Review/Outline

Recall: Looking for good codes

High info rate vs. high min distance
Want simple description, too
Linear, even cyclic, plausible
Gilbert-Varshamov bound for linear codes

Check matrix criterion for min dist

Convert to linear algebra issue
Proof of Gilbert-Varshamov bound

Vandermonde determinants
Primitive roots in $\mathbf{Z}/p$
Reed-Solomon (RS) codes

Check matrix criterion for min dist

Keep in mind that **a code with minimum distance $2e + 1$ can correct e errors.** The following fact is the starting point for constructions of linear codes.

Theorem: Let C be a *linear* code with check matrix H. Let d be the largest integer such that any d of the **columns** of H are linearly independent. Then C has minimum distance $d + 1$. (And, conversely, if C has minimum distance $d + 1$ then any d columns of H are linearly independent.)

Remark: An important point is that Hamming distance $d(,)$ is **translation invariant**:

$$d(x, y) = d(x + z, y + z)$$

In particular, $d(x, y) = d(x - y, 0)$. For a linear code with codewords x, y, $x - y$ is also a codeword. Thus, the minimum distance between codewords is the minimum distance from 0 to a codeword.

Proof: Let the check matrix be

$$H = (r_1 \quad r_2 \quad \ldots \quad r_n)$$

and $v = (c_1, \ldots, c_n)$ a codeword. Then

$$0 = v \cdot H^\top = \sum_i c_i r_i^\top$$

If *any* bunch of d of the columns r_is are linearly independent, then for *any* codeword v there must be at least $d + 1$ non-zero c_is.

Conversely, if *some* d of the r_is are linearly dependent, then for *some* codeword v there are at most d non-zero c_is. ///

Remark: Thus, to look for linear codes that correct many errors, look for check matrices H with many with any that any $2e$ columns are linearly independent. This motivates the specifics of the constructions of Hamming codes, BCH (Bose-Chaudhuri-Hocquengham) codes, RS (Reed-Solomon) codes, and Goppa codes.

Corollary: If any $2e$ columns of the check matrix are linearly independent, then the code can correct any e errors, and *vice versa.*

Corollary: For a binary linear code, if no 2 columns of a check matrix are the same, and if no column of the check matrix is 0, then the code can correct any single error.

Remark: The latter corollary is due to the fact that the scalars are just $\{0, 1\}$.

Corollary:

- A linear code can correct any 2 errors if and only if no 4 columns (or fewer) of a check matrix are linearly dependent.

- A linear code can correct any 3 errors if and only if no 6 columns (or fewer) of a check matrix are linearly dependent.

Proof of Gilbert-Varshamov bound

Consider linear codes with alphabet $\mathbf{F}_q$, block size n, dimension k, and minimum distance d. A **generating matrix** with linearly independent rows would be k-by-n. This would be an $[n,k,d]$ code. We can now prove

Theorem: (*Gilbert-Varshamov*) If

$$q^{n-k}-1 > (q-1)\binom{n-1}{1}+\ldots+(q-1)^{d-2}\binom{n-1}{d-2}$$

then an $[n,k,d]$ code over alphabet $\mathbf{F}_q$ exists.

The simple special case $q=2$:

Corollary: If

$$2^{n-k}-1 > \binom{n-1}{1}+\ldots+\binom{n-1}{d-3}+\binom{n-1}{d-2}$$

then a *binary* $[n,k,d]$-code exists. ///

Remark: The theorem assures that good codes *exist*, but does *not* give an *efficient* procedure to find them.

Remark: There is *no* assertion that this is the *best* that a code can do, only that we can *(in principle!)* expect *at least* this level of performance. Still, it is very hard to make even a single code that exceeds that GV bound.

Proof: Do the binary case for simplicity. Keep in mind that for a linear code the minimum distance is d if and only if any $d-1$ columns of a check matrix are linearly independent.

Consider the process of choosing n columns for a check matrix so that any $d-1$ of them are linearly independent. The code is the row space of a k-by-n generating matrix G (with linearly independent rows). Its check matrix is an $(n-k)$-by-n matrix of rank $n-k$. Suppose that in the construction of a check matrix we have successfully chosen ℓ columns with no $d-1$ of them linearly dependent. Now we want to choose an $(\ell+1)^{\text{th}}$ column.

The choice of $(\ell+1)^{\mathrm{th}}$ column must be made from among column vectors of size $n-k$. There are 2^{n-k} such vectors.

We must *exclude* the all-0-column, exclude any previous column, exclude the sum of any previous two columns, exclude the sum of any previous three columns, and so on up to excluding the sum of any previous $d-2$ columns.

In the worst case, all these things that we must exclude are *different*, leaving only

$$2^{n-k} - \left(1 + \binom{\ell}{1} + \binom{\ell}{2} + \ldots + \binom{\ell}{d-2}\right)$$

available vectors. Thus, to be *sure* that a choice is available, this number must be *positive*.

///

Remark: That is, if the Gilbert-Varshamov inequality holds then *in principle* we could make a code with the given $[n, k, d]$. However, as far as seems to be known, following the proof of Gilbert-Varshamov would give a construction no better than an extremely labor-intensive brute force search. That is, there does *not* seem to be any good algorithmic approach here.

Remark: The converse assertion is false. That is, there do exist some linear codes *exceeding* the Gilbert-Varshamov bound. Such codes are very good indeed. In fact, certain of the *geometric Goppa codes* were proven by Tsfasman, Vladut, and Zink to exceed the Gilbert-Varshamov bound.

Vandermonde determinants

We have shown that linear code can correct e errors if and only if any $2e$ columns of its check matrix are linearly independent.

How to make this happen?

From linear algebra, *ℓ vectors of length ℓ are linearly independent if and only if the determinant of the matrix made by stacking them up is not 0.*

But determinants are hard to evaluate in general, especially for *random* matrices which might have a chance of fulfilling the conclusion of Shannon's theorem.

We want a systematic trick to know that a whole class of determinants is non-zero, for **check matrices** for cyclic codes. Two standard types of Vandermonde matrix have non-zero determinants.

Recall that n vectors

$$\begin{gathered}(v_{11}, v_{12}, v_{13}, v_{14}, \ldots, v_{1n}) \\ (v_{21}, v_{22}, v_{23}, v_{24}, \ldots, v_{2n}) \\ (v_{31}, v_{32}, v_{33}, v_{34}, \ldots, v_{3n}) \\ \ldots \\ (v_{n1}, v_{n2}, v_{n3}, v_{n4}, \ldots, v_{nn})\end{gathered}$$

are linearly independent if and only if the **determinant** of the n-by-n matrix made by stacking them as rows is *non-zero*:

$$\det \begin{pmatrix} v_{11} & v_{12} & v_{13} & v_{14} & \ldots & v_{1n} \\ v_{21} & v_{22} & v_{23} & v_{24} & \ldots & v_{2n} \\ v_{31} & v_{32} & v_{33} & v_{34} & \ldots & v_{3n} \\ & & \ldots & & & \\ v_{n1} & v_{n2} & v_{n3} & v_{n4} & \ldots & v_{nn} \end{pmatrix} \neq 0$$

Remark: Whether or not you remember *how* to evaluate determinants, you should try to *avoid* evaluating big determinants.

Remark: We don't want its value, but the assurance that it's not 0.

One version of **Vandermonde matrix** is

$$M = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & x_4 & \dots & x_n \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 & \dots & x_n^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 & \dots & x_n^3 \\ x_1^4 & x_2^4 & x_3^4 & x_4^4 & \dots & x_n^4 \\ & & & \dots & & \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & x_4^{n-1} & \dots & x_n^{n-1} \end{pmatrix}$$

The determinant of a Vandermonde matrix is called a **Vandermonde determinant**.

Amazing Theorem:

$$\det M = (-1)^{n(n-1)/2} \prod_{i<j} (x_i - x_j)$$

Corollary: If the x_i's lie in a set of things with the property that a product of non-zero things cannot be zero, and if for all $i < j$ we have $x_i \neq x_j$ then the Vandermonde determinant is not 0.

Corollary: For example, if the x_i's are distinct and lie in $\mathbf{Z}/p$ with p prime, then the Vandermonde determinant is $\neq 0$.

Remark: Unfortunately, if we want to stay in the littlest finite field $\mathbf{F}_2$, it is hard to find many distinct x_i's to use in a Vandermonde determinant.

Remark: Keep in mind that in greatest generality the product of a bunch of non-zero things can nevertheless be 0. For example, in $\mathbf{Z}/6$, neither $\overline{2}$ nor $\overline{3}$ is $\overline{0}$, but their product is $\overline{0}$.

But this counter-intuitive phenomenon does not occur in $\mathbf{Z}/p$ with p prime.

Remark: More generally, in a **field**, by definition, every non-zero element has a multiplicative inverse. This prevents $ab = 0$ unless either a or b is 0. More generally, a *commutative ring* in which $ab = 0$ only when either a or b is 0 is an *integral domain.* Every field is an integral domain. The ordinary integers $\mathbf{Z}$ are an example of an integral domain which is not a field.

For example, for α in a field k, assuming that all the quantities 1, α, α^2, α^3, …, α^{n-1} are different from each other, by taking $x_i = \alpha^{i-1}$, we get a non-zero determinant

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \alpha & \alpha^2 & \dots & \alpha^{n-1} \\ 1 & \alpha^2 & (\alpha^2)^2 & \dots & (\alpha^{n-1})^2 \\ 1 & \alpha^3 & (\alpha^2)^3 & \dots & (\alpha^{n-1})^3 \\ 1 & \alpha^4 & (\alpha^2)^4 & \dots & (\alpha^{n-1})^4 \\ & & & \dots & \\ 1 & \alpha^{n-1} & (\alpha^2)^{n-1} & \dots & (\alpha^{n-1})^{n-1} \end{vmatrix} \neq 0$$

Remark: Since the x_i's in a Vandermonde determinant need not be consecutive powers of a common α, the different powers of α inside the parentheses don't have to be consecutive, only not *equal* to each other.

That is, for non-zero element α of a field and for integers $\ell_1, \ldots, \ell_n$ so that

$$\alpha^{\ell_1},\ \alpha^{\ell_2},\ \alpha^{\ell_3},\ \ldots,\ \alpha^{\ell_n}$$

are distinct, we have a non-zero determinant

$$\begin{vmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & \alpha^{\ell_1} & \alpha^{\ell_2} & \ldots & \alpha^{\ell_{n-1}} \\ 1 & (\alpha^{\ell_1})^2 & (\alpha^{\ell_2})^2 & \ldots & (\alpha^{\ell_{n-1}})^2 \\ 1 & (\alpha^{\ell_1})^3 & (\alpha^{\ell_2})^3 & \ldots & (\alpha^{\ell_{n-1}})^3 \\ 1 & (\alpha^{\ell_1})^4 & (\alpha^{\ell_2})^4 & \ldots & (\alpha^{\ell_{n-1}})^4 \\ & & & \ldots & \\ 1 & (\alpha^{\ell_1})^{n-1} & (\alpha^{\ell_2})^{n-1} & \ldots & (\alpha^{\ell_{n-1}})^{n-1} \end{vmatrix} \neq 0$$

If a row or column of a matrix is multiplied by β, then the determinant is multiplied by β.

This implies that a larger class of determinants is non-zero. From

$$\begin{vmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & x_4 & \dots & x_n \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 & \dots & x_n^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 & \dots & x_n^3 \\ x_1^4 & x_2^4 & x_3^4 & x_4^4 & \dots & x_n^4 \\ & & & \dots & & \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & x_4^{n-1} & \dots & x_n^{n-1} \end{vmatrix} \neq 0$$

for distinct $x_1, \dots, x_n$ we can multiply through the i^{th} column by x_i to obtain

$$\begin{vmatrix} x_1 & x_2 & x_3 & x_4 & \dots & x_n \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 & \dots & x_n^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 & \dots & x_n^3 \\ x_1^4 & x_2^4 & x_3^4 & x_4^4 & \dots & x_n^4 \\ x_1^5 & x_2^5 & x_3^5 & x_4^5 & \dots & x_n^5 \\ & & & \dots & & \\ x_1^n & x_2^n & x_3^n & x_4^n & \dots & x_n^n \end{vmatrix} \neq 0$$

for the $x_1, \dots, x_n$ all different from each other, and non-zero. This type of matrix is also called a **Vandermonde matrix**.

Primitive roots in $\mathbf{Z}/p$

From thinking about Vandermonde determinants, we want some set of numbers in which there is an element g such that

$$1, g, g^2, g^3, \ldots, g^N$$

are distinct for as large as possible exponent N. This should remind us of **primitive roots**. Recall

Theorem: For p prime, there exist primitive roots g modulo p. That is,

$$1, g, g^2, \ldots, g^{p-3}, g^{p-2}$$

are all different.

Remark: In more structural terms: the multiplicative group $\mathbf{Z}/p^\times$ of the finite field $\mathbf{Z}/p$ with p elements is a *cyclic group*.

Remark: Any generator of the cyclic group $\mathbf{Z}/p^\times$ is a *primitive root* for $\mathbf{Z}/p$. As a corollary of a study of *cyclotomic polynomials*, one would know that the multiplicative group $k^\times$ of any finite field k is cyclic. So all we would need do is check that $\mathbf{Z}/p$ is a field. That is, we must check that any non-zero element $b \in \mathbf{Z}/p$ has a multiplicative inverse.

Recall the (important!) explanation of why there is a multiplicative inverse for any non-zero thing b modulo p prime:

Since p is prime, if $b \neq 0 \bmod p$, then $\gcd(p, b) = 1$. Thus, by the peculiar characterization of gcd's, there are integers s, t so that

$$sp + tb = 1$$

Looking at the latter equation mod p, t is a multiplicative inverse to b modulo p.

///

Remark: Recall that g (relatively prime to p) is a primitive root modulo a *prime* p if and only if for every prime r dividing $p-1$

$$g^{\frac{p-1}{r}} \neq 1 \bmod p$$

Note that $g^{p-1} = 1 \bmod p$ by Fermat's Little Theorem, so we need not test this.

Remark: If p is very large, it may happen that it is infeasible to factor $p-1$.

Reed-Solomon (RS) codes

Let $\mathbf{F}_q = GF(q)$ be a finite field. Reed-Solomon codes over $\mathbf{F}_q$ are of block length $n = q - 1$.

Let β be a **primitive root** in $\mathbf{F}_q$, so

$$1, \beta, \beta^2, \beta^3, \beta^4, \ldots, \beta^{q-3}, \beta^{q-2}$$

are distinct, and every non-zero element of $\mathbf{F}_q$ is a power of β. So (!)

$$x^{q-1} - 1 = (x-1)(x-\beta)\ldots(x-\beta^{q-3})(x-\beta^{q-2})$$

Choose **design distance** t in the range $2 \leq t \leq q - 1$. Define

$$g(x) = (x-\beta)(x-\beta^2)\ldots(x-\beta^{t-2})(x-\beta^{t-1})$$

The cyclic code specified by this is a $[n, n-t+1]$-code using the alphabet $\mathbf{F}_q$, a **Reed-Solomon code**.

Theorem: The RS code C over $\mathbf{F}_q$ with generating polynomial

$$g(x) = (x-\beta)(x-\beta^2)\ldots(x-\beta^{t-2})(x-\beta^{t-1})$$

(β a primitive root in $\mathbf{F}_q$) has minimum distance *at least* its design distance t.

Remarks: The only possible sizes of finite fields are *prime powers* (meaning powers of prime numbers).

There are *no* finite fields with 6, 10, 12, 14, 15, 18, or other such non-prime-power number of elements.

Fields with *prime* numbers p of elements have easy models:

$$\mathbf{F}_p = GF(p) \approx \mathbf{Z}/p$$

Fields with *prime power* (but not prime) numbers of elements are less elementary. They **cannot be modeled by Z mod something!**

For any value of q other than primes, $\mathbf{Z}/q$ **is not a field**. Such $\mathbf{Z}/q$ have non-zero elements whose products are 0, so it is not possible that every non-zero element is prime.

$$GF(4) = \mathbf{F}_4 \neq \mathbf{Z}/4$$

$$GF(8) = \mathbf{F}_8 \neq \mathbf{Z}/8$$

$$GF(16) = \mathbf{F}_{16} \neq \mathbf{Z}/16$$

$$GF(9) = \mathbf{F}_9 \neq \mathbf{Z}/9$$