Vector spaces

A vector space $V$ of vectors over a field $F$ of scalars is a set with a binary operator $+$ on $V$ and a scalar-vector product $\cdot$ satisfying these axioms:

1. $(V, +)$ is a commutative group
2. Associative law: $(a_1 a_2) \cdot v = a_1 \cdot (a_2 \cdot v)$.
3. Distributive laws:
   
   $a \cdot (v_1 + v_2) = a \cdot v_1 + a \cdot v_2$, $(a_1 + a_2) \cdot v = a_1 \cdot v + a_2 \cdot v$

4. Unitary law: $1 \cdot v = v$ for every vector $v$ in $V$

Recall definition of field: a set $F$ with associative and commutative operators $+$ and $\cdot$ that satisfy the following axioms.

1. $(F, +)$ is a commutative group
2. $(F - \{0\}, \cdot)$ is a commutative group
3. Distributive law: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$

A field is a vector space over itself of dimension 1.

In general, vector spaces do not have vector-vector multiplication. An algebra is a vector space with an associative, distributive multiplication; e.g., $n \times n$ matrices over $F$. 
Span, basis, linear independence

**Definition:** The *span* of a set of vectors \( \{v_1, \ldots, v_n\} \) is the set of all *linear combinations* of those vectors:
\[
\left\{ a_1 \cdot v_1 + \cdots + a_n \cdot v_n : a_1, \ldots, a_n \in F \right\}.
\]

**Definition:** A set of vectors \( \{v_1, \ldots, v_n\} \) is *linearly independent (LI)* if no vector in the set is a linear combination of the other vectors; equivalently, there does not exist a set of scalars \( \{a_1, \ldots, a_n\} \), not all zero, such that
\[
a_1 \cdot v_1 + a_2 \cdot v_2 + \cdots + a_n \cdot v_n = 0.
\]

**Definition:** A *basis* for a vector space is a LI set that spans the vector space.

**Definition:** A vector space is *finite dimensional* if it has a finite basis.

**Lemma:** The number of vectors in a linearly independent set is \( \leq \) the number of vectors in a spanning set.

**Theorem:** All bases for a finite-dimensional vector space have the same number of elements. Thus the dimension of the vector space is well defined.
Vector coordinates

Theorem: Every vector can be expressed as a linear combination of basis vectors in exactly one way.

Proof: Suppose that \{v_1, \ldots, v_n\} is a basis and 

\[ a_1 \cdot v_1 + \cdots + a_n \cdot v_n = a'_1 \cdot v_1 + \cdots + a'_n \cdot v_n, \]

are two possibly different representations of the same vector. Then 

\[
(a_1 - a'_1) \cdot v_1 + \cdots + (a_n - a'_n) \cdot v_n = 0.
\]

By linear independence of the basis 

\[ a_1 - a'_1 = 0, \ldots, a_n - a'_n = 0 \implies a_1 = a'_1, \ldots, a_n = a'_n. \]

We say that \( V \) is isomorphic to vector space \( F^n = \{(a_1, \ldots, a_n) : a_i \in F\} \) with componentwise addition and scalar multiplication.

Theorem: If \( V \) is a vector space of dimension \( n \) over a field with \( q \) elements, then the number of elements of \( V \) is \( q^n \).

Proof: There are \( q^n \) \( n \)-tuples of scalars from \( F \).
Inner product and orthogonal complement

Definition: A vector subspace of a vector space \( V \) is a subset \( W \) that is a vector space, i.e., closed under vector addition and scalar multiplication.

Theorem: If \( W \) is a subspace of \( V \) then \( \dim W \leq \dim V \).

Proof: Any basis \( \{ w_1, \ldots, w_k \} \) for subspace \( W \) can be extended to a basis of \( V \) by adding \( n-k \) vectors \( \{ v_{k+1}, \ldots, v_n \} \). Therefore \( \dim W \leq \dim V \).

Definition: The inner product of two \( n \)-tuples over a field \( F \) is
\[
(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = a_1 b_1 + \cdots + a_n b_n .
\]

Definition: Two \( n \)-tuples are orthogonal if their inner product is zero.

Definition: The orthogonal complement of a subspace \( W \) of \( n \)-tuples is the set \( W^\perp \) of vectors orthogonal to every vector in \( W \); that is,
\[
v \in W^\perp \text{ if and only if } v \cdot w = 0 \text{ for every } w \in W .
\]

Theorem: (Dimension Theorem) If \( \dim W = k \) then \( \dim W^\perp = n - k \).

Theorem: If \( W \) is subspace of finite-dimensional \( V \), then \( W^{\perp \perp} = W \).

In infinite-dimensional vector spaces, \( W \) may be a proper subset of \( W^{\perp \perp} \).
Linear transformations and matrix multiplication

**Definition:** A linear transformation on a vector space $V$ is a function $T$ that maps vectors in $V$ to vectors in $V$ such that for all vectors $v_1, v_2$ and scalars $a_1, a_2$,

$$T(a_1v_1 + a_2v_2) = a_1T(v_1) + a_2T(v_2).$$

Every linear transformation $T$ can be described by an $n \times n$ matrix $M$. The coordinates of $T(v_1, \ldots, v_n)$ are the components of the vector-matrix product $(v_1, \ldots, v_n)M$.

Each component of $w = vM$ is inner product of $v$ with a column of $M$:

$$w_j = v_1m_{1j} + \cdots + v nm_{nj} = \sum_{i=1}^{n} v_im_{ij}. $$

Here is another way to look at $w = vM$. If $m_i$ is $i$-th row of $M$, then

$$(v_1, \ldots, v_n)M = v_1m_1 + \cdots + v nm_n.$$ 

Thus $vM$ is linear combination of rows of $M$ weighted by coordinates of $v$. As $v$ ranges over $n$-tuples, $vM$ ranges over subspace spanned by rows of $M$. 
Matrix rank

Definition: A square matrix is nonsingular if it has a multiplicative inverse.

Definition: The determinant of $A = (a_{ij})$ is

$$\det A = \sum_\sigma (-1)^\sigma a_{i_1\sigma(i_1)} \cdots a_{i_n\sigma(i_n)},$$

where $\sigma$ is a permutation of $\{1, \ldots, n\}$ and $(-1)^\sigma$ is the sign of $\sigma$.

Theorem: A matrix is nonsingular if and only if its determinant is nonzero.

Definition: The rank of a matrix is the size of largest nonsingular submatrix. The rank of a matrix can be computed efficiently:

1. Use Gaussian elimination to transform matrix to row-reduced echelon form.

2. The rank is the number of nonzero rows of the row-reduced echelon matrix.

Theorem: The rank of a matrix is the dimension of the row space, which equals the dimension of the column space.

Therefore the number of linearly independent rows is the same as the number of linearly independent columns.
Linear block codes and group codes

Definition: A linear block code over a field $F$ of blocklength $n$ is a linear subspace of $F^n$.

Facts about linear block codes:

- The sum and difference of codewords are codewords.
- Scalar multiples of codewords are also codewords.
- A binary block code that is closed under addition is linear because the only scalars are 0 and 1.
- Parity-check codes are linear block codes over $\text{GF}(2)$. Every PC code is defined by a set of homogeneous binary equations.
- If $\mathcal{C}$ is a LBC over $\text{GF}(Q)$ of dimension $k$, then its rate is

$$R = \frac{\log_Q Q^k}{n} = \frac{k}{n}.$$  

Note that the rate is determined by $k$ and $n$ and not by $Q$.

A group code is a subgroup of the $n$-tuples over an additive group.
Minimum weight

The **Hamming weight** $w_H(v)$ is the number of nonzero components of $v$.

Obvious facts:

- $w_H(v) = d_H(0, v)$
- $d_H(v_1, v_2) = w_H(v_1 - v_2) = w_H(v_2 - v_1)$
- $w_H(v) = 0$ if and only if $v = 0$

**Definition:** The *minimum (Hamming) weight* of a block code is the weight of the nonzero codeword with smallest weight:

$$w_{\text{min}} = w^* = \min\{w_H(c) : c \in C, c \neq 0\}$$

Examples of minimum weight:

- Simple parity-check codes: $w^* = 2$.
- Repetition codes: $w^* = n$.
- (7,4) Hamming code: $w^* = 3$. (There are 7 codewords of weight 3.)
  - Weight enumerator: $A(x) = 1 + 7x^3 + 7x^4 + x^7$.
- Simple product code: $w^* = 4$. 
**Theorem**: For every linear block code, \( d^* = w^* \).

**Proof**: We show that \( w^* \geq d^* \) and \( w^* \leq d^* \).

(\( \geq \)) Let \( c_0 \) be a nonzero minimum-weight codeword. Since \( \mathbf{0} \) is a codeword,

\[
    w^* = w_H (c_0) = d_H (0, c_0) \geq d^* .
\]

(\( \leq \)) Let \( c_1 \neq c_2 \) be closest codewords. Since \( c_1 - c_2 \) is a nonzero codeword,

\[
    d^* = d_H (c_1, c_2) = w_H (c_1 - c_2) \geq w^* .
\]

Combining these two inequalities, we obtain \( d^* = w^* \).

It is easier to find minimum weight than minimum distance because the weight minimization considers only a single parameter.

Computer search: test vectors of weight 1, 2, 3, \ldots until codeword is found.

It is also easier to determine the weight distribution of a linear code than the distance distribution of a general code.

The result \( d^* = w^* \) holds for group codes, since the proof used only subtraction.
Generator matrix

Definition: A generator matrix for a linear block code $C$ of blocklength $n$ and dimension $k$ is any $k \times n$ matrix $G$ whose rows form a basis for $C$.

Every codeword is a linear combination of the rows of a generator matrix $G$:

$$c = mG = \begin{bmatrix} m_0 & m_1 & \ldots & m_{k-1} \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{k-1} \end{bmatrix} = m_0g_0 + m_1g_1 + \cdots + m_{k-1}g_{k-1}.$$

Since $G$ has rank $k$, the representation of $c$ is unique.

Each component $c_j$ of $c$ is inner product of $m$ with $j$-th column of $G$:

$$c_j = m_0g_{0,j} + m_1g_{1,j} + \cdots + m_{k-1}g_{k-1,j}.$$

Both sets of equations can be used for encoding. Each codeword symbol requires $k$ multiplications (by constants) and $k - 1$ additions.
Parity-check matrix

Definition: The dual code of $C$ is the orthogonal complement $C^\perp$.

Definition: A parity-check matrix for a linear block code $C$ is any $r \times n$ matrix $H$ whose rows span $C^\perp$. (Obviously, $r \geq n - k$.)

Example: $G$ and $H$ for $(5, 4)$ simple parity-check code.

\[
G = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 1
\end{bmatrix} \implies H = \begin{bmatrix}
1 & 1 & 1 & 1 & 1
\end{bmatrix}
\]

($H$ is generator matrix for the $(5, 1)$ repetition code—the dual code.)

Example: $G$ and $H$ for $(7, 4)$ cyclic Hamming code.

\[
G = \begin{bmatrix}
1 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1
\end{bmatrix} \implies H = \begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1
\end{bmatrix}
\]

A cyclic code is a linear block code such that the cyclic shift of every codeword is also a codeword. It is not obvious that this property holds for the code generated by $G$. 

EE 387, October 7, 2015
Codewords of linear code are determined by $H$

*Theorem*: If $C$ is an $(n, k)$ linear block code with parity-check matrix $H$, then an $n$-tuple $c$ is a codeword if and only if $cH^T = 0$.

*Proof*:

$(\Rightarrow)$ Suppose $c$ is a codeword.

Each component of $cH^T$ is the inner product of $c$ and a column of $H^T$, which is a row of $H$.

Since every row of $H$ is in $C^\perp$, each row is $\perp$ to $c$.

Thus each component of $cH^T$ is $c \cdot h_i = 0$.

$(\Leftarrow)$ Since rows of $H$ span $C^\perp$, any $n$-tuple satisfying $cH^T = 0$ belongs to the orthogonal complement of $C^\perp$.

By the Dimension Theorem (Blahut Theorem 2.5.10), $C^{\perp\perp} = C$.

Therefore if $cH^T = 0$ then $c$ belongs to $C$.

$(\Leftrightarrow)$ Thus $C$ consists of vectors satisfying the check equations $cH^T = 0$. 
Generator vs. parity-check matrices

Usually $H$ consists of $n - k$ independent rows, so $H$ is $(n - k) \times n$.

Sometimes it is convenient or elegant to use a parity-check matrix with redundant rows (for example, binary BCH codes, to be discussed later).

Each row of $H$ corresponds to an equation satisfied by all codewords. Since each row of $G$ is a codeword, for any parity-check matrix $H$,

$$G_{k \times n} \cdot H_{r \times n}^T = 0_{k \times r} \quad (r \geq n - k)$$

Each $0$ is $0_{k \times r}$ corresponds to one codeword and one equation.

Conversely, if $GH^T = 0$ and $\text{rank } H = n - k$ then $H$ is a check matrix.

How do we find $H$ from $G$?

We can find $H$ from $G$ by finding $n - k$ linearly independent solutions of the linear equation $GH^T = 0$.

The equations $GH^T = 0$ are easy to solve when $G$ is systematic.
Systematic generator matrices

Definition: A systematic generator matrix is of the form

\[ G = [ P \mid I ] = \begin{bmatrix}
  p_{0,0} & \cdots & p_{0,n-k-1} & 1 & 0 & \cdots & 0 \\
  p_{1,0} & \cdots & p_{1,n-k-1} & 0 & 1 & \cdots & 0 \\
  \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
  p_{k-1,0} & \cdots & p_{k-1,n-k-1} & 0 & 0 & \cdots & 1
\end{bmatrix} \]

Advantages of systematic generator matrices:

- Message symbols appear unscrambled in each codeword, in the rightmost positions \( n - k, \ldots, n - 1 \).
- Encoder complexity is reduced; only check symbols need be computed:
  \[ c_j = m_0 g_{0,j} + m_1 g_{1,j} + \cdots + m_{k-1} g_{k-1,j} \quad (j = 0, \ldots, n - k - 1) \]
- Check symbol encoder equations easily yield parity-check equations:
  \[ c_j - c_{n-k} g_{0,j} - c_{n-k+1} g_{1,j} - \cdots - c_{n-1} g_{k-1,j} = 0 \quad (m_i = c_{n-k+i}) \]
- Systematic parity-check matrix is easy to find: \( H = [ I \mid -P^T ] \).
Systematic parity-check matrix

Let $G$ be a $k \times n$ systematic generator matrix:

$$G = \begin{bmatrix} P \mid I_k \end{bmatrix} = \begin{bmatrix} p_{0,0} & \cdots & p_{0,n-k-1} & 1 & 0 & \cdots & 0 \\ p_{1,0} & \cdots & p_{1,n-k-1} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{k-1,0} & \cdots & p_{k-1,n-k-1} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

The corresponding $(n-k) \times n$ systematic parity-check matrix is

$$H = \begin{bmatrix} I_{n-k} \mid -P^T \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & -p_{0,0} & \cdots & -p_{k-1,0} \\ 0 & 1 & \cdots & 0 & -p_{0,1} & \cdots & -p_{k-1,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -p_{0,n-k} & \cdots & -p_{k-1,n-k} \end{bmatrix}$$

(The minus signs are not needed for fields of characteristic 2, i.e., $\text{GF}(2^m)$.)

Each row of $H$ is corresponds to an equation satisfied by all codewords. These equations tell how to compute the check symbols $c_0, \ldots, c_{n-k-1}$ in terms of the information symbols $c_{n-k}, \ldots, c_{n-1}$. 
Minimum weight and columns of $H$

c$H^T = 0$ for every codeword $c = (c_0, c_1, \ldots, c_{n-1})$. Any nonzero codeword determines a linear dependence among a subset of rows of $H^T$. Then
\[
cH^T = 0 \implies 0 = (cH^T)^T = Hc^T = c_0 h^0 + c_1 h^1 + \cdots + c_{n-1} h^{n-1}
\]
is a linear dependence among a subset of the columns of $H$.

**Theorem:** The minimum weight of a linear block code is the smallest number of linearly dependent columns of any parity-check matrix.

**Proof:** Each linearly dependent subset of $w$ columns corresponds to a codeword of weight $w$.

Recall that a set of columns of $H$ is linearly dependent if one column is a linear combination of the other columns.

- A LBC has $w^* \leq 2$ iff one column of $H$ is a multiple of another column.
- For binary Hamming codes, $w^* = 3$ because no columns of $H$ are equal.

The Big Question: how to find $H$ such that no $2t + 1$ columns are LI?