$, tudíž máme k disposici 4 info bity.
2. Dimense ortogonálního doplňku $7 - 4 = 3$. Informace bude chráněna třemi bity (redundance).
3. Posílání zpráv: informace $\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ vytváří kódové slovo

$$\mathbf{G} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Pozorování: $\mathbf{G}$ je v blokovém tvaru $\left(\begin{array}{c} \mathbf{E}_4 \\ \hline \mathbf{B} \end{array}\right)$.[a]

[a] Takovým kódům se říká systematické: jsou v nich jasně odděleny informační a ochranné bity.

Příklad (Hammingův $(7,4)$-kód, pokrač.)

Kontrolní (Hammingova) matice

$$\mathbf{H}^T = \begin{pmatrix} \mathbf{h}_1^T \\ \mathbf{h}_2^T \\ \mathbf{h}_3^T \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Jde o ideální kód, pokud došlo nejvýše k jedné chybě:[a]

1. Odesláno $\mathbf{w}$, přijmeme $\mathbf{v}$ a předpokládáme, že došlo k nejvýše jedné chybě. Tj. $\mathbf{v} = \mathbf{w} + \mathbf{e}$ ($\mathbf{e}$ je error pattern). Víme, že $\mathbf{e}$ obsahuje nejvýše jednu jedničku.
2. Spočteme syndrom $\mathbf{s}$ slova $\mathbf{v}$: $\mathbf{s} = \mathbf{H}^T\mathbf{v} = \mathbf{H}^T\mathbf{e}$.
   1. Jestliže $\mathbf{s} = \mathbf{o}$, při přenosu nedošlo k chybě, tj. $\mathbf{e} = \mathbf{o}$, neboli $\mathbf{v} = \mathbf{w}$.
   2. Jestliže $\mathbf{s}$ je i-tý sloupec $\mathbf{H}^T$, je $\mathbf{e} = \mathbf{e}_i$. Došlo k chybě na i-tém místě. Opravíme ji: $\mathbf{w} = \mathbf{v} - \mathbf{e}_i$.
3. Isolujeme info bity.

[a] Jde dokonce o příklad tzv perfektního kódu pro opravu jedné chyby, viz příští přednášku.

Příklad (Hammingův $(7,4)$-kód, pokrač.)

Co se stane, pokud error pattern $\mathbf{e}$ obsahuje dvě jedničky? Tj., co nastane, pokud při přenosu došlo k právě dvěma chybám?

Spočteme syndrom $\mathbf{s}$ slova $\mathbf{v}$: $\mathbf{s} = \mathbf{H}^T\mathbf{v} = \mathbf{H}^T\mathbf{e}$.

Pokud jsou jedničky v $\mathbf{e}$ na místech i a j, je $\mathbf{H}^T\mathbf{e}$ součet i-tého a j-tého sloupce matice $\mathbf{H}^T$.

Připomeňme, že

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

Dvě chyby na první a druhé posici současně tedy nerozlišíme od jedné chyby na pouze třetí posici.

Problém návrhu lineárního kódu

Jak vyvážit následující požadavky? Chceme co největší opravné schopnosti kódu a co nejmenší počet kontrolních bitů.

Tyto požadavky jsou intuitivně protichůdné.

V plné obecnosti se nyní budeme věnovat dvěma tématům

1. Způsoby dekódování lineárních kódů.
2. Opravné a ochranné schopnosti lineárních kódů.

Obě témata mají jasnou geometrickou interpretaci.

Definice (lineární kód)

Ať p je prvočíslo. Lineární p-kód délky n a dimense k je lineární podprostor W prostoru $(\mathbb{Z}_p)^n$, $\dim(W) = k$, $0 \leq k \leq n$.
Terminologie:

1. Prvkům $(\mathbb{Z}_p)^n$ říkáme také slova, prvkům W kódová slova.[a]
2. Generující matice $\mathbf{G} : (\mathbb{Z}_p)^k \to (\mathbb{Z}_p)^n$ kódu W je (jakákoli) generující matice podprostoru W.
3. Vektoru $\mathbf{w} = \mathbf{G} \cdot \mathbf{a}$ říkáme kódové slovo určené vektorem informace $\mathbf{a}$ ze $(\mathbb{Z}_p)^k$.
4. Kontrolní matice $\mathbf{H}^T : (\mathbb{Z}_p)^n \to (\mathbb{Z}_p)^{n-k}$ je (jakákoli) kontrolní matice podprostoru W.
5. Součinu $\mathbf{s} = \mathbf{H}^T \cdot \mathbf{v}$ říkáme syndrom slova $\mathbf{v}$.

[a] Slova a kódová slova jsou vektory v $(\mathbb{Z}_p)^n$, píšeme je tedy do sloupce. Poznámky ke značení v jiné literatuře uvedeme na příští přednášce.

Geometrie error patternu a jeho syndromu

Ať W je lineární p-kód délky n a dimense k s generující maticí $\mathbf{G} = (\mathbf{g}_1, \ldots, \mathbf{g}_k)$ a kontrolní maticí $\mathbf{H}^T$. Zvolme jakékoli slovo $\mathbf{e}$ ze $(\mathbb{Z}_p)^n$ (tzv. error pattern) a označme $\mathbf{H}^T \cdot \mathbf{e} = \mathbf{s}$ syndrom slova $\mathbf{e}$.

Z lineární algebry okamžitě plyne:[a]

$\mathbf{e} + W$ je k-dimensionální plocha v $(\mathbb{Z}_p)^n$.

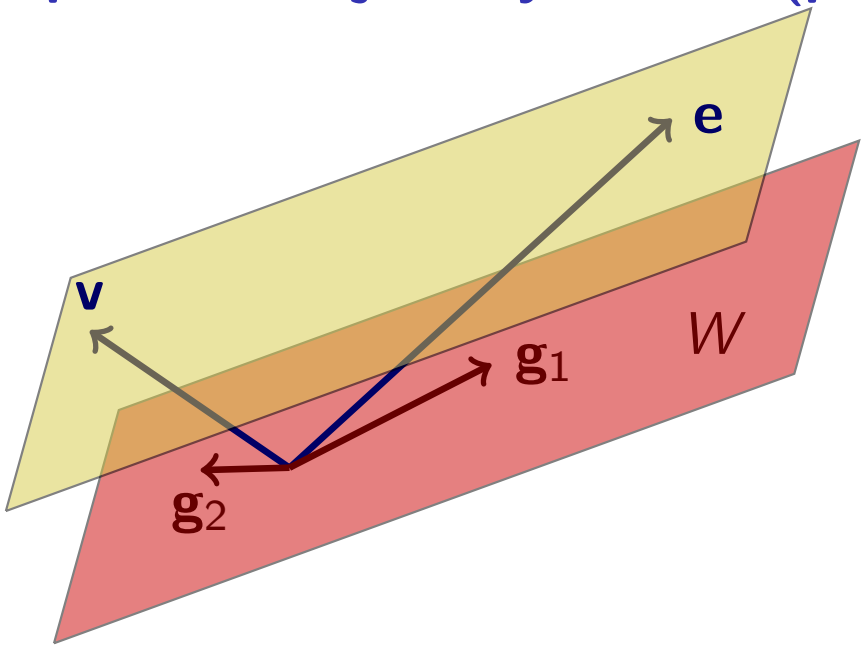
Tato plocha prochází bodem $\mathbf{e}$ a má směr $(\mathbf{g}_1, \ldots, \mathbf{g}_k)$.

Platí totiž, že $\mathbf{e} + W$ je přesně množina řešení soustavy $(\mathbf{H}^T \mid \mathbf{s})$.

Jiný pohled na totéž: $\mathbf{e} + W$ je přesně množina všech slov $\mathbf{v}$ ze $(\mathbb{Z}_p)^n$ se syndromem $\mathbf{s} = \mathbf{H}^T \cdot \mathbf{e}$.

[a] Připomeňte si přednášku 6B.

Geometrie error patternu a jeho syndromu (pokrač.)



k-dimensionální plocha $\mathbf{e} + W$ v $(\mathbb{Z}_p)^n$ je přesně množina všech slov $\mathbf{v}$ ze $(\mathbb{Z}_p)^n$ se syndromem $\mathbf{s}$.

Dekódovací strategie: přijmeme-li slovo $\mathbf{v}$, stačí nalézt $\mathbf{e}$ tak, aby $\mathbf{v}$ bylo v $\mathbf{e} + W$. Potom bylo odesláno slovo $\mathbf{w} = \mathbf{v} - \mathbf{e}$.

Problém této strategie: je pro $\mathbf{v}$ error pattern $\mathbf{e}$ určen jednoznačně? Není! Existuje ale „přirozená“ volba: ať $\mathbf{e}$ je „co nejblíže“ $\mathbf{o}$.

V úvahách o dekódování, opravách a detekci chyb budou hrát roli následující pojmy:

Definice (Hammingova váha slova a min. distance kódu)

1. Hammingova váha $w_H(\mathbf{v}) = d_H(\mathbf{v}, \mathbf{o})$ slova $\mathbf{v}$. Zjevně platí: $w_H(\mathbf{v}) =$ počet nenulových položek slova $\mathbf{v}$.
2. Minimální (Hammingova) distance kódu W
   $\mathrm{dist}(W) = \min\{w_H(\mathbf{w}) \mid \mathbf{w}$ je nenulové slovo ve $W\}$.

Poznámka (jiný vzorec pro minimální distanci kódu)

Platí: $\mathrm{dist}(W) = \min\{d_H(\mathbf{w}, \mathbf{w}') \mid \mathbf{w}, \mathbf{w}'$ jsou různá slova z $W\}$.

Opravdu: pro různá slova $\mathbf{w}$, $\mathbf{w}'$ z W je $\mathbf{w} - \mathbf{w}'$ nenulové slovo z W a platí $d_H(\mathbf{w}, \mathbf{w}') = w_H(\mathbf{w} - \mathbf{w}')$. Obráceně, pro nenulové slovo $\mathbf{w}$ z W je $w_H(\mathbf{w}) = d_H(\mathbf{w}, \mathbf{o})$ Hammingova vzdálenost dvou různých slov ve W.

Tvrzení

Ať $\mathbf{H}^T$ je kontrolní matice kódu W. Pro kladné přirozené číslo d jsou následující podmínky ekvivalentní:

1. Kód W má minimální distanci d.
2. Každých $d-1$ sloupců matice $\mathbf{H}^T$ je lineárně nezávislých a některých d sloupců matice $\mathbf{H}^T$ je lineárně závislých.

Důkaz.

Kód W obsahuje slovo $\mathbf{w}$ váhy $w > 0$ iff $\mathbf{H}^T \cdot \mathbf{w} = \mathbf{o}$ iff $w > 0$ sloupců $\mathbf{H}^T$ je lineárně závislých. ■

Důsledek (Singletonův odhad)

Ať W je kód délky n a dimense k. Potom $\mathrm{dist}(W) \leq n - k + 1$.

Důkaz.

Pro kontrolní matici $\mathbf{H}^T$ kódu W platí $\mathrm{rank}(\mathbf{H}^T) = n - k$. Tudíž $\mathrm{dist}(W) - 1 \leq n - k$. ■

Příklad (minimální distance Hammingova $(7,4)$-kódu)

Hammingův $(7,4)$-kód má kontrolní matici

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

Označme jako d minimální distanci tohoto kódu.

Singletonův odhad dává: $d - 1 \leq 7 - 4 + 1$, čili $d \leq 5$.

Ve skutečnosti platí $d = 3$.

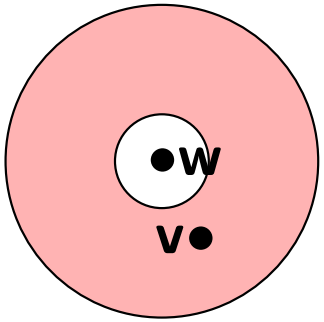
Proč? Například první tři sloupce matice $\mathbf{H}^T$ jsou lineárně závislé, jakákoli dvojice sloupců matice $\mathbf{H}^T$ je lineárně nezávislá.

Definice (detekce a oprava chyb)

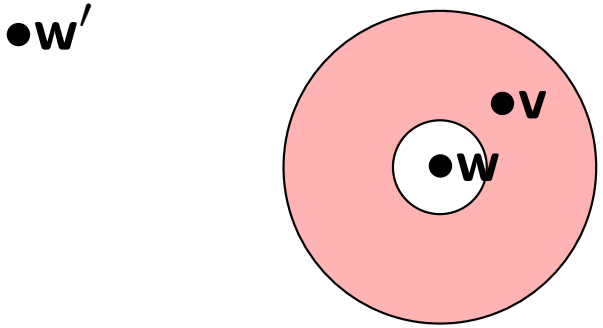
Ať W je lineární kód. Řekneme, že

1. W detekuje t chyb, pokud pro každé $\mathbf{w}$ ve W a každé $\mathbf{v}$ takové, že $1 \leq d_H(\mathbf{w}, \mathbf{v}) \leq t$, platí: $\mathbf{v}$ není ve W.
2. W opravuje t chyb, pokud pro každé $\mathbf{w}$ ve W a každé $\mathbf{v}$ takové, že $1 \leq d_H(\mathbf{w}, \mathbf{v}) \leq t$, platí: $d_H(\mathbf{w}, \mathbf{v}) < d_H(\mathbf{w}', \mathbf{v})$ pro všechna $\mathbf{w}'$ z W různá od $\mathbf{w}$.

Geometrie detekce a opravy chyb (slogany)



Detekce: slova nedaleko od kódového slova nejsou ve W.



Oprava: slova nedaleko od kódového slova jsou daleko od jiných slov z W.

Tvrzení

Kód W detekuje t chyb právě tehdy, když $\mathrm{dist}(W) > t$.

Důkaz.

1. Ať $\mathrm{dist}(W) \leq t$. Zvolme kódová slova $\mathbf{w}$, $\mathbf{w}'$ tak, že $d_H(\mathbf{w}, \mathbf{w}') = \mathrm{dist}(W)$. Potom $1 \leq d_H(\mathbf{w}, \mathbf{w}') \leq t$.

   To znamená, že nelze detekovat následujících t chyb: odesláno slovo $\mathbf{w}$, přijato slovo $\mathbf{w}'$.

2. Ať $\mathrm{dist}(W) > t$. Zvolme $\mathbf{w}$ ve W a $\mathbf{v}$ takové, že platí $1 \leq d_H(\mathbf{w}, \mathbf{v}) \leq t$.

   Potom $d_H(\mathbf{w}, \mathbf{v}) < \mathrm{dist}(W)$, takže $\mathbf{v}$ nemůže být kódové slovo.

■

Tvrzení

Kód W opravuje t chyb právě tehdy, když $\mathrm{dist}(W) > 2t$.

Důkaz.

1. Ať $\mathrm{dist}(W) \leq 2t$. Zvolme kódová slova $\mathbf{w}$, $\mathbf{w}'$ tak, že $d_H(\mathbf{w}, \mathbf{w}') = \mathrm{dist}(W)$ a označme jako $i_1, \ldots, i_r$ indexy položek, ve kterých se $\mathbf{w}$ a $\mathbf{w}'$ liší.
   Definujme $\mathbf{v}$ jako slovo, které má stejné položky jako $\mathbf{w}$, kromě položek $i_2, i_4, \ldots$, na kterých má stejné položky jako $\mathbf{w}'$.

   Potom $1 \leq d_H(\mathbf{w}, \mathbf{v}) \leq \frac{r}{2} \leq t$, ale $d_H(\mathbf{w}', \mathbf{v}) \leq d_H(\mathbf{w}, \mathbf{v})$.
2. Ať $\mathrm{dist}(W) > 2t$. Ať $\mathbf{w}$ je ve W a ať platí $1 \leq d_H(\mathbf{w}, \mathbf{v}) \leq t$.
   Pro jakékoli $\mathbf{w}'$ ve W platí $d_H(\mathbf{w}, \mathbf{w}') \geq \mathrm{dist}(W) > 2t$.
   To znamená, že $2t < d_H(\mathbf{w}, \mathbf{w}') \leq d_H(\mathbf{w}, \mathbf{v}) + d_H(\mathbf{v}, \mathbf{w}')$.

   Takže $d_H(\mathbf{w}', \mathbf{v}) > 2t - d_H(\mathbf{w}, \mathbf{v}) \geq 2t - t = t \geq d_H(\mathbf{w}, \mathbf{v})$.

Syndrome decoding

Ať W je lineární p-kód délky n a dimense k. Předpokládejme, že odeslané kódové slovo z W bylo přijato jako slovo $\mathbf{v}$ ze $(\mathbb{Z}_p)^n$.

Syndrome decoding je následující dekódovací procedura:

1. Spočteme syndrom $\mathbf{H}^T \cdot \mathbf{v} = \mathbf{s}$ slova $\mathbf{v}$.
2. Nalezneme takové řešení soustavy $(\mathbf{H}^T \mid \mathbf{s})$ tvaru $\mathbf{e} + W$, kde Hammingova váha $w_H(\mathbf{e})$ je nejmenší možná.[a]

   Předpokládáme, že bylo odesláno kódové slovo $\mathbf{w} = \mathbf{v} - \mathbf{e}$.

[a] Pokud je takových $\mathbf{e}$ více, vybereme některé z nich náhodně.

Kolik různých syndromů existuje?

Kontrolní matice W je lineární zobrazení $\mathbf{H}^T : (\mathbb{Z}_p)^n \to (\mathbb{Z}_p)^{n-k}$.
Slovo $\mathbf{s}$ je syndrom právě když $\mathbf{s}$ leží v $\mathrm{im}(\mathbf{H}^T)$.
Protože $\mathrm{rank}(\mathbf{H}^T) = n - k$, existuje celkem p^{n-k} různých syndromů.

Příklad (Hammingův $(7,4)$-kód a syndrome decoding)

Existuje 8 různých syndromů: $\mathbf{s}_0 = \begin{pmatrix}0\\0\\0\end{pmatrix}$, $\mathbf{s}_1 = \begin{pmatrix}0\\0\\1\end{pmatrix}$, $\mathbf{s}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}$,

$\mathbf{s}_3 = \begin{pmatrix}0\\1\\1\end{pmatrix}$, $\mathbf{s}_4 = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $\mathbf{s}_5 = \begin{pmatrix}1\\0\\1\end{pmatrix}$, $\mathbf{s}_6 = \begin{pmatrix}1\\1\\0\end{pmatrix}$, $\mathbf{s}_7 = \begin{pmatrix}1\\1\\1\end{pmatrix}$.

Vyřešením příslušných soustav nalezneme 8 „nejmenších“ error

paterns: $\mathbf{e}_0 = \begin{pmatrix}0\\0\\0\\0\\0\\0\\0\end{pmatrix}$, $\mathbf{e}_1 = \begin{pmatrix}1\\0\\0\\0\\0\\0\\0\end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix}0\\1\\0\\0\\0\\0\\0\end{pmatrix}$, $\ldots$, $\mathbf{e}_7 = \begin{pmatrix}0\\0\\0\\0\\0\\0\\1\end{pmatrix}$.

Vidíme, že syndrome decoding Hammingova $(7,4)$-kódu je korektní pro opravu nejvýše jedné chyby.[a]

[a] Dekódování již nelze nijak vylepšit, protože minimální distance tohoto kódu je 3.

Nearest neighbour decoding

Ať W je linární p-kód délky n a dimense k. Předpokládejme, že odeslané kódové slovo z W bylo přijato jako slovo $\mathbf{v}$ ze $(\mathbb{Z}_p)^n$.

Nearest neighbour decoding je následující dekódovací procedura:

1. Pokud je $\mathbf{v}$ kódové slovo, předpokládáme, že k žádné chybě nedošlo.

   Předpokládáme tedy, že bylo odesláno kódové slovo $\mathbf{v}$.

2. Pokud $\mathbf{v}$ kódové slovo není, nalezneme takové kódové slovo $\mathbf{w}$, pro které je Hammingova vzdálenost $d_H(\mathbf{v}, \mathbf{w})$ nejmenší.[a]

   Předpokládáme, že bylo odesláno kódové slovo $\mathbf{w}$.

[a] Pokud je takových kódových $\mathbf{w}$ slov více, vybereme některé z nich náhodně.

Tvrzení (syndrome decoding = nearest neighbour decoding)

Ať W je linární p-kód délky n a dimense k. Předpokládejme, že odeslané kódové slovo z W bylo přijato jako slovo $\mathbf{v}$ ze $(\mathbb{Z}_p)^n$. Potom množiny

$$\{\mathbf{w} \in W \mid \text{Hammingova vzdálenost } d_H(\mathbf{v}, \mathbf{w}) \text{ je nejmenší}\}$$

a

$$\{\mathbf{w} \in W \mid \text{Hammingova váha } w_H(\mathbf{v} - \mathbf{w}) \text{ je nejmenší}\}$$

jsou stejné.

To jest: syndrome decoding a nearest neighbour decoding jsou totožné procedury.

Důkaz.

Podle definice Hammingovy vzdálenosti platí pro libovolné vektory $\mathbf{x}$, $\mathbf{y}$ rovnost $d_H(\mathbf{x}, \mathbf{y}) = d_H(\mathbf{x} - \mathbf{y}, \mathbf{o}) = w_H(\mathbf{x} - \mathbf{y})$. ■