

# Encoding: Generator Matrix

The encoding procedure of a linear code is nicely captured by the generator matrix:

## Definition: Generator matrix

Given a  $[n, k, d]$  linear code, the generator matrix  $G^T$  is an  $n \times k$  matrix such that the columns  $c_1, \dots, c_k$  form a basis for  $C$ . The generator matrix can always be transformed into **systematic form**:

$$G^T = \left( \begin{array}{c|c} \mathbb{I}_k & P \end{array} \right),$$

where  $P$  is some  $(n - k) \times k$  matrix representing the parity bits of the code.

The generator matrix is used in the encoding function as follows:

$$\text{enc}(m) = G^T \cdot m.$$

The codebook  $C$  is the set  $\{G^T \cdot m \mid m \in \mathcal{X}^k\}$ .

The reason for the transposition in  $G^T$  is that historically, coding theorists prefer to use row vectors and matrix multiplication from the right instead of column vectors and multiplication from the left, which is more standard in other areas.

Notice that for row vectors  $c = m \cdot G$ , we equivalently have column vectors  $c^T = (m \cdot G)^T = G^T \cdot m^T$ .

## Example: Generator matrix of the $[7, 4]$ Hamming code

The following  $7 \times 4$  matrix generates the  $[7, 4]$  Hamming code:

$$G^T = \left( \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & & & & \\ 0 & 1 & 0 & 0 & & & & \\ 0 & 0 & 1 & 0 & & & & \\ 0 & 0 & 0 & 1 & & & & \\ \hline 1 & 1 & 1 & 0 & & & & \\ 0 & 1 & 1 & 1 & & & & \\ 1 & 0 & 1 & 1 & & & & \end{array} \right).$$

The generator matrix is given in systematic form. Encode the message 1010.

Show solution

To encode the message 1010, we compute

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

Note that due to the systematic form, the first  $k$  bits of the codeword is the actual message, whereas the rest are the parity-check bits.