Krociutki kurs teorii kodow.
(przepraszam za brak polskich znakow diakrytycznych!)

Wprowadzenie

Definicja Niech bedzie dany alfabet Q skladajacy sie z q symboli. *Kodem* o dlugosci n nazywamy niepusty zbior n-elementowych ciagow liter alfabetu Q (tzw. slow kodu).

Uwaga: Czasami taki kod nazywamy kodem blokowym, ale poniewaz bedziemy uzywac tylko takich kodow blokowych, slowo 'blokowy' bedziemy konsekwnetnie pomijac.

Poniewaz w dalszej czesci wykladu beda nas interesowac kody liniowe, od tego momentu bedziemy zakladac, ze alfabet Q sklada sie z elementow pewnego ciala (w szczegolnosci, Q zawiera element zerowy). Jesli $Q = Z_2$ to mowimy o kodzie binarnym, gdy $Q = Z_3$ ternarnym etc.

W teorii kodowania przy konstrukcji kodow nalezy zwykle pogodzic dwa przeciwstawne cele. Z jednej strony, chcemy w n przeslanych literach alfabetu zawrzec stosunkowo duzo informacji, tzn. chodzi o to by w kodzie bylo mozliwie duzo slow. Z drugiej strony, chcemy aby mozna bylo (szybko) zidentyfikowac slowo kodu nawet wtedy czesc liter slowa zostala przesylana blednie.

Jesli chcemy by wyrazy kodu daly sie odroznic nawet w przypadku, gdy sa przeslane z bledem, musimy zdefiniowac odleglosc miedzy tymi wyrazami i zadac, by zadne dwa wyrazy kodu nie lezaly blisko siebie. Naturalna odlegloscia w tym przypadku jest odleglosc Hamminga, w ktorej odleglosc miedzy dwoma wyrazami mierzymy liczba pozycji, na ktorych one sie roznia, tzn.

$$d(\mathbf{x}, \mathbf{y}) = |\{i : x_i \neq y_i\}| .$$

Rozstepem $r(C)$ kodu C nazywamy minimalna odleglosc miedzy jego wyrazami, tj.

$$r(C) = \min\{d(\mathbf{x}, \mathbf{y}) : \mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}\} .$$

Poniewaz majac zniekształcone slowo kodu $\mathbf{x} \in Q^n$ bedziemy je przyblizac slowem $\mathbf{y} \in C$ nalezacym do kodu lezacym najblizej $\mathbf{x}$, dlatego jesli kod ma rozstep $2r + 1$, mozemy liczyc w najlepszym przypadku na to, ze w kazdym slowie kodu jestesmy w stanie poprawic co najwyzej r bledow. Najwieksze kody, dla ktorych jest to mozliwe, nosza nazwe kodow doskonalych (*ang. perfect codes*).

Definicja Kod C o rozstepie $2r+1$ nazywamy *doskonalym*, gdy dla kazdego slowa alfabetu $\mathbf{x} \in Q^n$ istnieje dokladnie jedno slowo kodu $\mathbf{y} \in C$ takie, ze $d(\mathbf{x}, \mathbf{y}) \leqslant r$.

Slowa kodu doskonalego sa srodkami kul (Hamminga), ktore stanowia podzial kostki Q^n, a zatem dla takich kodow zachodzi warunek

$$|C| \sum_{i=0} \binom{n}{i} (q-1)^i = q^n \,.$$

Przyklad: 4-kod ternarny $C_4(9)$ skladajacy sie z 9 slow: 0000, 0111, 0222, 1021, 2012, 1102, 2201, 2120, 1210 jest doskonaly.

Kody liniowe

Niech F_q bedzie cialem takim, ze $|F_q|$. Przypomnijmy, ze cialo takie istnieje wtedy i tylko wtedy gdy $q = p^m$, gdzie p jest liczba pierwsza. Przykladem takiego ciala jest Z_p, zbior $\{0, 1, \ldots, p-1\}$ z dzialaniami dodawania i mnozenia modulo p, gdzie p jest liczba pierwsza.

Od tego momentu przyjmniemy $Q = F_q$, a slowa $\mathbf{x} \in Q^n$ bedziemy traktowac jako wektory n-wymiarowej przestrzeni liniowej nad cialem F_q.

Definicja Kodem $[n, k]$ bedziemy nazywac dowolna podprzestrzen liniowa C wymiaru k w przestrzeni Q^n. $[n, k]$ kod C nazywac bedziemy $[n, k, d]$ kodem, jesli rozstep C wynosi d.

Niech $\mathbf{x}_1, \ldots, \mathbf{x}_k$ beda wektorami bedacymi baza $[n, k]$ kodu C. Wtedy dla dowolnych stalych $a_1, \ldots, a_k$ wektor $\mathbf{y} = \sum_{i=1}^{k} a_i \mathbf{x}_i$ nalezy do kodu C. Chcac zapisac ten fakt w jezyku mnozenia macierzy, okreslmy *macierz* $\boldsymbol{G}$ *generujaca kod* C jako macierz, ktorej wierszami sa wektory $\mathbf{x}_1^T, \ldots, \mathbf{x}_k^T$ bazy kodu C. **(Zauwazmy, ze w tej czesci wykladu stosujemy kolumnowy zapis wektorow).** Wtedy

$$C = \{\mathbf{a}^T \mathbf{G} : \mathbf{a}^T \in Q^k\} \,.$$

Macierz generujaca $\mathbf{G}$ kod ma postac normalna gdy mozemy ja zapisac w postaci $G = [\mathbf{I}_k \mathbf{P}]$, gdzie $\mathbf{I}_k$ jest macierza identycznosciowa.

Przyklad: Macierzami generujacymi kod $C_4(9)$ sa np.

$$\mathbf{G}_1 = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad \text{i} \quad \mathbf{G}_2 = \begin{bmatrix} 1 & 1 & 0 & 2 \\ 2 & 0 & 1 & 2 \end{bmatrix},$$

ale tylko pierwsza z nich ma postac normalna. Oczywiscie, jesli, powiedzmy, $\mathbf{a}^T = 12$, to

$$\mathbf{a}^T\mathbf{G}_2 = 1 \cdot 1102 + 2 \cdot 2012 = 1102 + 1021 = 2120 \in C\,.$$

Waznym pojeciem w teorii kodow jest pojecie *macierzy parzystosci*. Jest to macierz $\mathbf{H}$ o maksymalnym rzedzie (dla $[n,k]$ kodow $\mathbf{H}$ jest macierza $(n-k) \times n$ o rzedzie rownym $n-k$) taka, ze $\mathbf{G}\mathbf{H}^T = \mathbf{0}$, gdzie $\mathbf{0}$ oznacza macierz zerowa, a $\mathbf{H}^T$ jest macierza transponowana. Jesli macierz $\mathbf{G}$ dana jest w postaci normalnej, tj. $\mathbf{G} = [\mathbf{I}_k, \mathbf{P}]$, to za macierz parzystosci mozemy przyjac $\mathbf{H} = [-\mathbf{P}^T, \mathbf{I}_{n-k}]$.

Zauwazmy, ze poniewaz $C = \{\mathbf{a}^T\mathbf{G} : \mathbf{a} \in Q^k\}$ i $\mathbf{G}\mathbf{H}^T = \mathbf{0}$, zatem dla wszystkich slow $\mathbf{y} \in C$ mamy $\mathbf{y}^T\mathbf{H}^T = \mathbf{0}$. Nietrudno sprawdzic, ze zachodzi rowniez implikacja odwrotna.

Fakt Slowo $\mathbf{y} \in Q^n$ nalezy do kodu C wtedy i tylko wtedy gdy $\mathbf{H}\mathbf{y} = \mathbf{0}$ (zero po lewej stronie oznacza wektor o wszystkich wspolrzednych rownych zero).

Przyklad: Wrocmy do kodu $C_4(9)$. Macierza parzystosci oparta na G_1 ma postac

$$\mathbf{H}_1 = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 2 & 0 & 1 \end{bmatrix}.$$

Nietrudno sprawdzic, ze, na przyklad, dla $\mathbf{x}^T = 2120 \in C_4(9)$ mamy $\mathbf{x}^T\mathbf{H}_1^T = 00$ bo zeruja sie oba iloczyny skalarne wektorow $\langle 2120, 1210\rangle$ i $\langle 2120, 2201\rangle$. Z drugiej strony, dla wektora $\mathbf{y} = 2200 \notin C_4(9)$, mamy $\mathbf{x}^T\mathbf{H}_1^T = 02 \neq \mathbf{0}$, bo chociaz $\langle 2200, 1210\rangle = 0$, ale $\langle 2200, 2201\rangle = 2 \neq 0$.

Kody liniowe sa wazne dlatego, ze stosunkowo latwo poprawiac w nich bledy, tzn. znajdowac slowa kodu bedace w najblizszej odleglosci od danego slowa. Cala przestrzen wektorowa Q^n mozemy bowiem podzielic na warstwy wzgledem podprzestrzeni C wyznaczonej przez nasz kod. Jesli dwa wektory $\mathbf{x}, \mathbf{y}$ naleza do jednej warstwy, to $\mathbf{x} - \mathbf{y} \in C$, a zatem $\mathbf{H}(\mathbf{x} - \mathbf{y}) = \mathbf{0}$ czyli $\mathbf{H}\mathbf{x} = H\mathbf{y}$. Przypuscmy, ze znajac kod C wyznaczylismy, dla kazdej warstwy o "indeksie" (lub "syndromie") $\mathbf{H}\mathbf{x}$, nalezacy do niej wektor $\mathbf{z}$ majacy, wsrod wektorow tej warstwy, jak najmniej wspolrzednych niezerowych Wtedy dla dowolnego slowa $\mathbf{y}$ o indeksie $\mathbf{H}\mathbf{x}$ mamy $H(\mathbf{y} - \mathbf{z}) = \mathbf{0}$, czyli $\mathbf{y} - \mathbf{z} \in C$ przy czym wektor $\mathbf{y} - \mathbf{z}$ rozni sie od $\mathbf{y}$ tylko na tylu miejscach, ile niezerowych wspolrzednych ma wektor $\mathbf{z}$. Zatem z wyboru $\mathbf{z}$ wynika, ze $\mathbf{y} - \mathbf{z}$ jest slowem kodu C bedacym najblizej slowa $\mathbf{y}$.

Zatem, to co musimy zrobic dla ustalonego $[n, k]$ kodu to wybrac dla kazdej warstwy (tj. dla kazdej wartosci wektora $\mathbf{Hx} \in Q^{n-k}$, $\mathbf{x} \in Q^n$), reprezentanta o najmniejszej wadze, tzn. o najmniejszej liczbie wspolrzednych niezerowych. Oczywiscie mozemy to uczynic zanim otrzymamy slowo $\mathbf{x} \in Q^n$, ktore bedziemy musieli skorygowac, tzn. znalezc slowo kodu najblizsze $\mathbf{x}$. W przypadku kodu $C_4(9)$ mozliwych wartosci (dwuwymiarowego) wektora $\mathbf{xH}_1^T$ jest dziewiec i dla kazdej z nich powinnismy znalezc odpowiedni wektor korygujacy $\mathbf{z}$. I tak np. wektorem o najmniejszej wadze nalezacym do warstwy $\{\mathbf{x} : \mathbf{H}_1\mathbf{x} = 02^T\}$ jest (oczywiscie) wektor 0002, o wadze 1. Poniewaz jak zauwazylismy w poprzednim przykladzie, dla wyrazu 2200 mamy $\mathbf{H}_1 2200^T = 02^T$, wyraz $2200 - 0002 = 2201 \in C$ jest wyrazem kodu C najblizszym 2200. Nawiasem mowiac zauwazmy, ze poniewaz kod $C_4(9)$ jest doskonaly, kazda warstwa $\mathbf{H}_1\mathbf{x}$ ma reprezentanta o co najwyzej jednej niezerowej wspolrzednej.

Kody Hamminga

Definicja Kodem Hamminga nazywamy $[n, n-\ell, 3]$ kod nad cialem Q, $|Q| = q$,

$$n = \frac{q^\ell - 1}{q - 1}\,.$$

Nietrudno zauwazyc, ze kazdy kod Hamminga jest doskonaly, tzn. jest doskonalym $[n, k, 3]$ kodem.

Macierz parzystosci kodu Hamminga mozna wyznaczyc w szczegolnie prosty sposob. Jest to macierz, ktorej kolumny sa parami niezalezne, tzn. zadna z kolumn nie jest iloczynem innej kolumny pomnozonej przez skalar.

Przyklad. Nietrudno zauwazyc, ze zadne dwie kolumny macierzy

$$\mathbf{H}_1 = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 2 & 0 & 1 \end{bmatrix}\,.$$

nie sa liniowo zalezne. Zatem kod $C_4(9)$ jest kodem Hamminga, tzn. doskonalym $[4, 2, 3]$ kodem.