Linear Codes

Jonathan L.F. King
University of Florida, Gainesville FL 32611-2082, USA
squash@ufl.edu
Webpage http://squash.1gainesville.com/
24 April, 2023 (at *19:47*)

Entrance. An N-*word* over alphabet F is a sequence $\mathbf{v} = v_1 v_2 \ldots v_N$ of N letters. With $Q := |\mathsf{F}| \geq 2$, the number of N-words is Q^N. A ***code(word)set*** is a subset $\mathcal{C} \subset \mathsf{F}^N$; there are $2^{[Q^N]}$ codesets.

The ***Hamming-distance*** between $\mathbf{v}$ and $\mathbf{w}$ is

$$\mathrm{HD}(\mathbf{v},\mathbf{w}) := {}^{\#}\{j \in [1 .. N] \mid v_j \neq w_j\}.$$

The ***(min-)Hamming distance*** of $\mathcal{C}$ is

$$\mathrm{minHD}(\mathcal{C}) := \mathrm{Min}\,\{\mathrm{HD}(\mathbf{v},\mathbf{w}) \mid \mathbf{v},\mathbf{w} \in \mathcal{C} \text{ with } \mathbf{v} \neq \mathbf{w}\}.$$

Detection/correction. Use $\boldsymbol{D}$ for $\mathrm{minHD}(\mathcal{C})$. A transmitted N-codeword $\mathbf{v}$ might be distorted into a N-word $\widetilde{\mathbf{v}}$. A position in $\mathbf{v}$ may have been: *Changed to a different letter* –an ***error***, or: *Replaced by noise* –an erasure. [E.g, an unreadable postion on a magtape because the local magnetic field has weakened over time].

A distance-$\boldsymbol{D}$ codeset $\mathcal{C}$ can:

i: *Correct* $\boldsymbol{D}-1$ erasures. [Knowing the erasure positions, there is but one consistent codeword.]

ii: *Detect* $\boldsymbol{D}-1$ errors. [Alas, we don't necessarily know the error-positions.]

iii: *Correct* $\lfloor \frac{\boldsymbol{D}-1}{2} \rfloor$ errors. [With $r := \lfloor \frac{\boldsymbol{D}-1}{2} \rfloor$, the radius-$r$ balls centered at codewords are mutually disjoint.]

Upperbnding $\mathcal{C}$. The volume of a radius-r ball in $\mathbf{v} \in \mathsf{F}^N$ is

$$V_r := \sum_{j=0}^{r} \binom{N}{j} \cdot [Q-1]^j .$$

Setting $r := \lfloor \frac{\boldsymbol{D}-1}{2} \rfloor$ yields the ***Hamming bound***

$*_H$:
$$ {}^{\#}\mathcal{C} \leq Q^N \big/ V_r \quad \textit{So}$$
$$\log_Q({}^{\#}\mathcal{C}) \leq N - \log_Q(V_r)$$

Singleton bnd. From each codeword, delete the first $\boldsymbol{D}-1$ letters; the reduced codewords remain distinct, but now have length $N - [\boldsymbol{D}-1]$ producing the ***Singleton bound***

$*_S$:
$$ {}^{\#}\mathcal{C} \leq Q^{N-[\boldsymbol{D}-1]} \quad \textit{So}$$
$$\log_Q({}^{\#}\mathcal{C}) \leq N - [\boldsymbol{D}-1]$$

The Setting. Henceforth, $K<N$ and M are posints, and p is prime. Our alphabet is finite field $\mathsf{F} := \mathsf{F}_Q$, where $Q = \mathsf{p}^M$. Both F^K and F^N are vectorspaces (VSes) over scalar field F.

A ***linear codeset*** is a dimension-K subspace $\mathcal{C} \subset \mathsf{F}^N$. As the ***weight*** of vector $\mathbf{v} \in \mathsf{F}^N$ is $\mathrm{HD}(\mathbf{v},\vec{\mathbf{0}})$, the $\mathrm{minWei}(\mathcal{C})$ is the minimum weight over all $\mathbf{v} \neq \vec{\mathbf{0}}$. Easily

$$\mathrm{minWei}(\mathcal{C}) = \mathrm{minHD}(\mathcal{C})$$

A ***(linear-)code(map)*** is an injective linear map $\mathscr{C}\colon \mathsf{F}^K \to \mathsf{F}^N$ whose range is $\mathcal{C}$. Realizing vectors as column-vectors, we'll define $\mathscr{C}$ by an $N \times K$ ***generator matrix*** G over F.