Föreläsning 7

MA1508

Vi pratade om Hammingavståndet förra gången. Nu ska vi använda det för lite intressanta saker.

Definition. Låt K vara en binär kod av längd n som innehåller med än ett kodord. *Minimaldistansen* hos koden K som det minsta avståndet som förekommer mellan två olika kodord i K.

Vi kommer i fortsättningen anta koder innehåller mer än ett kodord, annars kan man inte göra någon intressant kommunikation med dem.

Exempel 1. Minimaldistansen δ hos koden $K = \{10011, 11010, 10111, 11100\}$ är lika med 1 (kolla själv!). $\triangle$

Exempel 2. Minimaldistansen hos koden $\{000, 111\}$ är 3.

$\triangle$

Minimaldistansen har en nära koppling till koders felupptäckande och felrättande egenskaper.

Sats. *Låt K vara en kod med minimaldistansen δ. Då är koden en $(\delta - 1)$-felupptäckande kod. Koden är inte en δ- felupptäckande kod.*

Bevis. Vi antar att kodordet $x \in K$ skickas och att det blir något fel i överföringen, men inte fler än $\delta - 1$ fel. Anta strängen som Bob tar emot är y. Då är $d(x, y) < \delta$ vilket visar att y inte kan vara ett kodord. Alltså inser Bob att något blivit fel. Detta resonemang funkar oavsett vilket kodord som skickades, så $\delta - 1$ fel kan alltid upptäckas.

Omvänt, eftersom koden har minimaldistansen δ finns det två kodord, säg u och v, med avståndet δ. Om u skickas finns det alltså en möjlighet att δ stycken fel sker så att Bob istället tar emot v. Alltså är den inte δ-felupptäckande.

$\square$

Exempel 3. Vi betraktar den binära koden

$$K = \{00000, 10110, 01011, 11101\}$$

och beräknar

$$\begin{aligned} &d(00000, 10110) = 3 && d(10110, 01011) = 4 \\ &d(00000, 01011) = 3 && d(10110, 11101) = 3 \\ &d(00000, 11101) = 4 && d(01011, 11101) = 3. \end{aligned}$$

Detta visar att $\delta = 3$. Om kodordet 11101 utsätts för 1 eller 2=3-1 bitfel, då kan det aldrig ge oss ett kodord i K. Om vi däremot råkar ut för 3 bitfel, då kan 11101 övergå i $10110 \in K$ utan att vi märker något. △

Sats. *Låt K vara en kod med minimaldistansen δ. Om $2k+1 \leq \delta$, då är koden k-felrättande. Om $2\ell + 1 > \delta$ är den inte ℓ-felrättande.*

Om en kod har minimaldistansen 5 så säger alltså satsen att den är 2-felrättande eftersom $2 \cdot 2 + 1 \leq 5$. Däremot är den inte 3-felrättande för $2 \cdot 3 + 1 > 5$. Notera att om koden hade minimaldistans 6 så skulle den fortfarande vara 2-felrättande men inte 3-felrättande.

Bevis. Vi antar att $2k + 1 \leq \delta$ för något $k \geq 0$. Vi påstår att Bob så länge det inte sker mer än k fel i överföringen så kan Bob rätta dem genom att han väljer det närmaste kodordet (i Hammingavstånd) till det han tar emot.

Vi låter x vara ett kodord som skickas och y vara hur meddelandet mottas av Bob. Om y har högst k stycken fel, dvs $d(y, x) \leq k$, då gäller för alla $x \neq x' \in K$ att

$$d(x', y) + k \geq d(x', y) + d(y, x) \geq d(x', x) \geq \delta \geq 2k + 1.$$

Detta visar att $d(x', y) \geq k + 1$ för alla $x \neq x' \in K$, dvs att x ligger närmare y än alla andra kodord x'. Vi ser att x är den unika närmsta grannen till y, och vi kan *rätta* y genom att välja x som det korrekta kodordet.

Omvänt, välj ℓ som det minsta heltal så att $2\ell \geq \delta$. Det är samma sak som att kräva att ℓ ska vara det minsta heltal sådant att $2\ell + 1 > \delta$. Det finns två kodord, u och v, med $d(u, v) = \delta$. Ändra ℓ bitar i u så att de överensstämmer med motsvarande bitar i v. Kalla resultatet u'. Då är $d(u, u') = \ell$ och $d(u', v) = \delta - \ell$. På grund av hur vi valt ℓ så är $\delta - \ell \leq 2\ell - \ell = \ell$. Alltså kan vi skapa u' genom att göra högst ℓ fel i u eller genom att göra högst ℓ fel i v. Tar Bob emot u' kan han inte avgöra vilket av u och v som var det sända kodordet, och koden kan därför inte rätta ℓ fel alltid. □

Exempel 4. Vi betraktar återigen

$$K = \{00000, 10110, 01011, 11101\} \subseteq$$

med $\delta = 3$. Koden bör kunna rätta $k = 1$ bitfel: 11110 rättas till 10110, 10101 rättas till 11101, 01000 rättas till 00000 osv. Om däremot $k \leq 2$ bitfel inträffar, då klarar vi inte längre av att rätta felen. Det mottagna meddelandet 11000 kan ha sitt ursprung i båda 00000 och 11101. △

Målet med kodningsteori kan nu delvis omformuleras som att vi försöker skapa koder med många kodord och som ändå har hög minimaldistans.

Exempel 5. Sätt $K = \{000000, 010101, 101010, 111111\}$. Koden har minimaldistans 3. Vi har sett denna kod innan. $\triangle$

Exempel 6. Sätt $K = \{000000, 001110, 010011, 011101, 100101, 101011, 110110, 111000\}$. Det är en kod med minimaldistans 3 och längd 6. Den har 8 element. Verkar mycket bättre än repetitionskoden. $\triangle$

Minimaldistans för linjära koder

Det är lite bökigt att räkna minimaldistansen för koder med många kodord. Måste räkna avståndet för $\binom{n}{2}$ par av kodord. För linjära koder är det lite enklare.

Sats. *Låt K vara en linjär kod med minimaldistansen δ. Då gäller*

$$\delta = \min\{d(z,0) \,:\, 0 \neq z \in K\},$$

dvs minimaldistansen är lika med det minsta antalet ettor som förekommer i nollskilda vektoer/kodord.

Bevis. Det är inte svårt att se att för alla $x, y, z \in K$ gäller $d(x, y) = d(x + z, y + z)$. Detta ger

$$\begin{aligned}\delta &= \min\{d(x,y) \,:\, x \neq y\} \\ &= \min\{d(x-y, y-y) \,:\, x \neq y\} \\ &= \min\{d(z,0) \,:\, 0 \neq z\}\end{aligned}$$

$\square$

Gränser för koder

Det finns givetvis gränser för hur bra koder vi kan förvänta oss hitta. En sådan gräns finns i nästa sats.

Sats. *Anta att K är en (n, k)-kod som är t-felrättande. Då är*

$$2^k \leq \frac{2^n}{\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{t}}.$$

Bevis. För varje kodord, $x \in K$, kan vi skapa en mängd B_x som innehåller de strängar vi får från x om högst t fel inträffar. Om x och y är två olika kodord så måste det gälla att $B_x \cap B_y = \emptyset$, annars hade K inte kunnat vara t-felrättande. Låt oss kalla dessa mängder för *bollar*. Eftersom det finns $\binom{n}{0}$ sätt att göra 0 bitfel, $\binom{n}{1}$ sätt att göra 1 bitfel osv, så är antalet element i B_x lika med $\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{t}$. Antalet kodord i koden är 2^k. Det finns alltså 2^k disjunkta bollar, vara och en med $\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{t}$ stycken element.

Eftersom antalet element i bollarna totalt sett definitivt sett inte är fler än 2^n så är

$$2^k \cdot \left(\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{t} \right) \leq 2^n.$$

Genom en division följer olikheten vi ville bevisa. $\square$

Det finns också satser som garanterar att koder som är någorlunda bra existerar.

Sats. *Låt n vara ett positivt heltal och d ett positivt heltal som är mindre än n. Då finns det en (n, k)-kod med minimaldistans minst d för vilken det gäller att*

$$2^k \geq \frac{2^n}{\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{d-1}}.$$