Veaparanduskoodidest

Käesolevas tekstis on kasutatud järgmisi algebra mõisteid: jäägiklassikorpus, vektorruum, vektorruumi alamruum, baas, lineaarselt sõltumatu vektorite süsteem, lineaarkombinatsioon, lineaarne kate, maatriks, maatriksite korrutamine, transponeeritud maatriks, maatriksi astak, lineaarkujutus, lineaarkujutuse tuum ja kujutis, homogeenne lineaarvõrrandisüsteem, lahendite fundamentaalsüsteem.

Tänapäeval on kõikjal vaja edastada elektroonilisi signaale läbi erinevate kanalite, kus võivad tekkida vead. Näiteks signaali edastamisel kosmoselaevast maale võib vigu tekitada päikesekiirguse aktiivsus või faili salvestamisel kõvakettale (või mõnele teisele andmekandjale) võivad vead tekkida mikroskoopilistest vigastustest kõvaketta pinnal. Selliste probleemidega tegelemiseks oleks tore, kui me

a) oskaks kindlaks teha, kas edastatud signaal jõudis vigadeta pärale,

b) võimaluse korral oskaks tekkinud vigu parandada.

Nende probleemide lahendamisel tuleb appi algebra.

Enamasti on mugav edastatav signaal esitada bittide (0 ja 1) jadana. Matemaatilises mõttes tõlgendame me neid bitte jäägiklassikorpuse $\mathbb{Z}_2$ elementidena jättes kriipsud 0 ja 1 korral lihtsalt kirjutamata. Muuhulgas me võime sooritada tehteid: näiteks $1 + 1 = 0$ ja $0 \cdot 1 = 0$.

Kui me edastaksime bitte ükshaaval, siis puudub meil igasugune võimalus kindlaks teha, kas kohalejõudnud bitt on ikka see, mis teele saadeti. Asi on parem, kui me edastame bitte m-elemendiliste rühmadena ($m > 1$ on naturaalarv), mida me kutsume *blokkideks*. Niisiis edastatav bitijada jagatakse m-elemendilisteks blokkideks ja ka vastuvetavat sõnumit loetakse m-elemendiliste blokkidena. Neid blokke me tõlgendame vektoritena

$$\mathbf{a} = (a_1, a_2, \ldots, a_m)$$

vektorruumis $\mathbb{Z}_2^m$ (üle korpuse $\mathbb{Z}_2$). Nagu ikka, tehakse tehteid selles vektorruumis komponenthaaval.

Veaparanduse põhiidee on selles, et edastatavale blokile lisatakse (harilikult lõppu) üks ports lisasümboleid, mis ei kanna endas küll algse bloki sõnumit, mis aga aitavad tekkinud vigu parandada. Seega läbi mürarikka kanali saatmiseks suurendatakse bloki pikkust m-lt n-le, kus $n > m$.

Definitsioon 1 *Veaparanduskood* koosneb kodeerimiskujutusest

$$E : \mathbb{Z}_2^m \to \mathbb{Z}_2^n$$

ja dekodeerimiskujutusest

$$D : \mathbb{Z}_2^n \to \mathbb{Z}_2^m,$$

mis rahuldavad seost $D(E(\mathbf{a})) = \mathbf{a}$ iga $\mathbf{a} \in \mathbb{Z}_2^m$ korral. Sellist veaparanduskoodi $\mathcal{C} = (E, D)$ nimetatakse *lineaarseks* (m, n)*-koodiks*, kui E on lineaarkujutus.

Võrdusest $DE = 1_{\mathbb{Z}_2^m}$ järeldub (kuidas?), et kujutus E peab olema üksühene.

Lineaarkujutust $\mathbb{Z}_2^m \to \mathbb{Z}_2^n$ on võimalik defineerida maatriksi abil. Näiteks, kui $G \in \mathrm{Mat}_{m,n}(\mathbb{Z}_2)$, siis kujutus $E : \mathbb{Z}_2^m \to \mathbb{Z}_2^n$, mis on defineeritud võrdusega

$$E(\mathbf{a}) := \mathbf{a}G,$$

$\mathbf{a} \in \mathbb{Z}_2^m$, on lineaarkujutus. Sellist maatriksit G nimetatakse lineaarse koodi $\mathcal{C} = (E, D)$ *genereraatormaatriksiks*. (Siin ja edaspidi me samastame vektori $\mathbf{a} \in \mathbb{Z}_2^m$ ja $(1 \times m)$-maatriksi, mille reavektoriks on $\mathbf{a}$. Nad erinevad ainult komade olemasolu või puudumise poolest. Sellest ei tohiks segadust tekkida.)

Näide 1 Olgu $m = 4$, $n = 7$ ja

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

Oletame, et saatja tahab teele saata sõnumi $\mathbf{a}$. Ta kodeerib selle pikkusega 7 vektoriks $\mathbf{b}$ järgmiselt:

$$\mathbf{b} = \mathbf{a}G = (1\ 1\ 1\ 0) \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} = (1\ 1\ 1\ 0\ 0\ 0\ 0).$$

Nagu näha lisandus vektori $\mathbf{a}$ lõppu kolm nulli. Need on niinimetatud kontrollsümbolid. Hiljem näeme, kuidas nende abil saab parandada ära vea, mis vektori $\mathbf{b}$ edastamisel läbi kanali võib tekkida.

Kui $E : \mathbb{Z}_2^m \to \mathbb{Z}_2^n$ on lineaarkujutus, siis tema kujutis

$$E(\mathbb{Z}_2^m) = \{E(\mathbf{a}) \mid \mathbf{a} \in \mathbb{Z}_2^m\}$$

on vektorruumi $\mathbb{Z}_2^n$ alamruum. Selle alamruumi elemente nimetatakse *koodsõnadeks*.

Vektorruumis $\mathbb{Z}_2^m$ võib vaadelda baasi $\mathbf{e}_1, \ldots, \mathbf{e}_m$, kus vektori $\mathbf{e}_i$ komponent kohal i on 1 ja kõik ülejäänud komponendid on 0-d. Tähistades maatriksi G reavektoreid $\mathbf{g}_1, \ldots, \mathbf{g}_m$ näeme, et

$$E(\mathbf{e}_i) = \mathbf{e}_i G = \mathbf{e}_i \begin{pmatrix} \mathbf{g}_1 \\ \cdots \\ \mathbf{g}_m \end{pmatrix} = \mathbf{g}_i.$$

Kui vektor $\mathbf{a} \in \mathbb{Z}_2^m$ avaldub baasivektorite lineaarkombinatsioonina kujul $\mathbf{a} = k_1\mathbf{e}_1 + \ldots + k_m\mathbf{e}_m$, kus $k_1, \ldots, k_m \in \mathbb{Z}_2$, siis

$$E(\mathbf{a}) = E(k_1\mathbf{e}_1 + \ldots + k_m\mathbf{e}_m) = k_1E(\mathbf{e}_1) + \ldots + k_mE(\mathbf{e}_m) = k_1\mathbf{g}_1 + \ldots + k_m\mathbf{g}_m.$$

Seega näeme, et iga koodsõna avaldub maatriksi G reavektorite lineaarkombinatsioonina. Teiste sõnadega: $E(\mathbb{Z}_2^m)$ on vektorite $\mathbf{g}_1, \ldots, \mathbf{g}_m$ lineaarne kate,

$$E(\mathbb{Z}_2^m) = L(\mathbf{g}_1, \ldots, \mathbf{g}_m).$$

Lause 1 *Olgu $\mathcal{C} = (E, D)$ lineaarne (m, n)-kood. Selle koodi generaatormaatriksi G reavektorid on lineaarselt sõltumatud. Seega* $\operatorname{rank}(G) = m$ *ja* $\dim(E(\mathbb{Z}_2^m)) = m$.

Tõestus. Me tõestame, et G reavektorid $\mathbf{g}_1, \ldots, \mathbf{g}_m$ on lineaarselt sõltumatud, ülejäänud kaks väidet on selle ilmsed järeldused.

Oletame, et $k_1\mathbf{g}_1 + \ldots + k_m\mathbf{g}_m = \mathbf{0}$. Siis

$$\mathbf{0} = k_1\mathbf{g}_1 + \ldots + k_m\mathbf{g}_m = k_1E(\mathbf{e}_1) + \ldots + k_mE(\mathbf{e}_m) = E(k_1\mathbf{e}_1 + \ldots + k_m\mathbf{e}_m).$$

Lineaarkujutuse E üksühesuse tõttu $k_1\mathbf{e}_1 + \ldots + k_m\mathbf{e}_m = \mathbf{0}$. Et aga süsteem $\mathbf{e}_1, \ldots, \mathbf{e}_m$ on lineaarselt sõltumatu, siis ka $k_1 = \ldots = k_m = 0$. Definitsiooni põhjal on vektorid $\mathbf{g}_1, \ldots, \mathbf{g}_m$ lineaarselt sõltumatud. ■

Generaatormaatriks G on väga kasulik sõnumi saatjale. Selle abil saab sõnu kergesti kodeerida lihtsa maatrikskorrutamise abil. Aga vastuvõtjale ei ole sellest maatriksist abi. Vastuvõtjal on vaja ühte teist maatriksit, mis aitaks kontrollida, kas kohalejõudnud sõnum on ikka seesama, mis teele saadeti.

Definitsioon 2 Olgu $\mathcal{C} = (E, D)$ lineaarne (m, n)-kood. Maatriksit $H \in \mathrm{Mat}_{n-m,n}(\mathbb{Z}_2)$ nimetatakse selle koodi *kontrollmaatriksiks*, kui

$$\mathbf{x} \in E(\mathbb{Z}_2^m) \iff H\mathbf{x}^T = \mathbf{0}^T.$$

Seega kui vastuvõtja saab kätte vektori $\mathbf{x}$, leiab maatriksite korrutise $H\mathbf{x}^T$ ja kui see juhtub olema nullmaatriks (mõõtmetega $(n-m)\times 1$), siis on alust loota, et selline oli ka esialgne esialgne sõnum. Kui aga $H\mathbf{x}^T$ ei ole nullmaatriks, siis on selge, et sõnumi edastamisel on tekkinud viga.

Maatriks $H \in \mathrm{Mat}_{n-m,n}(\mathbb{Z}_2)$ tekitab lineaarkujutuse

$$\mathbb{Z}_2^n \to \mathbb{Z}_2^{n-m}, \;\; \mathbf{x} \mapsto (H\mathbf{x}^T)^T.$$

Tähistades selle lineearkujutuse tuuma $\mathrm{Null}(H)$ võime öelda, et H on kontrollmaatriks parajsti siis, kui

$$E(\mathbb{Z}_2^m) = \mathrm{Null}(H).$$

Paneme tähele, et $\mathrm{Null}(H)$ on ühtlasi ka maatriksiga H määratud homogeense lineaarvõrrandisüsteemi lahendite alamruum vektorruumis $\mathbb{Z}_2^n$.

Uurime nüüd, kuidas generaatormaatriksist lähtudes konstrueerida kontrollmaatriksit. Selleks eeldame, et generaatormaatriksis on m esimest veergu sellised nagu ühikmaatriksis. Seega

$$G = \begin{pmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \cdots \\ \mathbf{g}_m \end{pmatrix} = \begin{pmatrix} 1 & 0 & \ldots & 0 & a_{1,1} & \ldots & a_{1,n-m} \\ 0 & 1 & \ldots & 0 & a_{2,1} & \ldots & a_{2,n-m} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 1 & a_{m,1} & \ldots & a_{m,n-m} \end{pmatrix}.$$

Vaatleme homogeenset lineaarvõrrandisüsteemi, mille maatriks on G. Kuna see maatriks on astmelisel kujul ja astmeid on m tükki, siis on sellel m sõltuvat ja $n-m$ sõltumatut tundmatut. Sõltuvateks võime võtta $x_1, \ldots, x_m$ ja sõltumatuteks $x_{m+1}, \ldots, x_n$. Avaldame sõltuvad tundmatud vabade kaudu:

$$\begin{aligned} x_1 &= -a_{1,1}x_{m+1} - \ldots - a_{1,n-m}x_n \\ x_2 &= -a_{2,1}x_{m+1} - \ldots - a_{2,n-m}x_n \\ &\ldots \\ x_m &= -a_{m,1}x_{m+1} - \ldots - a_{m,n-m}x_n. \end{aligned}$$

Nii nagu harilikult, võime leida selle süsteemi lahendite fundamentaalsüsteemi vektorid

$$\begin{aligned} \mathbf{h}_1 &= (-a_{1,1}, -a_{2,1}, \ldots, -a_{m,1}, 1, 0, \ldots, 0), \\ \mathbf{h}_2 &= (-a_{1,2}, -a_{2,2}, \ldots, -a_{m,2}, 0, 1, \ldots, 0), \\ &\ldots \\ \mathbf{h}_{n-m} &= (-a_{1,n-m}, -a_{2,n-m}, \ldots, -a_{m,n-m}, 0, 0, \ldots, 1). \end{aligned}$$

(Kuna korpuses $\mathbb{Z}_2$ on iga element võrdne oma vastandelemendiga, siis võib viimases avaldises tegelikult miinusmärgid ära jätta.) Moodustame nüüd maatriksi $H \in \mathrm{Mat}_{n-m,n}(\mathbb{Z}_2)$ võttes tema reavektoriteks $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{n-m}$. Kuna $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{n-m}$ on lahenditeks homogeensele lineaarvõrrandisüsteemile, mille maatriks on G, siis $G\mathbf{h}_i^T = \mathbf{0}^T$ iga $i \in \{1, \ldots, n-m\}$ korral. Järelikult $\mathbf{g}_j\mathbf{h}_i^T = 0$ iga $j \in \{1, \ldots, m\}$ ja $i \in \{1, \ldots, n-m\}$ korral. Viimane võrdus on samaväärne võrdusega $\mathbf{h}_i\mathbf{g}_j^T = 0$. Nendest võrdustest saame järeldada, et

$$H\mathbf{g}_j^T = \mathbf{0}^T.$$

Näitame, et H on kontrollmaatriks. Selleks oletame, et $\mathbf{c} = k_1\mathbf{g}_1 + \ldots + k_m\mathbf{g}_m$ on mingi koodsõna. Siis

$$H\mathbf{c}^T = H(k_1\mathbf{g}_1^T) + \ldots + H(k_m\mathbf{g}_m^T) = k_1H\mathbf{g}_1^T + \ldots + k_mH\mathbf{g}_m^T = k_1\mathbf{0}^T + \ldots + k_m\mathbf{0}^T = \mathbf{0}^T.$$

Sellega oleme näidanud, et $E(\mathbb{Z}_2^m) \subseteq \mathrm{Null}(H)$. Kuna maatriksi H astak on $n-m$, siis talle vastaval homogeensel lineaarvõrrandisüsteemil on $n-(n-m) = m$ vaba tundmatut ja tema lahendite alamruum $\mathrm{Null}(H)$ (vektorruumis $\mathbb{Z}_2^n$) on m-mõõtmeline. Samas ka alamruum $E(\mathbb{Z}_2^m)$ on m-mõõtmeline tänu lausele 1. Kuna m-mõõtmelises vektorruumis ei saa olla m-mõõtmelist pärisalamruumi, siis peab kehtima võrdus $E(\mathbb{Z}_2^m) = \mathrm{Null}(H)$. Seega H on kontrollmaatriks ja me oleme tõestanud järgmise tulemuse.

Teoreem 1 *Olgu $\mathcal{C} = (E, D)$ lineaarne (m, n)-kood. Kui $G = (E_m|A)$ on selle koodi generaatormaatriks, siis $H = (-A^T|E_{n-m})$ on selle koodi kontrollmaatriks. (Siin E_m ja E_{n-m} on vastavat järku ühikmaatriksid.)*

Vektorit

$$S(\mathbf{b}) = H\mathbf{b}^T$$

nimetatakse vektori $\mathbf{b} \in \mathbb{Z}_2^n$ *sündroomiks.*

Teoreem 2 *Olgu $\mathcal{C} = (E, D)$ lineaarne (m, n)-kood, mille kontrollmaatriksi $H \in \mathrm{Mat}_{n-m,n}(\mathbb{Z}_2)$ kõik veerud on erinevad. Kui sõnumi edastamisel tekib üks viga kohal i, siis vastuvõetud vektori sündroom on võrdne maatriksi H i-nda veeruvektoriga.*

Tõestus. Olgu H veeruvektorid $\mathbf{h}^1, \ldots, \mathbf{h}^n \in \mathbb{Z}_2^{n-m}$, mida vaatleme üheveeruliste maatriksitena. Oletame, et saatja kodeerib oma sõnumi vektoriks $\mathbf{a}$. Läbi kanali saatmisel tekib sellesse vektorisse üks viga kohal i, see tähendab, et i-ndal kohal olev 1 muutub 0-ks või vastupidi. Sellist muutust modelleerib matemaatiliselt ühikvektori $\mathbf{e}_i$ liitmine: $\mathbf{a} + \mathbf{e}_i$. Niisiis vastuvõtja saab kätte vigase vektori $\mathbf{b} = \mathbf{a} + \mathbf{e}_i$. Ta sooviks kindlaks teha, kas $\mathbf{b}$ ongi see vektor, mis teele saadeti. Selleks ta arvutab sündroomi:

$$S(\mathbf{b}) = H\mathbf{b}^T = H(\mathbf{a} + \mathbf{e}_i)^T = H\mathbf{a}^T + H\mathbf{e}_i^T = \mathbf{0}^T + \mathbf{h}^i = \mathbf{h}^i.$$

Kuna tulemuseks on H i-s veeruvektor (mis ei saa olla nullvektor), siis vastuvõtja järeldab, et viga on tekkinud saadetava vektori i-ndal kohal ja ta saab selle vea parandada. ■

Näide 2 Vaatleme koodi generaatormaatriksiga

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} = (E_4|A).$$

näitest 1. Vastavalt teoreemile 1 saame moodustada selle koodi jaoks kontrollmaatriksi

$$H = (A^T|E_3) = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Nagu nägime kodeeritakse vektor $\mathbf{a} = (1, 1, 1, 0)$ vektoriks $\mathbf{b} = (1, 1, 1, 0, 0, 0, 0)$. Oletame, et vektori $\mathbf{b}$ saatmisel läbi kanali tekib viga teisel kohal (sümbol 1 muutub sümboliks 0) ja vastuvõtja saab kätte vektori

$$\mathbf{c} = (1, 0, 1, 0, 0, 0, 0) = \mathbf{b} + \mathbf{e}_2.$$

Vastuvõtja tahaks kindlaks teha, kas teelesaadetud sõnum oli $(1, 0, 1, 0)$ (4 esimest sümbolit vektorist $\mathbf{c}$) või midagi muud. Selleks arvutab ta sündroomi

$$S(\mathbf{c}) = H\mathbf{c}^T = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \mathbf{h}_2.$$

Teoreemi 2 kasutades saab vastuvõtja teada, et eeldusel, et vigu tekkis vaid üks, pidi see tekkima kohal 2. Seega teeb ta järelduse, et teelesaadetud sõnum oli mitte $(1, 0, 1, 0)$, vaid $(1, 1, 1, 0)$. Sellega on ta tekkinud vea edukalt ära parandanud.

Antud kood suudab ära parandada vaid üksikud vead. Kui kahe vea tekkimise tõenäosus pikkusega 7 vektoris on väga väikse tõenäosusega (see sõltub kanali omadustest), siis ongi selline kood praktikas kasutatav. Kui aga tahame, et kood suudaks ära parandada ka mitu viga, siis tuleb suurendada m-i ja n-i ning kasutada suuremaid maatrikseid.

Näites 2 kasutatud koodi nimetatakse $(4, 7)$ *binaarseks Hammingu*[1] *koodiks*. See on lihtsaim näide lõpmatust Hammingu koodide perest.

Meie kasutasime siin kodeerimisel lihtsaimat korpust $\mathbb{Z}_2$. Selle asemel võib kasutada ka teisi lõplikke korpusi (olgu siis $\mathbb{Z}_p$-d, kus p on algarv, või veel keerulisemad).

Kodeerimisel ei pea tingimata kasutama vektoreid. Või täpsemalt: vektorile $\mathbf{a} \in \mathbb{Z}_2^m$ võib vastavusse seada polünoomi, mille kordajateks on selle vektori komponendid ning kodeerimiseks võib saadud polünoomi korrutada teatud generaatorpolünoomiga — nii nagu ikka polünoome korrutatakse.

Kes soovib kodeerimisteooria kohta eesti keeles põhjalikumalt lugeda, võib tutvuda Peeter Puusempa (TTÜ) koostatud õppematerjaliga [2]. Käesolev tekst on koostatud Valdis Laane poolt 2017. aastal tuginedes Arkadii Slinko raamatule [1].

Viited

[1] A. Slinko, *Algebra for Applications*, Springer, 2015.

[2] P. Puusemp, *Kodeerimisteooria*, http://www.staff.ttu.ee/˜puusemp/Koot.pdf .

[1] Richard Wesley Hamming (1915–1998) — ameerika matemaatik