WIKIPEDIA

Decoding methods

In coding theory, **decoding** is the process of translating received messages into codewords of a given code. There have been many common methods of mapping messages to codewords. These are often used to recover messages sent over a noisy channel, such as a binary symmetric channel.

Contents

Notation

$C \subset \mathbb{F}_2^n$ is considered a binary code with the length n; x, y shall be elements of $\mathbb{F}_2^n$; and $d(x,y)$ is the distance between those elements.

Ideal observer decoding

One may be given the message $x \in \mathbb{F}_2^n$, then **ideal observer decoding** generates the codeword $y \in C$. The process results in this solution:

$$\mathbb{P}(y \text{ sent} \mid x \text{ received})$$

For example, a person can choose the codeword y that is most likely to be received as the message x after transmission.

Decoding conventions

Each codeword does not have an expected possibility: there may be more than one codeword with an equal likelihood of mutating into the received message. In such a case, the sender and receiver(s) must agree ahead of time on a decoding convention. Popular conventions include:

1. Request that the codeword be resent – automatic repeat-request.
2. Choose any random codeword from the set of most likely codewords which is nearer to that.

3. If another code follows, mark the ambiguous bits of the codeword as erasures and hope that the outer code disambiguates them

Maximum likelihood decoding

Given a received codeword $x \in \mathbb{F}_2^n$ **maximum likelihood decoding** picks a codeword $y \in C$ that maximizes

$$\mathbb{P}(x \text{ received} \mid y \text{ sent}),$$

that is, the codeword y that maximizes the probability that x was received, given that y was sent. If all codewords are equally likely to be sent then this scheme is equivalent to ideal observer decoding. In fact, by Bayes Theorem,

$$\begin{aligned}\mathbb{P}(x \text{ received} \mid y \text{ sent}) &= \frac{\mathbb{P}(x \text{ received}, y \text{ sent})}{\mathbb{P}(y \text{ sent})} \\ &= \mathbb{P}(y \text{ sent} \mid x \text{ received}) \cdot \frac{\mathbb{P}(x \text{ received})}{\mathbb{P}(y \text{ sent})}.\end{aligned}$$

Upon fixing $\mathbb{P}(x \text{ received})$, x is restructured and $\mathbb{P}(y \text{ sent})$ is constant as all codewords are equally likely to be sent. Therefore, $\mathbb{P}(x \text{ received} \mid y \text{ sent})$ is maximised as a function of the variable y precisely when $\mathbb{P}(y \text{ sent} \mid x \text{ received})$ is maximised, and the claim follows.

As with ideal observer decoding, a convention must be agreed to for non-unique decoding.

The maximum likelihood decoding problem can also be modeled as an integer programming problem.[1]

The maximum likelihood decoding algorithm is an instance of the "marginalize a product function" problem which is solved by applying the generalized distributive law.[2]

Minimum distance decoding

Given a received codeword $x \in \mathbb{F}_2^n$, **minimum distance decoding** picks a codeword $y \in C$ to minimise the Hamming distance:

$$d(x, y) = \#\{i : x_i \neq y_i\}$$

i.e. choose the codeword y that is as close as possible to x.

Note that if the probability of error on a discrete memoryless channel p is strictly less than one half, then *minimum distance decoding* is equivalent to *maximum likelihood decoding*, since if

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

then:

$$\begin{aligned}\mathbb{P}(y \text{ received} \mid x \text{ sent}) &= (1-p)^{n-d} \cdot p^d \\ &= (1-p)^n \cdot \left(\frac{p}{1-p}\right)^d\end{aligned}$$

which (since p is less than one half) is maximised by minimising d.

Minimum distance decoding is also known as *nearest neighbour decoding*. It can be assisted or automated by using a standard array. Minimum distance decoding is a reasonable decoding method when the following conditions are met:

1. The probability p that an error occurs is independent of the position of the symbol.
2. Errors are independent events – an error at one position in the message does not affect other positions.

These assumptions may be reasonable for transmissions over a binary symmetric channel. They may be unreasonable for other media, such as a DVD, where a single scratch on the disk can cause an error in many neighbouring symbols or codewords.

As with other decoding methods, a convention must be agreed to for non-unique decoding.

Syndrome decoding

Syndrome decoding is a highly efficient method of decoding a linear code over a *noisy channel*, i.e. one on which errors are made. In essence, syndrome decoding is *minimum distance decoding* using a reduced lookup table. This is allowed by the linearity of the code.[3]

Suppose that $C \subset \mathbb{F}_2^n$ is a linear code of length n and minimum distance d with parity-check matrix H. Then clearly C is capable of correcting up to

$$t = \left\lfloor \frac{d-1}{2} \right\rfloor$$

errors made by the channel (since if no more than t errors are made then minimum distance decoding will still correctly decode the incorrectly transmitted codeword).

Now suppose that a codeword $x \in \mathbb{F}_2^n$ is sent over the channel and the error pattern $e \in \mathbb{F}_2^n$ occurs. Then $z = x + e$ is received. Ordinary minimum distance decoding would lookup the vector z in a table of size $|C|$ for the nearest match - i.e. an element (not necessarily unique) $c \in C$ with

$$d(c, z) \leq d(y, z)$$

for all $y \in C$. Syndrome decoding takes advantage of the property of the parity matrix that:

$$Hx = 0$$

for all $x \in C$. The *syndrome* of the received $z = x + e$ is defined to be:

$$Hz = H(x + e) = Hx + He = 0 + He = He$$

To perform ML decoding in a binary symmetric channel, one has to look-up a precomputed table of size 2^{n-k}, mapping He to e.

Note that this is already of significantly less complexity than that of a standard array decoding.

However, under the assumption that no more than t errors were made during transmission, the receiver can look up the value He in a further reduced table of size

$$\textstyle\sum_{i=0}^{t} \binom{n}{i} < |C|$$

only (for a binary code). The table is against pre-computed values of He for all possible error patterns $e \in \mathbb{F}_2^n$.

Knowing what e is, it is then trivial to decode x as:

$$x = z - e$$

For **Binary** codes, if both k and $n - k$ are not too big, and assuming the code generating matrix is in standard form, syndrome decoding can be computed using 2 precomputed lookup tables and 2 XORs only. [4]

Let z be the received noisy codeword, i.e. $z = x \oplus e$. Using the encoding lookup table of size 2^k, the codeword z' that corresponds to the first k bits of z is found.

The syndrome is then computed as the last $n - k$ bits of $s = z \oplus z'$ (the first k bits of the XOR are zero [since the generating matrix is in standard form] and discarded). Using the syndrome, the error e is computed using the syndrome lookup table of size 2^{n-k}, and the decoding is then computed via $x = z \oplus e$ (for the codeword, or the first k bits of x for the original word).

The number of entries in the two lookup tables is $2^k + 2^{n-k}$, which is significantly smaller than 2^n required for standard array decoding that requires only 1 lookup. Additionally, the precomputed encoding lookup table can be used for the encoding, and is thus often useful to have.

Partial response maximum likelihood

Partial response maximum likelihood (PRML) is a method for converting the weak analog signal from the head of a magnetic disk or tape drive into a digital signal.

Viterbi decoder

A Viterbi decoder uses the Viterbi algorithm for decoding a bitstream that has been encoded using forward error correction based on a convolutional code. The Hamming distance is used as a metric for hard decision Viterbi decoders. The *squared* Euclidean distance is used as a metric for soft decision decoders.

See also

- Error detection and correction

Sources

- Hill, Raymond (1986). *A first course in coding theory* (https://archive.org/details/firstcourseincod0000hill). Oxford Applied Mathematics and Computing Science Series. Oxford University Press. ISBN 978-0-19-853803-5.
- Pless, Vera (1982). *Introduction to the theory of error-correcting codes*. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons. ISBN 978-0-471-08684-0.
- J.H. van Lint (1992). *Introduction to Coding Theory* (https://archive.org/details/introductiontoco0000lint). GTM. **86** (2nd ed.). Springer-Verlag. ISBN 978-3-540-54894-2.

References

1. Feldman, Jon; Wainwright, Martin J.; Karger, David R. (March 2005). "Using Linear Programming to Decode Binary Linear Codes". *IEEE Transactions on Information Theory*. **51** (3). pp. 954–972. doi:10.1109/TIT.2004.842696 (https://doi.org/10.1109%2FTIT.2004.842696).
2. Aji, Srinivas M.; McEliece, Robert J. (March 2000). "The Generalized Distributive Law". *IEEE Transactions on Information Theory*. **46** (2): 325–343. doi:10.1109/18.825794 (https://doi.org/10.1109%2F18.825794).
3. Albrecht Beutelspacher & Ute Rosenbaum (1998) *Projective Geometry*, page 190, Cambridge University Press ISBN 0-521-48277-1
4. Jack Keil Wolf (2008) *An Introduction to Error Correcting Codes*, Course: Communication Systems III, UCSD, http://circuit.ucsd.edu/~yhk/ece154c-spr17/pdfs/ErrorCorrectionI.pdf