Modul Diskrete Mathematik WiSe 2011/12

Ergänzungsskript zum Kapitel 3.4. Hinweis: Dieses Manuskript ist nur verständlich und von Nutzen für Personen, die regelmäßig und aktiv die zugehörige Vorlesung besuchen (also nicht nur körperlich anwesend sind), und es wurde auch nur für diesen Hörerkreis geschrieben. Es erhebt keinen Anspruch auf Vollständigkeit, Fehler sind zwar ärgerlich, aber leider nicht auszuschließen!

3 Codierungstheorie

4 Fehlerkorrigierende Codes

Wir beschäftigen uns mit der *Frage*, ob es Codes gibt, die nicht nur Fehler erkennen, sondern sogar in einem gewissen Umfang korrigieren können.

Beispiel: Wir definieren einen dreifachen Wiederholungscode über $A = \{0, 1\}$, jede der beiden Nachrichten 0 oder 1 werde dreifach übertragen: $C = \{000, 111\}$. Was macht der Empfänger, falls bei ihm die Ziffernfolge 1 0 1 ankommt?

Antwort: Unter Verwendung der „natürlichen Decodierregel“ *Mehrheit siegt* wird er 1 0 1 in 1 decodieren.

Allgemein: Empfänger decodiert $c_1\, c_2\, c_3$ in die Ziffer, die am häufigsten vorkommt. Dadurch ist in diesem Beispiel gewährleistet, dass jeder Einzelfehler nicht nur erkannt, sondern sogar korrigiert wird.

Wir übertragen diesen Wiederholungscode auf das Beispiel des zweifachen Münzwurfs und vergleichen ihn mit dem Verfahren ohne Codierung und mit dem Paritätscode $C(3, 2)$:

Ohne Codierung $C_1 = \{00, 01, 10, 11\}$, Paritätscode $C_2 = \{000, 011, 101, 110\}$,
dreifacher Wiederholungscode $C_3 = \{000000, 000111, 111000, 111111\}$.

	C_1	C_2	C_3
Richtig	81 %	72,9 %	94,48 %
Offen	0 %	24,4 %	0 %
Falsch	19 %	2,7 %	5,52 %
Informationsrate*	1	0.667	0.333

∗ Die *Informationsrate* misst die Effizienz eines Codes durch den Quotienten $\frac{\text{Länge der Nachricht}}{\text{Länge des Codewortes}}$.

Aus Kostengründen ist man an hohen Informationsraten interessiert. Somit ist unser *Problem*:
Finde Codes mit möglichst hoher Informationsrate und gleichzeitig guten Korrektureigenschaften.

Beispiel: Ein fünffacher Wiederholungcode würde sogar zwei Fehler verbessern, hätte aber die miserable Informationsrate 0.2.

Im Folgenden benutzen wir die natürliche Decodierregel, indem wir zu der empfangenen Nachricht y das Codewort c suchen, das sich an möglichst wenigen Stellen von y unterscheidet.

Def 3.4.1 Sei $A = \{0, 1\}$ und $V := A^n$ mit $n \in \mathbb{N}$. Für $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in V$ heißt

$$d(x, y) := |\{i \in \{1, \ldots, n\} \mid x_i \neq y_i\}|$$

Hamming–Abstand von x und y.

Beispiel: 1) Im Code C_3 ist $d((000111), (001011)) = 2$.

2) Im Paritätscode $C_2 = C(3, 2)$ haben unterschiedliche Codewörter den Hamming–Abstand 2.

Wir notieren einige Eigenschaften des Hamming–Abstandes:

Satz 3.4.1 Seien $x, y \in V$ wie oben. Dann gelten

(1) $d(x, y) \geq 0 \quad \forall x, y \in V, \qquad d(x, y) = 0 \iff x = y$

(2) $d(x, y) = d(y, x) \quad \forall x, y \in V$

(3) $d(x, z) \leq d(x, y) + d(y, z) \quad \forall x, y, z \in V$

Beweis: (1) und (2) sind offensichtlich richtig, (3) ist eventuell eine Übungsaufgabe.

Def 3.4.2 Sei $C \subseteq V$. Dann heißt

$$d_{\min}(C) := \min\{d(x, y) > 0 \mid x, y \in C\}$$

Minimalabstand des Codes C.

Beispiel: Die Codes C_1, C_2, C_3 von oben haben die Minimalabstände $d_{\min}(C_1) = 1$, $d_{\min}(C_2) = 2$, $d_{\min}(C_3) = 3$.

Je größer der Minimalabstand eines Codes ist, um so besser sind seine Korrekturmöglichkeiten:

Satz 3.4.2 Sei $C \subseteq V$ ein Code mit $d_{\min}(C) = d \in \mathbb{N}$.
Bei Anwendung der natürlichen Decodierregel werden bis zu $\frac{d-1}{2}$ Fehler korrigiert. (Passieren bei einer Übertragung bis zu $\frac{d-1}{2}$ Fehler, wird die korrekte Nachricht erkannt.)

Beweis: Das Codewort c wird als y empfangen. Die Decodierregel sucht nach $x \in C$ mit $d(x, y)$ minimal.

Beh.: Sind $f \leq \frac{d-1}{2}$ Fehler passiert, wird y korrekt als c erkannt.

Bew.: $\forall x \in C \backslash \{c\}$ gilt

$$\begin{aligned} & d \leq d(x, c) \leq d(x, y) + d(y, c) = d(x, y) + f \\ \Rightarrow \quad & d(x, y) \geq d - f \geq d - \frac{d-1}{2} = \frac{d+1}{2} > \frac{d-1}{2} \geq f = d(c, y) \ . \end{aligned}$$

Bem.: Zur Fehlerkorrektur ist ein Minimalabstand von mindestens 3 notwendig.

Problem: Der Anwender ist an Codes der Länge n interessiert, die möglichst viele Fehler korrigieren (also n möglichst groß) und aus Kostengründen eine hohe Informationsrate besitzen (also n möglichst klein)?!

Lösung: Benutze Erkenntnisse der Linearen Algebra!

Wir beschränken uns im Folgenden auf den Fall $V = \{0,1\}^n$, formulieren einige Aussagen aber allgemeiner. Man sollte aus dem zweiten Semester wissen, dass V ein n–dimensionaler Vektorraum über dem Körper $\mathbb{K} = \mathbb{Z}_2 = \{0,1\}$ mit Modulo 2 – Rechnung ist.

Def 3.4.3 Sei V ein n–dimensionaler Vektorraum über einem Körper $\mathbb{K}$. Dann heißt jeder k – dimensionale Untervektorraum *linearer* (n,k) – *Code*.

Bem.: (n,k) bei linearen Codes bitte nicht mit (n,q) bei Paritätscodes verwechseln!

Beispiele: 1) $C = \{00, 11\}$ ist ein linearer $(2,1)$ – Code über $\mathbb{Z}_2$.

2) $C = \{000, 102, 201\}$ ist ein linearer $(3,1)$ – Code über $\mathbb{Z}_3$ mit Basis $\{102\}$.

Wichtiges *Beispiel*:
$C := \{x_1x_2x_3x_4y_1y_2y_3 \mid x_i, y_i \in \{0,1\},\ y_1 = x_2 + x_3 + x_4,\ y_2 = x_1 + x_3 + x_4,\ y_3 = x_1 + x_2 + x_4\}$ ist ein linearer $(7,4)$ – Code über $\mathbb{Z}_2$. (Beachte: Mit + ist jeweils die Modulo 2 Addition gemeint.)

Wir überprüfen an dieser Stelle nicht die Untervektorraumeigenschaft von C und halten lediglich fest, dass die Vektoren 1000011, 0100101, 0010110, 0001111 $\in C$ eine Basis bilden. Wegen $|C| = 2^4 = 16$ hat dieser Code die Informationsrate $\frac{4}{7}$.
Uns interessiert $d_{\min}(C)$. Hierzu müssen wir zum Glück nicht alle Codewörter miteinander vergleichen:

Satz 3.4.3 Sei C ein Linearcode über einem Vektorraum mit Nullvektor o. Dann gilt

$$d_{\min}(C) = \min\{d(c,o) \mid c \in C \backslash \{o\}\}$$

Beweis: In jedem Untervektorraum C gilt $x, y \in C \Rightarrow x - y \in C$. Wegen $d(x,y) = d(x-y, y-y) = d(x-y, o)$ liefert der Vergleich aller Vektoren aus C mit dem Nullvektor (Zähle die Stellen $\neq 0$) den gesuchten Minimalabstand $d_{\min}(C)$.

Beispiel: Der lineare $(7,4)$ – Code C hat den Minimalabstand $d_{\min}(C) = 3$.

Was leistet dieser Code? Er verbessert jeden Einzelfehler bei einer guten Informationsrate, denn zur Übermittlung von vier Zeichen x_i werden lediglich drei Kontrollzeichen y_i benötigt. Bei unserem Münzwurfbeispiel könnten wir mit einem Codewort die Ergebnisse von vier Würfen mit Hilfe von sieben Zeichen übermitteln und wären sicher, dass jeder Einzelfehler verbessert wird. Zum Vergleich: Der Paritätscode $C(3,2)$ benötigt für vier Würfe zwei Codewörter mit Gesamtlänge „nur“ 6, leistet aber keine Fehlerkorrektur. Ein dreifacher Wiederholungscode verbessert zwar auch jeden Einzelfehler, dies wird aber mit einer Codewortlänge von $n = 4 \cdot 3 = 12$ erkauft.

Wir wollen am Beispiel des linearen $(7,4)$ – Codes zeigen, wie man mit ein wenig Mathematikeinsatz effizient (de)codieren kann. Beachte, dass alle folgenden Additionen Modulo 2 – Rechnungen sind.

Beispiel: Die Nachricht 0101 soll mit dem linearen $(7,4)$ – Code codiert werden: Es müssen (modulo 2) $y_1 = x_2 + x_3 + x_4 = 1 + 0 + 1 = 0$, $y_2 = x_1 + x_3 + x_4 = 0 + 0 + 1 = 1$, $y_3 = x_1 + x_2 + x_4 = 0 + 1 + 1 = 0$ berechnet werden, das zu versendende Codewort ist 0101010. Mit Hilfe der Matrizenrechnung kann das Codewort auch so ermittelt werden:

$$(0\,1\,0\,1) \circ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} = (0\,1\,0\,1\,0\,1\,0)$$

Allgemein gilt: Sei C ein linearer (n,k) – Code über einem Körper $\mathbb{K}$, die Nachricht $x = x_1 \dots x_k$ soll codiert werden. Mit Hilfe einer $(k \times n)$ – *Generatormatrix* G, deren Zeilen die Basisvektoren des zugehörigen Unterraums sind, gilt

$$C = \{x \circ G \mid x \in \mathbb{K}^k\}$$

Einzelheiten zur Gestalt von G werden wir in der Vorlesung kennenlernen.

Schwieriger ist eine effiziente Decodierung. Hier hilft uns eine weitere Matrix, die sogenannte *Kontrollmatrix*. Es gilt

Satz 3.4.4 Zu jedem (n,k) – Linearcode existiert eine $((n-k)\times n)$ Matrix H mit $Hx^T = o \iff x \in C$.

Dieser Satz wird aus Zeitgründen nicht bewiesen, als Stichworte seien nur die Begriffe Skalarprodukt und orthogonaler Vektor genannt.

Beispiel: Die Kontrollmatrix in unserem $(7,4)$ – Code ist die (3×7) – Matrix

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

Genaues Hinsehen zeigt, dass die i-te Spalte dieser Matrix die Dualdarstellung der Zahl i ist, zum Beispiel gilt in Spalte 3 (von oben nach unten) $0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 3$.

Warum hilft diese Matrix bei der Decodierung? Wenn ein Codewort c gesendet wird und an der Stelle i ein Fehler passiert, bedeutet dies an dieser Stelle die Umwandlung von 0 in 1 oder umgekehrt, d.h., zu c wird ein Vektor f mit 1 an der Stelle i und sonst überall 0 addiert. Für das empfangene Wort y gilt also $y = c + f$ oder

$$Hy^T = H(c+f)^T = Hc^T + Hf^T = o + Hf^T = Hf^T$$

Falls genau ein Fehler passiert ist, liefert diese Rechnung die Stelle des Fehlers.

Beispiel: Es geht um die Nachricht 1011.

a) Es wird $y = 1011010$ empfangen. Der Decodieralgorithmus liefert

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

Wir haben den Nullvektor o erhalten, es ist $y \in C$ und wir vermuten $c = y$ mit zugehöriger Nachricht 1011.

b) Es wird $y = 1011110$ empfangen. Der Decodieralgorithmus liefert

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

Falls genau ein Fehler passiert ist, ist er an der fünften Stelle geschehen, es ist $f = 0000100$. Das vermutete Codewort ist $c = y - f = 1011010$ mit zugehöriger Nachricht 1011.

c) Es wird $y = 0111010$ empfangen. Der Decodieralgorithmus liefert

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

Falls genau ein Fehler passiert ist, ist er an der dritten Stelle geschehen. Das vermutete Codewort lautet 0101010 mit Nachricht 0101. Da leider zwei Fehler passiert sind, wird falsch decodiert.

Der $(7, 4)$ – Linearcode hat eine weitere für Codierungstheoretiker angenehme Eigenschaft.

Def 3.4.4 Sei V ein n–dimensionaler Vektorraum über einem Körper $\mathbb{K}$, sei $t \in \mathbb{N}$ und $d(y, z)$ der Hammingabstand von y und z. Dann heißt $K_t(z) := \{y \in V \mid d(y, z) \leq t\}$ *Kugel* vom Radius t um Mittelpunkt z.

Für ein Codewort c besteht die Kugel $K_1(c)$ aus allen Vektoren, die sich an höchstens einer Stelle von c unterscheiden, im Fall des Paritätscode $C = \{000, 011, 101, 110\}$ ist beispielsweise $K_1(000) = \{000, 001, 010, 100\}$.

Unser $(7, 4)$ – Linearcode erfüllt für Kugeln vom Radius $t = 1$

Beh. 1: $K_1(c) \cap K_1(c') = \emptyset \quad \forall\, c, c' \in C,\ c \neq c'$

Bew.: Folgt direkt aus $d_{\min}(C) = 3$.

Beh. 2: $\bigcup_{c \in C} K_1(c) = V$

Bew.: Zu jedem $c \in C$ gibt es genau sieben $y \in V$, die sich an genau einer Stellle von c unterscheiden, damit ist $|K_1(c)| = 1 + 7 = 2^3$. Aus Beh. 1 folgt $\left| \bigcup_{c \in C} K_1(c) \right| = \sum_{c \in C} |K_1(c)| = 2^4 \cdot 2^3 = 2^7 = |V|$.

Codes, die diese beiden Eigenschaften für ein beliebiges t erfüllen, nennt man *perfekt*. Perfekte Codes sind interessant, weil sie den Raum V gleichmäßig ausfüllen und weil die Decodierung für jedes $y \in V$ eindeutig erklärt ist, da es zu jeder möglichen Nachricht genau ein Codewort mit kleinstem Hammingabstand gibt.
In den Übungen werden wir einen perfekten Code über $\mathbb{Z}_3$ kennenlernen, der bis zu zwei Fehler korrigiert.

Beispiel: Bei einem Code der Länge 11 über $\mathbb{Z}_3$, der zwei Fehler erkennen kann, wird die Nachricht $y = y_1 \dots y_{11}$ empfangen. Zur Decodierung wird y mit der Kontrollmatrix H multipliziert.
Die Rechnung möge $Hy^T = Hf^T = 2 \cdot s_3 + 1 \cdot s_5$ ergeben, wobei s_i die i-te Spalte von H ist. Als erfahrener Decodierer weiß der Empfänger jetzt, dass mindestens zwei Fehler passiert sind. Ausgehend von dem günstigsten Fall (genau zwei Fehler) wird mit Hilfe des aus der Rechnung bestimmten Fehlervektors $f = 00201000000$ das gesuchte Codewort $c = y - f$ (modulo 3 gerechnet) bestimmt. Wenn diese Annahme stimmt, wird korrekt decodiert. Wenn mehr als zwei Fehler passiert sind, hat man Pech gehabt.