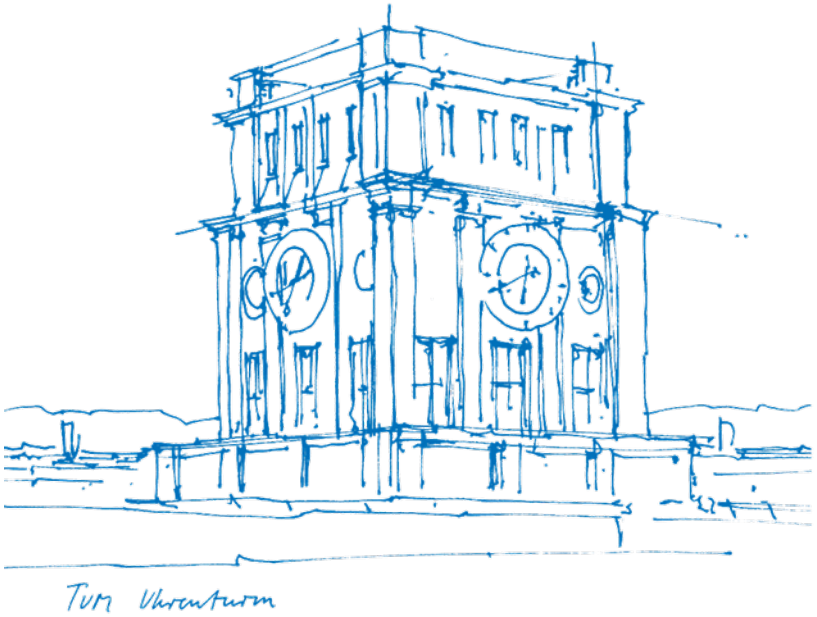
Lineare Algebra für Informatik - Woche 6

Cosmin Aprodu

Technische Universität München

Online, 20 Mai 2021



Warum Lineare Codes?

Motivation: Bild über einen verlustbehafteten Kanal versendet (*mit* und *ohne* Verwendung von Linearen Codes).

ohne Kanalkodierung mit Kanalkodierung

Lineare Codes

Sei K ein Körper und $(x_1, \ldots, x_k) \in K^k$ und $(c_1, \ldots, c_n) \in K^n$ zwei Vektoren (oder *Bitströme*, wenn wir Kanalkodierungen betrachten). Hierbei, gibt es eine Zuordnung $(x_1, \ldots, x_k) \mapsto (c_1, \ldots, c_n)$ in Form einer Matrix $G \in K^{n \times k}$, also:

$$G \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

$\rightarrow$ Der gesendete Vektor $(c_1, \ldots, c_n)$ heißt **Codewort**, $(x_1, \ldots, x_k)$ heißt **Informationswort** und G heißt **Generatormatrix**.

Ein **linearer Code** ist ein Unterraum $C = \left\{ G \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} \middle| \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} \in K^k \right\} \subseteq K^n$. Die **Länge** von C ist n.

$\rightarrow$ Die **Informationsrate** ist $\frac{k}{n}$. Die **Redundanz** ist $n - k$. Mithilfe von k und n, bezeichnen wir einen Code als (n, k)**-Code**.

Bemerkung: In der Regel muss $n > k$ gelten, insbesondere wenn wir über Kanalkodierungen sprechen.

Lineare Codes (2)

Für irgendein $x \in K^k$ muss das LGS $G \cdot x = c$, $\forall c \in C$ eindeutig lösbar sein. Dazu sind die Spalten von G linear unabhängig. Daraus folgt, dass:

$$\dim(C) = k$$

Sei $c, c' \in K^n$ und $C \subseteq K^n$ eine Teilmenge. Dann definieren wir folgendes:

- $w(c) := |\{i \in \{1, \ldots, n\} \mid c_i \neq 0\}|$ ist das **Hamming-Gewicht** von c.
- $d(c, c') := w(c - c') = |\{i \in \{1, \ldots, n\} \mid c_i \neq c'_i\}|$ ist der **Hamming-Abstand** von c und c'.
- $d(C) := \min\{d(c, c') \mid c, c' \in C, c \neq c'\}$ ist der **Hamming-Abstand** von C.
  $\rightarrow$ Falls *zusätzlich* C ein Unterraum ist, ergibt sich: $d(C) = \min\{w(c) \mid c \in C \setminus \{0\}\}$

Sei $C \subseteq K^n$ ein Code und $e \in \mathbb{N}_0$.

- Falls $d(C) = 2e + 1$, so ist C *e-fehler**korrigierend***.
- Falls $d(C) = e + 1$, so ist C *e-fehler**erkennend***.

Parity-Check-Matrix

In vielen Fällen betrachten wir eine Generatormatrix der Form $G = \begin{pmatrix} I_k \\ A \end{pmatrix}$, mit $I_k \in K^{k \times k}$ Identitätsmatrix und $A \in K^{(n-k) \times k}$. Aus diesem Grund, können wir eine neue Matrix definieren:

$$P := \begin{pmatrix} -A & I_{n-k} \end{pmatrix} \in K^{(n-k) \times n}$$

$\rightarrow$ P heißt **Parity-Check-Matrix**, hat den Rang $n-k$ und es gilt: $P \cdot G = \begin{pmatrix} -A & I_{n-k} \end{pmatrix} \cdot \begin{pmatrix} I_k \\ A \end{pmatrix} = 0$.

Wichtig: Heraus folgt, dass $P \cdot c = 0$, $\forall c \in C$. Daher, erhalten wir die folgende Äquivalenz:

$$c \in C \Leftrightarrow P \cdot c = 0$$

Sei nun $c' \in K^n$ ein empfangenes Wort. Den Unterschied von c und c' quantifizieren wir durch den **Fehlervektor** $f := c' - c \in K^n$. Es ergibt sich:

$$P \cdot c' = P \cdot (c + f) = 0 + P \cdot f = P \cdot f$$

Der Vektor $P \cdot c' \in K^{n-k}$ heißt das **Syndrom** von c' und misst, wie weit c' von einem gültigen Codewort abweicht.