Codes and related combinatorial structures

Relinde Jurrius
(joint work with Henry Crapo, Ruud Pellikaan)

The Netherlands Defence Academy

PhD seminar, UGent
February 22, 2018
encoding

message → codeword → channel → received word → message

noise
Alphabet $Q$
Length $n$
Hamming metric on $Q^n$:

$$d(x, y) = \text{number of positions in which vectors differ}$$
$$= |\{i \in [n] : x_i \neq y_i\}|$$

error-correcting code: $C \subseteq Q^n$
minimum distance = \lfloor d - \frac{1}{2} \rfloor
\[ d \] minimum distance  
\[ e \] error-correcting capacity  
\[ = \left\lfloor \frac{d - 1}{2} \right\rfloor \]
A **linear code** $C$ is a subspace of $\mathbb{F}_q^n$.

- $n$ length of the code
- $k$ dimension of the code

A $k \times n$ matrix whose rows span $C$ is called a generator matrix.
Running example
The $[7, 4]$ Hamming code over $\mathbb{F}_2$ has generator matrix

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

It has length $n = 7$, dimension $k = 4$ and minimum distance $d = 3$. It can correct 1 error.
Some research directions:

- Given two of $n, k, d$; optimize the third
- (Non)existence, constructions
- Efficient decoding
- Bounds on the parameters
- Links with other combinatorial objects
A linear code

\[ \text{supp}(\mathbf{c}) = \text{coordinates of } \mathbf{c} \text{ that are non-zero} \]
\[ \text{wt}_H(\mathbf{c}) = \text{size of support} \]

Weight enumerator

\[ W_C(X, Y) = \sum_{w=0}^{n} A_w X^{n-w} Y^w \]

with \( A_w = \text{number of words of weight } w \).
Example

Hamming code generated by \( G = \begin{pmatrix}
1 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 & 1
\end{pmatrix} \).

Some words and weights:

<table>
<thead>
<tr>
<th>( c )</th>
<th>( \text{supp}(c) )</th>
<th>( \text{wt}_H(c) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (0 0 0 0 0 0 0 0 0) )</td>
<td>( \emptyset )</td>
<td>0</td>
</tr>
<tr>
<td>( (1 0 0 0 1 1 1 0) )</td>
<td>( {1, 5, 6} )</td>
<td>3</td>
</tr>
<tr>
<td>( (0 1 1 1 0 0 1) )</td>
<td>( {2, 3, 4, 7} )</td>
<td>4</td>
</tr>
</tbody>
</table>

The weight enumerator is equal to

\( W_C(X, Y) = X^7 + 7X^4Y^3 + 7X^3Y^4 + Y^7 \).
$J$ subset of $[n]$

$$C(J) = \{ c \in C : \text{supp}(c) \subseteq J^c \}$$

Lemma

$C(J)$ is a subspace of $\mathbb{F}_q^n$

$$\ell(J) = \dim_{\mathbb{F}_q} C(J)$$

Theorem

$\ell(J)$ gives a nice formula for the weight enumerator.
Example

Hamming code generated by \( G = \left( \begin{array}{ccccccc}
1 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 & 1
\end{array} \right) \).

Some calculations of \( C(J) \):

\[
\begin{array}{ccc}
J & J^c & C(J) \\
\{ n \} & \emptyset & 0 \\
\{1, 5, 6\} & \{2, 3, 4, 7\} & \left< 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0 \ 1 \right> \\
\{1\} & \{2, 3, 4, 5, 6, 7\} & \left< 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1 \ 1 \right> \\
\end{array}
\]
**Matroid:** a pair \((E, \mathcal{I})\) with

- \(E\) finite set;
- \(\mathcal{I} \subseteq 2^E\) family of subsets of \(E\), the *independent sets*, with:
  1. \(\emptyset \in \mathcal{I}\)
  2. If \(A \in \mathcal{I}\) and \(B \subseteq A\) then \(B \in \mathcal{I}\).
  3. If \(A, B \in \mathcal{I}\) and \(|A| > |B|\) then there is an \(a \in A \setminus B\) such that \(B \cup \{a\} \in \mathcal{I}\).

**Examples:**

- Set of vectors; independence = linear independence
- Set of edges of a graph; independence = cycle free
Example

\[
\begin{pmatrix}
0 & 0 & 0 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 0
\end{pmatrix}
\]

Example

But: most matroids don't come from a matrix or graph.
Matroids are everywhere: graphs, linear algebra, optimization, tropical geometry, hyperplane arrangements, topology, . . .

Some research directions:

► Does a matroid come from a graph?
► Does a matroid come from a set of vectors? Over which field?
► How many matroids are there?
► Study concrete class of matroids
► Links with other combinatorial objects
A matroid is also a pair \((E, r)\) with

- \(E\) finite set;
- \(r : 2^E \to \mathbb{N}_0\) a function, the *rank function*, with for all \(A, B \in E\):
  - (r1) \(0 \leq r(A) \leq |A|\)
  - (r2) If \(A \subseteq B\) then \(r(A) \leq r(B)\).
  - (r3) \(r(A \cup B) + r(A \cap B) \leq r(A) + r(B)\) (semimodular)

\(r(A)\) = size of largest independent set contained in \(A\)

\(\mathcal{I} = \{\text{subsets whose size is equal to their rank}\}\)
Example

\[
\begin{pmatrix}
0 & 0 & 0 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 \\
\end{pmatrix}
\]

Example
Fact: a linear code gives a matroid with
\[ E = \text{index set for columns of generator matrix} \]
\[ r(J) = \text{dimension of subspace spanned by vectors indexed by } J \]

Recall: \[ \ell(J) = \dim\{c \in C : \text{supp}(c) \subseteq J^c\} \]

Theorem
\[ r(J) = \dim C - \ell(J) \]

Idea of proof: \( 0 \to C(J) \to C \to C_J \to 0 \) is an exact sequence.
Comercial break
$q$-Analogues

Finite set $\longrightarrow$ finite dimensional vectorspace over $\mathbb{F}_q$

Example

$$\binom{n}{k} = \text{number of sets of size } k \text{ contained in set of size } n$$

$$\left[ \begin{array}{c} n \\ k \end{array} \right]_q = \text{number of } k\text{-dim subspaces of } n\text{-dim vectorspace over } \mathbb{F}_q$$
q-Analogues

Example

t-(v, k, λ) design: pair (X, B) with

- X set with v elements (points)
- B family of subsets of X of size k (blocks)
- Every t-tuple of points is contained in exactly λ blocks

t-(v, k, λ; q) q-design: pair (X, B) with

- X v-dim vectorspace over \( \mathbb{F}_q \)
- B family of k-dim subspaces of X (blocks)
- Every t-dim subspace is contained in exactly λ blocks
**$q$-Analogues**

<table>
<thead>
<tr>
<th>finite set</th>
<th>finite space $\mathbb{F}^n_q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>element</td>
<td>1-dim subspace</td>
</tr>
<tr>
<td>size</td>
<td>dimension</td>
</tr>
<tr>
<td>$n$</td>
<td>$\frac{q^n-1}{q-1}$</td>
</tr>
<tr>
<td>intersection</td>
<td>intersection</td>
</tr>
<tr>
<td>union</td>
<td>sum</td>
</tr>
<tr>
<td>complement</td>
<td>?</td>
</tr>
</tbody>
</table>

From $q$-analogue to ‘normal’: let $q \to 1$. 
Candidates for complement $A^c$ of $A \subseteq \mathbb{F}_q^n$:

- All vectors outside $A$
  
  But: not a space

- Orthogonal complement
  
  But: $A \cap A^\perp$ can be nontrivial

- Quotient space $\mathbb{F}_q^n/A$
  
  But: changes ambient space

- Subspace such that $A \oplus A^c = \mathbb{F}_q^n$
  
  But: not unique
Network coding

![Diagram of network coding]
Network coding

Idea: send (rows of) matrices instead of vectors

Better idea: send (bases of) subspaces instead of matrices
Codewords are vectors:
   ‘Ordinary’ error-correcting codes

Codewords are matrices:
   Rank metric codes
   \( q \)-analogue of ‘ordinary’ codes

Codewords are subspaces:
   Subspace codes
   Constant dimension, constant weight: \( q \)-design
Rank metric codes

$\mathbb{F}_{q^m}/\mathbb{F}_q$ finite field extension with basis $\alpha_1, \ldots, \alpha_m$.

Write $c = c_1\alpha_1 + \cdots + c_m\alpha_m$.

\[
\begin{array}{c}
c_1 \\
c_2 \\
\vdots \\
c_m
\end{array}
\leftrightarrow
\begin{array}{c}
\mathbf{c}
\end{array}
\]

$m(\mathbf{x}) \in \mathbb{F}_{q^m}^{m \times n} \quad \mathbf{x} \in \mathbb{F}_q^n$

**Rank metric code** is subspace of $\mathbb{F}_{q^m}^n \rightarrow$ subspace of $\mathbb{F}_q^{m \times n}$. 
Running example

Let $\mathbb{F}_8/\mathbb{F}_2$ with basis $(1, \alpha, \alpha^2)$, where $\alpha^3 = \alpha + 1$.

Let $C$ be the code generated by

\[
G = \begin{pmatrix} 1 & \alpha & 0 & 0 \\ 0 & 1 & \alpha & 0 \end{pmatrix}.
\]

Some words and matrices:

<table>
<thead>
<tr>
<th>$c$</th>
<th>$(0,0,0,0)$</th>
<th>$(1,\alpha,0,0)$</th>
<th>$(1,\alpha^2,\alpha^5,0)$</th>
</tr>
</thead>
</table>
| $m(c)$    | \[
\begin{pmatrix} 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\] | \[
\begin{pmatrix} 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 \\
\end{pmatrix}
\] | \[
\begin{pmatrix} 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 \\
\end{pmatrix}
\] |
A linear code

\[ \text{supp}(c) = \text{coordinates of } c \text{ that are non-zero} \]
\[ \text{wt}_H(c) = \text{size of support} \]

**Weight enumerator**

\[ W_C(X, Y) = \sum_{w=0}^{n} A_w X^{n-w} Y^w \]

with \( A_w = \text{number of words of weight } w \).
A rank metric code

\[ \text{Rsupp}(c) = \text{row space of } m(c) \]
\[ \text{wt}_R(c) = \text{dimension of support} \]

Rank weight enumerator

\[ W_C^R(X, Y) = \sum_{w=0}^{n} A_w^R X^{n-w} Y^w \]

with \( A_w^R \) = number of words of rank weight \( w \).
Example

$C$ in $\mathbb{F}_8/\mathbb{F}_2$ generated by $G = \begin{pmatrix} 1 & \alpha & 0 & 0 \\ 0 & 1 & \alpha & 0 \end{pmatrix}$.

Some words and weights:

<table>
<thead>
<tr>
<th>$c$</th>
<th>$(0, 0, 0, 0)$</th>
<th>$(1, \alpha, 0, 0)$</th>
<th>$(1, \alpha^2, \alpha^5, 0)$</th>
</tr>
</thead>
</table>
| $m(c)$       | \[
\begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\] | \[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\] | \[
\begin{pmatrix}
1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\] |
| $R_{supp}(c)$  | $0$            | $\langle 1 \ 0 \ 0 \ 0 \rangle$ | $\langle 1 \ 0 \ 0 \ 0 \rangle$ |
| $\text{wt}_R(c)$ | $0$            | $2$                 | $3$                 |

There are no words of weight 1 or weight 4.
Lemma
$C(J)$ is a subspace of $\mathbb{F}_q^n$

$\ell(J) = \dim_{\mathbb{F}_q} C(J)$

Theorem
$\ell(J)$ gives a nice formula for the weight enumerator.
Theorem
\( \ell(J) \) gives a nice formula for the rank weight enumerator.
Example

$C$ in $\mathbb{F}_8/\mathbb{F}_2$ generated by $G = \begin{pmatrix} 1 & \alpha & 0 & 0 \\ 0 & 1 & \alpha & 0 \end{pmatrix}$.

Some calculations of $C(J)$:

<table>
<thead>
<tr>
<th>$J$</th>
<th>$J^\perp$</th>
<th>$C(J)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{F}_2^4$</td>
<td>$0$</td>
<td>$0$</td>
</tr>
<tr>
<td>dim = 3</td>
<td>dim = 1</td>
<td></td>
</tr>
<tr>
<td>$\langle 1 \ 0 \ 0 \ 0 \rangle$</td>
<td>$\langle 0 \ 0 \ 1 \ 0 \rangle$</td>
<td>$0$</td>
</tr>
<tr>
<td>$\langle 0 \ 1 \ 0 \ 0 \rangle$</td>
<td>$\langle 0 \ 0 \ 0 \ 1 \rangle$</td>
<td></td>
</tr>
<tr>
<td>$\langle 0 \ 0 \ 1 \ 0 \rangle$</td>
<td>$\langle 1 \ 0 \ 0 \ 0 \rangle$</td>
<td>$\langle 1 \ \alpha \ 0 \ 0 \rangle$</td>
</tr>
<tr>
<td>$\langle 0 \ 0 \ 0 \ 1 \rangle$</td>
<td>$\langle 0 \ 1 \ 0 \ 0 \rangle$</td>
<td></td>
</tr>
</tbody>
</table>

If dim $J = 2$ and $J^\perp \not\subseteq \langle 0 \ 0 \ 0 \ 1 \rangle^\perp$ then $C(J) = 0$. 
Recall: for linear codes and their associated matroids, we have proven \( r(J) = \dim C - \ell(J) \).

**Theorem**

Let \( r(J) = \dim C - \ell(J) \) for a rank metric code \( C \). Then \( r(J) \) satisfies:

1. \( 0 \leq r(A) \leq \dim A \)  
2. If \( A \subseteq B \) then \( r(A) \leq r(B) \).
3. \( r(A + B) + r(A \cap B) \leq r(A) + r(B) \) (semimodular)

This leads to the definition of the rank function of a \( q \)-matroid.
q-Matroid: a pair \((E, r)\) with

- \(E\) finite dimensional vector space;
- \(r : \{\text{subspaces of } E\} \rightarrow \mathbb{N}_0\) a function, the rank function, with for all \(A, B \subseteq E\):
  
  (r1) \(0 \leq r(A) \leq \dim A\)
  
  (r2) If \(A \subseteq B\) then \(r(A) \leq r(B)\).
  
  (r3) \(r(A + B) + r(A \cap B) \leq r(A) + r(B)\) (semimodular)
Example

$C$ in $\mathbb{F}_8/\mathbb{F}_2$ generated by $G = \begin{pmatrix} 1 & \alpha & 0 & 0 \\ 0 & 1 & \alpha & 0 \end{pmatrix}$.

Some calculations of $r(J)$:

<table>
<thead>
<tr>
<th>$J$</th>
<th>$C(J)$</th>
<th>$r(J)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{F}_2^4$</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$\dim = 3$</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$\langle 1 \ 0 \ 0 \ 0 \rangle$</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$\langle 0 \ 1 \ 0 \ 0 \rangle$</td>
<td>$\langle 1 \ \alpha \ 0 \ 0 \rangle$</td>
<td>1</td>
</tr>
<tr>
<td>$\langle 0 \ 0 \ 1 \ 0 \rangle$</td>
<td>$\langle 0 \ 0 \ 0 \ 1 \rangle$</td>
<td></td>
</tr>
</tbody>
</table>

If $\dim J = 2$ and $J^\perp \not\subseteq \langle 0 \ 0 \ 0 \ 1 \rangle^\perp$ then $r(J) = 2$.

These are exactly the bases of the $q$-matroid.
Theorem

A q-matroid is a pair \((E, \mathcal{I})\) with

- \(E\) finite dimensional vector space;
- \(\mathcal{I}\) family of subspaces of \(E\), the independent spaces, with:
  1. \(0 \in \mathcal{I}\).
  2. If \(J \in \mathcal{I}\) and \(I \subseteq J\), then \(I \in \mathcal{I}\).
  3. If \(I, J \in \mathcal{I}\) with \(\dim I < \dim J\), then there is some 1-dimensional subspace \(x \subseteq J\), \(x \not\subseteq I\) with \(I + x \in \mathcal{I}\).
  4. Let \(A, B \subseteq E\) and let \(I, J\) be maximal independent subspaces of \(A\) and \(B\), respectively. Then there is a maximal independent subspace of \(A + B\) that is contained in \(I + J\).

Note the extra axiom!
Other things that need a $q$-analogue:

- More axiom sets for matroids
- Relation between weight enumerator and Tutte polynomial
- Perfect matroid designs
- Graphs (??)

- How about rank metric cods that are not $\mathbb{F}_{q^m}$-linear?