Rank-metric codes and applications

Ernst M. Gabidulin
Moscow Institute of Physics and Technology (State University)
Email: ernst.gabidulin@gmail.com

Abstract
The State of Art for rank codes is represented. The theory and applications are considered.

1 Rank Codes

1.1 Definition
There exist two representations of Rank codes: matrix representation and vector representation.
In matrix representation, rank codes are defined as subsets of a normed space \( \{ \mathbb{F}_{q}^{N \times n}, \text{Rk} \} \) of \( N \times n \) matrices over a finite (base) field \( \mathbb{F}_{q} \), where the norm of a matrix \( M \in \mathbb{F}_{q}^{N \times n} \) is defined to be the algebraic rank \( \text{Rk}(M) \) of this matrix over \( \mathbb{F}_{q} \). The rank distance between two matrices \( M_1 \) and \( M_2 \) is the rank of their difference \( \text{Rk}(M_1 - M_2) \). The rank distance of a matrix rank code \( M \subset \mathbb{F}_{q}^{N \times n} \) is defined as the minimal pairwise distance:
\[
d(M) = d = \min(\text{Rk}(M_i - M_j) : M_i, M_j \in M, \ i \neq j).
\]
In vector representation, rank codes are defined as subsets of a normed \( n \)-dimensional space \( \{ \mathbb{F}_{q}^{n}, \text{Rk} \} \) of \( n \)-vectors over an extension field \( \mathbb{F}_{q^N} \), where the norm of a vector \( v \in \mathbb{F}_{q^N} \) is defined to be the column rank \( \text{Rk}(v | \mathbb{F}_{q}) \) of this vector over \( \mathbb{F}_{q} \), i.e., the maximal number of coordinates of \( v \) which are linearly independent over the base field \( \mathbb{F}_{q} \). The rank distance between two vectors \( v_1, v_2 \) is the column rank of their difference \( \text{Rk}(v_1 - v_2 | \mathbb{F}_{q}) \). The rank distance of a vector rank code \( V \subset \mathbb{F}_{q^N}^{n} \) is defined as the minimal pairwise distance:
\[
d(V) = d = \min(\text{Rk}(v_i - v_j) : v_i, v_j \in V, \ i \neq j).
\]

1.2 Background
Algebraic coding theory may be considered as the theory of subsets of a certain normed finite-dimensional space \( \Gamma \) over the finite field equipped with a norm function \( \mathcal{R} \). The most known norm in coding theory is the Hamming weight of a vector. It turns out that the rank function \( \text{Rk}(A) \) of matrices \( A \) over fields can be considered as the norm function. In particular, the well-known inequalities for sums of matrices \( |\text{Rk}(A) - \text{Rk}(B)| \leq \text{Rk}(A+B) \leq \text{Rk}(A) + \text{Rk}(B) \) hold.
Rk(A) + Rk(B) define implicitly the rank distance relations on the space of all matrices of identical size. Explicitly, the concept of the rank metric was introduced by Loo-Keng Hua [1] as "Arithmetic distance". Philippe Delsarte [2] defined the rank distance (or, q-distance) on the set of bilinear forms (equivalently, on the set of rectangular matrices) and proposed the construction of optimal codes in bilinear form representation. Ernst M. Gabidulin [3] introduced the rank distance for vector spaces over extension fields and found connections between rank codes in the vector representation and in the matrix representation. Optimal codes in vector representation were described. Fast coding and decoding algorithms were proposed for optimal codes.

2 Theory

The normed spaces \( \{ \mathbb{F}_q^{N \times n}, \operatorname{Rk} \} \) and \( \{ \mathbb{F}_q^n, \operatorname{Rk} \} \) are isomorphic isometrically. Let a basis \( \Omega = \{ \omega_1, \omega_2, \ldots, \omega_N \} \) of \( \mathbb{F}_q^N \) over \( \mathbb{F}_q \) be chosen. Then each vector \( v = [v_1 \ v_2 \ \ldots \ v_n] \in \mathbb{F}_q^n \) can be mapped into the \( N \times n \) matrix \( M \in \mathbb{F}_q^{N \times n} \) by replacing each coordinate \( v_j \) with the \( N \)-column consisting of coefficients in representing \( v_j \) by the basis \( \Omega \). This mapping is bijective and isometric.

Given a rank code \( M \) in matrix representation one can construct a rank code \( V \) in vector representation with the same size, code distance and pairwise distances, and vice versa. The size \( |M| = |V| \) of related codes with code distance \( d \) satisfy the Singleton bound \( |M| = |V| \leq \min(q^{N(n-d+1)}, q^N - d + 1) \). Codes reaching this bound are called maximum rank distance codes, or, MRD codes.

A rank code \( M \) in matrix representation is called \( \mathbb{F}_q \)-linear if \( M \) is a subspace of \( \mathbb{F}_q^{N \times n} \). A rank code \( V \) in vector representation is called \( \mathbb{F}_q^N \)-linear if \( V \) is a subspace of \( \mathbb{F}_q^N \).

Mapping a \( \mathbb{F}_q^N \)-linear code \( V \) in vector representation into related code \( M \) in matrix representation results in a \( \mathbb{F}_q \)-linear code.

Mapping a \( \mathbb{F}_q^N \)-linear code \( M \) in matrix representation into related code \( V \) in vector representation results in not necessary a \( \mathbb{F}_q^N \)-linear code.

Constructions of \( \mathbb{F}_q \)-linear rank codes in the matrix representation and \( \mathbb{F}_q^N \)-linear rank codes in the vector representation will be considered.

2.1 Delsarte’s optimal rank codes in matrix representation

Delsarte’s construction of rank codes in bilinear form representation is presented here in matrix representation.

Assume that \( n \leq N \). Let \( \operatorname{Tr}(x) = \sum_{i=0}^{N-1} x^q^i, \ x \in \mathbb{F}_q^N \), be the Trace function from \( \mathbb{F}_q^N \) into \( \mathbb{F}_q \). Let \( d \) be an integer in \( \{1, 2, \ldots, n\} \). Let \( u = [u_0 \ u_1 \ \ldots \ u_{n-d}] \in \mathbb{F}_q^{n-d+1} \). Let \( \mu_1, \mu_2, \ldots, \mu_n \) be linearly independent elements of \( \mathbb{F}_q^N \). Let \( \Omega = \{ \omega_1, \omega_2, \ldots, \omega_N \} \) be a basis for \( \mathbb{F}_q^N \).

Define a code in matrix representation as the set of \( N \times n \) matrices \( M = \{ M(u) = [M_{ij}(u)] : u \in \mathbb{F}_q^{n-d+1} \} \).
where

\[ M_{ij}(u) = \text{Tr} \left( \sum_{s=0}^{n-d} u_s \omega_i \mu_q^s \right). \]

Then \( \mathcal{M} \) is a rank code with code distance \( d \) reaching the Singleton bound \( |\mathcal{M}| = q^{N(n-d+1)} \). Let \( A_i(n, d) \), \( i = 0, 1, \ldots, n \), be the number of code matrices with rank \( i \). The weight distribution is as follows:

\[ A_0(n, d) = 1, \quad A_i(n, d) = 0, \quad i = 1, \ldots, d - 1. \]

If \( i = d, \ldots, n \), where \( \left[ \frac{n}{i} \right] = \prod_{j=0}^{i-1} \frac{q^n - q^j}{q^i - q^j} \) is the Gaussian binomial coefficient.

2.2 Optimal rank codes in vector representation

A \( F_{q^N} \)-linear vector code \( V \) is a subspace of the normed space \( \{ F_{q^N}^n, \text{Rk} \} \). Denote by \( (n, k, d) \) a code \( V \) of dimension \( k \leq n \) and rank distance \( d \). Such a code can be described in terms of a full rank generator matrix \( G_k \) over the extension field \( F_{q^N} \) of size \( k \times n \). Code vectors \( \{v\} \) are all linear combinations of this matrix. Thus the size of a code is equal to

\[ |V| = q^{nk}. \]

Equivalently, a rank code \( V \) can be described in terms of a full rank parity-check matrix \( H_{n-k} \) over \( F_{q^N} \) of size \( (n-k) \times n \). It satisfies the condition \( G_k H_{n-k}^\top = \mathbf{0} \), where \( \mathbf{0} \) is the all zero \( k \times (n-k) \) matrix. Code vectors \( \{v\} \) are all solutions of the linear system of equation \( v H_{n-k}^\top = \mathbf{0} \).

For optimal (MRD) codes, it must be \( k = n - d + 1 \), or, \( n - k = d - 1 \).

General constructions of MRD codes in terms of parity-check matrices can be described as follows. Let \( h_1, h_2, \ldots, h_n \) be a set of elements from the extension field \( F_{q^n} \) linearly independent over the base field \( F \). Let \( s \) be a positive integer such that \( \gcd(s, N) = 1 \).

Then a parity matrix of the form

\[
H_{d-1} = \begin{bmatrix}
    h_1 & h_2 & \ldots & h_n \\
    h_1^q & h_2^q & \ldots & h_n^q \\
    h_1^{q^2} & h_2^{q^2} & \ldots & h_n^{q^2} \\
    \vdots & \vdots & \ddots & \vdots \\
    h_1^{q^{(d-2)s}} & h_2^{q^{(d-2)s}} & \ldots & h_n^{q^{(d-2)s}}
\end{bmatrix}
\]

defines an MRD \( (n, k, d) \) code with code length \( n \leq N \), dimension \( k = n - d + 1 \) and rank distance \( d = n - k + 1 \).

Equivalently, general constructions of MRD codes can be described in terms of generator matrices. Let \( g_1, g_2, \ldots, g_n \) be a set of elements from the extension field \( F_{q^n} \) linearly
independent over the base field $F$. Then a generator matrix of the form

$$G_k = \begin{bmatrix}
g_1 & g_2 & \ldots & g_n \\
g_1^{q^s} & g_2^{q^s} & \ldots & g_n^{q^s} \\
g_1^{q^{s_1}} & g_2^{q^{s_1}} & \ldots & g_n^{q^{s_1}} \\
\vdots & \vdots & \ddots & \vdots \\
g_1^{q^{(k-1)s}} & g_2^{q^{(k-1)s}} & \ldots & g_n^{q^{(k-1)s}}
\end{bmatrix}.$$ 

defines an MRD $(n, k, d)$ code with code length $n \leq N$, dimension $k = n - d + 1$ and rank distance $d = n - k + 1$. The weight distribution of vector MRD codes coincides for a given $d$ with the weight distribution of Delsarte’s codes above.

The case $s = 1$ is used mostly.
No other constructions of MRD codes are known (2009).

### 2.3 Correcting rank errors and rank erasures

Let a MRD $(n, k, d = n - k + 1)$ code $V$ be given. Let a transmitted signal be $v$ and received signal be $y = v + e_{\text{total}}$, where $e_{\text{total}}$ is an error. The code $V$ can correct in general vector errors of the form

$$e_{\text{total}} = e + e_{\text{row}} + e_{\text{col}} = e_1 u_1 + e_2 u_2 + \ldots + e_t u_t + a_1 r_1 + a_2 r_2 + \ldots + a_v r_v + w_1 c_1 + w_2 c_2 + \ldots + w_l c_l$$

provided that $2t + v + l \leq d - 1$.

The part $e = e_1 u_1 + e_2 u_2 + \ldots + e_t u_t$ is called a random rank error of rank $t$ under assumption that elements $e_i \in F_{q^N}$ are linearly independent over the base field $F_q$ and unknown to the decoder; $n$-vectors $u_1, u_2, \ldots, u_t$ have coordinates in the base field $F_q$, are linearly independent over the base field $F_q$ and also unknown to the decoder. The rank $t$ is unknown to the decoder.

The part $e_{\text{row}} = a_1 r_1 + a_2 r_2 + \ldots + a_v r_v$ is called a vector rank row erasure with side information under assumption that elements $a_i \in F_{q^N}$ are linearly independent over the base field $F_q$ and known to the decoder; $n$-vectors $r_1, r_2, \ldots, r_v$ have coordinates in the base field $F_q$, are linearly independent over the base field $F_q$ and unknown to the decoder.

The part $e_{\text{col}} = w_1 c_1 + w_2 c_2 + \ldots + w_l c_l$ is called a vector rank column erasure with side information under assumption that elements $w_i \in F_{q^N}$ are linearly independent over the base field $F_q$ and are unknown to the decoder; $n$-vectors $c_1, c_2, \ldots, c_l$ have coordinates in the base field $F_q$, are linearly independent over the base field $F_q$ and known to the decoder.

First fast correcting random rank errors only was proposed in [3]. The algorithm is based on the extended Euclidean division algorithm for linearized polynomials. There exist several further modifications.

Algorithms for correcting random rank errors and rank erasures simultaneously are proposed in [4], [5], [7].
3 Applications

3.1 Rank codes as space-time codes

Space-time codes are introduced by Tarokh, Jafarkhani, and Calderbank in 1998 [10]. Codes are designed to simultaneously take advantage of two dimensions, namely the spatial diversity of antenna elements, and coding gain introduced by designed redundancy in the time dimension.

It is assumed that the base station is equipped with $T$ transmit antennas and the terminal is equipped with $m$ receiving antennas. A signal is transmitted in time slots $1, 2, \ldots, n$.

The received signal can be written as a $m \times n$ matrix

$$Y = AC + N.$$ 

$C$ is a $T \times n$ signal code matrix with entries in some constellation. $A$ is a $m \times T$ is a matrix of complex transfer coefficients from the $j$th transmit antenna to the $i$th receiving antenna. $N$ is an AWGN $m \times n$ matrix.

The Full rank criterion was proposed:

Choose a matrix code $C = \{C_i\}$ in such a manner that the difference $C_{i,j} = \{C_i - C_j\}$ has full rank.

$MRD$ codes over finite fields can not be used directly as space-time codes. However, it is still possible to use $MRD$ codes over finite fields as templates.

Non-constructive statement. There exist infinitely many of finite complex or real constellations such that differential $MRD$ codes (i.e. full rank of difference matrices) can be constructed.

Constructive statement. $MRD$ codes over the binary field can be transformed to differential $MRD$ codes over the complex or real constellations of size 2.

Constructive statement. For $T = 2, 4, 8$ and $q = 2^s$, $MRD$ codes over $\mathbb{F}_q$ can be transformed to differential $MRD$ codes over complex constellations of size $q$.

Constructive statement. $MRD$ codes over $\mathbb{F}_p$, $p \equiv 1 \mod 4$, can be transformed to differential $MRD$ codes over the complex Gaussian field of size $p$.

Decoding differential $MRD$ codes is reduced to decoding in finite field by hard-decision algorithms.

Problem: For known matrix $A$, use Gauss elimination in $A$ for getting rank erasures instead of rank errors.

3.2 Rank codes in network coding

Consider a communication network, where a single source transmits information to a single destination. The model of a network was proposed and investigated in [4]. The source
formats the information to be transmitted into $N$ packets $X(1), \ldots, X(N)$ of length $N+n$ over the finite field $\mathbb{F}_q$ and constructs a $(N \times (N + n))$ matrix $X$ with these packets as rows. The source chooses a code $\mathcal{X}$ consisting of matrices $X$ which can be transmitted.

Each intermediate node calculates random linear combinations of ingoing packets, where a packet is represented as an element of a finite field $\mathbb{F}_q$. The node retransmits randomly calculated packets. Therefore, the destination collects a random number $N_r$ of packets $Y(1), \ldots, Y(N_r)$ of length $N + n$ and creates a $N_r \times (N + n)$ matrix $Y$.

The problem is to recover the original packets $X(1), \ldots, X(N)$, or the matrix $X$ from the received matrix $Y$.

Koetter and Kschischang [6] introduce the concept of subspace codes when an alphabet of transmitted messages consists of subspaces, not symbols. They constructed a family of subspace codes. Silva, Kschischang, Koetter, proposed a rank-metric approach to error control in random network coding [7]. Silva and Kschischang proposed their algorithms of fast encoding and decoding of Gabidulin codes [8].

The problem is to recover the original packets $X(1), \ldots, X(N)$, or the matrix $X$ from the received matrix $Y$.

Koetter and Kschischang [6] introduce the concept of subspace codes when an alphabet of transmitted messages consists of subspaces, not symbols. They constructed a family of subspace codes. Silva, Kschischang, Koetter, proposed a rank-metric approach to error control in random network coding [7]. Silva and Kschischang proposed their algorithms of fast encoding and decoding of Gabidulin codes [8].

The basic model of a channel induced by random network coding is described as follows. The transmitted matrix $X$ and the received matrix $Y$ are connected by the relation $Y = AX + BZ$, where $A$ is an $N_r \times N$ matrix corresponding to the overall linear transformation applied by intermediate nodes of the network; $Z$ is an $l \times (N + n)$ matrix whose rows are the error packets $z_1, \ldots, z_l$; $B$ is an $N_r \times l$ matrix corresponding to the overall linear transformation applied to $z_1, \ldots, z_l$ on route to the destination. The number of nonzero rows of $Z$ gives the total number of corrupt packets injected in the network. Random matrices $A, B, Z$ are unknown to the destination.

It is proposed to apply so called lifting construction for constructing a code $\mathcal{X}$. Each matrix $X \in \mathcal{X}$ has the form $X = [I_N \ M]$, where $I_N$ is the identity matrix of order $N$ while $M \in \mathcal{M}$ is a code matrix of some matrix code $\mathcal{M}$ consisting of $N \times n$ matrices over the field $\mathbb{F}_q$. A code $\mathcal{M}$ is assumed to be a MRD rank code with rank distance $d$, if $n \leq N$, or a transposed MRD rank code, if $N < n$. The corresponding code $\mathcal{X}$ were analyzed in The basic model of a channel induced by random network coding is described as follows. The transmitted matrix $X$ and the received matrix $Y$ are connected by the relation $Y = AX + BZ$, where $A$ is an $N_r \times N$ matrix corresponding to the overall linear transformation applied by intermediate nodes of the network; $Z$ is an $l \times (N + n)$ matrix whose rows are the error packets $z_1, \ldots, z_l$; $B$ is an $N_r \times l$ matrix corresponding to the overall linear transformation applied to $z_1, \ldots, z_l$ on route to the destination. The number of nonzero rows of $Z$ gives the total number of corrupt packets injected in the network. Random matrices $A, B, Z$ are unknown to the destination.

It is proposed in [7] to apply so called lifting construction for constructing a code $\mathcal{X}$. Each matrix $X \in \mathcal{X}$ has the form $X = [I_N \ M]$, where $I_N$ is the identity matrix of order $N$ while $M \in \mathcal{M}$ is a code matrix of some matrix code $\mathcal{M}$ consisting of $N \times n$ matrices over the field $\mathbb{F}_q$. A code $\mathcal{M}$ is assumed to be a MRD rank code with rank distance $d$, if $n \leq N$, or a transposed MRD rank code, if $N < n$. The corresponding code $\mathcal{X}$ were analyzed in [7]. Decoding codes $\mathcal{X}$ can be reduced to decoding embedded rank codes $\mathcal{M}$. If the matrix $X = [I_N \ M]$ is transmitted, then the matrix $Y = AX + BZ = [A + BZ_1 \ AM + BZ_2]$ is received with unknown to the destination matrices $A, B, Z_1, Z_2$. By a linear transformation
of rows and injecting all zero rows, the part $[A + BZ_1]$ can be reduced to the upper triangular matrix of order $N$. Elements of the main diagonal are "0"'s or "1"'s. The number of "1"'s is equal to the rank of $[A + BZ_1]$. The same operations over the matrix $[AM + BZ_2]$ allows to extract the submatrix of the form $R = M + LM + DC$, where $R$, $L$ and $C$ are known matrices. Thus, the result is a matrix $M$ of the rank code $M$ corrupted by a row rank erasure $LM$ and a column rank erasure $DC$. The unknown matrix $M$ can be uniquely recovered from $R$ provided that $\text{Rk}(L) + \text{Rk}(C) \leq d - 1$.

Other network codes are known generalizing constructions above (see, GabBoss:2009). It is shown that decoding this subspace code is equivalent to correcting random errors and generalized erasures in the rank code [5].

3.3 Rank codes in cryptography

The McEliece like public key cryptosystem but based on rank error correcting codes was proposed by Gabidulin, Paramonov, Tretjakov in 1991 [?]. The cryptosystem is described as follows. The public key $G_{\text{pub}}$ is a left and right scrambled generator matrix of a rank code:

$$G_{\text{pub}} = SG_kP.$$  

The matrix $G_k$ is used to correct rank errors of rank not greater than $t = \left\lfloor \frac{n-k}{2} \right\rfloor$. A matrix $S$ over the extension field $\mathbb{F}_{q^n}$ is called the row scrambling matrix. It is used to destroy any visible structure of the matrix $G_k$ by mixing its rows.

A matrix $P = \begin{bmatrix} p_{ij} \end{bmatrix}$ is called the column scrambler. This matrix is a non singular square matrix of order $n$. It is used to mix columns of $G_k$.

Another generator matrix has the form

$$G_{\text{pub}} = S \begin{bmatrix} X & G_k \end{bmatrix} P.$$  

Plaintext is any $k$-vector

$$m = (m_1, m_2, \ldots, m_k), m_s \in \mathbb{F}_{q^n}, \ s = 1, 2, \ldots, k.$$  

The Private keys are matrices $S, G_k, X, P$ separately and (explicitly) a fast decoding algorithm of an MRD code.
Encryption. The ciphertext is given by
\[ c = mG_{pub} + e = mS[X|G_k]P + e, \]
where \( e \) is an artificial vector of errors of rank \( t_2 \).

Decryption. The legitimate receiver upon receiving \( c \) calculates
\[ c' = cP^{-1} = mS[X|G_k] + eP^{-1}. \]
Then he extracts from \( c' \) the plaintext \( m \) using decoding algorithms and properties of public keys.

Attacks and counter-attacks

Rank codes are well structured. It makes easier creation of attacks.
Subsequently in a series of works, Gibson [12], [13], [14] developed attacks that break the system for practical instances.
Several variants of PKC were introduced to withstand Gibson’s attacks [15].
Recently, R. Overbeck [16], [17] proposed a new attack which is more effective than Gibson’s attacks. His method is based on the fact that a column scrambler \( P \) is defined over the base field.

It was found [18], [19] that a cryptographer can define a proper column scrambler over the extension field without violation of the standard mode of the PKC. Overbeck’s attack fails in this case.

Conclusion

- Theory of rank codes is of great interest for many researchers.
- Fast decoding algorithms are developed.
- Applications in many areas are possible and recommended. Space-time coding, random network coding, public key cryptosystems are areas of such applications.

References


