Códigos Reed-Solomon

Corpos finitos com aplicações em teoria de códigos

CNPq/UFRGS
Apoio

Lucas Gabriel Mota da Silveira (lucasgmsilveira@gmail.com)
Bolsista

Vilmar Trevisan (trevisan@mat.ufrgs.br)
Orientador

Introdução

Quando transmitimos alguma informação, queremos ter certeza de que a informação enviada é igual à recebida, mas essa transmissão está sujeita a ruídos que causam erros na mensagem, como perda ou deslocamento de símbolos. A Teoria de Códigos estuda técnicas para detectar e corrigir tais erros, utilizando para isto resultados de aritmética de corpos finitos, polinômios irredutíveis e álgebra linear.

Os códigos Reed-Solomon são um tipo específico de códigos cíclicos que corrigem até t erros, amplamente utilizados na troca de dados, com usos em telefonia, CDs, transmissão de satélites etc.

Códigos Lineares

Como o próprio nome sugere, os *códigos lineares* estão baseados em várias ferramentas da álgebra linear. Assim, é importante visualizar $GF(q)^n$ não apenas como um corpo finito, mas também como um espaço vetorial sobre $GF(q)$.

Definição 1 Seja H uma matriz $(n\text{-}k)_x n$ com entradas em $GF(q)$. Dizemos que o conjunto

$$C = \{c \in GF(q)^n \; ; \; Hc^T = 0\}$$

é um *código linear* sobre $GF(q)$, também denotado por $C(n,k)$, enquanto que n é o *comprimento* e k é a *dimensão* do código. Os elementos de C são chamados *palavras-códigos* e H a *matriz de paridade* de C. Se G é uma matriz $k_x n$ cujo espaço gerado pelas linhas é igual a C, então dizemos que G é a *matriz geradora* de C.

Pela definição de G, qualquer $c = a.G$ com $a \in GF(q)^n$ é uma palavra-código. Portanto, ambas matrizes H e G podem ser usadas para obter C.

Distância de Hamming

Definição 2 Sejam $x,y \in GF(q)^n$. A *distância de Hamming* $d(x,y)$ é o número de coordenadas nas quais x e y diferem.

Definição 3 Se c é uma palavra-código e y uma palavra recebida, então o *erro* é a diferença $e = y\text{-}c$.

Definição 4 Seja t um inteiro positivo. Um código $C \in GF(q)^n$ *corrige t erros* se para cada palavra recebida $y \in GF(q)^n$ há no máximo um $c \in C$ tal que $d(c,y) \leq t$.

Definição 5 A *distância mínima* de um código C é definida por $d_C = \min(u,v)$, com $u,v \in C$, $u \neq v$.

Teorema 1 Um código C com distância mínima d_C pode corrigir até t erros se $d_C \geq 2t+1$.

Para sabermos se um código corrige t erros, devemos investigar d_C. Calcular d_C percorrendo as q^k palavras-código pode ser muito custoso. Uma alternativa é proposta a seguir.

Teorema 2 Um código linear C com matriz de paridade H tem $d_C \geq s+1$ se e somente se quaisquer s colunas de H são linearmente independentes.

Códigos Cíclicos

Definição 6 Um código linear $C(n,k)$ sobre $GF(q)$ é *cíclico* se $(c_0,c_1,..,c_{n-1}) \in C$ implica que $(c_{n-1},c_0,..,c_{n-2}) \in C$.

Consideramos $GF(q)^n$ representado por polinômios de grau menor que n em $GF(q)[x]/(x^n\text{-}1)$, associando o vetor $c=(c_0,c_1,..,c_{n-1})$ ao polinômio $c(x)=c_0+c_1x+...+c_{n-1}x^{n-1}$. O deslocamento cíclico em c corresponde à multiplicação de $c(x)$ por x.

Teorema 3 Um código linear $C(n,k)$ sobre $GF(q)$ é cíclico se e somente se C é um ideal de $GF(q)[x]/(x^n\text{-}1)$.

Como $GF(q)[x]$ é um domínio de ideais principais, todo ideal não-nulo em $GF(q)[x]/(x^n\text{-}1)$ é gerado por um polinômio mônico g de grau mínimo no ideal. Portanto todo código cíclico é gerado por um polinômio.

Definição 7 Seja $C=(g)$ um código cíclico. Dizemos que g é o *polinômio gerador* de C e $h=(x^n\text{-}1)/g$ é o *polinômio verificador* de C.

Teorema 4 Seja C um ideal não-nulo em $GF(q)[x]/(x^n\text{-}1)$, isto é, C é um código cíclico de comprimento n.

1. O código C é gerado por um único polinômio g de grau mínimo em C.
2. O polinômio gerador g de C é um fator de $x^n\text{-}1$.
3. Em $GF(q)[x]$, qualquer $c \in C$ pode ser escrito unicamente como $c=f.g$, onde $grau(f)<n\text{-}r$ e $grau(g)=r$. Além disso, a dimensão de C é $n\text{-}r$.
4. Se $g(x)=g_0+g_1x+...+g_rx^r$, então C é gerado como um subespaço de $GF(q)^n$ pelas linhas da matriz geradora

$$G = \begin{pmatrix} g_0 & g_1 & \cdots & & g_r & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & & g_r & 0 & \cdots & 0 \\ & & & & \vdots & & & & \\ 0 & \cdots & & 0 & g_0 & g_1 & & \cdots & g_r \end{pmatrix}$$

Códigos Reed-Solomon

Códigos Reed-Solomon tem, por definição, comprimento $n=q\text{-}1$, onde $q\neq 2$.

Exemplo Sejam $q=5$, $n=4$. Um elemento primitivo em $GF(5)$ é $\alpha=2$. Assim, $g(x) = (x\text{-}\alpha)(x\text{-}\alpha^2) = x^2+4x+3$. Como $k=n\text{-}grau(g)=2$, a dimensão do código é 2. Assim, temos $q^k = 5^2 = 25$ palavras-código. A matriz geradora é

$$G = \begin{pmatrix} 3 & 4 & 1 & 0 \\ 0 & 3 & 4 & 1 \end{pmatrix}$$

Alguns exemplos de palavras-código são:

(0 0)G = (0 0 0 0)
(1 0)G = (3 4 1 0)
(2 0)G = (1 3 2 0)

Os códigos Reed-Solomon possuem $d_C = n\text{-}k+1$, sendo classificados como códigos MDS (distância máxima separável), significando que conseguem atingir a distância mínima mais alta possível.

Referências

MACWILLIAMS, F.J.; SLOANE, N.J.A. (1978). The Theory of Error-Correcting Codes.
MASUDA, A.M.; PANARIO, D. (2007), Tópicos de Corpos Finitos com Aplicações em Criptografia e Teoria de Códigos.
MULLEN, G.L.; MUMMENT, C. (2007), Finite Fields and Applications.
CAMPELO, D.G. (2012), Decodificação de Códigos Não Sistemáticos de Reed-Solomon.