誤り訂正符号の歴史と展望

丸山 直昌　データ科学研究系　准教授

1　誤り訂正符号とは

雑音の多い電話回線を通じて、例えば「たちかわ」という言葉を相手に伝えたい時、「たけのこ」の「た」、「ちくわ」の「ち」、「からし」の「か」、「わらび」の「わ」、というふうに伝える方法があります。聞き手が「たけのこ」を「らけのこ」に聞き違えたとしても、「らけのこ」という食物はありませんから、元の単語「たけのこ」を容易に想像でき、最初の文字「た」を認識できます。このやり方は非常に雑音が多い電話での通信などに用いられていた方法ですが、この手法を数学的な枠組で語ると誤り訂正符号の理論になります。やっていることは、伝えたい情報に「余分な情報」を付け加えて送り、受信側で余分な情報を除去する、ということですが、通信路の途中で雑音によって内容が歪められても、「余分な情報」に助けられて、送った情報を再現できる、という仕掛けです。この仕掛けがうまく機能するためには、送信側と受信側が共通の辞書を持っていることが重要な鍵となります。

2　誤り訂正符号で用いる「辞書」

雑音が多い電話での通信を「単語」を繰返し送る、というふうに見ることができます。ここで「単語」は文字の列ですが、どんな文字の羅列でも良いかというと、そうではなく、ある特別の文字列だけを「単語」として扱うわけです。つまり、「たけのこ」はここでは「単語」として扱いますが、「らけのこ」は単語とは扱いません。数学的には、あらゆる文字の羅列の中で「単語」というのはある部分集合をなしている、と考えます。もう少し抽象化して言うと、通信路を通る可能性がある信号の全体(あらゆる文字の羅列に相当)の中で、都合が良い部分集合(単語の全体に相当)を決めておいて、それによって上記のような「雑音対策」を成功させようとするわけです。成功するには、歪められているかもしれない受信情報から真の送信情報(これは部分集合の元)を捜し出す方法が簡単であることが必要で、その手法を「復号法」と呼びます。

3　線形符号

有限体上のベクトル空間を全体集合として、都合の良い部分空間を使ってこの誤り訂正がうまく機能する場合が数多く知られています。qを素数巾、$\mathbb{F}_q$をq個の元から成る有限体、Cを${\mathbb{F}_q}^n$のk次元部分空間とするとき、$(x_1, x_2, \ldots, x_k)$に$n-k$個の成分を付加して$(x_1, x_2, \ldots, x_n)$がCの元となるようにします。射影

$$p : (x_1, x_2, \ldots, x_n) \mapsto (x_1, x_2, \ldots, x_k)$$

によりCが$\mathbb{F}_q^k$に全射で写っていればそれは可能で、$(x_1, x_2, \ldots, x_n)$　は一意的です。この場合、送信したい情報$(x_1, x_2, \ldots, x_k)$　に対して$(x_1, x_2, \ldots, x_n)$　を送信し、受信した情報$(y_1, y_2, \ldots, y_n)$　に一番「近い」Cの元$(z_1, z_2, \ldots, z_n)$を探して$(z_1, z_2, \ldots, z_k)$を$(x_1, x_2, \ldots, x_k)$の推定値とします。$(y_1, y_2, \ldots, y_n)$　から$(z_1, z_2, \ldots, z_n)$を見つける手続きが復号法となります。

4　代数幾何的符号

20世紀初頭から1980年頃まで、誤り訂正符号は、組み合わせ論や、代数的な手法を使って研究されてきましたが、1980年代中ごろに代数幾何学の枠組を使ってとらえ直す研究が出てきました。

Xを$\mathbb{F}_q$上の特異点が無い代数曲線、$\overline{X}$をその非特異完備化とし、X上のn個の点集合$\{p_1, p_2, \ldots, p_n\}$を考え、$\overline{X} - X$から点 p_0 を選びます。$\overline{X}$上の有理関数でp_0で$k-1$位以下の極を持ち、他に極を持たないもの全体が成す線形空間をLとし、写像$\psi : L \to {\mathbb{F}_q}^n$を、$f \in L$に対して$\psi(f) = (f(p_1), f(p_2), \ldots, f(p_n))$と決めると、$\psi(L) \subset {\mathbb{F}_q}^n$が符号の役割を果たします。19世紀に遡るリーマン・ロッホの定理がこの部分集合の誤り訂正符号としての都合が良い性質を証明することに役立ちます。

このようにして構成される符号は一般に代数幾何的符号と呼ばれます。

5　リードソロモン符号

リードソロモン符号は元々巡回符号の理論の枠組の中で、代数的な手法で定義されていましたが、代数幾何的符号として表現できます。X　が射影直線でp_0が無限遠点とした場合、リードソロモン符号と等価です。具体的に書きますと、$n = q - 1$とし、有限体$\mathbb{F}_q$のq個の元の一個を除いて$p_1, p_2, \ldots, p_n$とし、

$$RS(n,k) = \{(f(p_1), f(p_2), \ldots, f(p_n)) | f(x) \text{は}$$

$$\text{次数} k-1 \text{以下の} \mathbb{F}_q \text{係数多項式}\}$$

とすると、$RS(n,k)$は$n(=q-1)$次元ベクトル空間の中のk次元部分空間で、これがリードソロモン符号の代数幾何学的な表示を与えます。

6　誤り訂正符号の理論の今後

リードソロモン符号は非常に有力な誤り訂正符号で、CD, DVDなどにも、衛星通信にも用いられていますが、復号法に関してはより強力な方法があるのではないか、という方向での研究は続けられています。また、これまで知られていなかった符号を探す研究もされています。単に「探す」といっても、適切な復号法が見つからなければ使い道がないので、新しい符号を探すということと、新しい復号法を探す、という研究は相互に結び付いています。近年になって代数幾何的符号の復号を多項式の因数分解に結び付ける研究が出てきて、数式処理やグレーブナー基底の理論とも結び付いて、興味が尽きない研究対象となっています。