

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA
DEPARTAMENTO DE COMUNICAÇÕES

CODIFICAÇÃO ÓTIMA DE PADRÕES E ROTULAÇÃO DE SINAIS
VIA ANÁLISE DE CORRESPONDÊNCIA, E HOMEOMORFISMO ENTRE
SUBESPAÇOS DE HAMMING, EUCLIDEANO E EUCLIDEANO PONDERADO

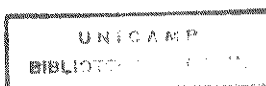
BLANCA ROSA MAQUERA SOSA

Este exemplar corresponde à redação final da tese
defendida por Blanca Rosa Maquera
Sosa. aprovada pela Comissão
Julgadora em 21. 09. 93.
Reginaldo Palazzo Jr.
Orientador

Orientador : PROF. DR. REGINALDO PALAZZO JR.

Tese de Doutorado apresentada
à Faculdade de Engenharia Elé-
trica da Universidade Estadual
de Campinas, como parte dos re-
quisitos exigidos para a ob-
tenção do título de *DOUTOR EM*
ENGENHARIA ELÉTRICA

Setembro/1993



LA MAYOR NECESIDAD DEL MUNDO ES LA DE:

Hombres que no se vendan ni se compren;

*Hombres que sean sinceros y honrados en lo más íntimo
de sus almas;*

*Hombres que no teman dar al pecado el nombre que le
corresponde;*

*Hombres cuya conciencia sea tan leal al deber como la
brújula al polo;*

*Hombres que se mantengan de parte de la justicia aunque
se desplomen los cielos.*

Elena de White

Dedico esta tese aos meus pais,

Arnaldo e Florencia,
pelo amor e exemplo.

E aos meus irmãos,

Manuel, Selma, Jorge e Gladys
por tudo o que eles significam para mim.

AGRADECIMENTOS

- A Deus porque permitiu que fizesse a minha Pós-Graduação no Brasil. Um país que me deu a oportunidade de participar de suas vivências, costumes e alegrias que muito contribuíram para meu desenvolvimento pessoal e profissional.
- Ao meu orientador Prof. Dr. Reginaldo Palazzo Júnior pela sua orientação segura e valiosa, seu constante apoio e sua amizade.
- Aos Profs. Drs. Pushpa N. Rathie do Departamento de Estatística da Universidade de Brasília - UnB, Regina C.C.P. Moran do IMECC-UNICAMP, Lee L. Ling e Yuzo Iano ambos da FEE-UNICAMP, por terem aceito fazer parte da Banca Examinadora.
- Ao Jorge, meu irmão e amigo, por ter-me acompanhado durante este doutorado e acima de tudo por ter representado meu lar.
- Aos amigos da FEE, do IMECC e conterrâneos pelas experiências compartilhadas. Especialmente quero agradecer ao Carlos Eduardo Câmara (Dinho), ao José Carmelo Interlando e ao Alexandre L.T. Carvalho pelas discussões e comentários a este trabalho; à minha companheira de sala de estudos, Myriam, pela sua alegria e otimismo, à Andrea, pelo grupo de oração, à Ana e Ofélia e aos casais Jorge e Vasty, Fernando e Vaneide, Carlos e Nely e Élvio e Cristina pelo apreço manifesto em muitas vezes.
- À Elza Aoki porque com muito zelo, digitou o presente trabalho e aos desenhistas Lúcia e Airton pela confecção dos desenhos.
- À UNICAMP (Universidade Estadual de Campinas) por proporcionar-me condições para a conclusão deste trabalho.
- À CAPES pela ajuda financeira.

RESUMO

Nesta tese desenvolvemos itens das áreas de Estatística, Matemática e Comunicações dentro da Engenharia Elétrica. Da área de Estatística, consideramos os itens da Análise de Correspondência, Reconhecimento de Padrões e Teoria de Decisão. Na área de Matemática, a Teoria de Grupos, Corpos de Galois e Topologia, e da área de Comunicações, Codificação, Modulação e Modulação-Codificada. Todas estas áreas possibilitaram um encaminhamento estruturado nos campos de Codificação de Fonte e Codificação de Canal.

No contexto de Codificação de Fonte, mostramos uma forma ótima de codificar os padrões Via Análise de Correspondência (AC) até então, não utilizada dentro da área de comunicações.

Com respeito a Codificação de Canal, mostramos que os conjuntos de sinais simplex cujo número de elementos é da forma 2^n , $n \in \mathbb{Z}^+$, são rotulados com elementos de grupos, cujas características são comutativas, não cíclico e cada elemento operado por ele mesmo, resulta na identidade do grupo. Também mostramos que através da AC, é possível rotular os elementos do conjunto de sinais com elementos de um grupo.

Finalmente, como uma consequência desta rotulação, mostramos que os subespaços métricos dos espaços de Hamming, Euclideano e Euclideano Ponderado são homeomorfos.

ABSTRACT

This thesis encompasses subjects from the fields of Statistics, Mathematics and Communication within the fields of Electrical Engineering. Correspondence Analysis, Pattern Recognition and Decision Theory are some areas of statistics which are considered. In Mathematics, the Groups Theory, Galois Fields and Topology it is utilized, and Communication, Codification, Coded Modulation. All these areas make it possible direction structure in the field of Source Codification and Channel Codification.

In Source Coding context, it is presented an optimum form of coding the patterns by use of Correspondence Analysis. To the best of our knowledge, this is shown for the first time. This technique generalizes the Huffman coding technique for discrete-time source coding.

In terms of Channel Coding it is show that the set of simplex signals, whose number of elements is of the form 2^n , $n \in \mathbb{Z}^+$, are labeled as group elements whose characteristics are: commutative, non cyclic and idempotent. It is also shown that it is possible to label the signal set elements as group elements through Correspondence Analysis.

Finally, as a consequence of this labeling it is shown that these exists a homeomorphism between each pair of the following spaces: Hamming, Euclidian, and weighted Euclidian.

ÍNDICE

CAPÍTULO 1

INTRODUÇÃO	1
1.1. O QUE É RECONHECIMENTO DE PADRÕES?	3
1.2. COMO SERÁ ROTULADO UM CONJUNTO DE SINAIS SIMPLEX POR ELEMENTOS DE UM GRUPO?	6

CAPÍTULO 2

ANÁLISE DE CORRESPONDÊNCIA	8
2.1. INTRODUÇÃO	9
2.2. ELEMENTOS DA ANÁLISE DE CORRESPONDÊNCIA	10
2.2.1. Matriz de Correspondência, Perfis e Centróide	10
2.2.2. Distância χ^2 , Inércia e Equivalência Distribucional	13
2.3. GERAÇÃO DE SUBESPAÇOS ÓTIMOS	15
2.4. COORDENADAS PRINCIPAIS DOS PERFIS DAS LINHAS, COM RESPEITO AO CENTRÓIDE	21
2.5. DECOMPOSIÇÃO DA INÉRCIA E INTERPRETAÇÃO DOS PERFIS NOS SUBESPAÇOS	23
2.5.1. Contribuição Absoluta	24
2.5.2. Contribuição Relativa	25
2.6. PERFIS SUPLEMENTARES	26

CAPÍTULO 3

RECONHECIMENTO DE PADRÕES MATEMÁTICO E CLASSIFICAÇÃO	28
---	----

3.1. INTRODUÇÃO	29
3.2. RECONHECIMENTO DE PADRÕES MATEMÁTICO.....	30
3.2.1. Espaço Padrão	31
3.2.2. Espaço de Características	32
3.2.3. Espaço de Classificação	33
3.3. SELEÇÃO E EXTRAÇÃO DE CARACTERÍSTICAS	33
3.3.1. Transformação Diagonal	35
3.3.2. Transformação Rotacional	37
3.4. EXTRAÇÃO DE CARACTERÍSTICAS ATRAVÉS DA ANÁLISE DE CORRESPONDÊNCIA.....	40
3.5. CLASSIFICAÇÃO DE PADRÕES.....	41
3.5.1. Classificação de Padrões Utilizando a Teoria de Decisão de Bayes.....	43
3.5.1.1. Função Perda	44
3.5.1.2. Função discriminante de Bayes	46
3.5.1.3. Decisão de máxima verossimilhança	48
3.5.2. Classificação de Padrões Quando a Função de Densidade é Gaussiana.....	51

CAPÍTULO 4

CASAMENTO ENTRE CONJUNTO DE SINAIS E GRUPOS	55
4.1. INTRODUÇÃO	56
4.2. GRUPOS	56
4.2.1. Definições e Alguns Exemplos.....	56
4.2.2. Grupos Quocientes e Homomorfismo de Grupos	60
4.2.3. Anéis e Corps	61
4.3. REVISÃO DE SISTEMAS DE COMUNICAÇÕES	63
4.3.1. Codificação de Fonte	64
4.3.2. Canais de Comunicação	66
4.3.3. Codificação de Canal	67
4.3.2.1. Codificação	67

4.3.3.2. Modulação	72
4.3.3.3. Probabilidade de erro em conjuntos de sinais	76
• Sinais binários	78
• Sinais simplex	79
• Sinais M-PSK	81
4.3.3.4. Modulação-codificada	82
4.4. CASAMENTO DE CONJUNTO DE SINAIS COM GRUPOS	82
4.5. CASAMENTO ENTRE CONJUNTO DE SINAIS SLEPIAN E GRUPOS	87

CAPÍTULO 5

CODIFICAÇÃO E RECONHECIMENTO DE PADRÕES VIA ANÁLISE DE CORRESPONDÊNCIA	92
5.1. INTRODUÇÃO	93
5.2. ELEMENTOS DA TEORIA DE GALOIS PARA CORPOS FINITOS	94
5.2.1. Anel de Polinômios	94
5.2.1.1. Corpos de Galois	97
5.3. GERAÇÃO DE CÓDIGOS COM ELEMENTOS EM $\mathbb{Z}_p$	99
5.3.1. Geração de Códigos Lineares Binários	101
5.3.1.1. Códigos Reed-Muller de primeira ordem	103
5.3.2. Geração de Códigos p-ários	104
5.4. CODIFICAÇÃO DE PADRÕES VIA ANÁLISE DE CORRESPONDÊNCIA ...	106
5.4.1. Construção da "Tabela de Contingência" e Obtenção dos Vetores Sinais	106
5.4.2. Obtenção de Conjuntos de Sinais Simplex	109
5.4.2.1. Obtenção de conjuntos de sinais simplex através de códigos binários quando o número de sinais é 2^n , $n \in \mathbb{Z}^+$	110
5.4.2.2. Obtenção de conjuntos de sinais simplex cujos elementos das palavras código estão em $\mathbb{Z}_3$	119
5.5. CODIFICAÇÃO DE CLASSES DE PADRÕES	124

CAPÍTULO 6

HOMEOMORFISMO ENTRE SUBESPAÇOS MÉTRICOS DE HAMMING, EUCLIDEANO E EUCLIDEANO PONDERADO PARA O CONJUNTO DE SINAIS SIMPLEX	126
6.1. INTRODUÇÃO	127
6.2. CASAMENTO E ROTULAÇÃO DE CONJUNTOS DE SINAIS SIMPLEX A GRUPOS	128
6.2.1. Isomorfismo entre um Sub-grupo de SO_{2^n-1} e um Subgrupo de $(\mathbb{Z}_2)^{2^n-1}$	128
6.2.2. Mapeamento Casado e Rotulamento do Conjunto de Sinais Simplex em $\mathbb{R}^{2^n-1}$, com Elementos do Grupo $Z \subset (\mathbb{Z}_2)^{2^n-1}$	136
6.3. HOMEOMORFISMO ENTRE SUBESPAÇOS DOS ESPAÇOS MÉTRICOS DE HAMMING, EUCLIDEANO E EUCLIDEANO PONDERADO	140
6.3.1. Homeomorfismo entre Subespaços de Hamming e Euclideano ...	140
6.3.2. Homeomorfismo entre Subespaços dos Espaços Métricos Euclideano e Euclideano Ponderado	143

CAPÍTULO 7

CONCLUSÕES	148
------------------	-----

REFERÊNCIAS BIBLIOGRÁFICAS	153
----------------------------------	-----

GLOSSÁRIO

A	: matriz
A_n	: grupo alternante de grau n
(a, b)	: elementos de $U \times T$
AWGN	: Additive White Gaussian Noise
$[a]$	: classe de equivalência contendo a
$a \equiv b \text{ mod } H$	: a é congruente a b módulo H
aH	: classe lateral à esquerda
(a)	: ideal gerado por a
AC	: Análise de Correspondência
b_i	: i -ésimo vetor base de espaço J -dimensional
BSC	: Binary Symmetric Channel
B_i	: i -ésima matriz do grupo $\mathcal{B}$
$\mathcal{B}$	: grupo de matrizes ortogonais
λ	: autovalores da matriz Φ_s
C	: matriz dos perfis das colunas
C_s	: matriz de co-variância da classe ω_s quando a função de densidade é gaussiana
$\underline{c}$	: vetor massa dos perfis das colunas e vetor centróide dos perfis das linhas
CTR(i)	: contribuição absoluta do i -ésimo perfil linha
$\mathcal{C}$	: código
$\mathcal{C}'$	: subcódigo de $\mathcal{C}$
D_c	: matriz diagonal com elementos de c

D_r	: matriz diagonal com elementos de r
DVSG	: Decomposição de Valores Singulares Generalizada
D_μ	: matriz diagonal composta pelos valores singulares generalizados de A ou $(R - I \underline{c}^T)$
D_w	: matriz diagonal com elementos de w
$D^{(s)}$	: matriz diagonal que contém os autovalores de Φ_s
D_s^2	: distância quadrática média dos protótipos da classe ω_s
d_{χ^2}	: distância Chi-quadrado
$d_s(x)$ e $d'_s(x)$	: forma alternativa de função discriminante
$d_H(\cdot, \cdot)$	: distância de Hamming
$d_E^2(\cdot, \cdot)$	: distância Euclideana quadrada
d_{ij}	: distância entre os sinais s_i e s_j
$d_p(\cdot, \cdot)$	: distância de fase
$d_{\min}$	: distância mínima
$\mathcal{D}$	: grupo de matrizes ortogonais obtidas pelo uso da AC
E	: operador linear valor esperado
e	: vetor erro
$\text{erfc}(x)$	: função erro complementar
e	: elemento identidade do grupo
$\mathcal{E}$	: energia média
F	: matriz que contém as coordenadas dos perfis das linhas
f_s	: vetor de coordenadas do perfil suplementar
f	: função
F	: corpo
$F[x]$	: anel de polinômios
$f_r(r m_i)$	: função densidade condicional do vetor r dado m_i
$g_s(x)$	: função discriminante

G ou $(G, \cdot)$	: grupo algébrico
$G \cong H$	: G é isomorfo a H
G/N	: grupo quociente
G	: matriz geradora ou matriz de codificação do código
$GF(\cdot)$	: corpo de Galois
$GF(\cdot)^*$	: conjunto de elementos não nulos no $GF(\cdot)$
GL_n	: grupo linear geral
H	: matriz verificação de paridade
H ou $(H, \cdot)$	: subgrupo
$H(\cdot)$	: entropia
H_a	: classe lateral à direita
$\mathcal{H}$ e $\mathcal{H}'$	: subcódigo
h	: mapeamento casado ou mapeamento de rotulação
H_{2^n-1}	: espaço de Hamming de dimensão 2^n-1
I	: matriz identidade
I	: ideal
$In(I)$	: inércia dos perfis das linhas
$In(J)$	: inércia dos perfis das colunas
$Im f$	: imagem do homomorfismo f
$I(\cdot)$	: quantidade de informação
$Ker f$	: núcleo de homomorfismo f
K	: corpo
$[K:F]$	: grau da extensão K de um corpo F
L	: matriz perda
L	: eixo de simetria
L_{ij}	: função perda

$\ell_{ij}(x)$	: razão de verossimilhança
ℓ_k	: comprimento da palavra-código atribuída ao símbolo ζ_k
$\bar{\ell}$	: comprimento médio das palavras-código
M	: matriz ortogonal relativa a Λ
$M^{(s)}$	: matriz ortogonal que contém os autovetores de Φ_s
M	: número de padrões (Cap. 1 e 5); número de sinais ou formas de onda (Cap. 4)
M_s	: número de protótipos da classe padrão ω_s
m_s	: vetor de média de x
m_i	: i -ésima mensagem
M	: espaço métrico
N	: matriz ortogonal relativa a θ
n	: vetor ruído
$\mathbb{N}$	: conjunto dos números naturais
N/M	: N divide M
N	: eixo de simetrias
N	: espaço métrico
O_n	: grupo ortogonal
P	: matriz de correspondência
P	: ideal gerado por $p(x)$
$p(\omega_i)$	: probabilidade da classe ω_i
$p(x \omega_i)$	: probabilidade do vetor x condicionado a classe ω_i
$p(\omega_i x)$	: probabilidade da classe ω_i condicionado ao vetor x
$p(e)$	: probabilidade de erro
$p(c)$	: probabilidade de decisão correta
$p(x)$	: polinômio no indeterminado x

RP	: Reconhecimento de Padrões
R	: matriz que contém os perfis das linhas
$\underline{r}$	: vetor massa dos perfis das linhas e vetor centróide dos perfis das colunas
$\underline{r}_i$	: i-ésimo perfil linha
$\underline{r}_i$	: massa do i-ésimo perfil linha
R_P^I	: espaço ponderado I-dimensional
R_P^J	: espaço ponderado J-dimensional
R^n	: espaço Euclideano n-dimensional
$R_s(x)$	: perda média de má classificação de x , quando $x \in \omega_s$
$\underline{r}_s$	: perfil suplementar
R	: relação
R ou $(R, +, \circ)$	: anel
$R[x]$	: anel de polinômio com coeficiente no anel R
R/I	: anel quociente de R por I
r	: vetor aleatório gaussiano com média s_{ij} e variância $N_o/2$
$r(t)$	: sinal recebido pelo demodulador
r	: taxa do código
r	: seqüência que recebe o decodificador
$\mathcal{R}(1, m)$	: código Reed-Muller de primeira ordem
$\mathcal{R}_i$	: i-ésima região do espaço de características ou espaço de sinais
S	: número de classes (Cap. 3) e conjunto de sinais (Caps. 4, 5 e 6)
S'	: conjunto de sinais no espaço Euclideano ponderado
$s_i(t)$	: i-ésimo sinal em função do tempo
s_i	: vetor sinal

SO_n	: grupo ortogonal especial
S_n	: grupo simétrico de grau n ou grupo de permutações
$\mathcal{S}$	: subespaço de dimensão $K < J$
$\mathcal{S}^*$	: subespaço ótimo de dimensão $K^* < K$
T	: conjunto
TC	: Tabela de Contingência
t	: capacidade de correção do código
T_{B_i}	: transformação linear
$T_{\mathcal{B}}(S)$	: grupo de transformações ortogonais determinado por $\mathcal{B}$
T_f	: transformação ortogonal de $\mathbb{R}^n$ em $\mathbb{R}^n$
U	: conjunto
$U \times T$	: conjunto produto cartesiano de U e T
u	: seqüência de k dígitos de informação
U/R	: conjunto quociente de U por R
v ou v_i	: palavra-código ou i -ésima palavra-código
w	: vetor peso
W	: matriz ortogonal relativa a D_r
$W(\cdot)$	: peso de Hamming
$W_p(\cdot)$	: peso de fase
X	: matriz de freqüências de ocorrência
x_{ij}	: freqüência de ocorrência na i -ésima linha e na j -ésima coluna
$x_{i\cdot}$	: soma dos elementos da i -ésima linha
$x_{\cdot j}$	: soma dos elementos da j -ésima coluna

$\underline{x}$	: vetor de características
$\underline{X}$	: matriz padrão
$\underline{x}$	: vetor padrão
$\underline{x}_i$	: componente do vetor padrão
Y	: matriz ortonormal relativa a D_c^{-1}
y_i	: i-ésimo vetor que pertence a R_p^J
$\hat{y}_i$	: aproximação do y_i ao subespaço $\mathcal{P}$
$y^{(s)}$	: variável aleatória protótipo
$y_m^{(s)}$	: m-ésimo protótipo da classe ω_s
$\bar{y}$	: centróide dos y_i
$z^{(s)}$	: variável aleatória protótipo
$z_m^{(s)}$	: valor da variável aleatória $z^{(s)}$ para o m-ésimo protótipo
$\mathbb{Z}$	: conjunto dos números inteiros
$\mathbb{Z}^+$	: inteiros positivos
$\mathbb{Z}_n$	: anel dos inteiros módulo n
$\mathbb{Z}_q^n$	: conjunto de todas as n-uplas
$\mathbb{Z}_p[x]$	: anel do polinômios com coeficientes em $\mathbb{Z}_p$
$\mathbb{Z}_q^n/\mathcal{C}$	: grupo quociente de $\mathbb{Z}_q^n$ em $\mathcal{C}$
$\mathcal{Z}$	: subgrupo de $(\mathbb{Z}_2)^{2^n-1}$
α	: elemento primitivo de $GF(p^m)$
χ^2	: estatística Chi-quadrada
η ou η_H	: transformação linear de $\mathbb{Z}_q^n$ em $\mathbb{Z}_q^{n-k}$
γ ou γ_G	: codificador ou transformação linear de $\mathbb{Z}_q^k$ em $\mathbb{Z}_q^n$
Δ	: grupo transitivo de isometrias
$\Delta(s)$	: órbita de s sob Δ
$\Gamma(S)$	: grupo de simetrias de S

$\Gamma'(S)$	: grupo isomorfo a $\Gamma(S)$
$\delta(\cdot)$	: função delta de Kronecker
Λ	: matriz simétrica
Ω	: conjunto de S classes de padrões
ω_i	: i-ésimo elemento do conjunto Ω
φ	: isomorfismo entre $\mathcal{Z}$ e $\mathcal{B}$
Φ_s	: matriz de co-variância da classe ω_s
Φ	: matriz de co-variância total
ϕ_i	: i-ésima função da base ortonormal
Ψ	: isomorfismo entre G/K e $\text{Im } f$
ψ	: função que depende dos subespaços $\mathcal{S}$ e $y_{i's}$
ρ	: isomorfismo entre $(0, 1)$ e $(1, -1)$
σ_i^2	: variância de i-ésima dimensão
τ	: percentagem da soma de quadrados
θ_{ij}	: valor limiar entre as classes ω_i e ω_j
θ_k	: ângulo do vetor perfil com respeito ao k-ésimo eixo (Cap. 2) e ângulo de fase (Cap. 4)
θ	: matriz simétrica
$\mathcal{E}_k^{(s)}$	: erro de aproximação da dimensão I a dimensão K
$\mathcal{E}$	: energia média
ζ	: variável aleatória discreta
μ_i	: i-ésimo valor singular generalizado
$\#G$	: ordem do grupo G
$\oplus$	: adição módulo p
$ a $	: valor absoluto de a
$\ a\ $	: norma do vetor a
$\ominus$	: diferença módulo M
$+$	: soma de vetores

CAPÍTULO 1

INTRODUÇÃO

Em comunicação digital, o teste de hipóteses estatístico é fundamental num modelamento de problemas tais como delineamento de sinais e obtenção do receptor ótimo baseados em um critério específico; por exemplo, probabilidade mínima de erro e o conhecimento de estatística do canal.

No problema de delineamento de sinais, para um dado conjunto de formas de onda, o problema é encontrar o melhor arranjo de pontos de sinais pertencentes ao espaço de sinal tal que a probabilidade de erro seja mínima. Sob o ponto de vista geométrico este problema faz uso do procedimento de ortogonalização de Gram-Schmidt via a decomposição de um sinal no tempo através de uma base ortogonal com a correspondente representação vetorial do mesmo.

Para este problema clássico o melhor arranjo do vetor sinal é o simplex quando nenhuma restrição sobre a largura de faixa é imposta e o ruído é assumido ser o ruído Gaussiano aditivo branco. Por outro lado, quando existe restrição na largura de faixa então, uma forma combinada de modulação e codificação deve ser imposta.

Transmitido o sinal através do canal Gaussiano efetua-se o processo de demodulação e decodificação que tem como objetivo estimar a palavra-código e a sequência de informação respectivamente. Estes processos baseiam-se na decisão de máxima verossimilhança a qual minimiza a probabilidade de erro.

Nosso trabalho compreende duas partes: a primeira trata sobre a Codificação de Padrões a qual está no contexto de Codificação de Fonte e a segunda sobre o homeomorfismo entre subespaços de Hamming e Euclideano os quais estão baseados na rotulação de conjuntos de sinais simplex quando o número de elementos é 2^n , $n \in \mathbb{Z}^+$, que está dentro do contexto de Codificação de Canal quando temos uma forma combinada de modulação e codificação.

Primeiramente, introduziremos o conceito de reconhecimento de padrões.

1.1. O QUE É RECONHECIMENTO DE PADRÕES?

Como seres humanos em nosso dia a dia fazemos um constante reconhecimento de padrões através de nossos sentidos, lógica, etc....

O problema de reconhecimento de padrões (RP) usualmente denota a classificação e/ou descrição de um conjunto de processos e/ou eventos. Estes conjuntos podem ser de objetos físicos ou algo mais abstrato como de estado mental.

Uma aproximação intuitiva do RP é o "casamento-protótipo", que consiste num conjunto de protótipos onde é atribuído um para cada classe padrão a qual é armazenada na máquina. O padrão de entrada (classificação desconhecida) é comparado com os protótipos de cada classe, e a classificação é baseada sobre um critério de pré-seleção. Se o padrão de entrada é melhor casado com o protótipo da i -ésima classe padrão, então a entrada é classificada como da i -ésima classe padrão.

Na Teoria da Decisão, em lugar de um simples casamento do padrão de entrada com o protótipo, a classificação é baseada em um conjunto de extração de características do padrão de entrada.

A extração de características é uma parte fundamental dentro do processo de RP, alguns consideram isto como uma codificação. As medidas das características são supostas ser invariantes ou menos sensíveis com respeito a variações ou distorções e de menor redundância.

Geralmente, considera-se o RP como sendo composto de duas partes. A primeira está relacionada com que tipos de medidas podem ser padrões de entrada. Usualmente a decisão de que medidas serão consideradas depende de situações práticas tais como a disponibilidade de medidas, o custo delas, etc.... A sele-

ção de características é importante porque através delas os padrões são caracterizados para o desempenho ou precisão do reconhecimento. O segundo problema em RP é a classificação (ou fazer decisão sobre a classe atribuída ao padrão de entrada) baseado nas medidas tomadas pela seleção ou extração de características. Assim, em termos gerais a aproximação "protótipo-casamento" pode ser interpretada como um caso especial de aproximação da Teoria da Decisão, onde os protótipos são armazenados em termos de medidas de características e um critério de classificação especial é utilizada pelo classificador.

Atualmente existem diferentes técnicas e algoritmos para a extração de características dos padrões e a classificação de um padrão de entrada. Estas técnicas variam dependendo do tipo de padrão. Por exemplo, para o processamento de imagens um dos algoritmos utilizado é o de Thinning [1]. O uso de momentos invariantes para características sensíveis dos padrões [2] é uma outra técnica utilizada em RP. Apesar da existência e proposição de novas técnicas a transformada de Fourier continua tendo considerável aplicação em processamento digital.

Nos últimos anos, as pesquisas em RP tem empregado a técnica de decomposição de valores singulares (DVS) em diferentes campos tais como compressão de dados, processamento de sinais, análise de padrões, extração de Características [3]. Utilizando DVS com poucos valores dominantes, obtém-se alto desempenho na separação de classes as quais são suficientes para classificação e, portanto, possibilitando a construção de planos de discriminante ótimo [4].

Das características do processo de RP apresentadas e o uso de DVS para a classificação, advém a seguinte pergunta: Será que os padrões podem ser codificados com códigos de bloco de tal maneira que no espaço de características formem conjuntos de sinais bem definidos? Sendo mais específico quanto à questão levantada, considere o seguinte problema: Suponha que temos um conjunto

de M padrões, onde um dos M , num determinado tempo é enviado através de um canal com ruído aditivo gaussiano. A seguinte questão vem naturalmente. Sob o ponto de vista geométrico, Qual é o melhor conjunto de sinais correspondentes a um dado conjunto de padrões? Onde por "melhor" queremos dizer que os pontos de sinais sejam arranjados tal que a probabilidade de erro seja mínima.

Neste trabalho nos propomos utilizar o conceito de Análise de Correspondência (AC), até então, não utilizado na área de Teoria de Comunicações, como um procedimento para encontrar a melhor correspondência 1-a-1 entre o conjunto de padrões e o conjunto de pontos de sinais no espaço Euclídeano, isto é, definindo a melhor correspondência 1-a-1 no processo de codificação ótima.

A base algébrica da AC é a Decomposição de Valores Singulares Generalizados (DVSG), onde a aproximação de uma matriz de posto menor é feita através da distância χ^2 .

Utilizando a definição de perfil suplementar da AC e o conceito de classe lateral em Teoria de Grupos, é possível codificar classes de padrões e desta maneira, obter um classificador ótimo de padrões.

Na segunda parte desta pesquisa iremos nos concentrar nos temas casamento e rotulação entre conjunto de sinais e grupos.

O conceito de Modulação-Codificada na sua forma generalizada, envolve o conceito de casamento entre conjunto de sinais e grupos, o qual está intimamente ligado com a restrição de largura de faixa. Na literatura técnica encontramos esparsos exemplos de casamento entre conjuntos de sinais e grupos. Um desses exemplos é descrito em [24].

Na Seção 1.2, apresentamos o encaminhamento utilizado na Rotulação de Conjuntos de Sinais Simplex.

1.2. COMO SERÁ ROTULADO UM CONJUNTO DE SINAIS SIMPLEX POR ELEMENTOS DE UM GRUPO?

A resposta a esta pergunta vem com: 1) a construção de um isomorfismo entre dois grupos: um deles associado ao espaço de Hamming e o outro associado ao espaço Euclideano (Ponderado); 2) Como decorrência deste isomorfismo, temos o casamento e a rotulação do conjunto de sinais; 3) Uma vez rotulados os sinais, o seguinte passo é obter o homeomorfismo entre os subespaços métricos dos espaços de Hamming, Euclideano e Euclideano Ponderado.

Para atingir estes objetivos, no Capítulo 2 apresentamos uma técnica estatística que é utilizada como uma análise exploratória de dados de uma tabela de contingência a qual denomina-se Análise de Correspondência (AC). A abordagem da AC está enfatizada em representações geométricas, apresentando a estrutura dos dados da tabela em forma ótima sem a necessidade de assumir modelos nem distribuições fundamentais. Estas características permitem que este ferramenta possa ser utilizado para a codificação de Padrões.

No Capítulo, 3 introduzimos o Reconhecimento de Padrões Matemático através da Teoria de Decisão de Bayes, dando ênfase na classificação de Padrões. Esta teoria será utilizada no Capítulo 4 para a obtenção das regiões de decisão e probabilidade de erro do conjunto de sinais.

O Capítulo 4 trata sobre o Casamento entre Conjunto de Sinais e Grupos. Para desenvolver este tópico apresentamos os conceitos de Teoria de Grupos dando ênfase em grupos de simetria. E da área de Comunicações, Codificação, Modulação, Transmissão de Sinais através do canal Gaussiano e o conjunto de sinais com suas respectivas probabilidades de erro.

No Capítulo 5, apresentamos a Codificação Ótima de Padrões Via AC.

Mas, antes desenvolvemos a Teoria de Corpos de Galois e Geração de Códigos em $\mathbb{Z}_p$. Uma vez gerados os códigos mostramos como, a partir das classes laterais do código, é possível gerar conjuntos de sinais através da AC. E, particularmente, pela utilização das classes laterais de códigos Reed-Muller de primeira ordem a geração de conjuntos sinais simplex.

Finalmente, no Capítulo 6 mostramos o homeomorfismo entre os subespaços métricos de Hamming, Euclideano e Euclideano Ponderado para o conjunto de sinais simplex.

CAPÍTULO 2

ANÁLISE DE CORRESPONDÊNCIA

2.1. INTRODUÇÃO

A Análise de Correspondência (AC) é uma técnica gráfica multidimensional que tem por objetivo mostrar, em forma gráfica, as relações entre as linhas e colunas de uma Tabela de Contingência (TC) as quais são representadas simultaneamente como pontos nos subespaços. Estes pontos representam as projeções das linhas ou colunas da Tabela de Contingência.

O objetivo central da AC é encontrar o conjunto de coordenadas nos respectivos subespaços, tais que neles esteja maximizada a inércia, sem a necessidade de assumir modelos nem distribuições fundamentais.

A AC é considerada como uma técnica de análise exploratória de dados multidimensionais, mais destinada a gerar hipóteses do que testá-las. A base desta forma de estudo está baseada na decomposição de valores singulares generalizada (DVSG), aproximação de uma matriz de posto menor e distância Qui-quadrada, $d\chi^2$.

História da Análise de Correspondência

O método tem uma longa e variada história o qual apareceu em diferentes formas na literatura ecológica e psicométrica, entre outros. A primeira descrição matemática foi mostrada por Hierchfeld [5] consistindo de uma formulação algébrica de "correlação", entre as linhas e colunas de uma TC, embora Richarson e Kuder [6] e Horst [7] independentemente tenham utilizado idéias similares não matemáticas na literatura psicométrica. Desde então, os mesmos procedimentos algébricos e numéricos têm sido redescobertos em diferentes contextos notavelmente em ecologia (média recíproca) e psicologia (escalonamento dual).

O termo "Análise de Correspondência" e a forma geométrica tiveram sua origem na França. O primeiro a apresentar este método com este título foi Jean-Paul Benzécri [8].

Desde então, um grande número de trabalhos tem sido realizados em AC, por exemplo [9] e [10].

Neste capítulo desenvolvemos a base teórica da Análise de Correspondência Simples (ACS). A ACS distingue-se da Análise de Correspondência Múltipla (ACM) na forma da construção das matrizes [14].

2.2. ELEMENTOS DA ANÁLISE DE CORRESPONDÊNCIA

2.2.1. Matriz de Correspondência, Perfis e Centróide

Antes de apresentarmos a matriz de correspondência, se faz necessário apresentarmos algumas definições e conceitos sobre tabelas de contingência que serão utilizados posteriormente (para maiores informações ver [11]).

Uma *Tabela de Contingência* é essencialmente uma amostra de uma população multivariada com as probabilidades e partições de categorias, sujeitas às restrições da distribuição multinomial.

Geralmente, os dados são representados sob a forma de uma Tabela de Contingência de dupla entrada.

Suponha que temos uma amostra de observações independentes e que cada uma dessas observações seja classificada segundo duas variáveis categóricas: nas linhas temos I categorias de uma variável categórica, e nas colunas J cate-

gorias de outra variável categórica. Dessa forma, a tabela de contingência conterá $I \times J$ células.

A tabela de contingência pode então ser representada por uma matriz X de ordem $I \times J$, onde: x_{ij} denota a frequência de ocorrência na i -ésima linha e na j -ésima coluna; $x_{i.}$ denota a soma dos elementos da i -ésima linha; $x_{.j}$ denota a soma dos elementos da j -ésima coluna e $x_{..}$ denota o número total de elementos de X .

Uma *Matriz de Correspondência* P , é definida como a matriz de elementos de X , divididos pelo número total de elementos de X , isto é;

$$P = X/x_{..} \quad (2.2.1)$$

Para se encontrar as associações entre linhas e colunas de uma TC, utiliza-se vetores de frequências relativas, chamados perfis. Além disso associa-se uma "massa" a cada um destes perfis.

Os perfis das linhas e colunas de uma matriz X são definidos como os vetores das linhas e colunas de X , divididos pela soma de suas coordenadas, respectivamente.

A soma dos elementos das linhas de P , são elementos de um novo vetor que é chamado *vetor de massa* dos perfis das linhas e é definido como:

$$\underline{r} = P \underline{1}_J \quad (2.2.2)$$

onde:

$\underline{1}_J$ é o vetor de J elementos composto por uns.

Denotaremos por R a matriz que contém os perfis das linhas. Então R é definido como

$$\mathbf{R} = \mathbf{D}_r^{-1} \mathbf{P} \quad (2.2.3)$$

onde:

$\mathbf{D}_r^{-1}$ é a matriz inversa diagonal positiva cujos elementos são as massas dos perfis das linhas de $\mathbf{P}$.

Similarmente, a soma dos elementos de cada uma das colunas de $\mathbf{P}$ compõe os elementos do vetor de massa dos perfis das colunas de $\mathbf{P}$ definido como:

$$\underline{c} = \mathbf{P}^T \mathbf{1}_I \quad (2.2.4)$$

onde:

$\mathbf{1}_I$ é o vetor de I elementos composto por uns.

Denotaremos por $\mathbf{C}$ a matriz que contém os perfis das colunas de $\mathbf{P}$. Então $\mathbf{C}$ é definida como:

$$\mathbf{C} = \mathbf{D}_c^{-1} \mathbf{P}^T \quad (2.2.5)$$

onde:

$\mathbf{D}_c$ é matriz diagonal positiva cujos elementos são as massas dos perfis das colunas.

Observe que os perfis das linhas e perfis das colunas definem duas nuvens de pontos no espaço Euclideano ponderado $\mathbb{R}_P^J$ e $\mathbb{R}_P^I$, respectivamente.

O centróide, $\underline{c}$, dos perfis das linhas, $\underline{r}_1, \dots, \underline{r}_I$ é definido por:

$$\underline{c} = \sum_{i=1}^I \underline{r}_i \cdot \underline{r}_i \quad \text{onde} \quad \sum_{i=1}^I \underline{r}_i = 1 \quad (2.2.6)$$

Similarmente, o centróide, $\underline{r}$, dos perfis das colunas, $\underline{c}_1, \dots, \underline{c}_J$ é definido por:

$$\underline{r} = \sum_{j=1}^J \underline{c}_j \cdot \underline{c}_j, \quad \text{onde} \quad \sum_{j=1}^J \underline{c}_j = 1 \quad (2.2.7)$$

Estes dois centróides estão relacionados, devido ao fato de que o centróide dos perfis das colunas é igual ao vetor de massa dos perfis das linhas, e vice-versa.

O fato de trabalhar com distribuições condicionais ou vetores de frequência relativa, nos conduz a utilizar uma distância distinta da Euclideana. A AC utiliza uma distância apropriada para distribuições condicionais a distância Chi-quadrada à qual denotaremos por d_{χ^2} . Esta difere da distância Euclideana unicamente na ponderação.

2.2.2. Distância χ^2 , Inércia e Equivalência Distribucional

Sejam $\underline{r}_i$ e $\underline{r}_{i'}$, dois perfis de linha, a distância χ^2 entre estes dois perfis no espaço $\mathbb{R}_p^J$ ponderado pela matriz $\mathbf{D}_c^{-1}$ é definida por:

$$d_{\chi^2}^2(\underline{r}_i, \underline{r}_{i'}) = (\underline{r}_i - \underline{r}_{i'})^T \mathbf{D}_c^{-1} (\underline{r}_i - \underline{r}_{i'}) \quad (2.2.8)$$

Similarmente, define-se a distância χ^2 entre os perfis das colunas no espaço $\mathbb{R}_p^I$ ponderado pela matriz $\mathbf{D}_r^{-1}$.

Agora considerando (2.2.8), a distância χ^2 , pode ser transformada para uma distância Euclideana por

$$d_{\chi^2}^2(\underline{r}_i, \underline{r}_{i'}) = (\underline{r}_i - \underline{r}_{i'})^T \underline{D}_c^{-1}(\underline{r}_i - \underline{r}_{i'}) = (\tilde{\underline{r}}_i - \tilde{\underline{r}}_{i'})^T (\tilde{\underline{r}}_i - \tilde{\underline{r}}_{i'}) \quad (2.2.9)$$

onde:

$$\tilde{\underline{r}} = \underline{D}_c^{1/2} \underline{r}_i$$

Em conseqüência, realizar uma análise considerando a distância χ^2 é o mesmo que realizar uma análise com a distância Euclideana aplicando com os novos pontos das linhas da matriz $\underline{R} \underline{D}_c^{-1/2}$ mas sem deixar de considerar as massas correspondentes.

Utilizando a distância χ^2 , estaremos dando um efeito estabilizador sobre os dados, uma vez que satisfaz a propriedade denominada Equivalência Distribucional, que veremos mais adiante.

A inércia dos perfis das linhas é definido por

$$\begin{aligned} \ln(I) &= \sum_{i=1}^I \underline{r}_i (\underline{r}_i - \underline{c})^T \underline{D}_c^{-1}(\underline{r}_i - \underline{c}) \\ &= \text{Traço}(\underline{D}_r (\underline{R} - \underline{1} \underline{C}^T) \underline{D}_c^{-1} (\underline{R} - \underline{1} \underline{C}^T)) \end{aligned} \quad (2.2.10)$$

A inércia é a quantização da variação espacial total de cada nuvem de pontos. Claramente esta é uma medida de variação total ou variância total dos perfis com respeito ao centróide.

Da mesma maneira, define-se o momento de inércia dos perfis das colunas $\ln(J)$ por:

$$\begin{aligned} \ln(J) &= \sum_{j=1}^I \underline{c}_j (\underline{c}_j - \underline{r})^T \underline{D}_r^{-1}(\underline{c}_j - \underline{r}) \\ &= \text{Traço}(\underline{D}_c (\underline{C} - \underline{1} \underline{r}^T) \underline{D}_r^{-1} (\underline{C} - \underline{1} \underline{r}^T)) \end{aligned} \quad (2.2.11)$$

Note que as inércias totais de ambos perfis são iguais. Além disso são iguais à estatística χ^2 com $(r-1).(c-1)$ graus de liberdade. Note que esta estatística tem a característica de testar a independência das linhas e colunas de uma TC. Portanto, se normalizarmos a estatística χ^2 por $x_{..}$ obteremos as inércias totais, isto é,

$$\ln(I) = \ln(J) = \chi^2/x_{..} \quad (2.2.12)$$

As razões pelas quais se utiliza a distância χ^2 , para medir a distância entre os perfis são as seguintes:

- porque utilizando a distância χ^2 , as inércias totais dos dois conjuntos de perfis com respeito aos seus centróides são iguais;
- Pelo efeito estabilizador sobre os dados, visto que satisfaz a propriedade denominada "Princípio de Equivalência Distribucional" [8]. O princípio diz: "Se dois perfis de linhas (colunas) são idênticos, então as duas linhas (colunas) correspondentes da matriz original de dados podem ser substituídas pela soma delas sem afetar a geometria dos perfis das colunas (linhas)".

A seguir mostraremos como são gerados os correspondentes subespaços, para isto consideraremos a distância χ^2 e a massa de cada um dos perfis, respectivamente.

2.3. GERAÇÃO DE SUBESPAÇOS ÓTIMOS

Definimos o termo subespaço em um sentido diferente ao da matemática usual.

Definição 2.3.1. : O conjunto de vetores $b + a_1 b_1 + \dots + a_k b_k$ é chamado de *Subespaço K-dimensional* de um espaço J-dimensional, onde b é qualquer vetor J-dimensional fixo, e $b_1, \dots, b_k$, são vetores J-dimensionais linearmente independentes e $a_1, a_2, \dots, a_k$, são números reais.

Em 1901, Karl Pearson [12] publicou um trabalho sobre o ajuste de um sistema de pontos em um multiespaço a uma linha ou um plano utilizando para tal a geometria Euclideana. Este enfoque foi retomado por Hotelling [13], que foi o primeiro a formular a análise por componentes principais que perdura até nossos dias.

O trabalho original de Pearson (1901) centrava naquelas componentes e combinações lineares de variáveis originais, para as quais a variância não explicada fosse mínima. Estas combinações geram um plano, que está em função das variáveis originais na qual o ajuste do sistema de pontos é "o melhor", por ser mínima a soma das distâncias de cada ponto ao plano de ajuste.

A seguir com este mesmo critério, mas utilizando a distância χ^2 , apresentaremos um método de como ajustar os pontos do espaço J-dimensional ao subespaço K-dimensional.

Seja $\mathcal{S}$ o subespaço de dimensão K, denotemos por y_i aos pontos que pertencem ao espaço Euclideano J-dimensional ponderado por D_q ; denotaremos por $\hat{y}_i$ um ponto do subespaço "mais próximo" a y_i ; e por d_i a distância mínima entre estes pontos.

Se y_i é ponderado pela massa w_i ($i=1, \dots, I$), então nossa definição de proximidade de todo o conjunto de pontos para o subespaço $\mathcal{S}$ é,

$$\psi(\mathcal{S}; y_1, y_2, \dots, y_I) = \sum_{i=1}^I w_i d_i^2 \quad (2.3.1)$$

onde:

$$d_i^2 = \|y_i - \hat{y}_i\|_{D_q}^2 = (y_i - \hat{y}_i)^T D_q (y_i - \hat{y}_i)$$

D_q é matriz diagonal de ordem $J \times J$.

Claramente ψ depende do subespaço $\mathcal{S}$, então temos que encontrar o subespaço $\mathcal{S}^*$ tal que minimize a função ψ em (2.3.1).

Pela Definição 2.3.1. podemos considerar um ponto Δ , como um subespaço de dimensão zero, logo a função (2.3.1) torna-se

$$\psi(\Delta; y_1, \dots, y_I) = \sum_{i=1}^I w_i (y_i - \Delta)^T D_q (y_i - \Delta) \quad (2.3.2)$$

desde que $\hat{y}_i$ seja igual a Δ para qualquer i .

Derivando (2.3.2) com respeito a Δ , temos

$$\partial\psi/\partial\Delta = \sum_{i=1}^I w_i 2 D_q (y_i - \Delta) = 0$$

$$\Delta = \sum_{i=1}^I w_i y_i = \bar{y}$$

O centróide $\bar{y}$ é o ponto que minimiza esta função, assim neste sentido, o centroide $\bar{y}$ é o ponto mais próximo a todos os pontos $y_1, \dots, y_I$.

Greenacre [14], mostra que qualquer subespaço $\mathcal{S}$ ótimo no sentido de minimizar (2.3.1) necessariamente contém o centróide. Com este resultado restringiremos as aproximações $\hat{y}_i$, dos pontos y_i à forma:

$$\hat{y}_i = \bar{y} + \sum_{k=1}^K f_{ik} b_k \quad i=1,2,\dots,I$$

onde $b_1, b_2, \dots, b_K$ são vetores bases do subespaço. Logo a função (2.3.1) pode ser escrita como:

$$\psi(\mathcal{Y}; y_1, \dots, y_I) = \sum_{i=1}^I w_i (y_i - \bar{y} - \sum_k f_{ik} b_k)^T D_q (y_i - \bar{y} - \sum_k f_{ik} b_k) \quad (2.3.3)$$

As variáveis desta função objetivo são os K eixos, $b_1, \dots, b_K$ implicando um total de I vezes K variáveis escalares.

Pela escolha particular da distância χ^2 , a solução da função ψ é consideravelmente simplificada.

A solução completa teórica para a minimização de (2.3.3), que especifica a dimensão de K , está envolta no conceito de DVSG e na aproximação de uma matriz de posto menor que serão apresentados nos teoremas que se seguem, respectivamente.

Teorema 2.3.1. [14] : Sejam duas matrizes simétricas positivas θ e Λ de ordem $I \times I$ e $J \times J$ respectivamente, então qualquer matriz A de ordem $I \times J$ com posto $K \leq \min(I, J)$ pode ser expressa como:

$$A = N D_\mu M^T = \sum_{k=1}^K \mu_k n_k m_k^T \quad (2.3.4)$$

onde D_μ é uma matriz diagonal de ordem $K \times K$ com elementos positivos μ_k e N e M são matrizes ortonormais relativamente a θ e Λ de ordens $I \times K$ e $J \times K$ respectivamente, i.e.

$$N^T \Theta N = M^T \Lambda M = I \quad (2.3.5)$$

e n_k, m_k são as colunas das matrizes N e M respectivamente.

A decomposição (2.3.4) é chamada de decomposição de valores singulares generalizada, os números $\mu_1, \mu_2, \dots, \mu_k$ são os valores singulares generalizados da matriz A e as matrizes N e M são formadas pelos vetores singulares generalizados $n_1, \dots, n_k$ e $m_1, \dots, m_k$ respectivamente da matriz A . Eles são bases ortonormais para as colunas e linhas de A no espaço Euclidiano I, J -dimensional ponderado nas matrizes Θ e Λ , respectivamente. Além disso, as coordenadas das linhas e colunas de A com respeito aos vetores base M e N são $N D_\mu$ e $M D_\mu$, respectivamente.

Esta decomposição em valores singulares generalizada tem aplicação na aproximação de uma matriz de posto menor [15].

Uma matriz $A_k^* : I \times J$ de posto k^* é a melhor aproximação de uma matriz $A : I \times J$ de posto k , $k \geq k^*$, se

$$\|A - A_k^*\|^2 = \inf_{\text{posto}(X)=k^*} \left\{ \|A - X\|_{\Theta\Lambda} \right\}$$

onde:

$$\|A - X\|_{\Theta\Lambda}^2 = \sum_{i=1}^I \sum_{j=1}^J \theta_i \lambda_j (a_{ij} - x_{ij})^2 = \sum_{i=1}^I \theta_i (a_i - x_i)^T \Lambda (a_i - x_i) \quad (2.3.6)$$

O teorema que enunciaremos a seguir é importante para o problema de aproximação de nosso interesse.

Teorema 2.3.2. [14] : Seja $A : I \times J$ de posto K , e seja

$$A = N D_{\mu} M^T = \sum_{k=1}^K \mu_k n_k m_k^T \quad (2.3.7)$$

sua decomposição em valores singulares generalizados. Então para $K^* \leq K$, a melhor aproximação de posto K^* da matriz A é dado por:

$$A_{K^*} = \sum_{k=1}^{K^*} \mu_k n_k m_k^T$$

A aproximação é única se $\mu_k > \mu_{k+1}$.

A melhor aproximação dada pelo Teorema 2.3.2. é no sentido de mínimos quadrados.

Pelo Teorema 2.3.2., vemos que a DVSG fornece a solução para qualquer subespaço de dimensão K^* . Por exemplo: para $K^*=1$ utilizamos o primeiro vetor singular associado ao seu valor singular e , assim sucessivamente, podemos adicionar os demais vetores singulares generalizados (eixos principais) associados a seus respectivos valores singulares generalizados.

O quadrado dos valores singulares generalizados dão uma idéia de como uma matriz é representada sobre os eixos principais (vetores singulares generalizados).

A variação total de uma matriz é quantificada pela sua norma quadrada (2.3.6)

$$\|A\|_{\theta\Lambda}^2 = \sum_{i=1}^I \theta_i a_i^T \Lambda a_i = \text{traço}(\theta A^T \Lambda A)$$

$$\|A\|_{\theta\Lambda}^2 = \text{traço}(\theta M D_{\mu} N^T \phi N D_{\mu} M^T) = \text{traço}(\theta M D_{\mu} D_{\mu} M^T)$$

$$= \text{traço}(D_{\mu}^2 M^T \theta M) = \sum_{k=1}^K \mu_k^2$$

Similarmente a variação da aproximação de A_K^* é :

$$\|A_K^*\|_{\theta\Lambda}^2 = \sum_{k=1}^{K^*} \mu_k^2$$

e a variação inexplicada é:

$$\|A - A_K^*\|_{\theta\Lambda}^2 = \sum_{k=K^*+1}^K \mu_k^2.$$

A medida tradicional da qualidade de aproximação de A por A_K^* é a percentagem da soma de quadrados, definida como

$$\tau = 100 \sum_{k=1}^{K^*} \mu_k^2 / \sum_{k=1}^K \mu_k^2$$

Baseados nestes conceitos desenvolveremos a seguir, somente a correspondência entre linhas de uma TC por ser de nosso interesse.

2.4. COORDENADAS PRINCIPAIS DOS PERFIS DAS LINHAS, COM RESPEITO AO CENTRÓIDE

Seja $(R - 1\underline{c}^T)$ a matriz dos perfis das linhas centradas no seu centróide de ordem $I \times J$. Aplicando a DVSG a esta matriz temos:

$$(R - 1\underline{c}^T) = W D_{\mu} Y^T \quad (2.4.1)$$

onde W e Y são matrizes ortonormais relativamente a D_r e D_c^{-1} , de ordem $I \times K$ e

$K \times J$, respectivamente, i.e.

$$W^T D_r W = Y^T D_c^{-1} Y = I \quad (2.4.2)$$

e pela DVSG, a matriz que contém as coordenadas dos perfis de R com respeito aos vetores base Y é $W D_\alpha$ ao qual denotaremos por F

Esta matriz contém as coordenadas dos perfis com respeito ao centróide em todos os subespaços, à qual denotaremos por F . Logo:

Considerando (2.4.1) e (2.4.2)

$$F = (R - \underline{c}^T) D_c^{-1} Y \quad (2.4.3)$$

O centróide das coordenadas dos perfis das linhas, F , coincide com a origem dos subespaços, isto é,

$$\underline{r}^T F = \underline{r}^T (R - \underline{1} \underline{c}^T) D_c^{-1} Y = (\underline{r}^T R - \underline{c}^T) D_c^{-1} Y = (\underline{c}^T - \underline{c}^T) D_c^{-1} Y = 0^T$$

Para observar graficamente as relações entre as linhas da matriz X , alocamos as coordenadas F nos subespaços de interesse. Mas para interpretar estes gráficos precisamos de uma informação adicional, isto porque, estes pontos mostram a projeção dos perfis sobre os subespaços, e, logicamente, este gráfico não mostra quais dos pontos estão perto ou longe do subespaço que estamos considerando. Por isso é que no próximo item desenvolveremos esta informação adicional para saber onde a exposição é precisa e onde não é. Isto é análogo para muitas outras áreas de Estatística. Por exemplo, na construção de um modelo para dados, estudamos a qualidade do ajuste do modelo nos dados classificando-o como satisfatório ou não.

2.5. DECOMPOSIÇÃO DA INÉRCIA E INTERPRETAÇÃO DOS PERFIS NOS SUB-ESPAÇOS

A inércia total de uma matriz de dados pode ser decomposta sobre os eixos principais, a qual é análoga à decomposição da soma de quadrados na análise de variância.

Na Tabela 2.5.1 é apresentada a decomposição da inércia para cada conjunto de pontos que representam as linhas. Na última linha da tabela onde está o quadrado dos valores singulares generalizados, temos as inércias decompostas sobre cada eixo, na última coluna da tabela, temos a soma da inércia total de cada ponto com respeito aos eixos, e na intersecção da última linha com a última coluna, temos a inércia total da matriz de dados.

TABELA 2.5.1

Decomposição de Inércia das Linhas de $(R - 1\bar{c}^T)$

	EIXOS						TOTAL
	1	2	.	.	.	k	
1	$r_1 f_{11}^2$	$r_1 f_{12}^2$	.	.	.	$r_1 f_{1k}^2$	$r_1 \sum f_{1k}^2$
2	$r_2 f_{21}^2$	$r_2 f_{22}^2$	.	.	.	$r_2 f_{2k}^2$	$r_2 \sum f_{2k}^2$
.	.	.	.	.	.	.	.
l	$r_l f_{l1}^2$	$r_l f_{l2}^2$	.	.	.	$r_l f_{lk}^2$	$r_l \sum f_{lk}^2$
TOTAL	μ_1^2	μ_2^2	.	.	.	μ_k^2	$\ln(I)$

Observação: $r_i f_{ik}^2$ é a Contribuição do i-ésimo perfil de linha para a inércia do k-ésimo eixo (f_{ik} elementos de F).

Para interpretar as relações entre as linhas e colunas de uma TC sobre os subespaços gerados pelos eixos obtidos através da análise de correspondência, é necessário analisar dois conjuntos de coeficientes para cada um dos perfis em correspondência:

- 1) contribuição absoluta, que exprime o aporte de um perfil para a inércia "explicada" em um subespaço;
- 2) contribuição relativa ou correlação entre um eixo e o perfil, que exprime o aporte de um subespaço à explicação da dispersão do perfil.

2.5.1. Contribuição Absoluta

Cada perfil contribui para a inércia decomposta em cada eixo, o qual depende de sua massa e da distância ao quadrado de sua projeção sobre o subespaço e o centróide.

Assim, o quociente

$$CTR(i) = \frac{r_{ik}^2}{\mu_k^2}$$

mede a proporção de contribuição do i-ésimo perfil de linha para a inércia explicada no k-ésimo eixo. Este quociente é chamado de *Contribuição Absoluta* do i-ésimo perfil linha para o k-ésimo eixo.

Observe que :

$$\sum_{i=1}^I CTR_K(i) = 1 \text{ para todos os eixos.}$$

2.5.2. Contribuição Relativa

A qualidade de um perfil à ser representado pelos eixos é obtida pelo $\cos^2 \theta_k$, onde θ_k é o ângulo entre o perfil e a projeção do perfil sobre o k-ésimo eixo que mede a proximidade entre eles. Assim, por exemplo, se a matriz $(R - 1\bar{c}^T)$ é de posto $k=2$, teremos dois subespaços. Observe na Figura 2.1, a projeção do perfil $\underline{r}_i$ sobre os dois eixos é dado por $r_{i1} \cdot \cos \theta_1$ e $r_{i2} \cdot \cos \theta_2$, onde θ_1 e θ_2 são ângulos entre o perfil e os eixos 1 e 2, respectivamente. Logo a distância do perfil $\underline{r}_i$ ao centróide é decomposta sobre seus eixos como $(r_{i1} \cdot \cos \theta_1, r_{i2} \cdot \cos \theta_2)$, e pelo Teorema de Pitágoras temos que:

$$\cos^2 \theta_1 + \cos^2 \theta_2 = 1$$

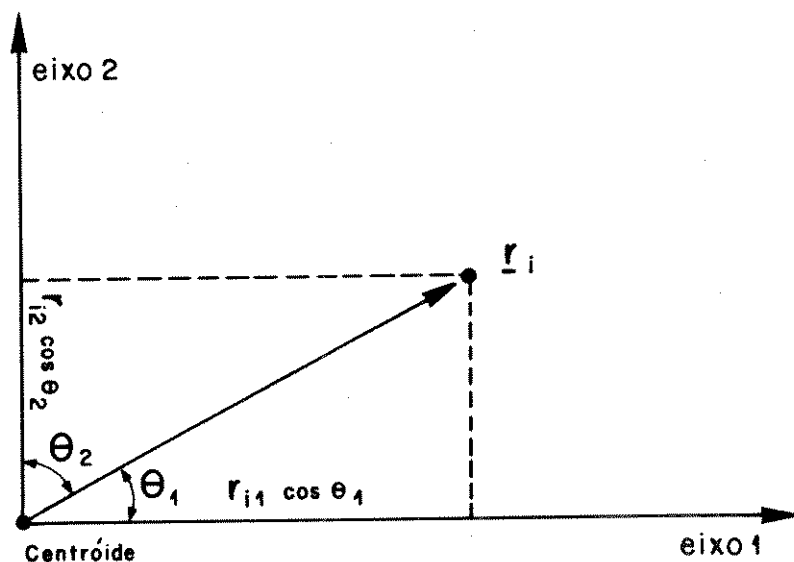


Figura 2.1 - Posição do Perfil $\underline{r}_i$ sobre um Espaço Bi-Dimensional.

Agora, se a matriz $(R - 1\bar{c}^T)$ é de posto $K = K^*$ a posição do perfil no espaço K^* -dimensional é dado por:

$$\sum_{k=1}^{K^*} \cos^2 \theta_k = 1 \quad (2.5.1)$$

onde:

$$\cos \theta_k = f_{ik} / d_i$$

Cada elemento de (2.5.1) nos mostrará a qualidade da representação do perfil $\underline{r}_i$ pelo k-ésimo eixo. Este valor é denominado *Contribuição Relativa* do k-ésimo eixo para o i-ésimo perfil de linha. Observe que a contribuição relativa não depende das massas.

Agora, se quisermos saber a qualidade da representação do perfil $\underline{r}_i$ pelo plano gerado pelo ℓ -ésimo eixo e m-ésimo eixo, a qualidade será obtida pela soma de $\cos^2 \theta_\ell + \cos^2 \theta_m$.

Para finalizar este capítulo, descreveremos o que vem a ser um perfil suplementar e a sua importância para o contexto de codificação de padrões.

2.6. PERFIS SUPLEMENTARES

Os perfis suplementares são de grande interesse na análise de dados.

Um perfil suplementar é aquele que não participa ativamente nas análises, isto é, que não é utilizado para o cálculo do centróide, inércia total e base dos subespaços, porque a massa associada a ele é zero. Mas as coordenadas deste perfil são calculadas em função das bases e do centróide dos demais perfis ativos representados nos subespaços ótimos. Por exemplo se $\underline{r}_s$ é um perfil suplementar linha, as coordenadas deste perfil nos subespaços são determinadas utilizando-se a equação (2.4.3), isto é,

$$\mathbf{f}_s = (\mathbf{r}_s - \mathbf{c})^T \mathbf{D}_c^{-1} \mathbf{Y}$$

Os perfis suplementares são utilizados na análise de uma TC quando:

- Existem observações aberrantes que são observadas através de uma análise preliminar;
- Temos um caso novo;
- Quando os elementos são de natureza diferente (variáveis quantitativas, qualitativas).

As variáveis cujas categorias são consideradas como perfis suplementares são denominadas *Variáveis Não Ativas*, caso contrário *Variáveis Ativas*.

CAPÍTULO 3

RECONHECIMENTO DE PADRÕES MATEMÁTICO E CLASSIFICAÇÃO

3.1. INTRODUÇÃO

É frequentemente declarado que o processo de reconhecimento e classificação são atividades humanas fundamentais.

Com o advento dos computadores um substancial impulso se fez sentir em todos os ramos de atividades de pesquisa. Em particular na área do reconhecimento de Padrões. Basicamente, em reconhecimento de padrões estudam-se os mecanismos artificiais bem como naturais com o objetivo de analisar, detectar, reconhecer e descrever padrões de dados sensoriais e/ou numéricos.

O estudo de reconhecimento de padrões tem apresentado um crescimento em campos disciplinares incluindo desde as áreas de Engenharia, Ciência da Computação, como as de Biologia, Psicologia e Medicina.

Como exemplos de aplicações das técnicas de reconhecimento e classificação de padrões podemos citar: assistência computadorizada em diagnósticos médicos, processamento de imagens, previsão do tempo, controle de processos, processamento de sinais neurobiológicos, reconhecimento de voz, impressão digital e muito mais.

Neste capítulo, trataremos especificamente de reconhecimento de padrões matemático. Na Seção 3.2 apresentamos a conceituação matemática do reconhecimento de padrões. Na Seção 3.3 desenvolveremos dois métodos que são utilizados para a Extração de características bem como referenciamos outros métodos. Na Seção 3.4 mostramos como a análise de correspondência pode ser utilizada como Extrator de características, e finalmente na Seção 3.5 mostraremos formas de classificação dos padrões. Descrevemos brevemente a classificação não paramétrica e desenvolveremos a classificação Estatística segundo a Teoria de Bayes, visto que a mesma é necessária no desenvolvimento dos capítulos subsequentes.

3.2. RECONHECIMENTO DE PADRÕES MATEMÁTICO

O reconhecimento de padrão matemático é o estudo de técnicas matemáticas com o propósito de construir máquinas de reconhecimento que ajudem na sistematização e/ou automação das tarefas corriqueiras.

Uma das principais características frequentemente utilizadas em reconhecimento de padrões relaciona-se com a classificação ou descrição de um conjunto de processos ou eventos, onde os processos ou eventos são agrupados em classes segundo algumas propriedades.

Para um melhor entendimento do processo de reconhecimento de padrões matemático, é conveniente que se visualize como a composição do Espaço de Padrões, do Espaço de Características e do Espaço de Classificação fazem parte deste processo.

Na Figura 3.1 ilustramos este processo.

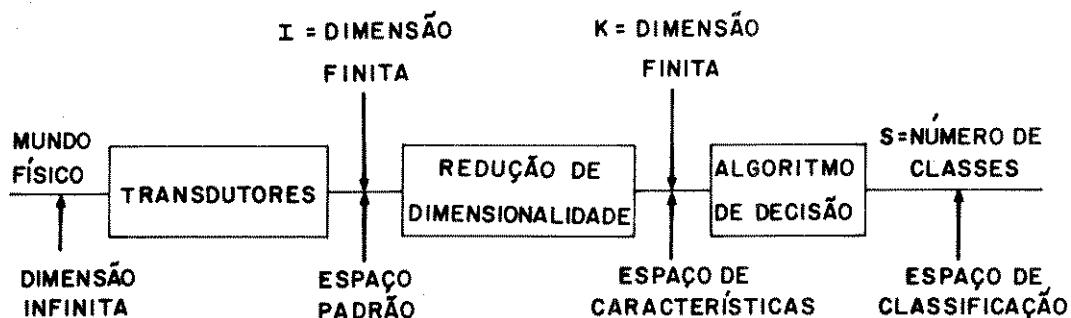


Figura 3.1 - Diagrama de Blocos da Conceptualização de Reconhecimento de Padrões.

Observe que segundo este modelo, o mundo físico é sentido por um sistema transdutor que descreve objetos em I valores escalares, onde I é a dimensão do espaço padrão que geralmente é muito grande. Esta operação de processamento transforma um padrão de sua forma original para uma forma conveniente para novo processamento.

3.2.1. Espaço Padrão

O espaço padrão é essencialmente definido pelos dados observados do mundo real por um transdutor ou dispositivo sensorial.

No espaço padrão, cada padrão é representado por um vetor.

Seja $\underline{x}$ um vetor que representa um padrão no espaço de padrões I-dimensional, onde x_i , $i=1, \dots, I$, representa o valor particular associado com a i-ésima dimensão, ou valor associado com a i-ésima característica do padrão que pode ser um valor amostrado, de tempo, cor ou outras medidas quantitativas.

Como os padrões, são representados por um vetor, cujos elementos são do tipo quantitativo, é possível definir métricas e assim o espaço de padrões torna-se um espaço métrico [16].

Seja $\Omega = \{\omega_1, \omega_2, \dots, \omega_s\}$ o conjunto de S classes de padrões. Aos vetores que definem estas classes de padrões segundo algumas propriedades chamaremos de protótipos, e denotaremos pelo vetor coluna $y_m^{(s)}$

$$y_m^{(s)} = \begin{bmatrix} y_{1m}^{(s)} \\ \vdots \\ y_{im}^{(s)} \end{bmatrix} \quad \begin{array}{l} 1 \leq s \leq S \\ 1 \leq m \leq M_s ; \\ 1 \leq i \leq I \end{array} \quad (3.1)$$

onde:

s indexa a s-ésima classe;

m indexa o m-ésimo protótipo da classe padrão ω_s ;

i indexa as componentes do padrão na i-ésima dimensão.

Dado que a dimensão no espaço padrão é elevada, é desejável reduzir a dimensionalidade enquanto se mantém a característica discriminatória para os

propósitos de classificação. Assim, é necessário um novo espaço que é chamado de espaço de características, no qual as regras de classificação podem ser melhor computadas e em tempo razoável.

3.2.2. Espaço de Características

O espaço de características no processo de reconhecimento de padrões é fundamental. Este espaço está entre o espaço de padrões e o espaço de classificação.

As razões, pelas quais é necessário um espaço de características podem ser assim descritas:

- Neste espaço pode-se definir a característica discriminatória dos dados presentes no espaço padrão, a qual é otimizada para problemas específicos, isto é, eliminar a redundância dos dados;
- Devido à redução da dimensão, os algoritmos de classificação podem ser eficientemente implementados.

Assim, cada padrão é representado por um vetor de características do padrão, cuja dimensão K , é muito menor que I . Desse modo, um vetor padrão x no espaço de características é representado por K elementos, $K \ll I$.

Uma vez definido o espaço de características, é necessário definir um espaço onde possamos classificar os padrões, isto é, para cada novo padrão de entrada, como classificá-lo corretamente.

3.2.3. Espaço de Classificação

O espaço de classificação é descrito no sentido de encontrar regiões disjuntas no espaço de características K -dimensional. A cada região está associada uma e uma só classe.

As regiões são obtidas através de algoritmos de classificação que estão baseados em critérios estatísticos, e/ou determinísticos, e/ou de Teoria da Informação, tal que proporcionem um grau de confiabilidade na classificação correta de padrões desconhecidos.

A seguir apresentaremos uma das partes principais no processo de reconhecimento de padrões qual seja, a seleção e extração de características dos padrões.

3.3. SELEÇÃO E EXTRAÇÃO DE CARACTERÍSTICAS

Um dos problemas fundamentais em reconhecimento de padrões matemático é determinar quais características são as mais importantes para uma melhor classificação de padrões. O propósito da seleção e extração de características é identificar características "úteis" e "boas" para delinear um classificador. Este processo de extração de características é uma forma de codificação e representação dos quadros originais.

Seja $\underline{X}$ uma matriz padrão de ordem $I \times J$, em cujas colunas estão os J padrões de dimensão I . A redução do espaço de padrões à dimensão K , é feita por duas estratégias. A primeira é a seleção de características e a segunda a extração de características, que geram uma matriz de ordem $K \times J$, $K \ll I$.

A seleção de características é a retenção ótima de um número mínimo de K características $K \ll I$, enquanto mantém e/ou maximiza a probabilidade de classificação correta. Este termo, seleção de características, é mais útil quando o objetivo é minimizar o custo médio de aquisição de medidas.

A extração de características é a aplicação de uma transformação à matriz padrão, obtendo-se dessa maneira um espaço transformado de onde seleciona-se as K melhores características. Assim, cada nova característica é uma combinação de todas as I características, dita de outra forma, a extração de características é considerada como um processo de mapeamento de características originais sobre características efetivas. Estas técnicas de mapeamento podem ser lineares ou não lineares.

Os índices de desempenho para medir a separabilidade de classes de subconjuntos de características podem ser divididos em duas principais categorias a saber:

- Probabilidade de erro: O erro pode ser estimado sob um modelo paramétrico ou não paramétrico;
- Medidas de distância: Um número de distâncias probabilísticas e medidas de informação entre as densidades de classe-condicional que estão "relacionados" com a probabilidade de erro. Para medir estas distâncias pode-se utilizar o conjunto de características efetivas.

A seguir desenvolveremos dois métodos, o de transformação diagonal e o de transformação rotacional para a extração de características uma vez que os mesmos, de uma certa maneira, indicarão as razões pelas quais utilizamos a análise de correspondência para extração de características.

3.3.1. Transformação Diagonal

Esta técnica tem como objetivo aumentar a similaridade ou agrupamento de protótipos na mesma classe, através da minimização de uma métrica entre os pontos que definem a classe. Sebestyen [17] define medidas dentro de uma classe, todas elas baseadas na métrica Euclideana.

Seja a s -ésima classe ω_s com M_s protótipos $y_m^{(s)}$, $m=1, \dots, M_s$, e cada protótipo com a mesma probabilidade. A distância quadrática média dos protótipos de classe ω_s é definida como:

$$D_s^2 = \frac{1}{M_s(M_s - 1)} \sum_{p=1}^{M_s} \sum_{m=1}^{M_s} d^2(y_p^{(s)}, y_m^{(s)}) \quad (3.2)$$

onde:

$$d^2(y_p^{(s)}, y_m^{(s)}) = (y_p^{(s)} - y_m^{(s)})^T (y_p^{(s)} - y_m^{(s)}) \quad (3.3)$$

e $(y_p^{(s)} - y_m^{(s)})$ é um vetor coluna.

Esta distância D_s^2 , será minimizada considerando uma variável peso $w = [w_1, w_2, \dots, w_I]$. Cada dimensão no espaço de padrões é ponderado por w_i , isto é, equivalente a multiplicar o espaço de padrões por uma matriz diagonal D_w cujos elementos são os w_i . Portanto, a distância quadrática média entre os protótipos da classe ω_s , torna-se

$$D_s^2 = \frac{1}{M_s(M_s - 1)} \sum_{p=1}^{M_s} \sum_{m=1}^{M_s} \sum_{i=1}^I w_i^2 (y_{ip}^{(s)} - y_{im}^{(s)})^2 \quad (3.4)$$

Para que D_s^2 permaneça como métrica, observe que os w_i devem ser positivos, isto é $w_i > 0$.

As restrições para a obtenção dos pesos de cada dimensão são:

$$\sum_{i=1}^I w_i = 1 \quad (3.5)$$

ou

$$\prod_{i=1}^I w_i = 1 \quad (3.6)$$

utilizando o método de multiplicadores de Lagrange para minimizar (3.4) temos:

- Sob a restrição (3.5).

$$w_i = \frac{\left(\sum_{s=1}^I \frac{1}{\sigma_s^2} \right)^{-1}}{\sigma_i^2} \quad (3.7)$$

- Sob a restrição (3.6)

$$w_i = \frac{\left(\prod_{s=1}^I \sigma_s \right)^{1/I}}{\sigma_i} \quad (3.8)$$

onde:

σ_i^2 é a variância para a i -ésima dimensão.

Os pesos são ordenados em forma decrescente e a matriz diagonal D_w é definida por estes pesos. O trabalho da seleção de características para a s -ésima classe consiste em uma permutação ordenada, operação de matriz sobre o

espaço padrão 1-dimensional, seguida da aplicação de uma matriz transformação diagonal, onde o espaço de características pode ser definido pelos K maiores pesos, com $K \ll I$.

Se as funções de probabilidade são conhecidas, os resultados das análises anteriores estão baseados em propriedades de segunda ordem.

Por outro lado, considerando o conceito de entropia desenvolve-se a obtenção dos pesos w_i para cada dimensão do espaço padrão, considerando para isto todas as classes com seus respectivos protótipos [18].

A outra transformação utilizada na extração de características é denominada de Transformação Rotacional.

3.3.2. Transformação Rotacional

A transformação rotacional seleciona as características através de uma transformação linear, descrita por uma matriz ortogonal. Isto é equivalente à rotação do espaço de padrões original para um novo conjunto de vetores coordenados os quais são ortogonais e proporcionam uma redução de dimensionalidade que facilita a identificação de características.

Geralmente, para selecionar características através da transformação ortogonal utiliza-se a matriz de co-variância.

Consideremos a classe ω_s , de M_s protótipos $y_m^{(s)}$. Denotaremos por $y^{(s)}$ a variável aleatória protótipo. A matriz de co-variância é uma matriz simétrica que denotaremos por Φ_s , esta matriz é de ordem $J \times J$, e é definida como

$$\Phi_s = E \left(\mathbf{z}^{(s)} \mathbf{z}^{(s)T} \right)$$

onde:

$$\mathbf{z}^{(s)} = \left(\mathbf{y}^{(s)} - E\left(\mathbf{y}^{(s)}\right) \right)$$

E é o operador linear, valor esperado.

Aplicando o Teorema 2.3.1. à matriz simétrica Φ_s e fazendo $\phi = \theta = I$ resulta que:

$$\Phi_s = \mathbf{M}^{(s)} \mathbf{D}^{(s)} \mathbf{M}^{(s)T} \quad (3.9)$$

onde:

$$\mathbf{M}^{(s)T} \mathbf{M}^{(s)} = \mathbf{I} \quad (3.10)$$

$\mathbf{D}^{(s)}$ é a matriz diagonal que contém os autovalores $\lambda_1, \dots, \lambda_I$

$\mathbf{M}^{(s)}$ é uma matriz ortogonal que contém os autovetores de Φ_s .

Esta transformação ortogonal é utilizada para obter o vetor de características, $\mathbf{n}_m^{(s)}$, do protótipo $\mathbf{y}_m^{(s)}$ da seguinte maneira:

$$\mathbf{y}_m^{(s)} \longrightarrow \mathbf{n}_m^{(s)} = \mathbf{M}^{(s)T} \mathbf{z}_m^{(s)} \quad (3.11)$$

Considerando (3.10) e (3.11), cada $\mathbf{z}_m^{(s)}$ pode ser expresso por

$$\mathbf{z}_m^{(s)} = \sum_{i=1}^I \mathbf{n}_{mi}^{(s)} \mathbf{m}_i^{(s)}$$

O erro de aproximação do $\mathbf{z}_m^{(s)}$ de dimensão I à dimensão K é dado por:

$$\begin{aligned}
\mathcal{E}_K^{(s)} &= E \left(\left| \mathbf{z}_m^{(s)} - \sum_{i=1}^K n_{mi} \mathbf{m}_i^{(s)} \right|^2 \right) \\
&= E \left(\left| \sum_{i=K+1}^I n_{mi} \mathbf{m}_i^{(s)} \right|^2 \right) \\
&= E \left(\sum_{i=K+1}^I \mathbf{m}_i^{(s)T} \mathbf{z}_m^{(s)} \mathbf{z}_m^{(s)T} \mathbf{m}_i^{(s)} \right) \\
&= \left(\sum_{i=K+1}^I \mathbf{m}_i^{(s)} \Phi_s \mathbf{m}_i^{(s)} \right) \\
&= \sum_{i=K+1}^I \mathcal{E}_i
\end{aligned} \tag{3.12}$$

$\mathcal{E}_K^{(s)}$, será minimizado se forem escolhidos os K autovalores correspondentes aos K maiores autovalores.

Este problema é essencialmente o de análise de componentes principais [19], conseqüentemente contido na classe de expansões de Karhunen-Loève.

A transformação (3.9) é uma rotação do espaço padrão I-dimensional ao espaço de características onde a quantidade máxima de energia está sobre K eixos.

Considere as S classes onde à cada classe será atribuída uma determinada probabilidade. Agora, seja Φ a matriz de co-variância total. Então, adotaremos Φ como sendo o valor médio das matrizes co-variâncias individuais, isto é,

$$\Phi = \sum_{s=1}^S p(\omega_s) \Phi_s \tag{3.13}$$

Para a obtenção das melhores características de cada classe e protótipo, utilizaremos novamente o conceito de Decomposição de Valores Singulares. Admitiremos também que os K autovetores correspondem aos K maiores autovalores.

Observe que as características não somente dependem das estatísticas de segunda ordem como também das funções de distribuição das classes.

Uma das inconveniências, advinda do uso das técnicas de transformações rotacionais apresentadas, é o fato que as rotações, por si só não são suficientes para separar as classes.

3.4. EXTRAÇÃO DE CARACTERÍSTICAS ATRAVÉS DA ANÁLISE DE CORRESPONDÊNCIA

Pelas propriedades que possui a Análise de Correspondência, podemos utilizá-la como um algoritmo para extrair as características dos padrões.

Seja $\underline{X}$ a matriz de padrões de ordem $I \times J$, em cujas colunas estão representados os padrões. Aplicando a AC a $\underline{X}$, esta transforma um espaço J-dimensional em um espaço K-dimensional onde $K \leq \min(I, J) - 1$. Visto desta maneira, um vetor J-dimensional é reduzido a um vetor K-dimensional.

Considerando os métodos acima expostos, a AC engloba os dois métodos e ainda fornece maiores informações como veremos a seguir.

- Com respeito à transformação diagonal, a AC utiliza a distância χ^2 , que é uma distância ponderada onde a j-ésima dimensão é ponderada pelo j-ésimo elemento do vetor centróide;

- Com respeito a transformação ortogonal, a AC tem vantagens porque além de ponderar cada uma das dimensões, também pondera cada padrão. Depois de aplicado o algoritmo de AC, escolhe-se os K vetores singulares generalizados associados aos K maiores valores singulares e desta maneira obtém-se os vetores de características K dimensionais.

Finalizaremos este capítulo mostrando a classificação de padrões matemático.

3.5. CLASSIFICAÇÃO DE PADRÕES

Basicamente a Teoria da Classificação consiste no particionamento do espaço de características em regiões, onde cada região pertence a uma classe padrão. Este particionamento é obtido por um classificador que tem como base teórica, os campos estatístico, probabilístico, Teoria da Informação e outras áreas de matemática aplicada.

O tópico de Teoria de Classificação de Padrões evoluiu muito mais com respeito à seleção e extração de características.

Existem muitas estratégias para delinear um classificador, porém todas elas dependem da informação que se possui.

Na Figura 3.2, mostramos em forma de estrutura de árvore, as várias dicotomias no processo de classificação de padrões. Utilizando esta figura descreveremos brevemente cada um dos "nós" e "ramificações". O nó incompleto, refere-se aos dados para a classificação que não são completamente conhecidas. Por exemplo, parâmetros, funções de densidade, etc..

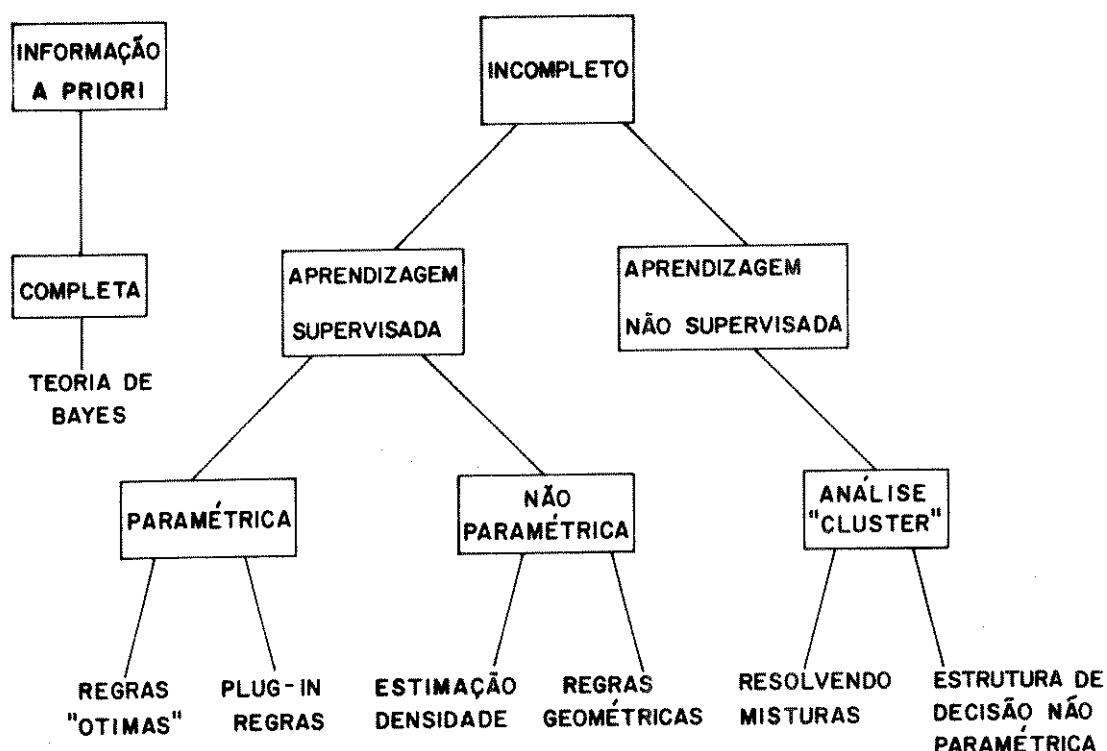


Figura 3.2 – Processo de Classificação de Padrões Quando os Dados são Completos e Incompletos.

Consideremos a ramificação que conduz ao nó, aprendizagem não supervisionada, temos que a classificação não depende de uma informação a priori e que a mesma leva a uma análise de "cluster". Esta análise consiste basicamente na confecção de algoritmos que agrupam os padrões em classes sem que se tenha um conhecimento a priori dos mesmos.

No aprendizado supervisionado tem-se um supervisor ensinando ao sistema como classificar um conjunto de padrões conhecido. Uma vez efetuado o aprendizado o sistema reconhecerá outros padrões. Nestes sistemas necessita-se de informação a priori para formar as bases do ensinamento. O processo de aprendizado supervisionado divide-se em duas classes de classificação, as do tipo paramétrico e as do tipo não paramétrico. Uma aproximação paramétrica é definida através de uma função discriminante (classe de densidade de probabili-

dade) e por um número pequeno de parâmetros. Muitas vezes estes parâmetros são desconhecidos, neste caso eles são substituídos pelos seus valores estimados. Quando a forma de densidade não é conhecida e nenhuma hipótese pode ser feita acerca da caracterização dos parâmetros estamos no caso de aproximação não paramétrica. Funções discriminante parametrizadas (i.e. os coeficientes de um polinômio multivariado de algum grau) são utilizadas em métodos não paramétricos, porque nenhuma forma de distribuição é assumida.

A seguir, apresentamos de maneira detalhada uma parte da Figura 3.2, quando a informação é conhecida a priori e a classificação de padrões está baseada na Teoria da Decisão de Bayes.

3.5.1. Classificação de Padrões Utilizando a Teoria da Decisão de Bayes

A Teoria da Decisão de Bayes é fundamental na aproximação estatística, para o problema de classificação de padrões. Esta aproximação é baseada na hipótese de que o problema de decisão está caracterizado em termos probabilísticos e que todos os valores probabilísticos conhecidos são relevantes.

Basicamente o reconhecimento de padrões estatístico segundo a Teoria da Decisão de Bayes, significa que: Dado um vetor de características $\mathbf{x} = (x_1, \dots, x_K)^T$ de um padrão a ele será atribuída uma das S classes de padrão. Este processo está baseado na densidade de probabilidade do vetor $\mathbf{x}$ condicionado a uma classe padrão ω_i , a qual denota-se por $p(\mathbf{x}|\omega_i)$, e na probabilidade a priori da classe ω_i , denotado por $p(\omega_i)$. A $p(\mathbf{x}|\omega_i)$ está também associada a função de verossimilhança da classe ω_i . Com estas descrições, introduziremos a função perda, a função discriminante de Bayes, e a decisão por verossimilhança.

3.5.1.1. Função Perda

O uso da função perda em Teoria da Decisão é totalmente relevante.

Para a obtenção da função perda é preciso supor que:

- 1) $p(\omega_i)$ é conhecida ou estimada;
- 2) $p(x|\omega_i)$ é conhecida ou será estimada diretamente do conjunto de treinamento;
- 3) $p(\omega_i|x)$ geralmente não é conhecida.

A função perda L_{ij} será definida como a perda, ou o custo, ou a penalidade devido a uma decisão errônea de que, $x \in \omega_j$; quando de fato $x \in \omega_i$. Assim, o objetivo é procurar minimizar a perda média. A perda média condicional ou risco médio condicional $R_s(x)$, é definido como:

$$R_s(x) = \sum_{i=1}^S L_{is} p(\omega_i|x) \quad (3.14)$$

que é a perda média ou perda esperada de má classificação de x , quando $x \in \omega_s$.

Assim, o classificador encontrará uma decisão ótima, que minimiza o risco médio ou o custo médio.

A regra de decisão consiste nos seguintes passos:

PASSO 1 : Calcular as perdas esperadas, $R_i(x)$, para decidir que $x \in \omega_i$,
 $\forall i=1, \dots, S$;

PASSO 2 : Decidir que $x \in \omega_s$ se

$$R_s(x) \leq R_i(x) \quad \forall i, i \neq s$$

A função discriminante $g_s(x)$, é definida por

$$g_s(x) = -R_s(x) \quad (3.15)$$

O sinal negativo é escolhido para que $g_s(x)$ represente a classe mais provável. Isto é, o menor valor $R_s(x)$, indica uma maior probabilidade de que $x \in \omega_s$.

As funções perda podem variar, dependendo da informação que se possua. Mostraremos duas matrizes perda L de ordem $S \times S$. A primeira é dada pela seguinte forma:

$$L = \begin{bmatrix} 0 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ 1 & & & & 0 \end{bmatrix} \quad (3.16)$$

onde devemos considerar dois casos:

- 1) $L_{ii} = 0$, $i=1, \dots, s$ visto que nenhuma má classificação ocorre em tais casos.
- 2) $L_{is} = 1$, que é uma penalidade de má classificação, isto é classifica-se como $x \in \omega_s$, sendo que $x \in \omega_i$ $i=1, \dots, s$, $i \neq s$

Neste caso, temos uma função perda simétrica, isto é,

$$L_{is} = 1 - \delta(s - i)$$

onde $\delta(\cdot)$ é a função delta de Kronecker, isto é,

$$\delta(s - i) = \begin{cases} 1 & \text{se } s=i \\ 0 & \text{se } s \neq i \end{cases}$$

A segunda matriz função perda é:

$$L_{\text{Neg}} = \begin{bmatrix} -h_1 & & & \\ & -h_2 & & 0 \\ & & \ddots & \\ 0 & & & -h_s \end{bmatrix} \quad (3.17)$$

onde:

$$L_{is} = \begin{cases} -h_i & s=i \\ 0 & s \neq i \end{cases}$$

O significado desta matriz função perda negativa é que uma perda negativa (i.e. ganho positivo) é atribuída a uma decisão correta e nenhuma perda para uma decisão errônea. Em outras palavras, a perda atribuída para uma decisão é maior para uma decisão errônea que uma correta.

Dadas estas funções perda, descreveremos a seguir a função discriminante de Bayes.

3.5.1.2. Função discriminante de Bayes

Pela regra de Bayes temos que,

$$p(\omega_i | x) = \frac{p(x | \omega_i) p(\omega_i)}{p(x)} \quad (3.18)$$

onde:

$p(\mathbf{x}) = \sum_{i=1}^S p(\mathbf{x}|\omega_i) p(\omega_i)$ é a probabilidade de que $\mathbf{x}$ ocorre sem considerar a classe a qual pertence.

Substituindo (3.18) em (3.14), temos

$$R_s(\mathbf{x}) = \frac{1}{p(\mathbf{x})} \sum_{i=1}^S L_{is} p(\mathbf{x}|\omega_i) p(\omega_i) \quad (3.19)$$

observe, que em (3.19) $R_s(\mathbf{x})$, $s = 1, \dots, S$, independe de $p(\mathbf{x})$. Portanto, $p(\mathbf{x})$ pode ser eliminado da equação de risco médio condicional. Assim sendo, nosso objetivo então passa a ser

$$\min_s \{R_s(\mathbf{x})\} = \min_s \left\{ \sum_{i=1}^S L_{is} p(\mathbf{x}|\omega_i) p(\omega_i) \right\}. \quad (3.20)$$

Para obtermos o melhor resultado entre todas as possíveis decisões, definiremos

$$g_s(\mathbf{x}) = -R_s(\mathbf{x}) \quad (3.21)$$

Desse modo, $g_s(\mathbf{x})$ é denominado a função discriminante de Bayes. O classificador baseado nesta minimização é chamado de classificador de Bayes tendo, portanto, um desempenho ótimo sob o ponto de vista estatístico.

Com o conceito de função discriminante de Bayes introduzimos a decisão de máxima verossimilhança.

3.5.1.3. Decisão de máxima verossimilhança

Para o desenvolvimento da decisão de máxima verossimilhança, consideraremos primeiramente duas classes de padrões ω_1 e ω_2 , para depois generalizá-la. Para isto utilizaremos a definição de função perda média $R_s(x)$.

Seja, $R_1(x)$ a perda média quando decidimos que $x \in \omega_1$, isto é,

$$R_1(x) = L_{11} p(x|\omega_1) p(\omega_1) + L_{21} p(x|\omega_2) p(\omega_2) \quad (3.22)$$

Similarmente, a perda de decidir $x \in \omega_2$ é

$$R_2(x) = L_{12} p(x|\omega_1) p(\omega_1) + L_{22} p(x|\omega_2) p(\omega_2) \quad (3.23)$$

A decisão de que $x \in \omega_1$ é feita se:

$$R_1(x) < R_2(x) \quad (3.24)$$

Substituindo (3.22) e (3.23), respectivamente, em (3.24) resulta em

$$\begin{aligned} L_{11} p(x|\omega_1) p(\omega_1) + L_{21} p(x|\omega_2) p(\omega_2) \\ < L_{12} p(x|\omega_1) p(\omega_1) + L_{22} p(x|\omega_2) p(\omega_2) \end{aligned} \quad (3.25)$$

desenvolvendo esta desigualdade temos,

$$(L_{21} - L_{22}) p(x|\omega_2) p(\omega_2) < (L_{12} - L_{11}) p(x|\omega_1) p(\omega_1) \quad (3.26)$$

$$\frac{p(x|\omega_1)}{p(x|\omega_2)} > \frac{(L_{21} - L_{22}) p(\omega_2)}{(L_{12} - L_{11}) p(\omega_1)} \quad (3.27)$$

Denotando:

$$\frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} = \ell_{12}(\mathbf{x}) \quad (3.28)$$

e

$$\frac{(L_{21} - L_{22}) p(\omega_2)}{(L_{12} - L_{11}) p(\omega_1)} = \theta_{12} \quad (3.29)$$

e substituindo em (3.27), a expressão (3.27) torna-se

$$\ell_{12}(\mathbf{x}) > \theta_{12} \quad (3.30)$$

e, o critério de decisão será

$$\mathbf{x} \in \omega_1, \quad \text{se} \quad \ell_{12}(\mathbf{x}) > \theta_{12} \quad (3.31)$$

As expressões (3.28) e (3.29) são denominadas de razão de verossimilhança e valor do limiar, respectivamente.

Estes resultados podem ser generalizados para um problema de múltiplas classes, isto é, para as S classes. Neste caso (3.28) e (3.29) podem ser expressadas como:

$$\ell_{si} = \frac{p(\mathbf{x}|\omega_s)}{p(\mathbf{x}|\omega_i)} \quad (3.32)$$

e

$$\theta_{si} = \frac{(L_{is} - L_{ii}) p(\omega_i)}{(L_{si} - L_{ss}) p(\omega_s)} \quad (3.33)$$

Logo, o critério de decisão para as múltiplas classes será

$$x \in \omega_s \quad \text{se} \quad \ell_{si} > \theta_{si} \quad \forall i=1, \dots, S \quad (3.34)$$

A expressão (3.34) é denominada regra de máxima verossimilhança.

Sob estas regras de decisão não é difícil implementar-se o classificador. Observe que em (3.33) o limiar depende dos valores da função perda. Considerando a matriz função perda (3.16), a expressão (3.33) transforma-se em:

$$\theta_{si} = \frac{p(\omega_i)}{p(\omega_s)}$$

Substituindo o valor de θ_{si} em (3.34) temos:

$$\frac{p(x|\omega_s)}{p(x|\omega_i)} > \frac{p(\omega_i)}{p(\omega_s)} \quad (3.35)$$

Se $p(\omega_i) = p(\omega_s)$. A regra de máxima verossimilhança será então

$$x \in \omega_s \quad \text{se} \quad \ell_{si} > 1 \quad (3.36)$$

O caso mais geral ocorre quando $p(\omega_i) \neq p(\omega_s)$. Considerando estas possibilidades, formularemos então as funções discriminante, quando a função perda é simétrica.

Sabemos que:

$$\ell_{si} = \frac{p(x|\omega_s)}{p(x|\omega_i)} > \frac{p(\omega_i)}{p(\omega_s)}$$

então:

$$p(x|\omega_{s1}) p(\omega_s) > p(x|\omega_i) p(\omega_i) \quad (3.37)$$

em outras palavras,

$$x \in \omega_s \quad \text{se}$$

$$p(x|\omega_s) p(\omega_s) > p(x|\omega_i) p(\omega_i) \quad \forall i=1, \dots, S$$

Uma das formas assumidas pela função discriminante é,

$$d_s(x) = p(x|\omega_s) p(\omega_s) \quad (3.38)$$

Todavia, a forma alternativa mais utilizada para a função discriminante é,

$$d'_s(x) = \log p(x|\omega_s) + \log p(\omega_s) \quad (3.39)$$

Em seguida, desenvolveremos a classificação de padrões para funções densidade normal ou gaussiana.

3.5.2. Classificação de Padrões Quando a Função Densidade é Gaussiana

A função densidade gaussiana multivariada para S classes de padrões é representada matematicamente por

$$p(x|\omega_s) = \frac{1}{(2\pi)^{k/2} |C_s|^{1/2}} \exp \left[-\frac{1}{2} (x - m_s)^T C_s^{-1} (x - m_s) \right] \quad s=1, \dots, S \quad (3.40)$$

onde:

- k é a dimensionalidade do vetor de características do padrão;
- $\mathbf{m}_s = E(\mathbf{x})$ é o vetor esperado de $\mathbf{x}$ ou vetor de médias;
- $\mathbf{C}_s = E_s[(\mathbf{x} - \mathbf{m}_s)(\mathbf{x} - \mathbf{m}_s)^T]$ é a matriz covariância da classe ω_s .

Para a obtenção da função discriminante consideremos (3.39), isto é,

$$d'_s(\mathbf{x}) = \log p(\mathbf{x}|\omega_s) + \log p(\omega_s)$$

Substituindo (3.47) nesta igualdade, a função discriminante fica

$$d''_s(\mathbf{x}) = -\frac{k}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{C}_s| - \frac{1}{2} (\mathbf{x} - \mathbf{m}_s)^T \mathbf{C}_s^{-1} (\mathbf{x} - \mathbf{m}_s) + \log p(\omega_s) \quad (3.41)$$

Observemos que o primeiro termo do lado direito é comum para as S classes. Portanto, podemos desconsiderá-lo sem perda de generalidade. Logo a função discriminante reduz-se a:

$$d'''_s(\mathbf{x}) = -\frac{1}{2} (\mathbf{x} - \mathbf{m}_s)^T \mathbf{C}_s^{-1} (\mathbf{x} - \mathbf{m}_s) + \left[\log p(\omega_s) - \frac{1}{2} \log |\mathbf{C}_s| \right] \quad (3.42)$$

Pelo primeiro termo do lado direito, vemos que $d'''_s(\mathbf{x})$ é uma função discriminante quadrática.

A seguir consideraremos, esta função discriminante para dois casos diferentes:

CAPÍTULO 1: Quando as Matrizes Covariância são Iguais para Diferentes Classes

$$(i.e. \mathbf{C}_1 = \mathbf{C}_2 = \dots = \mathbf{C}_s = \mathbf{C})$$

O significado físico é; que as classes separadas são de igual tamanho e forma similar e os "cluster" estão centrados em diferentes médias. Expandindo a equação geral (3.42) temos

$$d_s(x) = -\frac{1}{2} x^T C^{-1} x - \frac{1}{2} m_s^T C^{-1} m_s + x^T C^{-1} m_s + \log p(\omega_s) - \frac{1}{2} \log |C| \quad s=1, \dots, S \quad (3.43)$$

Aqui, observamos que o primeiro e último termo são comuns em todas as classes. Então a função de discriminante mais compacta é como segue:

$$d_s(x) = x^T C^{-1} m_s + \left(\log p(\omega_s) - \frac{1}{2} m_s^T C^{-1} m_s \right) \quad s=1, \dots, S \quad (3.44)$$

Obviamente, esta é uma função discriminante linear, devido a x .

A função discriminante para um problema de duas classes $S = 2$, é definida como:

$$d(x) = d_1(x) - d_2(x) \quad (3.45)$$

$$d(x) = x^T C^{-1} (m_1 - m_2) + \log \frac{p(\omega_1)}{p(\omega_2)} - \frac{1}{2} (m_1^T C^{-1} m_1 - m_2^T C^{-1} m_2) \quad (3.46)$$

ou

$$d(x) = x^T C^{-1} (m_1 - m_2) + \log \frac{p(\omega_1)}{p(\omega_2)} - \frac{1}{2} \left[(m_1 - m_2)^T C^{-1} (m_1 + m_2) \right] \quad (3.47)$$

CAPÍTULO 2: Quando a Matriz C_2 é uma Matriz $\sigma_s^2 I$, onde $|C_s| = \sigma_s^2$

Fisicamente significa que o "cluster" tem componentes iguais ao longo dos eixos principais e a distribuição é de forma esférica. Substituindo $\sigma_s^2 I$ por C_s em (3.42), resulta,

$$d_s(\mathbf{x}) = -\frac{1}{2} \frac{(\mathbf{x} - \mathbf{m}_s)^T (\mathbf{x} - \mathbf{m}_s)}{\sigma_s^2} + \log \left[p(\omega_s) - \frac{1}{2} \log \sigma_s^2 \right] \quad (3.48)$$

$$\text{Se, } \sigma_s = \sigma_j \quad \forall s, j$$

$$\text{isto é, } \sigma_s^2 \mathbf{I} = \sigma_j^2 \mathbf{I} = \sigma^2 \mathbf{I}$$

(3.48), transforma-se em

$$d_s(\mathbf{x}) = -\frac{1}{2} \frac{(\mathbf{x}^T \mathbf{x} - 2\mathbf{x}^T \mathbf{m}_s + \mathbf{m}_s^T \mathbf{m}_s)}{\sigma^2} + \log p(\omega_s) - \frac{1}{2} \log \sigma^2 \quad (3.49)$$

novamente, $\mathbf{x}^T \mathbf{x}$ e $\frac{1}{2} \log \sigma^2$ são comuns a todas as classes de padrões.

Portanto, eliminando estes termos em $d_s(\mathbf{x})$ temos

$$d_s(\mathbf{x}) = \frac{1}{\sigma^2} \mathbf{x}^T \mathbf{m}_s + \left(\log p(\omega_s) - \frac{1}{2\sigma^2} \mathbf{m}_s^T \mathbf{m}_s \right) \quad (3.50)$$

que pode ser tratada como função discriminante linear. Esta função discriminante poderá ser simplificada se $p(\omega_s) = \frac{1}{S}$, $\forall s$. Como S é constante, o termo $\log p(\omega_s)$ pode ser eliminado de (3.50), e assim $d_s(\mathbf{x})$ simplifica-se para

$$d_s(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_s - \frac{1}{2} \|\mathbf{m}_s\|^2 \quad (3.51)$$

que também é uma função linear.

As funções discriminante linear (3.44), (3.50) e (3.51) podem ser implementadas por máquinas lineares.

Através destas funções discriminantes pode-se determinar as regiões de decisão, e assim classificar os padrões.

CAPÍTULO 4

CASAMENTO ENTRE CONJUNTO DE SINAIS E GRUPOS

4.1. INTRODUÇÃO

Dentro do contexto de codificação de canal o casamento de conjuntos de sinais com a estrutura algébrica de grupo é um dos temas atuais de pesquisa na área de Teoria de Comunicações. Através deste casamento é possível diminuir a potência transmitida para um mesmo valor da probabilidade de erro, sem que com isso, haja uma expansão da largura de faixa.

O casamento consiste na rotulação dos vetores sinais por elementos de um grupo, implicando em uma correspondência entre linearidade e distância quando a distância de Hamming é substituída pela distância Euclideana.

Massey e Mittlholzer [20] - [23] introduziram o casamento de códigos lineares sobre $\mathbb{Z}_M$ com modulação de fase M-ários.

Loeliger [24] analisa este casamento entre conjunto de sinais e grupos. Como resultado principal mostra que os conjuntos de sinais casados a grupos são equivalentes a conjuntos de sinais do tipo Slepian [25].

Neste capítulo desenvolveremos o casamento entre sinais e estrutura de grupo. Na Seção 4.2, serão fornecidos os conceitos básicos de Teoria de Grupos. Na Seção 4.3, uma breve revisão de Sistemas de Comunicação. Finalmente na Seção 4.4 apresentamos o casamento de conjuntos de sinais e grupos.

4.2. GRUPOS

4.2.1. Definições e Alguns Exemplos

Definição 4.2.1. : Seja G um conjunto não vazio com a operação

$$\begin{aligned} * : G \times G &\longrightarrow G \\ (a,b) &\longrightarrow a * b \end{aligned}$$

G é um grupo se as seguintes condições são satisfeitas:

i) Se $a, b \in G$, então, $a * b \in G$ (propriedade de fechamento)

ii) A operação $*$ é *Associativa*; isto é,

$$(a * b) * c = a * (b * c), \quad \forall a, b, c \in G;$$

iii) Existe o *Elemento Identidade*, isto é,

$$\exists e \in G \text{ tal que } a * e = e * a = e, \quad \forall a \in G;$$

iv) Cada elemento $a \in G$ tem um *Elemento Inverso*, isto é,

$$\forall a \in G, \exists b \in G \text{ tal que } a * b = b * a = e$$

- Se em um grupo G , verifica-se a propriedade:

$$a * b = b * a$$

dizemos que o grupo é *Comutativo ou Abelliano*.

Exemplo 4.2.1.: Seja S um conjunto não vazio e seja

$$G = \{f: S \rightarrow S : f \text{ bijetiva}\}$$

Se $*$ é a operação composição de funções, isto é,

$$\begin{aligned} * : G \times G &\longrightarrow G \\ (g, f) &\longrightarrow g * f \end{aligned}$$

então $(G, *)$ é claramente um grupo tendo

$$\begin{aligned} I_s : S &\longrightarrow S \\ x &\longrightarrow x \end{aligned}$$

como identidade.

Esse grupo é chamado de *grupo das permutações ou grupo das simetrias do conjunto* S . Se $S = \{1, 2, \dots, n\}$ denotaremos esse grupo por S_n , e temos que o número de elementos de S_n é exatamente $n!$

O número de elementos em um grupo G é chamado *ordem do grupo* denotado por $\#G$. G é chamado um *grupo finito* se $\#G$ é finito e é chamado um *grupo infinito* caso contrário.

Um grupo G é chamado cíclico se existe um elemento $g \in G$ tal que $G = \{g^n \mid n \in \mathbb{Z}\}$. O elemento g é chamado *gerador* do grupo cíclico.

A *ordem* de um elemento g em um grupo G é o menor inteiro positivo, r , tal que $g^r = e$. Se tal r não existe, a ordem dos elementos é dita infinita.

Uma *relação* R de um conjunto U em T é um subconjunto de $U \times T$. Se diz que a está relacionado a b sob R se o par (a, b) pertence ao subconjunto e denota-se por aRb , caso contrário $a \not R b$.

Definição 4.2.2.: Uma relação R sobre um conjunto U é chamada de uma *relação de equivalência* se

- i) $aRa \quad \forall a \in U$ (Reflexiva)
- ii) Se aRb , então bRa (Simétrica)
- iii) Se aRb e bRc , então aRc (Transitiva)

Se R é uma relação de equivalência sobre U e $a \in U$ então

$$[a] = \{x \in U \mid xRa\}$$

é chamado de *classe de equivalência* contendo a . O conjunto de todas as classes é chamado o *conjunto quociente* de U por R e denotado por U/R , isto é

$$U/R = \{[a] \mid a \in U\}.$$

Para cada $n \in \mathbb{Z}^+$, tem-se uma classe muito importante de relação de equivalência sobre $\mathbb{Z}$, que é a *congruência módulo n* .

Se n é um inteiro positivo fixo e a e b inteiros quaisquer, se diz que a é *congruente* a b *módulo n* , se n divide $(a-b)$, comumente denotado por $a \equiv b \pmod{n}$. Esta relação de congruência módulo n é uma relação de equivalência sobre $\mathbb{Z}$.

Definição 4.2.3.: Se $(G, \cdot)$ é um grupo e H é um subconjunto não vazio de G , então $(H, \cdot)$ é chamado um *subgrupo* de $(G, \cdot)$ se

- i) $a \cdot b \in H \quad \forall a, b, \in H$ (Fechamento);
- ii) $a^{-1} \in H \quad \forall a \in H$ (Existência de Inversos).

Denota-se $H \leq G$.

Exemplo 4.2.2.: O grupo composto pelas permutações par em S_n forma um subgrupo de ordem $n!/2$ chamado também o grupo alternante A_n de grau n .

Definição 4.2.4.: Seja H um subgrupo de um grupo G . Para $a, b, \in G$, se diz que a é *congruente a b módulo H* denotado por $a \equiv b \pmod{H}$, se e somente se $ab^{-1} \in H$.

A relação $a \equiv b \pmod{H}$ é uma relação de equivalência sobre G . Se denotarmos

$$Ha = \{ha \mid h \in H\}$$

então, temos a chamada classe lateral à *direita* de H em G . O elemento a é chamado um *representativo* da classe lateral Ha . A classe lateral à *esquerda* é dada por

$$aH = \{ah \mid h \in H\}.$$

Teorema 4.2.1.: *Teorema de Lagrange.* Se G é um grupo finito e H é um subgrupo de G , então a ordem de H divide a ordem de G .

4.2.2. Grupos Quocientes e Homomorfismo de Grupos

Definição 4.2.5.: Se $(G, \circ)$ e $(H, *)$ são dois grupos, a função $f : G \rightarrow H$ é chamada homomorfismo de grupo se

$$f(a \circ b) = f(a) * f(b).$$

Um *isomorfismo de grupo* é um homomorfismo de grupos que é injetivo. Se existe um isomorfismo entre os grupos $(G, \circ)$ e $(H, *)$, se diz que $(G, \circ)$ e $(H, *)$ são *isomórfos* e denota-se por $(G, \circ) \cong (H, *)$. Dois grupos são isomórfos se as suas estruturas são essencialmente as mesmas.

Definição 4.2.6.: Um subgrupo H de um grupo G é chamado um *subgrupo normal* de G se

$$g^{-1}Hg \in H, \quad \forall g \in G, \quad \forall h \in H.$$

Teorema 4.2.2.: Se N é um subgrupo normal de $(G, \circ)$, o conjunto de classes laterais $G/N = \{Ng \mid g \in G\}$ forma um grupo $(G/N, \circ)$ onde a operação é dada por

$$(Ng_1)(Ng_2) = Ng_1 \cdot g_2.$$

Definição 4.2.7.: O grupo G/N é chamado *grupo quociente* ou *grupo fator* de G em N .

Definição 4.2.8.: Se $f : G \rightarrow H$ é um homomorfismo de grupo, então o *núcleo* de f denotado por $\text{Ker } f$, é o conjunto de elementos de G mapeados por f na identidade de H . Isto é,

$$\text{Ker } f = \{g \in G \mid f(g) = e_H\}.$$

Proposição 4.2.1.: Seja $f : G \rightarrow H$ um homomorfismo de grupo. Então,

- i) $\text{Ker } f$ é um subgrupo normal de G ;
- ii) f é injetiva $\Leftrightarrow \text{Ker } f = \{e_G\}$.

Proposição 4.2.2.: Para qualquer homomorfismo de grupo $f : G \rightarrow H$, a *imagem* de f , $\text{Im } f = \{f(g) | g \in G\}$ é um subgrupo de H (embora não necessariamente normal).

Teorema 4.2.3.: *Teorema de Homomorfismo para Grupos.* Seja K o núcleo do homomorfismo de grupo $f : G \rightarrow H$. Então G/K é isomorfo à imagem de f , e o isomorfismo $\Psi : G/K \rightarrow \text{Im } f$ é definido por

$$\Psi(Kg) = f(g)$$

4.2.3. Anéis e Corpos

Definição 4.2.9.: Um anel $(R, +, \circ)$ é um conjunto R , juntamente com duas operações binárias $+$ e $\circ$ sobre R satisfazendo os seguintes axiomas. Para quaisquer elementos $a, b, c \in R$.

- i) $(a + b) + c = a + (b + c)$ (Associativa)
- ii) $a + b = b + a$ (Comutativa)
- iii) $\exists 0 \in R$, chamado zero, tal que $a + 0 = a$ (Identidade Aditiva)
- iv) $\exists (-a) \in R$ tal que $a + (-a) = 0$ (Inverso Aditivo)
- v) $(a \circ b) \circ c = a \circ (b \circ c)$ (Associativa)
- vi) $\exists 1 \in R$ tal que $1 \circ a = a \circ 1 = a$ (Identidade Multiplicativa)
- vii) $a \circ (b + c) = a \circ b + a \circ c$
 $(b + c)a = b \circ a + c \circ a$ (Distributiva)

O anel $(R, +, \circ)$ é chamado *Comutativo* se viii) $a \circ b = b \circ a \quad \forall a, b \in R$.

Se $(R, +, \cdot)$ é um anel comutativo, um elemento não nulo $a \in R$ é chamado de *divisor de zero* se $\exists$ um elemento não nulo $b \in R$ tal que $a \cdot b = 0$.

Um anel comutativo não trivial contendo a unidade é chamado de *domínio integral* se não contém divisores de zero. Portanto, um anel comutativo não trivial é um domínio integral se $a \cdot b = 0$ sempre que $a = 0$ ou $b = 0$.

Definição 4.2.10.: Um corpo é um anel no qual os elementos não nulos formam um grupo abeliano sob a operação de multiplicação, isto é, além de satisfazer os axiomas de Definição 4.2.9 satisfaz

ix) $a \in R, a \neq 0 \exists a^{-1} \in R$ tal que $a^{-1} \cdot a = 1$.

Teorema 4.2.4.: *Algoritmo de Divisão para Inteiros*. Se a e b são inteiros e b é não nulo, então existem inteiros q e r únicos tais que

$$a = qb + r \quad 0 \leq r < |b|$$

O inteiro r é chamado o *resíduo* na divisão de a por b , e q de *quociente*.

Definição 4.2.11.: Um domínio integral R é chamado um *anel Euclidiano* se para cada elemento não nulo $a \in R$, existe um inteiro não negativo $|a|$.

- i) Se a e b são elementos não nulos de R , então $|a| \leq |ab|$;
- ii) Para cada par de elementos $a, b \in R$ com $b \neq 0 \exists$ elementos $q, r \in R$ tais que $a = qb + r$; $r = 0$ ou $|r| < |b|$ (algoritmo de divisão).

Definição 4.2.12.: Um subconjunto não vazio I de um anel R é chamado *ideal* de R , se $\forall x, y \in I, e r \in R$

- i) $x - y \in I$
- ii) $x \cdot r$ e $r \cdot x \in I$

Proposição 4.2.3.: Seja a um elemento de um anel comutativo R . O conjunto $\{ar \mid r \in R\}$ de todos os múltiplos de a é um ideal de R chamado *ideal principal* gerado por a . Este ideal é denotado por (a) .

Teorema 4.2.5.: Seja I um ideal no anel R . Então o conjunto de classes laterais formam um anel $(R/I, +, \cdot)$ sob as operações definidas por:

$$(I + r_1) + (I + r_2) = I + (r_1 + r_2)$$

$$(I + r_1) \cdot (I + r_2) = I + (r_1 \cdot r_2)$$

Definição 4.2.13.: O anel $(R/I, +, \cdot)$ é chamado *anel quociente de R por I* .

Um elemento p não inversível em um anel Euclideo R é chamado *irreduzível* se, $p = ab$, então a ou b é inversível em R .

Teorema 4.2.6.: Seja a um elemento de um anel Euclideo. O anel quociente $R/(a)$ é um corpo $\Leftrightarrow a$ é irreduzível em R .

Corolário 4.2.1.: $\mathbb{Z}_p = \mathbb{Z}/(p)$ é um corpo $\Leftrightarrow p$ é primo.

4.3. REVISÃO DE SISTEMAS DE COMUNICAÇÕES

Na Figura 4.1 mostramos o diagrama de bloco de um sistema de comunicação digital. Neste diagrama, três operações básicas de processamento de sinais são identificados: *Codificação de Fonte*, *Codificação do Canal* e *Modulação*.

Desenvolveremos brevemente cada uma destas operações básicas.

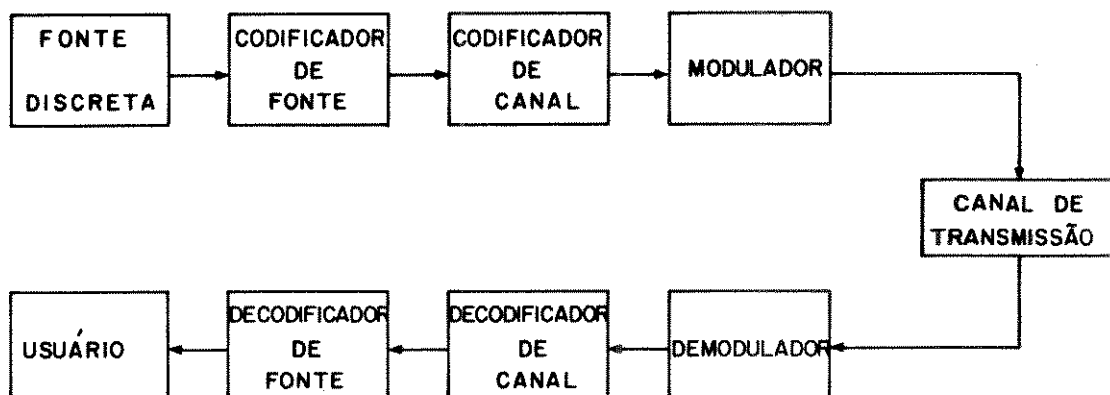


Figura 4.1 - Diagrama de Bloco de um Sistema de Comunicação Digital.

4.3.1. Codificação de Fonte

O objetivo em codificação de fonte é eliminar ou reduzir a redundância das palavras-código que são associadas aos elementos do alfabeto respectivamente. Como consequência disto obtém-se uma representação eficiente da saída da fonte. O benefício da aplicação de codificação de fonte é reduzir a largura de faixa requerida.

Uma forma de avaliar o desempenho de um sistema de comunicação digital diz respeito à eficiência com que a informação de uma dada fonte pode ser representada.

Suponha que um experimento probabilístico envolva a observação da saída emitida por uma fonte discreta em cada unidade de tempo. A saída da fonte é modelada como uma variável aleatória discreta, ζ , onde os símbolos assumem valores pertencentes à um alfabeto finito fixo. Seja ζ representado por

$$\zeta = \{\zeta_0, \dots, \zeta_{K-1}\}$$

onde a cada elemento ζ_k , temos associado uma medida, $p(\cdot)$, de tal forma que

$$p(\zeta = \zeta_k) = p_k \quad k=0,1,\dots,K-1$$

satisfazendo,

$$\sum_{k=0}^{K-1} p_k = 1$$

assuma que os símbolos emitidos pela fonte durante intervalos de sinais sucessivos são estatisticamente independentes.

A quantidade de informação associada ao evento $\zeta = \zeta_k$, é definida como

$$I(\zeta_k) = \log_2(1/p_k)$$

onde, a unidade resultante de informação é chamada bit. Um bit é a quantidade de informação associada a cada um dos dois eventos possíveis de ocorrer quando as mesmas são igualmente prováveis. A quantidade de informação, $I(\zeta_k)$, só depende do símbolo ζ_k emitido pela fonte.

O valor médio de $I(\zeta_k)$ sobre o alfabeto da fonte ζ é dado por

$$H(\zeta) = E(I(\zeta_k)) = \sum_{k=0}^{K-1} p_k \log_2(1/p_k)$$

onde $H(\zeta)$, é chamada de *entropia* de uma fonte sem memória discreta com alfabeto finito ζ . Isto é, uma medida da *informação média contida por símbolo da fonte*.

Teorema 4.3.1. Codificação da Fonte: Dada uma fonte sem memória discreta de entropia $H(\zeta)$, a média $\mathcal{L}$ do comprimento das palavras-código para qualquer codificação de fonte é limitada inferiormente por

$$\mathcal{L} > H(\zeta)$$

onde:

$$\mathcal{L} = \sum_{k=0}^{K-1} p_k \ell_k$$

ℓ_k é o comprimento da palavra código atribuída ao símbolo ζ_k .

4.3.2. Canais de Comunicação

Os canais mais comumente utilizados na transmissão de informação são classificados como pertencentes a um dos dois tipos, descritos a seguir:

- *Canais do Tipo Discreto*, são aqueles que se apresentam entre a saída do codificador de canal, e a entrada do decodificador. O canal que é mais utilizado é o *Canal Simétrico Binário (BSC)* que tem a particularidade de que a probabilidade de mudar o dígito 1 para 0 é igual a probabilidade de mudar 0 a 1 (ver Figura 4.2).

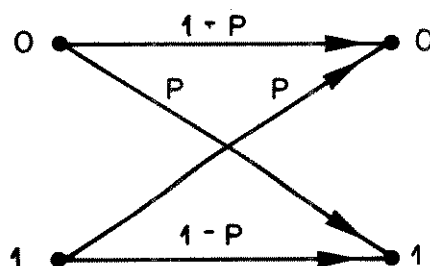


Figura 4.2 – Canal Simétrico Binário.

• *Canais do Tipo Contínuo*, são aqueles que se apresentam entre a saída do modulador e entrada do demodulador. Na sua forma mais simples, é representado como um canal aditivo. O canal é completamente caracterizado pela distribuição de probabilidade do ruído. O canal de tipo contínuo que utilizaremos para transmitir os sinais $s_i(t)$ (ver Figura 4.3) é o *canal Gaussiano Branco e Aditivo (AWGN)* com potência espectral $N_0/2$.

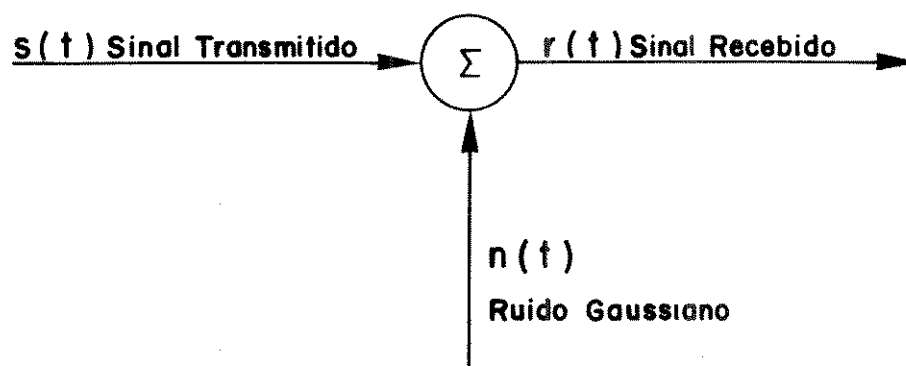


Figura 4.3 – Canal Gaussiano Aditivo.

4.3.3. Codificação de Canal

4.3.3.1. Codificação

Códigos são utilizados em sistemas de comunicações com o objetivo de corrigir e detectar erros ocorridos nas mensagens transmitidas através de um canal de comunicação. Apesar de existirem duas classes de códigos, nos restringiremos à somente uma delas, uma vez que estaremos interessados neste trabalho com os códigos de bloco. A seguir, apresentaremos os conceitos básicos necessários para um melhor entendimento do material contido nos capítulos subsequentes. Iremos considerar somente o caso de códigos q -ários, com elementos em $\mathbb{Z}_q$.

A transmissão de mensagens sobre um canal é feita através do particionamento destas mensagens em blocos de k dígitos e a cada um desses blocos são adicionados $n-k$ dígitos de paridade obtendo-se assim uma *palavra codificada* de comprimento n dígitos. Tal código é denotado por (n, k) .

Como cada bloco da mensagem u consiste de k dígitos de informação, temos então um total de q^k possíveis mensagens. O codificador, transforma cada mensagem u em uma n -upla q -ária v , onde $n > k$. Dessa forma, às q^k possíveis mensagens, temos associadas q^k palavras codificadas de um conjunto de q^n possíveis palavras.

Os $(n-k)$ dígitos adicionados aos dígitos de informação são chamados *dígitos de verificação de paridade* pois os mesmos não "carregam" informação; mas permitem que a informação existente seja transmitida de uma forma mais confiável.

As palavras código serão representadas por $v = (v_1, v_2, \dots, v_n)$ pertencentes ao espaço vetorial $\mathbb{Z}_q^n = \mathbb{Z}_q \times \mathbb{Z}_q \times \dots \times \mathbb{Z}_q$ de dimensão n , onde q é um primo. Por outro lado, as q^k possíveis mensagens de comprimento k , $u = (u_1, u_2, \dots, u_k)$, pertencem ao espaço vetorial $\mathbb{Z}_q^k$.

Definição 4.3.1. : Um *codificador* é uma função injetiva

$$\gamma : \mathbb{Z}_q^k \rightarrow \mathbb{Z}_q^n$$

que atribui a cada k dígitos de informação, uma palavra-código de n dígitos.

Definição 4.3.2. : Se G é uma matriz de ordem $k \times n$ cujos elementos estão em $\mathbb{Z}_q$. Define-se a transformação linear

$$\gamma_G : \mathbb{Z}_q^k \rightarrow \mathbb{Z}_q^n \quad \text{por}$$

$$u \rightarrow uG$$

onde:

u é um vetor linha que representa os k dígitos de informação de uma mensagem.

A matriz G de ordem $k \times n$ é denominada *matriz geradora* ou *matriz de codificação do código*.

Definição 4.3.3.: Um código (n, k) é chamado *código linear* se a função de codificação é uma transformação linear de $\mathbb{Z}_q^k$ em $\mathbb{Z}_q^n$.

Se u é a mensagem, então a palavra codificada via G é dada por

$$v = u G \quad (4.3.1)$$

onde os vetores código são os vetores da imagem de γ_G (proposição 4.2.2), formando assim um subespaço vetorial de $\mathbb{Z}_q^n$ de dimensão k . Logo, as linhas de G formam uma base para este subespaço. Portanto, um vetor é uma palavra-código se e somente se este vetor resulta da combinação linear das linhas de G .

Como o processo de decodificação realiza a operação inversa ao da codificação, então se faz necessário definir esta transformação. Para isso, temos

Definição 4.3.4.: Se H é uma matriz $(n-k) \times n$ cujos elementos pertencem a $\mathbb{Z}_q$ então, define-se a transformação linear

$$\eta_H : \mathbb{Z}_q^n \rightarrow \mathbb{Z}_q^{n-k} \quad \text{por}$$

$$v^T \rightarrow H_{(n-k) \times n} v^T$$

onde:

v^T é um vetor coluna, T indica o transposto de v .

A matriz H é chamada de *matriz verificação de paridade* cujo posto é $(n-k)$.

Então, $v = (v_1, v_2, \dots, v_n)$ é uma palavra-código se e somente se

$$H v^T = 0 \quad (4.3.2)$$

Por outro lado, um vetor recebido $r = (r_1, \dots, r_n)$ é um vetor código se e somente se $H r^T = 0$.

A relação entre as matrizes H e G obtém-se de (4.3.1) e (4.3.2)

$$v = u G \rightarrow v^T = G^T u^T$$

Logo,

$$H v^T = H G^T u^T = 0$$

onde,

$$H G^T = 0 \text{ ou } G H^T = 0 \quad (4.3.3)$$

• Decodificação e Correção de Erros

Consideremos um código linear (n,k) como sendo uma função de codificação $\gamma_G : \mathbb{Z}_q^k \rightarrow \mathbb{Z}_q^n$. Seja $\mathcal{C} = \text{Im}(\gamma_G)$ o subespaço de vetores-código. Se $v \in \mathcal{C}$ é enviado através do canal ruidoso e um erro $e \in \mathbb{Z}_q^n$ ocorre durante a transmissão, o vetor recebido será

$$r = v + e \quad (4.3.4)$$

O decodificador recebe r , e determinará o mais provável vetor transmitido v através do padrão de erro e . De (4.3.4) $e = v - r$. O decodificador não conhece o vetor-código v , mas reconhece que o erro e está na classe lateral $\mathcal{C} + v$.

O mais provável padrão de erro em cada classe lateral de $\mathbb{Z}_q^n/\mathcal{C}$ é chamado líder da classe lateral.

As classes laterais de $\mathcal{C}$ em $\mathbb{Z}_q^n$ podem ser caracterizadas pelo uso da matriz H .

O subespaço $\mathcal{C}$ é o núcleo da transformação $\eta : \mathbb{Z}_q^n \rightarrow \mathbb{Z}_q^{n-k}$. Portanto, pelo teorema do homomorfismo (teorema 4.2.3.) o conjunto de classes laterais $\mathbb{Z}_q^n/\mathcal{C}$ é isomorfo a $\text{Im}(\eta)$, isto é, o isomorfismo envia a classe lateral $\mathcal{C} + v$ para $\eta(v^T) = H v^T$. Portanto, a classe lateral $\mathcal{C} + v$ é caracterizada pelo vetor $H v^T$.

Se H é uma matriz verificação de paridade e $r \in \mathbb{Z}_q^n$, então o vetor $(n-k)$ -dimensional, $H r^T$ é chamado *síndrome* de r .

Cada elemento de $\mathbb{Z}_q^{(n-k)}$ é uma síndrome. Desse modo, temos q^{n-k} diferentes classes laterais e q^{n-k} diferentes síndromes.

Definição 4.3.5.: A *distância de Hamming* entre dois vetores r e v é o número de posições nos quais elas diferem; e é denotada por $d_H(r, v)$.

Definição 4.3.6.: O *peso de Hamming* de um vetor r é o número de elementos r_i diferente de zero; e é denotado por $w(r)$.

Das definições temos que a distância de Hamming e o peso de Hamming estão relacionados por

$$d_H(r, v) = w(r - v) \quad (4.3.5)$$

O terceiro parâmetro de um código $\mathcal{C}$, além do comprimento e dimensão, é a distância de Hamming mínima entre suas palavras codificadas, isto é,

$$\begin{aligned} d &= \min d_H(r, v) \\ &= \min w(r - v); \quad \ell = r - v \in \mathcal{C} \end{aligned} \quad (4.3.6)$$

onde d é chamada distância mínima do código, significando que qualquer duas palavras código diferem ao menos em d lugares.

Definição 4.3.7.: A taxa η , de um código é definido por

$$\eta = \frac{k}{n}$$

onde:

k é o número de dígitos de informação;

n é o comprimento da palavra código.

A capacidade de detectar e corrigir um erro é dado pelo teorema 4.3.1

Teorema 4.3.1. [26]: Um código com distância mínima d pode corrigir $t = \left\lfloor \frac{d-1}{2} \right\rfloor$ erros*. Se d é par, o código pode simultaneamente corrigir $t = \frac{d-2}{2}$ erros e detectar $d/2$ erros.

• Probabilidade de Erro Quando a Mensagem é Transmitida por um Canal do Tipo BSC

Um código de bloco com capacidade de corrigir t erros é usualmente capaz de corrigir alguns padrões de erro, $t + i$, com $i \geq 1$.

Se um código bloco de t -correção de erros é utilizado estritamente para corrigir erros sobre um canal BSC com probabilidade de transição p , a probabilidade de que o decodificador cometa uma decisão errônea é limitada superiormente por

$$p(e) \leq \sum_{i=t+1}^n \binom{n}{i} p^i (1-p)^{n-i} \quad (4.3.7)$$

Uma maneira de obter a probabilidade de erro exata é através do Arranjo Padrão [26]

4.3.3.2. Modulação

O mapeamento de seqüências digitais em um conjunto de formas de onda é chamado de *modulação digital*.

* $\lfloor x \rfloor$ denota o maior inteiro menor ou igual que x .

O *modulador digital* é o que faz tal mapeamento. Quando o modulador mapeia dígitos binários em um conjunto de duas formas de onda tem-se, então, a *modulação binária*. Alternativamente o modulador pode ter $M = 2^n$ diferentes formas de onda pelo mapeamento de n dígitos a um tempo sobre um conjunto de M formas de onda ou sinais. Este tipo de transmissão é dito *modulação M-ária* ou *modulação multinível*.

O canal para a transmissão dos sinais será do tipo canal com ruído gaussiano aditivo branco (AWGN) com potência espectral $N_0/2$.

A análise de desempenho do sistema de comunicações, quando um dado conjunto de formas de ondas é selecionado, fica bastante simplificada se este conjunto de sinais puder ser representado geometricamente em um espaço de sinais.

A tarefa de transformar uma mensagem de entrada m_i , $i = 1, \dots, M$, na forma de onda modulada $s_i(t)$ pode ser dividida em operações de tempo discreto e contínuo. A justificação para esta operação está no *processo de ortogonalização de Gram-Schmidt*, o qual permite a representação de qualquer conjunto de M sinais $\{s_i(t)\}$ como combinações lineares de N *funções base ortonormais*, $\phi_1(t), \dots, \phi_N(t)$, onde $N \leq M$, isto é,

$$s_i(t) = \sum_{j=1}^N s_{ij} \phi_j(t) \quad 0 \leq t \leq T \quad i=1, 2, \dots, M \quad (4.3.8)$$

onde os coeficientes desta expansão são definidos por

$$s_{ij} = \int_0^T s_i(t) \phi_j(t) dt \quad i=1, \dots, M \quad j=1, \dots, N$$

uma vez que as funções base de valores reais $\phi_i(t)$, são *ortonormais*, ou seja,

$$\int_0^T \phi_i(t) \phi_j(t) dt = \begin{cases} 1 & \text{se } i=j \\ 0 & \text{se } i \neq j \end{cases}$$

Por conseguinte, cada sinal no conjunto $\{s_i(t)\}$ é completamente determinado pelos coeficientes, isto é:

$$s_i = \begin{bmatrix} s_{i1} \\ s_{i2} \\ \vdots \\ s_{iN} \end{bmatrix} \quad i=1,2,\dots,M \quad (4.3.9)$$

O vetor s_i é chamado de *vetor sinal*, que geometricamente pertence ao *espaço Euclidiano N-dimensional*. Assim, temos M pontos num espaço Euclidiano N-dimensional correspondentes aos M sinais. Estes sinais estão representados sobre o espaço Euclidiano N-dimensional gerado por $\phi_1, \dots, \phi_N$. A este espaço Euclidiano denomina-se *espaço de sinais*.

Quando $s_i(t)$ passa pelo canal aditivo temos

$$r(t) = n(t) + s_i(t) \quad (4.3.10)$$

Expressando cada um destes sinais (4.3.10) em função das bases $\phi_1, \dots, \phi_N$ temos

$$r_j = s_{ij} + n_j \quad i=1,\dots,M \quad j=1,\dots,N \quad (4.3.11)$$

onde n_j tem distribuição gaussiana com média 0 e variância $N_o/2$.

Como r_j está em função de n_j também terá uma distribuição gaussiana com média s_{ij} e variância $N_o/2$.

Seja $\mathbf{r}$ o vetor de N variáveis aleatórias r_j definido por:

$$\mathbf{r} = \begin{bmatrix} r_1 \\ \vdots \\ r_N \end{bmatrix}$$

Dado que os elementos do vetor $\mathbf{r}$ são estatisticamente independentes a função densidade de probabilidade condicional do vetor $\mathbf{r}$, dado o sinal $s_i(t)$ ou correspondente ao símbolo m_i transmitido pode ser expressa como o produto das funções densidade de cada um dos seus elementos individualmente

$$f_{\mathbf{r}}(\mathbf{r}|\mathbf{m}_i) = \prod_{j=1}^N f_{r_j}(r_j|\mathbf{m}_i) \quad i=1,\dots,M \quad (4.3.12)$$

onde os r_j 's são valores amostrais do sinal recebido $r(t)$. A função densidade condicional $f_{\mathbf{r}}(\mathbf{r}|\mathbf{m}_i)$, para cada mensagem transmitida m_i , $i = 1,\dots,M$ é chamada de *função de verossimilhança*. Esta função é a que caracteriza o canal. O canal cuja função de máxima verossimilhança satisfaz (4.3.12) é chamado um *canal sem memória*.

Como r_j é uma v.a. Gaussiana com média s_{ij} e variância $N_0/2$, temos que:

$$f_{r_i}(r_i|\mathbf{m}_i) = \frac{1}{\sqrt{\pi N_0}} \exp\left\{-\frac{1}{N_0}(r_j - s_{ij})^2\right\} \quad i=1,\dots,M \quad j=1,\dots,N \quad (4.3.13)$$

Pela substituição de (4.3.13) e (4.3.12), temos que a função de verossimilhança de um canal AWGN é dada por

$$f_{\mathbf{r}}(\mathbf{r}|\mathbf{m}_i) = (\pi N_0)^{-N/2} \exp\left\{-\frac{1}{N_0} \sum_{j=1}^N (r_j - s_{ij})^2\right\} \quad (4.3.14)$$

Assim, a distância entre s_i e $\mathbf{r}$ é

$$d^2(s_i, \mathbf{r}) = \|\mathbf{r} - s_i\|^2 \quad (4.3.15)$$

onde $\|\mathbf{a}\|$ é a norma do vetor $\mathbf{a}$.

Do item 3.5.1.3 temos, então, que o demodulador ótimo selecionará o

vetor sinal s_i mais próximo ao vetor recebido r . Para tal, o espaço de sinais deverá ser dividido em M regiões distintas $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_M$ correspondentes a cada mensagem $m_1, \dots, m_M$. Se r está na região $\mathcal{R}_k$ a decisão será m_k .

O problema de determinação do receptor ótimo então se reduz à escolha destas regiões limitadas, tais que a probabilidade de erro em tomar estas decisões seja mínima.

4.3.3.3. Probabilidade de erro em conjuntos de sinais

Consideraremos o caso quando os símbolos são igualmente prováveis. A probabilidade de erro do símbolo pode ser escrita como

$$\begin{aligned} p(e) &= 1 - p(c) \\ &= 1 - \frac{1}{M} \sum_{j=1}^M p(c|s_j) \end{aligned} \quad (4.3.16)$$

onde:

$p(c|s_j)$ é a probabilidade de decisão correta dado que o vetor sinal s_j , correspondente ao símbolo m_j , foi transmitido. Do final do item 4.3.3.2, temos que

$$p(e) = 1 - \frac{1}{M} \sum_{j=1}^M p(r \in \mathcal{R}_j | s_j). \quad (4.3.17)$$

Sob o ponto de vista geométrico, concluímos que a escolha das bases ortonormais $\{\phi_j\}_{j=1}^M$ não afetam a probabilidade de erro. Notamos que as regiões de decisão dependem só do conjunto de pontos sinais. Portanto, qualquer rotação ou translação de pontos de sinais não muda o valor de probabilidade de erro. Embora a probabilidade de erro seja invariante com respeito às translações do ponto de sinal, tal modificação altera a energia requerida para transmitir cada sinal.

A energia média requerida para transmitir um conjunto de sinais é definida como:

$$\mathcal{E} = \frac{1}{M} \sum_{i=1}^M |s_i|^2. \quad (4.3.18)$$

A equação (4.3.18) é precisamente a definição de momento de inércia em torno da origem para um conjunto de M pontos com massas iguais cujo centro de gravidade é a origem.

Matematicamente, esta afirmação pode ser expressa como

$$\sum_{i=1}^M s_i = 0 \quad (4.3.19)$$

Portanto, um conjunto de sinais satisfazendo (4.3.19) para uma dada probabilidade de erro requer energia média mínima para transmissão.

Para concluir estas considerações geométricas relacionadas com a equação (4.3.17) serão considerados conjuntos de sinais que apresentam uma *simetria completa*. Uma configuração de pontos de sinais que apresenta uma simetria completa é aquela em que suas regiões de decisão são congruentes. Nesta condição (4.3.17) transforma-se em

$$p(e) = 1 - p(r \in \mathcal{R}_j | s_j) \quad (4.3.20)$$

Para determinarmos a probabilidade de erro precisamos conhecer a função de probabilidade do vetor recebido r para um dado sinal transmitido s_j . Visto que o canal é AWGN e o vetor ruído n é independente do sinal s_j , a função densidade de probabilidade condicional do vetor recebido r com média s_j é dado por

$$f_{r|s_j}(\alpha | s_j) = f_n(\alpha - s_j) \quad (4.3.21)$$

Substituindo (4.3.21) em (4.3.17) obtém-se que:

$$p(e) = 1 - \frac{1}{M} \sum_{j=1}^M \int_{\mathcal{R}_j} f_n(\alpha - s_j) d\alpha \quad (4.3.22)$$

Este resultado é utilizado para calcular a probabilidade de erro para conjuntos de sinais.

• Sinais Binários

O caso geral dos sinais binários é mostrado na Figura 4.4a e o seu equivalente na Figura 4.4b.

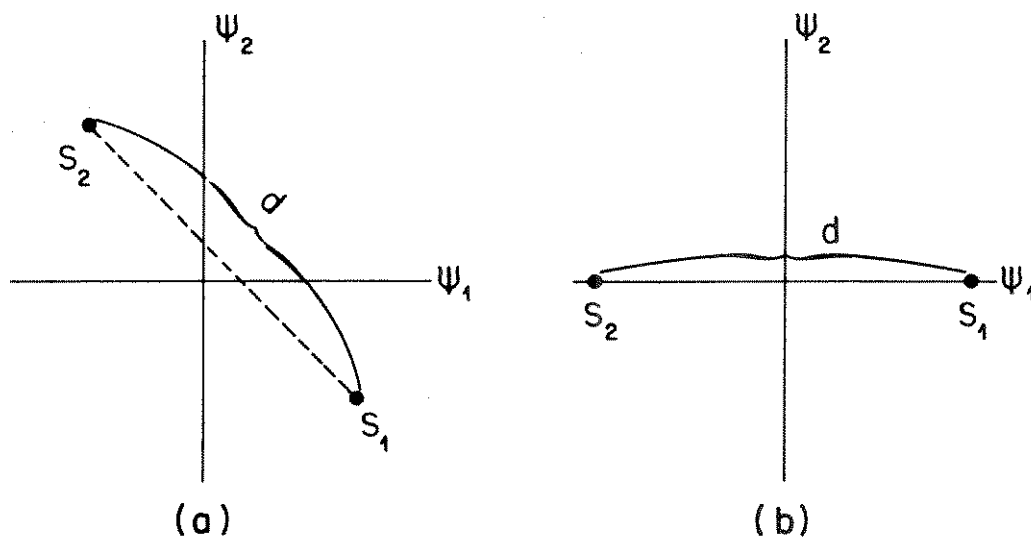


Figura 4.4 - Representação Geométrica de Sinais Binários:

(a) Caso geral;

(b) Sinal equivalente com diminuição de energia.

Dado que os sinais são completamente simétricos, utiliza-se (4.3.21) e a Figura 4.4b. Assim,

$$p(e) = 1 - p(r \in \mathcal{R}_1 | s_1) = p(r_1 < 0 | s_1)$$

$$p(e) = p\left(n_1 < -\frac{d}{2}\right) \quad (4.3.23)$$

onde:

$d \triangleq d(s_1, s_2)$ é a distância entre os sinais s_1 e s_2 .

Utilizando a função erro complementar, isto é,

$$\text{erfc}(x) \triangleq \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt$$

em (4.3.23), resulta que:

$$p(e) = \frac{1}{2} \text{erfc}\left(\frac{d}{2\sqrt{N_o}}\right) \quad (4.3.24)$$

• Sinais Simplex

É o conjunto de M sinais com as características de que cada sinal é equidistante dos demais e estão sobre uma esfera centrada na origem num espaço de sinais $(M - 1)$ -dimensional.

Pode-se observar que os sinais binários (Figura 4.4b) é um caso particular dos sinais simplex quando $M = 2$.

Muitas vezes a probabilidade de erro não pode ser calculada na forma fechada para sinais com $M \geq 3$ é por isso que se calcula a cota superior da probabilidade de erro, cujo cálculo é simplificado como veremos.

Nesta direção, um evento de erro e_{ij} ocorre se o vetor sinal transmitido s_i é transformado pelo ruído no vetor sinal recebido r que está mais próximo do vetor s_j diferente de s_i

$$e_{ij} \triangleq \left\{ r : \|r - s_j\| < \|r - s_i\| \mid s_i \right\}$$

Quando s_i é transmitido, o evento de erro é a união de todos os eventos e_{ij} . Assim,

$$p(e|s_i) = p\left(\bigcup_{\substack{j=1 \\ j \neq i}}^M e_{ij}\right) \leq \sum_{\substack{j=1 \\ j \neq i}}^M p(e_{ij}) \quad (4.3.25)$$

esta é a cota superior da probabilidade de erro condicional $p(e|s_i)$.

Observe que os eventos de erro dependem somente dos pares de sinais i e j , com $i \neq j$. Portanto, utilizando (4.3.24) e (4.3.25) resulta em

$$p(e|s_i) \leq \sum_{\substack{j=1 \\ j \neq i}}^M \frac{1}{2} \operatorname{erfc}\left(\frac{d_{ij}}{2\sqrt{N_o}}\right) \quad (4.3.26)$$

onde:

$d_{ij} \triangleq d(s_i, s_j)$ é a distância entre os sinais s_i e s_j .

Descondicionando (4.3.26), temos

$$\begin{aligned} P(e) &= \frac{1}{M} \sum_{i=1}^M P(e|s_i) \\ &\leq \frac{1}{M} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{1}{2} \operatorname{erfc}\left(\frac{d_{ij}}{2\sqrt{N_o}}\right) \end{aligned} \quad (4.3.27)$$

Mas como os sinais são do tipo simplex (4.3.27) reduz a

$$\begin{aligned} p(e) &\leq \frac{1}{2} \sum_{j=1}^{M-1} \operatorname{erfc}\left(\frac{d_j}{2\sqrt{N_o}}\right) \\ &= \frac{M-1}{2} \operatorname{erfc}\left(\frac{d_j}{2\sqrt{N_o}}\right) \end{aligned}$$

uma vez que a distância entre quaisquer dois sinais é a mesma.

• Sinais M-PSK

Em sistemas de comunicações M-ária um dos M possíveis sinais $s_1(t), \dots, s_M(t)$, é enviado a cada intervalo de tempo T segundos. As $M = 2^h$ diferentes formas de onda pelo mapeamento de h dígitos, são gerados pela mudança de amplitude, fase ou frequência de uma portadora em M passos discretos.

Um sinal M-PSK (Phase-Shift Keying) é caracterizado pela fase da portadora poder assumir um dos M possíveis valores, $\theta_i = \frac{2\pi i}{M}$, $i = 0, 1, \dots, M-1$. Por conseguinte, a cada intervalo de sinalização de duração T , um dos M possíveis sinais

$$s_i(t) = \sqrt{\frac{2\mathcal{E}}{T}} \cos\left(2\pi f_c t + \frac{2\pi i}{M}\right) \quad i=0,1,\dots,M-1$$

é enviado, onde $\mathcal{E}$ é a energia do sinal.

A frequência da portadora é dada por $f_c = \frac{h_c}{T}$, para algum inteiro fixo h_c . Cada sinal $s_i(t)$ pode ser expandido em termos de funções base $\phi_1(t)$ e $\phi_2(t)$ definidas como

$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t) \quad 0 \leq t \leq T$$

$$\phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t) \quad 0 \leq t \leq T$$

Ambos $\phi_1(t)$ e $\phi_2(t)$ tem energia unitária. Dessa forma, a constelação de sinais neste caso é bidimensional, com a característica de que as M mensagens estão igualmente espaçadas sobre uma circunferência de raio $\sqrt{\mathcal{E}}$ e centro na origem.

4.3.3.4. Modulação-codificada

Em 1982 Ungerboeck [27] propôs um esquema de codificação combinada com modulação para o caso binário extremamente eficiente tanto no aspecto de potência como de faixa.

Mostrou que é possível construir sistemas de comunicações com ganhos assintóticos de até 6dB para altas relações sinal-ruído (SNR) quando comparado com os esquemas convencionais de transmissão; sem a necessidade de aumentar a faixa ou sacrificar a taxa de informação, utilizando-se de um esquema novo de codificação e modulação como uma só entidade, a qual denomina-se Modulação-Codificada.

Este novo esquema de modulação-codificada é composto de um codificador de taxa $\frac{k}{k+1}$ responsável pela expansão da constelação de 2^k para 2^{k+1} elementos para o caso binário e de um mapeador que se utiliza de um processo denominado *mapeamento por partição de conjuntos*. Este mapeamento tem como objetivo básico prover um aumento na distância Euclideana entre sinais a serem transmitidos através do canal.

A finalidade da partição de conjuntos é a de rotular os pontos da constelação de sinais de modo a conseguir um aumento na distância entre pontos pertencentes a um mesmo sub-conjunto.

4.4. CASAMENTO DE CONJUNTO DE SINAIS COM GRUPOS

O problema que se considera no casamento entre conjunto de sinais e grupos é a correspondência entre linearidade e distância, quando a distância de Hamming é substituída pela distância Euclideana.

Massey e Mittelholzer [20]-[23] introduziram o casamento de códigos lineares sobre anéis $\mathbb{Z}_M$ e o conjunto de sinais M-PSK onde cada ponto desse conjunto pode ser identificado como um ponto sobre uma circunferência. Consequentemente, os M pontos podem ser representados geometricamente num plano complexo sob a forma $e^{j2\pi i/M}$, $i=0,\dots,M-1$ (ver Figura 4.5).

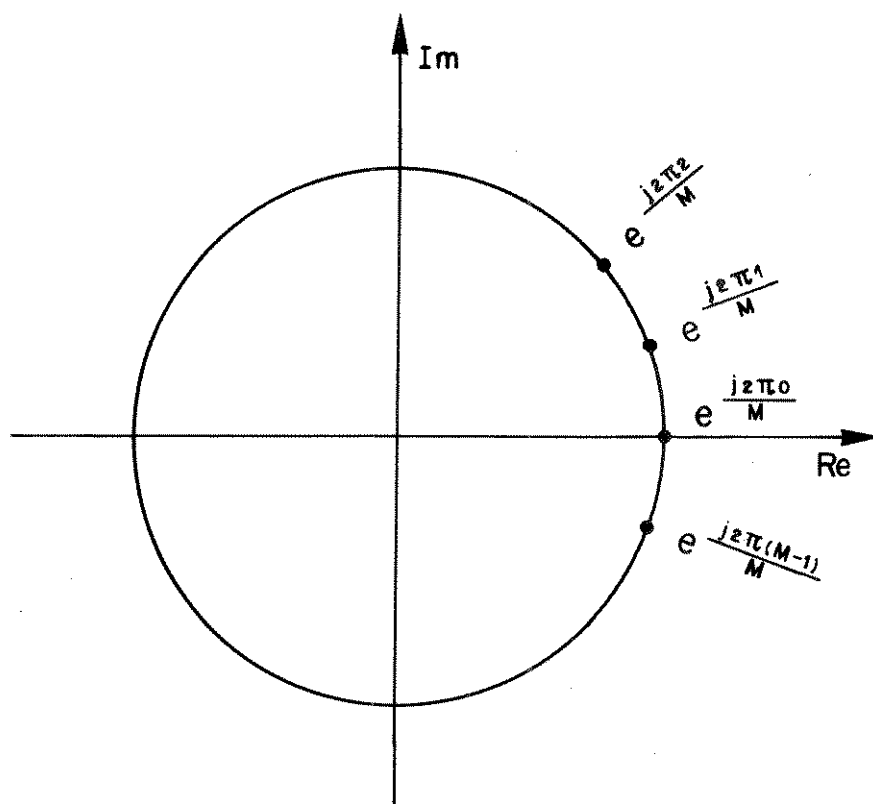


Figura 4.5 - Representação Geométrica do Conjunto de Sinais M-PSK.

A distância quadrática Euclideana entre quaisquer dois pontos de sinais, por exemplo h e i , é dada por:

$$d_E^2(h, i) = \left\| e^{\frac{j2\pi i}{M}} - e^{\frac{j2\pi h}{M}} \right\|^2 = \left\| 1 - e^{\frac{j2\pi(h-i)}{M}} \right\|^2 = \left\| 1 - e^{\frac{j2\pi(h \ominus i)}{M}} \right\|^2 \quad (4.4.1)$$

onde $\ominus$ é a diferença módulo M .

Deste resultado observa-se que a distância, então, depende somente da diferença módulo M.

Se h e i são elementos do anel de inteiros módulo M, $\mathbb{Z}_M$, define-se o *peso de fase* do elemento i por

$$W_p(i) = d_E^2(i, 0) = \left\| 1 - e^{\frac{j2\pi i}{M}} \right\|^2 \quad (4.4.2)$$

Considerando (4.4.2), a *distância de fase* entre os elementos h e i é definida como:

$$d_p(h, i) = W_p(h \ominus i) \quad (4.4.3)$$

De (4.4.1), (4.4.2) e (4.4.3) vemos que a distância Euclideana quadrática entre quaisquer dois elementos de $\mathbb{Z}_M$ é igual ao peso de sua diferença.

$$d_p(h, i) = d_E^2(h, i) = W_p(h \ominus i) \quad (4.4.4)$$

Para palavras codificadas de comprimento n com elementos em $\mathbb{Z}_M$, a distância de fase entre as mesmas, é a soma correspondente das distâncias de fase dígito a dígito. Então, a distância Euclideana quadrática entre os sinais correspondentes às palavras codificadas v_1 e v_2 com elementos em $\mathbb{Z}_M$ é:

$$\begin{aligned} d_p(v_1, v_2) &= \sum_{i=1}^n W_p(v_{1i} \ominus v_{2i}) \\ &= \sum_{i=1}^n \left\| 1 - e^{\frac{j2\pi(v_{1i} \ominus v_{2i})}{M}} \right\|^2 \\ &= W_p(v_1 \ominus v_2) \\ d_p(v_1, v_2) &= W_p(v_1 \ominus v_2) \end{aligned} \quad (4.4.5)$$

As palavras-código de comprimento n , cujos elementos pertencem à $\mathbb{Z}_M$, são mapeado no espaço Euclidiano $2n$ -dimensional. Por exemplo, o código de comprimento 1 está no espaço de sinais de dimensão 2.

Agora considerando o espaço Euclidiano $\mathbb{R}^3$, Massey [22] observou que o conjunto de sinais tridimensionais (Figura 4.6) é "casado" com o grupo $\mathbb{Z}_8$, no sentido de que a distância entre os vetores de sinais, depende somente da diferença dos elementos do grupo. Esta forma de construção de sinais no grupo $\mathbb{R}^3$ é possível, para conjuntos de sinais cujo número de elementos é par; tal conjunto de sinais são polítopos regulares e conhecidos em geometria como "anti-prismas" [28].

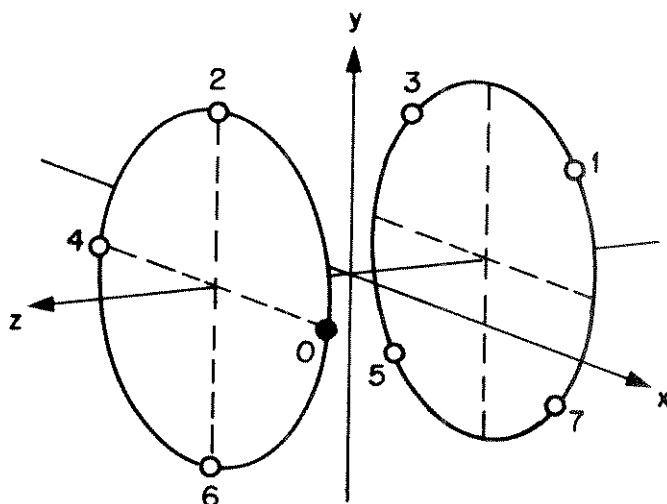


Figura 4.6 - Conjunto de Sinais Tridimensional Casado com o Grupo $\mathbb{Z}_8$.

A seguir apresentamos as definições de casamento de um conjunto de sinais a grupo e o peso de um elemento do grupo.

Definição 4.4.1. [24] : Um conjunto de sinais S é *casado* a um grupo G se existe um mapeamento sobrejetor $h : G \rightarrow S$ tal que para qualquer $g, g' \in G$

$$d(h(g), h(g')) = d(h(g^{-1} * g'), h(e))$$

onde e denota a unidade de G . Um mapeamento, h satisfazendo esta condição é chamado de *mapeamento casado*. Se h é um a um então h^{-1} é chamado *rotulamento casado*.

Definição 4.4.2. [24] : Se h é um mapeamento casado de um grupo G com um conjunto de sinais S , então o *peso* $W(g)$ de um elemento g de G é definido como a distância Euclideana quadrática entre $h(g)$ e $h(e)$, i.e.

$$W(g) = d^2(h(g), h(e)).$$

A Definição 4.4.1. não permite que muitos pontos do espaço de sinais estejam associados com o mesmo elemento do grupo; mas permite que muitos elementos do grupo estejam associados com o mesmo ponto do espaço de sinais.

Exemplos de casamento entre conjuntos de sinais e grupos temos:

- i) O grupo $\mathbb{Z}_M$ com o conjunto de sinais M-PSK;
- ii) O grupo $\mathbb{Z}_2$ com o conjunto de sinais binário.

Considerando-se a definição de casamento temos o seguinte lema.

Lema 4.4.1. [24] : Seja h um mapeamento casado de um grupo G com um conjunto de sinais S , seja s_e a imagem da unidade de G sob h , e seja H definida como $h^{-1}(s_e)$. Então H é um subgrupo de G , e $h(g) = h(g')$ se e somente se $gH = g'H$, i.e. se e somente se g e g' estão na mesma classe lateral à esquerda de H em G .

O Lema 4.4.1. implica que existe uma correspondência um a um entre pontos de sinais e classes laterais à esquerda de H em G . Além disso, pelo Teo-

rema 4.2.2., se H é normal em G o conjunto de classes laterais à esquerda G/H é um grupo ao qual S está casado.

Os seguintes teoremas caracterizam as mais importantes consequências imediatas da Definição 4.4.2.

Teorema 4.4.1. [24]: Se S é um conjunto de sinais em $\mathbb{R}^N$ que está casado a um grupo G e se f é uma transformação que preserva distâncias, i.e. uma isometria do $\mathbb{R}^N$, então $f(S)$ está casado a G .

A extensão de h para um mapeamento $G^n \rightarrow \mathbb{R}^{nN}$ (para um inteiro n positivo) é feita componente a componente.

Teorema 4.4.2. [24]: Seja h um mapeamento casado de um grupo G com um conjunto de sinais, e seja $\mathcal{C}$ um código linear sobre G . Então a extensão do conjunto de sinais $v(\mathcal{C})$ está casada com $\mathcal{C}$ e $c \rightarrow h(c)$ é um mapeamento casado.

4.5. CASAMENTO ENTRE CONJUNTO DE SINAIS SLEPIAN E GRUPOS

Slepian definiu uma classe de conjunto de sinais a qual denominou "*Group Codes for the Gaussian Channel*" [25]. Atualmente esta classe de conjuntos de sinais é denominada o "*Conjunto de Sinais Slepian*" [24].

Este conjunto de sinais está baseado num grupo de transformações ortogonais e, portanto, tem a propriedade de exibir uma forte simetria.

No que se segue, utilizaremos conceitos elementares da Teoria de Grupos desenvolvida na Seção 4.2.

Um conjunto de matrizes inversíveis de ordem n cujos elementos são números reais formam um grupo sob a multiplicação de matrizes. Cada matriz A , de ordem $n \times n$, do grupo determina uma transformação linear inversível

$$f_A : \mathbb{R}^N \longrightarrow \mathbb{R}^N \quad \text{definida por}$$

$$x \longrightarrow x A^t$$

onde:

$$x = (x_1, \dots, x_n) \in \mathbb{R}^N.$$

Agora, considerando esta definição

$$f_{AB}(x) = x(AB)^t = x B^t A^t = f_A(f_B(x))$$

da qual observa-se que o produto de matrizes AB determina a *composição de transformações lineares* $f_A \circ f_B$. Note que sob esta consideração as transformações lineares formam um grupo sob composição.

Reciprocamente, se $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ é uma transformação linear inversível, e se A é a matriz a qual é representada com respeito as bases canônicas em $\mathbb{R}^N$, então A é inversível e $f = f_A$. Por esta razão este grupo é chamado o *Grupo Linear Geral*, GL_n .

A coleção de *matrizes ortogonais* é um subgrupo de GL_n . Este subgrupo é chamado *grupo ortogonal* O_n . Aqueles elementos de O_n cujo determinante é igual a $+1$ formam um subgrupo de O_n chamado o *grupo ortogonal especial*, SO_n .

Se $A \in O_n$, a correspondente transformação linear f_A preserva distância e preserva ortogonalidade.

Seja S um subconjunto de $\mathbb{R}^n$. Então uma *isometria* de S é uma aplicação $f : S \longrightarrow S$ que preserva distâncias. O conjunto das isometrias de S forma um grupo sob composição que é chamado *grupo de simetrias* de S e será denotado por $\Gamma(S)$.

Seja S um conjunto e $f : S \rightarrow S$ uma isometria. Se Δ é o grupo de transformações do conjunto S e s é um elemento de S , a *órbita* de s sob Δ é o conjunto $\Delta(s) = \{f(s) : f \in \Delta\}$. O grupo de transformações Δ é chamado *transitivo* se $\Delta(s) = S$ para algum $s \in S$ (portanto, para todo $s \in S$).

As transformações ortogonais são isometrias em $\mathbb{R}^N$.

O seguinte lema é necessário para definir o conjunto de sinais de Slepian.

Lema 4.5.1. [24] : Seja S um subconjunto de $\mathbb{R}^N$ que satisfaz $\sum_{s \in S} s = 0$. Se S gera $\mathbb{R}^N$, então cada isometria f de S tem uma única extensão para uma transformação ortogonal de $\mathbb{R}^N$, i.e., existe uma única transformação ortogonal $T_f : \mathbb{R}^N \rightarrow \mathbb{R}^N$ tal que $T_f s = f(s) \quad \forall s \in S$ e o conjunto $\Gamma'(s) = \{T_f : f \in \Gamma(s)\}$ é um grupo sob composição que é isomorfo a $\Gamma(s)$.

Definição 4.5.3. [24] : Um conjunto de sinais Slepian em $\mathbb{R}^N$ é a órbita de um ponto de $\mathbb{R}^N$ sob um grupo finito de transformações ortogonais em $\mathbb{R}^N$.

Pela definição acima, o conjunto de sinais Slepian exibe:

- i) Um forte tipo de simetrias. Todos os pontos são completamente equivalentes exceto por suas localizações absolutas no espaço;
- ii) $\sum_{s \in S} s = 0$
- iii) Todos os pontos de um conjunto de sinais Slepian deverão estar sobre uma esfera em torno da origem. As distâncias entre os sinais são as mesmas e as regiões de máxima verossimilhança são congruentes e todas as palavras tem a mesma probabilidade de erro.

A violação da condição ii) significa perda de energia, e também a diminuição de uma dimensão.

Uma condição necessária e suficiente para que um conjunto de sinais do tipo Slepian esteja casado a um grupo é dada pelo corolário do seguinte teorema.

Teorema 4.5.2. [24]: Se Δ é um grupo transitivo de isometrias, então S está casado a Δ e, para qualquer $s \in S$, o mapeamento

$$\begin{aligned} h_s : \Delta &\longrightarrow S && \text{é um mapeamento casado.} \\ f &\longrightarrow f(s) \end{aligned}$$

Reciprocamente, se o conjunto de sinais S está casado a um grupo G , então existe um homomorfismo de G sobre o subgrupo transitivo de $\Gamma(S)$.

Do Teorema 4.5.2. e do Lema 4.5.1. temos o seguinte corolário.

Corolário 4.5.1. [24] : Um conjunto de sinais é casado a um grupo se e somente se o mesmo é uma translação do conjunto de sinais Slepian.

Prova :

$\Rightarrow$) Pelo Teorema 4.5.2. temos que o conjunto de sinais S está casado com um grupo G , então existe um homomorfismo de G sobre o subgrupo transitivo de $\Gamma(S)$ (simetrias). Um subgrupo transitivo de $\Gamma(S)$ são transformações ortogonais e por definição de subgrupo transitivo, os elementos deste subgrupo são órbitas para algum s . Logo é um conjunto de sinais Slepian.

⇐) Um conjunto de sinais Slepian por definição é um grupo finito de transformações ortogonais e pelo Lema 4.5.1. este grupo de transformações é isomorfo a um grupo de isometrias. Logo pelo Teorema 4.5.2. o conjunto de sinais Slepian está casado a um grupo. ■

CAPÍTULO 5

CODIFICAÇÃO E RECONHECIMENTO DE PADRÕES VIA ANÁLISE DE CORRESPONDÊNCIA

5.1. INTRODUÇÃO

No Capítulo 2 apresentamos a Análise de Correspondência, no Capítulo 3 o Reconhecimento de Padrões Matemático e na Seção 4.3.1 a Codificação de Fonte com o objetivo de serem utilizados neste capítulo.

Este capítulo pretende mostrar uma codificação eficiente e o reconhecimento de padrões via a Análise de Correspondência, até então uma técnica não utilizada na área de Teoria de Comunicações [30]–[31]. Mostramos que através da AC é possível obter um conjunto de sinais (vetor de características) a partir de palavras-código, resultado este que torna possível a codificação de Padrões. A codificação é feita tal que a probabilidade de erro seja mínima. Esta codificação obtém-se quando as palavras-código associadas a cada padrão provêm de códigos Reed-Muller de primeira ordem; código simplex; ou de códigos lineares sobre $\mathbb{Z}_q$, onde q é um primo. Note que esta codificação via a AC gera conjuntos de sinais simplex no espaço Euclidiano Ponderado (espaço de características).

Utilizando as definições de classe lateral e perfil suplementar (conceito da AC) é possível codificar outros padrões que pertencem a uma determinada classe padrão.

Como o canal de transmissão é caracterizado por um ruído do tipo Gaussiano, o processo de classificação dos padrões desconhecidos estão baseados na distribuição Gaussiana. Os classificadores são então determinados pelos protótipos de cada classe padrão.

Assim, sendo necessária a geração de códigos lineares, na Seção 5.2 damos uma breve revisão de corpos finitos por ser base para a geração de códigos. Na Seção 5.3 mostramos o processo de geração dos códigos lineares. Na Seção 5.4 mostramos a codificação de padrões via AC e finalmente, na Seção 5.5 a codificação de classes de padrões baseado nas definições de perfil suplementar e classes laterais.

5.2. ELEMENTOS DA TEORIA DE GALOIS PARA CORPOS FINITOS

Nesta seção faremos uma breve revisão de conceitos da Teoria de Galois para corpos finitos uma vez que as extensões destes corpos formam a base para a geração de códigos. Algumas definições e teoremas apresentados na Seção 4.2 para o anel de inteiros serão utilizados para o anel de polinômios.

5.2.1. Anel de Polinômios

Se R é um anel. Um polinômio $p(x)$, sobre o anel R é uma expressão da forma

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

onde:

$a_0, a_1, \dots, a_n \in R$ e $n \in \mathbb{N}$. O elemento a_i é chamado o *coeficiente* de x^i em $p(x)$. Se n é o maior inteiro para o qual $a_n \neq 0$, $p(x)$ tem *grau* n e denota-se $\text{grau}(p(x)) = n$. Se todos os coeficientes de $p(x)$ são zero, então $p(x)$ é chamado de *polinômio Zero*. O polinômio zero e os polinômios de grau 0 são chamados *polinômios constantes* porque eles não contem o termo x .

Definição 5.2.1. : O conjunto de todos os polinômios em x com coeficientes sobre o anel comutativo R é denotado por $R[x]$, e é definido por

$$R[x] = \left\{ a_0 + a_1x + a_2x^2 + \dots + a_nx^n \mid a_i \in R, n \in \mathbb{N} \right\}$$

A terna $[R[x], +, \cdot]$ é chamada de *anel de polinômios com coeficientes em R* , se a adição e multiplicação de dois polinômios

$$p(x) = \sum_{i=0}^n a_i x^i \quad \text{e} \quad q(x) = \sum_{i=0}^m b_i x^i$$

são definidos por

$$p(x) + q(x) = \sum_{i=0}^{\max(m,n)} (a_i + b_i) x^i$$

e

$$p(x).q(x) = \sum_{k=0} c_k x^k$$

onde

$$c_k = \sum_{i+j=k} a_i b_j$$

Teorema 5.2.1. (*Algoritmo de Divisão para Polinômios*) : Seja $f(x)$ e $g(x)$ elementos do anel de polinômios $F[x]$, com coeficientes no corpo F . Se $g(x)$ não é o polinômio zero, existe um único polinômio $q(x)$, $r(x) \in F[x]$, tal que

$$f(x) = q(x) g(x) + r(x)$$

onde:

$r(x)$ é o polinômio zero ou $\text{grau}(r(x)) < \text{grau}(g(x))$.

Um polinômio $f(x)$ de grau positivo é dito ser *reduzível* sobre $F[x]$, se o mesmo pode ser fatorado em dois polinômios de grau positivo sobre $F[x]$. Caso contrário é *irreduzível*.

Teorema 5.2.2.: $F[x]$ é um anel ideal principal se F é um corpo.

Se F é corpo, o anel quociente do anel de polinômios $F[x]$ forma uma classe importante de anéis que são utilizadas para construir novos corpos.

Pelo Teorema 5.2.2. $F[x]$ é um anel ideal, assim qualquer anel quociente é da forma $F[x]/(p(x))$ para algum polinômio $p(x) \in F[x]$.

Os elementos $F[x]/(p(x))$ formam classes de equivalência sob a relação em $F[x]$ definida por

$$f(x) \equiv g(x) \pmod{(p(x))} \iff f(x) - g(x) \in (p(x))$$

Cada classe lateral de $F[x]$ por $(p(x))$ contém um polinômio zero ou um polinômio de grau menor que $p(x)$.

Teorema 5.2.3.: Seja P o ideal $(p(x))$, gerado pelo polinômio $p(x)$ de grau $n > 0$. Os elementos diferentes de $F[x]/(p(x))$ são da forma

$$P + a_0 + a_1x + \dots + a_{n-1}x^{n-1}; \quad a_0, a_1, \dots, a_{n-1} \in F$$

Teorema 5.2.4.: O anel $F[x]/p(x)$ é um corpo $\iff (p(x))$ é irredutível sobre o corpo F . Além disso, o anel $F[x]/(p(x))$ sempre contém um subanel isomorfo ao corpo F .

O grau da extensão K de um corpo F , denotado por $[K : F]$ é a dimensão de K como um espaço vetorial sobre F . K é chamado uma *extensão finita* se $[K : F]$ é finita.

Exemplo 5.2.1.: Se $K = \mathbb{Z}_5[x]/(x^3 + x + 1)$, então $[K : \mathbb{Z}_5] = 3$. Isto porque $\{1, x, x^2\}$ é uma base para K sobre $\mathbb{Z}_5$, e pelo Teorema 5.2.3. cada elemento de K pode ser escrito unicamente como a classe lateral contendo $a_0 + a_1x + a_2x^2$, onde $a_i \in \mathbb{Z}_5$. Portanto, $[K : \mathbb{Z}_5] = 3$.

Seja K um corpo estendido de F e seja $a \in K$. O *subcorpo de* K contendo F e a é denotado por $F(a)$. $F(a)$ é chamado o corpo obtido através da adjunção de a em F .

Se K é um corpo estendido de F , o elemento $k \in K$ é chamado *Algébrico* sobre F se existe $a_0, a_1, \dots, a_n \in F$, não todos nulos, tal que

$$a_0 + a_1 k + \dots + a_n k^n = 0$$

Elementos que não são algébricos sobre F são chamados *transcedentais* sobre F .

Teorema 5.2.5.: Seja α algébrico sobre F e seja $p(x)$ um polinômio irredutível de grau n sobre F com α uma raiz. Então

$$F(\alpha) \cong F[x]/(p(x)).$$

Corolário 5.2.1.: Se α é uma raiz do polinômio $p(x)$ de grau n , irredutível sobre F , então $[F(\alpha) : F] = n$.

Lema 5.2.1.: Seja $p(x)$ um polinômio irredutível sobre o corpo F . Então F tem um corpo de extensão finita K no qual $p(x)$ tem uma raiz.

5.2.1.1. Corpos de Galois

Definição 5.2.2.: Se para um anel R existe um inteiro positivo n tal que $n \cdot a = 0$ para todo $a \in R$, então o menor inteiro positivo é a *característica do anel* R .

Proposição 5.2.1.: Se o corpo F tem característica prima p , então F contém um subcorpo isomorfo a $\mathbb{Z}_p$.

Teorema 5.2.6.: Se F é um corpo finito, este tem p^m elementos para algum primo p e algum inteiro m .

Definição 5.2.3.: Um corpo finito com p^m elementos é chamado um *Corpo de Galois* de ordem p^m e é denotado por $GF(p^m)$.

Pelo Teorema 5.2.6. $GF(p^m)$ é um corpo de extensão de $\mathbb{Z}_p$ de grau m . Pelo Teorema 5.2.4. cada corpo finito $GF(p^m)$ pode ser construído através de um polinômio $q(x)$ de grau m , irredutível em $\mathbb{Z}_p[x]$, e definido como

$$GF(p^m) = \mathbb{Z}_p[x]/(q(x))$$

Pelo Lema 5.2.1. existe um elemento x em $GF(p^m)$, tal que $q(x) = 0$, e $GF(p^m) = \mathbb{Z}_p(\alpha)$, o corpo obtido por adjunção de α em $\mathbb{Z}_p$.

Exemplo 5.2.3.: $GF(4) = \mathbb{Z}_2[x]/(x^2+x+1) = \mathbb{Z}_2(\alpha) = \{0, 1, \alpha, \alpha+1\}$,
onde $\alpha^2 = \alpha+1$

Os elementos de um corpo de Galois $GF(p^m)$ são escritos como

$$\left\{ a_0 + a_1\alpha + \dots + a_{m-1}\alpha^{m-1} \mid a_i \in \mathbb{Z}_p \right\}$$

onde α é uma raiz de um polinômio irredutível $q(x)$ de grau m sobre $\mathbb{Z}_p$.

A adição é facilmente executada utilizando esta representação porque é soma de polinômios em $\mathbb{Z}_p[\alpha]$. Entretanto, a multiplicação é mais complicada e requer repetido uso da relação $q(\alpha) = 0$.

Através de uma escolha apropriada de α , os elementos de $\text{GF}(p^m)$ podem ser escritos como

$$\left\{0, 1, \alpha, \alpha^2, \alpha^3, \dots, \alpha^{p^m-2}\right\} \quad \text{onde} \quad \alpha^{p^m-1} = 1$$

O elemento α é chamado um elemento primitivo de $\text{GF}(p^m)$, e a multiplicação é facilmente calculada utilizando a potência de α .

Teorema 5.2.8.: Seja $\text{GF}(q)^*$ o conjunto de elementos não nulos no corpo de Galois $\text{GF}(q)$. Então $\text{GF}(q)^*$ é um grupo cíclico de ordem $q - 1$.

Um gerador de um grupo cíclico $(\text{GF}(q)^*, \circ)$ é chamado um *elemento primitivo* de $\text{GF}(q)$.

Baseados nesta teoria passamos a gerar códigos com elementos em $\mathbb{Z}_p$.

5.3. GERAÇÃO DE CÓDIGOS COM ELEMENTOS EM $\mathbb{Z}_p$

A geração das palavras código sobre $\mathbb{Z}_p$, estão baseadas no corpo de Galois $\text{GF}(p^m)$, onde m é o grau do polinômio $p(x)$ irredutível em $\mathbb{Z}_p[x]$.

Seja $\mathbb{Z}_p = \{0, 1, \dots, p-1\}$ um anel comutativo módulo p . Seja $\mathbb{Z}_p^n$ o conjunto dos n -uplas $(x_1, x_2, x_3, \dots, x_n)$; $x_i \in \mathbb{Z}_p$.

Definição 5.3.1. Um código de bloco linear de comprimento n sobre $\mathbb{Z}_p$ é um sub-anel de $\mathbb{Z}_p^n$, onde $\mathbb{Z}_p^n$ é um anel de n -uplas de elementos em $\mathbb{Z}_p$ com a operação adição módulo p .

Para a geração do código, primeiramente define-se o corpo de Galois $GF(p^m)$, em função de polinômios em $\mathbb{Z}_p[x]$ e elemento primitivo α , i.é..

i) Polinômios em $\mathbb{Z}_p[\alpha]$

$$GF(p^m) = \frac{\mathbb{Z}_p[x]}{p(x)} = \mathbb{Z}_p[\alpha] = \left\{ a_0 + a_1\alpha + \dots + a_{m-1}\alpha^{m-1} \mid a_i \in \mathbb{Z}_p \right\}$$

onde:

- $\mathbb{Z}_p[x]$ é o anel de polinômios na variável indeterminada x com coeficientes em $\mathbb{Z}_p$;
- $p(x)$ é o polinômio irredutível em $\mathbb{Z}_p[x]$ de grau m ;
- $\mathbb{Z}_p[\alpha]$ é o anel de polinômios na variável indeterminada α ;
- α é uma raiz de $p(\alpha)$ em $\mathbb{Z}_p[\alpha]$.

ii) Elemento Primitivo

Os elementos de $GF(p^m)$, com ajuda de $p(\alpha)$ podem ser escritos em função do elemento primitivo α como: $\left\{ 0, 1, \alpha, \dots, \alpha^{p^m-2} \right\}$ onde $\alpha^{p^m-1} = 1$.

A dimensão de $GF(p^m)$ visto como espaço vetorial com coeficiente em $\mathbb{Z}_p$ é m (corolário 5.2.1.).

Cada polinômio de $\mathbb{Z}_p[\alpha]$ pode ser expressado como uma m -upla. Assim,

$$a_0\alpha + a_1\alpha^2 + \dots + a_{m-1}\alpha^{m-1}$$

é representado pela m -upla $(a_0, a_1, \dots, a_{m-1})$.

Com esta representação dos elementos de $\mathbb{Z}_p[\alpha]$, pode ser construída a matriz geradora G de ordem $m \times n$, onde as colunas de G são os elementos de $\mathbb{Z}_p[\alpha]$ representados por m -uplas. Em outras palavras cada m -upla de $\mathbb{Z}_p[\alpha]$ é uma coluna de G . Isto é,

$$G = \left\{ g_0, g_1, \dots, g_{n-1} \right\}$$

onde:

g_i é um vetor coluna que representa um polinômio de $\mathbb{Z}_p[\alpha]$.

O conjunto de palavras codificadas serão todas as combinações das linhas de G .

Os parâmetros deste código são:

- comprimento do código $n = p^m$
- dimensão do código $k = m$
- distância mínima $d_{\min} = p^m - p^{m-1}$
- número de palavras codificadas p^m

5.3.1. Geração de Códigos Lineares Binários

Através de um exemplo, mostraremos a geração de códigos lineares binários, baseados no corpo $GF(2^m)$ onde m é o grau do polinômio irredutível em $\mathbb{Z}_2[x]$.

Exemplo 5.3.1.: Geração de um código de parâmetros $(n=2^m, k=m, d_H=2^m - 2^{m-1}) = (n=8, k=3, d_H=4)$. O código está baseado no corpo de Galois $GF(2^3)$. Seja $p(x) = x^3 + x + 1$ o polinômio irredutível em $\mathbb{Z}_2[x]$ então

$$\text{GF}(2^3) = \frac{\mathbb{Z}_2[x]}{p(x)} = \mathbb{Z}_2[\alpha] = \left\{ a_0 + a_1\alpha + a_2\alpha^2 \mid a_i \in \mathbb{Z}_2 \right\}$$

$$p(\alpha) = 0 \text{ em } \mathbb{Z}_2[\alpha], \text{ logo } \alpha^3 = \alpha + 1$$

Na Tabela 5.3.1 mostramos os elementos de $\text{GF}(2^3)$ na sua forma potência do elemento primitivo, polinomial e vetorial.

TABELA 5.3.1

$\text{GF}(2^3)$ com $\alpha^3 + \alpha + 1 = 0$

Forma Potência do Elemento Primitivo	Forma Polinomial	Forma Vetorial
0	0	0 0 0
1	1	0 0 1
α	α	0 1 0
α^2	α^2	1 0 0
α^3	$\alpha + 1$	0 1 1
α^4	$\alpha + \alpha^2$	1 1 0
α^5	$1 + \alpha + \alpha^2$	1 1 1
α^6	$1 + \alpha^2$	1 0 1

Arranjaremos os elementos de $\mathbb{Z}_2[\alpha]$ da seguinte maneira

$$\{0, 1, \alpha, \alpha+1, \alpha^2, \alpha^2+1, \alpha^2+\alpha, \alpha^2+\alpha+1\}$$

A matriz geradora G é construída com os elementos de $\mathbb{Z}_2[\alpha]$ expressados na sua forma vetorial, isto é:

$$G = \begin{bmatrix} 0 & 1 & \alpha & \alpha+1 & \alpha^2 & \alpha^2+1 & \alpha+\alpha^2 & \alpha^2+\alpha+1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Pela Definição 5.3.1. um código linear é um subanel. Então o código é composto

por todas as possíveis combinações lineares das linhas de G .

As palavras codificadas são:

$$\begin{aligned}
 v_0 &= 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\
 v_1 &= 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1 \\
 v_2 &= 0\ 0\ 1\ 1\ 0\ 0\ 1\ 1 \\
 v_3 &= 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1 \\
 v_1 \oplus v_2 &= 0\ 0\ 1\ 1\ 1\ 1\ 0\ 0 \\
 v_1 \oplus v_3 &= 0\ 1\ 0\ 1\ 1\ 0\ 1\ 0 \\
 v_2 \oplus v_3 &= 0\ 1\ 1\ 0\ 0\ 1\ 1\ 0 \\
 v_1 \oplus v_2 \oplus v_3 &= 0\ 1\ 1\ 0\ 1\ 0\ 0\ 1
 \end{aligned}$$

Os códigos binários gerados desta maneira podem ser considerados como subcódigos dos códigos Reed-Muller de primeira ordem como veremos posteriormente

5.3.1.1. Códigos Reed-Muller de primeira ordem

Os códigos Reed-Muller de primeira ordem denotados por $\mathcal{R}(1, m)$, onde $m \in \mathbb{Z}^+$, podem ser obtidos de diferentes formas, uma delas é através das funções booleanas [29]. Uma das características deste código é que todas as palavras código com exceção de 0 e 1 tem peso 2^{m-1} .

Os parâmetros deste código são:

- comprimento da palavras código $n = p^m$
- dimensão do código $k = 1 + m$
- distância mínima $d_{\min} = 2^{m-1} = 2^m - 2^{m-1}$

Comparando os parâmetros do código Reed-Muller de primeira ordem e do código gerado anteriormente, vemos que o comprimento da palavra-código e a distância mínima são iguais. A diferença está na dimensão dos códigos; o código $\mathcal{R}(1, m)$ com respeito ao gerado anteriormente varia numa unidade a mais. Isto se deve a que a matriz geradora de $\mathcal{R}(1, m)$ contém a linha com elementos todos um, a qual aumenta a cardinalidade do código.

Logo, o código gerado anteriormente pode ser considerado como um sub-código do código $\mathcal{R}(1, m)$.

5.3.2. Geração de Códigos p-ários

O processo da geração de códigos p-ários é o mesmo que aquele apresentado na Seção 5.3.1.

Também, através de um exemplo mostraremos esta geração.

Exemplo 5.3.2.: Geração do código de parâmetros, $m = 2$

$$(n=3^m, \quad k=m, \quad d_H=3^m-3^{m-1}) = (n=9, \quad k=2, \quad d_H=6)$$

O corpo de Galois é $GF(3^2)$, e o polinômio $p(x) = x^2 + x + 2$ é irredutível em $\mathbb{Z}_3[x]$

$$GF(3^2) = \frac{\mathbb{Z}_3[x]}{p(x)} = \mathbb{Z}_3[\alpha] = \left\{ a_0 + a_1 \alpha \mid a_i \in \mathbb{Z}_3 \right\}$$

$$p(\alpha) = 0 \quad \text{em} \quad \mathbb{Z}_3[\alpha], \quad \text{logo} \quad \alpha^2 = -\alpha - 2 = 2\alpha + 1.$$

Na Tabela 5.3.2 mostramos os elementos de $GF(3^2)$ em suas formas diferentes.

TABELA 5.3.2

 $GF(3^2)$ com $\alpha^2 + \alpha + 2 = 0$

Potência do Elemento Primitivo	Polinomial	Vetorial 2-upla
0	0	00
1	1	10
α	α	01
α^2	$1+2\alpha$	12
α^3	$2+2\alpha$	22
α^4	2	20
α^5	2α	02
α^6	$2+\alpha$	21
α^7	$1+\alpha$	11

Os elementos de $\mathbb{Z}_2[\alpha]$ são arranjados da seguinte maneira:

$$GF(3^2) = \{0, 1, 2, \alpha, 1+\alpha, 2+\alpha, 2\alpha, 1+2\alpha, 2+2\alpha\}$$

e a matriz geradora de G é:

$$G = \begin{bmatrix} 0 & 1 & 2 & \alpha & 1+\alpha & 2+\alpha & 2\alpha & 1+2\alpha & 2+2\alpha \\ 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \end{bmatrix}$$

as palavras-código são:

$$\begin{aligned} v_0 &= 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \\ v_1 &= 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 2 \ 2 \ 2 \\ v_2 &= 0 \ 1 \ 2 \ 0 \ 1 \ 2 \ 0 \ 1 \ 2 \\ 2v_1 &= 0 \ 0 \ 0 \ 2 \ 2 \ 2 \ 1 \ 1 \ 1 \\ 2v_2 &= 0 \ 2 \ 1 \ 0 \ 2 \ 1 \ 0 \ 2 \ 1 \\ v_1 \oplus v_2 &= 0 \ 1 \ 2 \ 1 \ 2 \ 0 \ 2 \ 0 \ 1 \end{aligned}$$

$$2v_1 \oplus v_2 = 0 \ 1 \ 2 \ 2 \ 0 \ 1 \ 1 \ 2 \ 0$$

$$2v_2 \oplus v_1 = 0 \ 2 \ 1 \ 1 \ 0 \ 2 \ 2 \ 1 \ 0$$

$$2v_2 \oplus 2v_1 = 0 \ 2 \ 1 \ 2 \ 1 \ 0 \ 1 \ 0 \ 2$$

Deste modo podemos gerar códigos sobre corpos de Galois.

Tomando como base a Teoria de Códigos e Análise de Correspondência (Capítulo 2), introduziremos a Codificação de Padrões Via Análise de Correspondência.

5.4. CODIFICAÇÃO DE PADRÕES VIA ANÁLISE DE CORRESPONDÊNCIA

Agora, a questão é: como a AC pode ser utilizada em reconhecimento de padrões?

Sob o ponto de vista de codificação de padrões temos os seguintes passos; 1) as propriedades da AC propiciam uma flexibilidade para fazermos os melhores arranjos ou distribuições de zeros e uns na "tabela de Contingência" e assim, obter os melhores arranjos dos vetores sinais no espaço de sinais pela aplicação da AC sobre a TC; 2) uma vez obtida a melhor constelação de sinais, as palavras-código que gerarão estes sinais são determinadas; 3) uma vez de posse destas palavras-código os padrões são codificados.

5.4.1. Construção da "Tabela de Contingência" e Obtenção dos Vetores Sinais

Para construir a TC devemos considerar em forma conjunta o conceito

de código, propriedades de codificação de fonte, e de AC. Para isso, faremos três observações:

- 1) A palavra código 0 cujo peso de Hamming é zero, não tem significado na AC, porque a "massa" é zero; portanto, não contribui na geração do espaço de sinais e não é incluída na TC;
- 2) A distância χ^2 utilizada na AC é indefinida quando pelo menos um elemento do vetor centróide é zero; portanto, se a j -ésima componente de todas as palavras são zero, estas devem ser eliminadas para evitar a divisão por zero;
- 3) Dado que o objetivo na codificação de fonte é eliminar a redundância das palavras, então as componentes redundantes serão eliminadas, isto é, se todas as palavras possuem a mesma k -ésima componente, então esta componente é eliminada. As componentes toda zero, referidas no item 2), também são consideradas redundantes.

Destas observações, concluímos que não pode existir nenhuma linha ou coluna composta por zeros. Como o objetivo é codificar os padrões com um código, mostraremos que as linhas ou colunas da TC deve conter os elementos que provém de uma classe lateral para que assim, depois de transladadas, tenhamos o código.

Seja $\mathcal{C}$ o código composto por N palavras e $\mathcal{C}'$ um subcódigo de $\mathcal{C}$ composto por M palavras. Pelo Teorema de Lagrange (Teorema 4.2.1), o número de classes laterais $\mathcal{C}' + v_j$, é dado por N/M , onde N/M significa que M divide N para $j = 1, 2, \dots, N/M$. A classe lateral utilizada para construir a TC será qualquer $\mathcal{C}' + v_i$, tal que $v_i \neq 0$.

Seja M o número de padrões a serem codificados, e sejam $v_1, v_2, \dots, v_M$ as palavras-código de comprimento n de uma classe lateral de um código linear.

Considerando as três observações denotaremos as novas palavras-código por v'_i , $i = 1, \dots, M$ de comprimento $n' < n$, onde $(n - n')$ é o número de elementos redundantes. Logo, a "Tabela de Contingência" (Tabela 5.4.1) é construída segundo duas variáveis. Nas linhas a variável categórica *Padrão* e nas colunas a variável categórica *Posição do Dígitto* em cada dimensão, são associadas, respectivamente.

TABELA 5.4.1
"Tabela de Contingência" de Duas Entradas

Posição do dígito em cada dimensão Padrão	1	2	3	...	n'
	v'_{11}	v'_{12}	v'_{13}	...	$v'_{1n'}$
v'_2	v'_{21}	v'_{22}	v'_{23}	...	$v'_{2n'}$
...	...	...	...	...	...
v'_M	v'_{M1}	v'_{M2}	v'_{M3}	...	$v'_{Mn'}$

Uma vez obtida a "Tabela de Contingência", aplica-se o algoritmo da AC (Seção 2.4), através do qual obteremos os subespaços ótimos e as coordenadas perfis destas palavras, isto é, os vetores sinais.

A quantidade de energia média do conjunto de sinais associada a cada subespaço é mostrada na Tabela 2.5.1 (a qual reproduzimos para facilidade do leitor). Note também que a distribuição da energia média total também está presente. A última coluna da tabela mostra a soma da energia média de cada vetor sinal com respeito aos eixos, e na interseção da última linha com a última coluna, temos a energia média total do conjunto de sinais.

A interpretação dos perfis das palavras-código baseia-se em dois conjuntos de coeficientes.

A primeira é a contribuição absoluta, esta contribuição sob o ponto de vista de sinais mede a proporção da contribuição do i -ésimo sinal para a energia em cada eixo (ver Seção 2.5.1). O segundo conjunto de coeficientes é a contribuição relativa que mostra a correlação de cada um dos perfis com cada um dos subespaços (ver Seção 2.5.2).

Nas seguintes subseções mostraremos como otimizar a codificação de padrões (segundo o critério de mínima probabilidade de erro) via o balanceamento da TC através da AC resultando na geração de sinais simplex como a "configuração" dos pontos no espaço de sinais.

5.4.2. Obtenção de Conjuntos de Sinais Simplex

A geração de conjuntos de sinais simplex depende do arranjo das palavras feitas na "Tabela de Contingência". O conjunto das M palavras-código associado aos M padrões deve possuir características do conjunto de sinais simplex, isto é, que a distância de Hamming entre as palavras-código sejam iguais e o peso de Hamming seja igual para todas elas. Além disso, vistas as palavras-código (elementos de uma cada classe lateral) como vetores num espaço Euclidiano M -dimensional eles devem ser Linearmente Independentes (LI) neste espaço.

Uma vez construída a TC com estas características aplica-se a AC e obtém-se o conjunto de sinais cujas coordenadas s_i , são vértices de um poliedro e os quais estão sobre uma esfera de raio $\|s_i\| = \sqrt{s_i^T s_i} = 1$ centrada na origem, num espaço Euclidiano ponderado de dimensão $M - 1$. Os pontos sobre a esfera são eqüidistantes entre si, e também com respeito ao centróide. Estas caracte-

rísticas advém do balanceamento de linhas e colunas da TC via a AC.

Dos Teoremas 2.3.1. e 2.3.2. vemos que a energia média total do conjunto de sinais está num espaço Euclidiano de dimensão $M - 1$. Isto porque a matriz $(\mathbf{R} - \mathbf{1} \underline{\mathbf{c}}^T)$ (2.4.1) tem posto $M - 1$.

Finalmente, os M padrões são codificados e os sinais associados estão sobre uma esfera centrada na origem uniformemente distribuídos num espaço Euclidiano ponderado de dimensão $M - 1$.

5.4.2.1. Obtenção de conjuntos de sinais simplex através de códigos binários quando o número de sinais é 2^n , $n \in \mathbb{Z}^+$

Dentro da classe de códigos binários, os códigos que satisfazem as condições para gerar o conjunto de sinais simplex são os códigos Reed-Muller de primeira ordem e Simplex.

A seguir apresentamos através de um exemplo, o processo de codificação de quatro padrões. Para em seguida, generalizar os conjuntos de sinais cujo número de elementos é da forma 2^n , $n \in \mathbb{Z}^+$.

Exemplo 5.4.1.: Seja $\mathcal{C}_0$ um subcódigo do código Reed-Muller $\mathcal{R}(1, 3)$ com parâmetros $(n=2^3, k=3+1, d_H=2^2)$, cujas palavras, mostramos no Exemplo 5.3.1.

O Código $\mathcal{C}$ obtido de $\mathcal{C}_0$ é um código de parâmetros $(n = 2^3-1, k = 3, d_H = 2^2)$. Com respeito a $\mathcal{R}(1, 3)$, a dimensão do código diminui de 4 para 3 porque não foi considerada a palavra composta por uns devido ao balanceamento. A distância de Hamming é a mesma. O comprimento da palavra diminui de 8 para 7, porque foi eliminada a componente redundante "0" em cada palavra.

Sejam $\mathcal{H}$ e $\mathcal{H}'$ dois subcódigos de

$$\mathcal{C} = \{v_0, v_1, v_2, v_3, v_1 \oplus v_2, v_1 \oplus v_3, v_2 \oplus v_3, v_1 \oplus v_2 \oplus v_3\}$$

definidos como:

$$\mathcal{H} = \{v_0, v_2, v_3, v_2 \oplus v_3\} \quad \text{e} \quad \mathcal{H}' = \{v_0, v_1\}$$

As classes laterais deste código em $\mathcal{C}$, apresentamos na Tabela 5.4.2.

TABELA 5.4.2

Classes Laterais

Classes Laterais do Código $\mathcal{H}$	Classes Laterais do Código $\mathcal{H}'$				
		$\mathcal{H}' \oplus v_0$	$\mathcal{H}' \oplus v_2$	$\mathcal{H}' \oplus v_3$	$\mathcal{H}' \oplus v_2 \oplus v_3$
$\mathcal{H} \oplus v_0$		v_0	v_2	v_3	$v_2 \oplus v_3$
$\mathcal{H} \oplus v_1$		v_1	$v_2 \oplus v_1$	$v_3 \oplus v_1$	$v_1 \oplus v_2 \oplus v_3$

Com os elementos da classe lateral $\mathcal{H} \oplus v_1$ construiremos a "Tabela de Contingência"; antes verificaremos se os elementos desta classe lateral vistos como vetores num espaço Euclidiano 4-dimensional são linearmente independentes. Por definição de vetores linearmente independentes, os coeficientes da combinação linear destas palavras codificadas devem ser zero, isto é,

$$\beta_1 v_1 + \beta_2 (v_1 + v_2) + \beta_3 (v_1 + v_3) + \beta_4 (v_1 + v_2 + v_3) = 0$$

se e somente se

$$\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

Substituindo os correspondentes valores dos v_i 's, temos

$$\beta_1(000111) + \beta_2(011100) + \beta_3(101010) + \beta_4(110001) = 0$$

$$\beta_3 + \beta_4 = 0 \quad (\text{I})$$

$$\beta_2 + \beta_4 = 0 \quad (\text{II})$$

$$\beta_2 + \beta_3 = 0 \quad (\text{III})$$

$$\beta_1 + \beta_2 = 0 \quad (\text{IV})$$

$$\beta_1 + \beta_3 = 0 \quad (\text{V})$$

$$\beta_1 + \beta_4 = 0 \quad (\text{VI})$$

Considerando (I) e (II), temos que $\beta_2 = \beta_3$. Substituindo este resultado em (III), $2\beta_3 = 0$, então $\beta_3 = 0$, e $\beta_2 = 0$. Como $\beta_3 = 0$, por (V), $\beta_1 = 0$, e substituindo β_1 em (VI), temos $\beta_4 = 0$.

Por conseguinte, os elementos de $\mathcal{H}\mathcal{O}\mathcal{V}_1$ são linearmente independentes no Espaço Euclidiano $\mathbb{R}^4$.

A Tabela 5.4.3 mostra a TC cujos elementos são da classe lateral $\mathcal{H}\mathcal{O}\mathcal{V}_1$ depois de eliminar as componentes redundantes, as notações utilizadas são aquelas da AC (ver Cap. 2).

TABELA 5.4.3

"Tabela de Contingência" juntamente com o Vetor de Massas e
o Centróide dos Perfis das Linhas

Posição do dígito Padrão	1	2	3	4	5	6	$x_{.j}$	$x_{.j}/x_{..}$
"A"	0	0	0	1	1	1	3	3/12
"B"	0	1	1	1	0	0	3	3/12
"C"	1	0	1	0	1	0	3	3/12
"D"	1	1	0	0	0	1	3	3/12
$x_{i.}$	2	2	2	2	2	2	$x_{..}=12$	
$x_{i.}/x_{..}$	2/12	2/12	2/12	2/12	2/12	2/12		

Chamaremos esta "TC" de balanceada devido ao fato de que a soma de uns em todas as linhas são iguais e de igual maneira nas colunas. Considerando a Seção 2.2.1, temos que $x_{.j}$ é a quantidade de números 1 na j -ésima coluna, x_i a quantidade de números 1 na i -ésima linha e o número total de números 1 é x ...

A matriz que contém os perfis das palavras-código é:

$$R = \begin{bmatrix} 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 & 0 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 & 0 \\ 1/3 & 1/3 & 0 & 0 & 0 & 1/3 \end{bmatrix}$$

Na última coluna e última linha da Tabela 5.4.3 temos o vetor massa $\underline{r}$, e o centróide $\underline{c}$, dos perfis das linhas, respectivamente.

A distância χ^2 entre os perfis e o centróide $\underline{c}$ são respectivamente

$$d\chi^2(\underline{r}_1, \underline{c}) = \sum_{i=1}^6 \frac{(r_{1i} - c_i)^2}{c_i} = \frac{3 \cdot \left(\frac{2}{12}\right)^2}{\frac{2}{12}} + \frac{3 \cdot \left(\frac{2}{12} - \frac{1}{3}\right)^2}{\frac{2}{12}}$$

$$d\chi^2(\underline{r}_1, \underline{c}) = \frac{12}{12} = 1$$

$$d\chi^2(\underline{r}_2, \underline{c}) = d\chi^2(\underline{r}_3, \underline{c}) = d\chi^2(\underline{r}_4, \underline{c}) = d\chi^2(\underline{r}_1, \underline{c}) = 1$$

Note que todas as distâncias são iguais.

Para a obtenção das coordenadas de cada palavra-codificada, isto é, os sinais para cada dimensão no espaço Euclidiano ponderado, a decomposição de valores singulares generalizado deve ser aplicado à matriz $(R - 1\underline{c}^T)$. Disto decorre que,

$$(R - 1\underline{c}^T) = L D_{\mu} Y^T$$

onde

$$L^T D_r L = Y^T D_c^{-1} Y = I$$

cujos resultados mostramos na Tabela 5.4.4.

TABELA 5.4.4

Coordenadas dos Sinais em cada Dimensão

Sinais \ Eixo	Eixo		
	1	2	3
s_1	(-0.5774)	(-0.8165)	(0.0000)
s_2	(-0.5774)	(0.8165)	(0.0000)
s_3	(0.5774)	(0.0000)	(-0.8165)
s_4	(0.5774)	(0.0000)	(0.8165)

O correspondente gráfico dos sinais, está exposto na Figura 5.1.

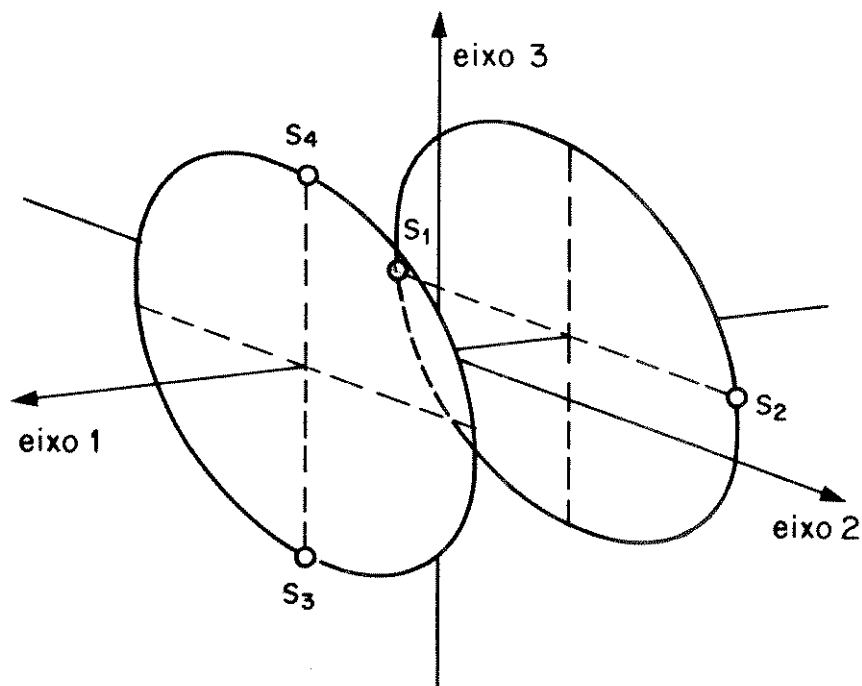


Figura 5.1 – Conjunto de Sinais Simplex sobre um espaço Euclidiano de Dimensão 3.

Observando a Figura 5.1 vemos que os 4 pontos de sinais estão sobre uma esfera de raio 1 centrada na origem de um espaço Euclidiano de dimensão 3. Os sinais

são equidistantes entre si formando assim, um conjunto de sinais simplex.

Na Tabela 5.4.5 mostramos a decomposição da Energia média do conjunto de sinais segundo a Tabela 2.5.1.

TABELA 5.4.5

Decomposição de Energia do Conjunto de Sinais

Sinais \ Eixo	1	2	3
s_1	(0.0833)	(0.1667)	0
s_2	(0.0833)	(0.1667)	0
s_3	(0.0833)	0	(0.1667)
s_4	(0.0833)	0	(0.1667)

Note que os valores resultantes da soma dos valores em cada linha são iguais o que significa que a decomposição da energia do conjunto de sinais em cada sub-espaço é a mesma e como a soma das linhas também são iguais, implica que todos os sinais tem a mesma energia.

Finalmente, na Tabela 5.4.6 mostramos as correlações entre os subespaços e os vetores sinais quando consideramos o conteúdo da Seção 2.5.2.

TABELA 5.4.6

Correlações de cada Sinal com Respeito aos Subespaços

Sinais \ Eixo	1	2	3
s_1	(0.3333)	(0.6667)	(0.0000)
s_2	(0.3333)	(0.6667)	(0.0000)
s_3	(0.3333)	(0.0000)	(0.6667)
s_4	(0.3333)	(0.0000)	(0.6667)

Da Tabela 5.4.6 observamos que os sinais s_1 e s_2 estão totalmente correlaciona-

dos com o subespaço gerado pelo eixo 1 e eixo 2. O que significa que estes pontos de sinais estão sobre o plano gerado por estes eixos. Os sinais s_3 e s_4 estão totalmente correlacionados com o subespaço gerado pelos eixos 1 e 3 o que significa que estes sinais estão no plano gerado pelos eixos 1 e 3. Isto é, cada dois sinais estão correlacionados totalmente com planos diferentes.

Como resultado deste processo temos que a classe lateral $\mathcal{H} \oplus v_1$ deverá ser transladada, de v_1 , resultando no código $\mathcal{H}$. É este mesmo código, menos os elementos redundantes, que será utilizado para codificar os 4 padrões.

Considerando a Tabela 5.4.3, a Figura 5.1 e o código $\mathcal{H}$, vemos que os vetores sinais s_1 e s_2 que estão sobre um mesmo plano correspondem às palavras código v_0 e v_2 , de igual maneira, o s_3 e s_4 correspondem as palavras-código v_3 e $v_2 \oplus v_3$ que estão no outro plano. Sob o ponto de vista algébrico, observamos que o particionamento do código $\mathcal{H}$ em duas classes laterais apresenta uma correspondência com os dois planos em $\mathbb{R}^3$.

A seguir baseados nestes resultados, generalizaremos a codificação de Padrões quando o número de elementos é 2^n , $n \in \mathbb{Z}^+$.

A generalização baseia-se na existência de códigos Reed-Muller $\mathcal{R}(1, m)$, e no Teorema de Lagrange. Pela Seção 5.3.1.1 sabemos que as cardinalidades dos códigos Reed-Muller e sub-códigos são potência de 2; esta propriedade faz com que utilizemos o Teorema de Lagrange e assim garantiremos a existência de classes laterais.

Seja $\mathcal{H}$ um código de cardinalidade 2^m e $\mathcal{H}'$ o sub-código de $\mathcal{H}$ de cardinalidade 2^n , com $n < m$, pelo Teorema de Lagrange o número de classes laterais é 2^{m-n} cada uma com cardinalidade 2^n .

Com os elementos de qualquer classe lateral obtidos anteriormente,

tal que, não contém a palavra código 0 mais a eliminação dos elementos redundantes, construímos a "Tabela de Contingência" onde o número de linhas é 2^n (número de padrões) e o de colunas $2^n - 2$. Finalmente, aplicamos a AC a esta tabela e, dessa maneira, obtemos o conjunto de 2^n sinais simplex.

Como resultado deste processo, a classe lateral considerada para a construção da TC deverá ser transladada a uma palavra-código de $\mathcal{H}$, tal que resulte no código $\mathcal{H}'$. Com as palavras-código deste código mais a eliminação dos elementos redundantes codificaremos os 2^n padrões.

Como estaremos interessados, no próximo capítulo, com a relação existente entre as diferentes métricas utilizadas é que passaremos a apresentar uma análise comparativa entre a distância de Hamming e distância χ^2 .

Lembremos que na AC a distância χ^2 é definida em função de perfis ou vetores de frequência relativas que tem a particularidade de que a soma das componentes do vetor de frequências relativas é sempre 1; então, utilizando qualquer código Reed-Muller cujos parâmetros variam de tal maneira que o comprimento da palavra-código aumenta ou diminui, sempre as componentes dos perfis das palavras somarão 1. Por isso, é necessário fazer uma padronização com respeito à distância χ^2 .

Denotaremos por $d_s \chi^2$ a distância χ^2 padronizada, então a distância χ^2 padronizada da i -ésima palavra codificada com respeito ao centróide é dada por:

$$d_s \chi^2 = d\chi^2 \cdot \underline{r}_i \cdot N$$

onde:

$d\chi^2$ é a distância χ^2 entre os perfis de linha;

$\underline{r}_i$ é a massa do i -ésimo perfil;

N é a soma de valores dos dígitos de todas as palavras codificadas.

Na Tabela 5.4.7, mostramos a relação entre as distâncias de Hamming e a distância Euclideana ponderada entre as palavras-código e vetores sinais cujas notações damos a seguir:

m = parâmetro do código $\mathcal{R}(1, m)$

M = número de padrões

n = comprimento da palavra ou do vetor em cada espaço

K = dimensão do espaço

d_H = distância de Hamming

$d_e \chi^2$ = distância chi-quadrado entre os sinais e o centróide

$d_s \chi^2$ = distância chi-quadrado padronizada entre os sinais. Observe que esta distância é duas vezes a distância Euclideana.

TABELA 5.4.7

Comparação entre os Parâmetros de cada Espaço Variando
o Número de Padrões

m	$\mathcal{R}(1, m)$	$M=2^{m-1}$	Espaço de Hamming			Espaço Euclídiano Ponderado			
			$n=2^m-2$	$K=m-1$	$d_H=2^{m-1}$	$n=2^m-2$	$K=M-1$	$d_e \chi^2$	$d_s \chi^2$
3	$\mathcal{R}(1, 3)$	4	6	2	4	6	3	1	8
4	$\mathcal{R}(1, 4)$	8	14	3	8	17	7	1	16
5	$\mathcal{R}(1, 5)$	16	30	4	16	30	15	1	32

Da Tabela 5.4.7, observamos que:

◦ A relação entre a distância χ^2 e de Hamming é dada por

$$d_s \chi^2 = 2d_H$$

- O conjunto dos $M = 2^{m-1}$ sinais está contido em um espaço Euclidiano ponderado de dimensão $M-1$.

Na Tabela 5.4.8 mostramos o comportamento que existe entre as distâncias d_H e $d_s \chi^2$, quando fixamos o número de padrões.

Considerando os resultados apresentados nesta tabela, a relação entre as d_H e $d_s \chi^2$, mantendo fixo o número de padrões, é também

$$d_s \chi^2 = 2d_H$$

TABELA 5.4.8

Comparação entre os Parâmetros de Cada Espaço Fixando
o Número de Padrões

m	$\mathcal{R}(1, m)$	M=4	Espaço de Hamming			Espaço Euclidiano Ponderado			
			$n=2^m-2^{m-1}$	k	$d_H=2^{m-1}$	$n=2^m-2^{m-1}$	$K=M-1$	$d_c \chi^2$	$d_s \chi^2$
3	$\mathcal{R}(1, 3)$	4	6	2	4	6	3	1	8
4	$\mathcal{R}(1, 4)$	4	12	2	8	12	3	1	16
5	$\mathcal{R}(1, 5)$	4	24	2	16	24	3	1	32

5.4.2.2. Obtenção de conjuntos de sinais simplex cujos elementos das palavras código estão em $\mathbb{Z}_3$

A geração de sinais simplex a ser considerada é similar ao da Seção 5.4.2.1. As considerações para a construção das "Tabelas de Contingência" continuam válidas. A única diferença é que os elementos das palavras-código pertencem ao grupo $\mathbb{Z}_3$.

Exemplo 5.4.2.: Consideremos o Exemplo 5.3.2, seja

$$\mathcal{H} = \{v_0, v_2, 2v_2\} \quad \text{e} \quad \mathcal{H}' = \{v_0, v_1, 2v_1\}$$

dois códigos de parâmetros ($m = 3^2$, $k = 2$, $d_H = 3^3 - 3 = 6$). As classes laterais de cada um deste código, são mostradas na Tabela 5.4.9.

TABELA 5.4.9

Classes Laterais de $\mathcal{H}$ e $\mathcal{H}'$

Classes Laterais do Código $\mathcal{H}$	Classes Laterais do Código $\mathcal{H}'$			
		$\mathcal{H}' \oplus v_0$	$\mathcal{H}' \oplus v_1$	$\mathcal{H}' \oplus 2v_0$
$\mathcal{H} \oplus v_0$		v_0	v_2	$2v_2$
$\mathcal{H} \oplus v_1$		$v_0 \oplus v_1$	$v_2 \oplus v_1$	$2v_2 \oplus v_1$
$\mathcal{H} \oplus 2v_1$		$v_0 \oplus 2v_1$	$v_2 \oplus 2v_1$	$2v_2 \oplus 2v_1$

Através da classe lateral $\mathcal{H} \oplus 2v_1$, construímos a Tabela de Contingência como mostra a Tabela 5.4.10.

TABELA 5.4.10

"Tabela de Contingência" Composta pelos Elementos

da Classe Lateral $\mathcal{H} \oplus 2v_1$

Dimensão Padrão									$x_{.j}$	$x_{.j}/x_{..}$
	1	2	3	4	5	6	7	8		
"A"	0	0	2	2	2	1	1	1	9	9/27
"B"	1	2	2	0	1	1	2	0	9	9/27
"C"	2	1	2	1	0	1	0	2	9	9/27
$x_{i.}$	3	3	6	3	3	3	3	3	27=x _{..}	
$x_{i.}/x_{..}$	3/27	3/27	6/27	3/27	3/27	3/27	3/27	3/27		

A matriz dos perfis das linhas estão na matriz R ,

$$R = \begin{bmatrix} 0 & 0 & 2/9 & 2/9 & 2/9 & 1/9 & 1/9 & 1/9 \\ 1/9 & 2/9 & 2/9 & 0 & 1/9 & 1/9 & 2/9 & 0 \\ 2/9 & 1/9 & 2/9 & 1/9 & 0 & 1/9 & 0 & 2/9 \end{bmatrix}$$

O vetor de massa $\underline{r}$, e o centróide $\underline{c}$ dos perfis das linhas está na última coluna e linha da tabela, respectivamente.

A distância χ^2 , entre os perfis de linha e o centróide $\underline{c}$ são iguais como veremos a seguir:

$$\begin{aligned} d\chi^2(\underline{r}_4, \underline{c}) &= \sum_{i=1}^8 \frac{(r_{i4} - c_i)^2}{c_i} \\ &= 2 \times \left(\frac{3}{27}\right) + 2 \times \left(\frac{3}{27}\right) = \frac{12}{27} = \frac{4}{9} \end{aligned}$$

$$d\chi^2(\underline{r}_1, \underline{c}) = d\chi^2(\underline{r}_2, \underline{c}) = d\chi^2(\underline{r}_3, \underline{c}) = \frac{4}{9}$$

Aplicando a AC à matriz $(R - 1\underline{c}^T)$ as coordenadas dos perfis para cada dimensão são obtidas. Isto é mostrado na Tabela 5.4.11.

TABELA 5.4.11

Coordenadas dos Sinais em cada Dimensão

Eixo Sinais	1	2
s_1	0.6667	0
s_2	-0.3333	0.5774
s_3	-0.3333	-0.5774

O espaço de sinais, e os sinais correspondentes à Tabela 5.4.11 são mostrados na Figura 5.2.

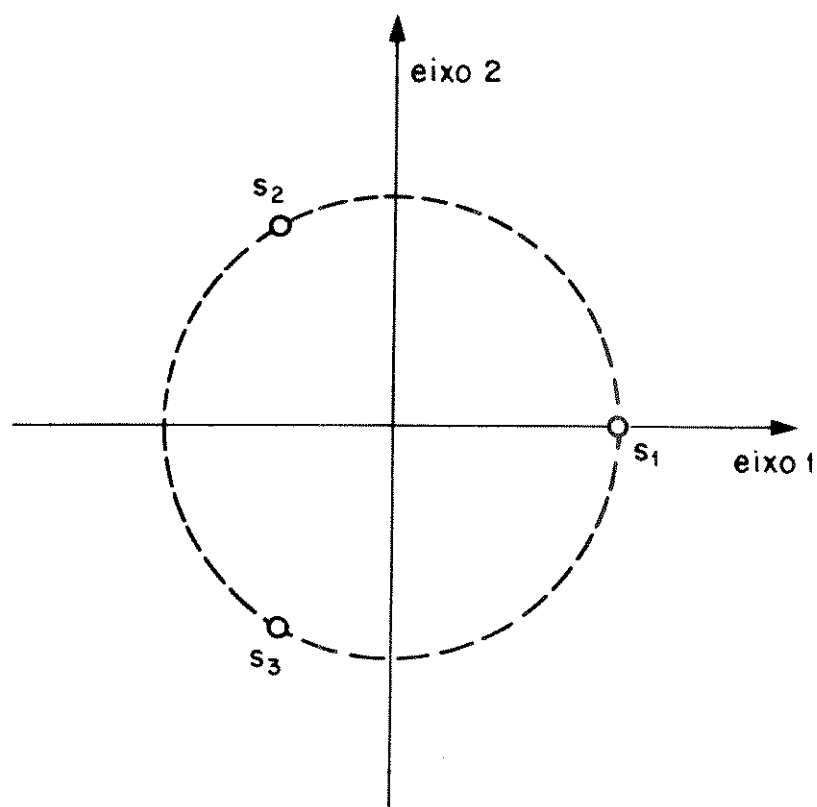


Figura 5.2 - Conjunto de Sinais Simplex para $M = 3$.

Note que os 3 sinais estão igualmente distribuídos sobre a circunferência de raio 1.

A decomposição da energia deste conjunto de sinais é ilustrada na Tabela 5.4.12, ao passo que na Tabela 5.4.13 mostramos as correlações entre os subespaços e os sinais.

TABELA 5.4.12

Decomposição de Energia do Conjunto de Sinais

Sinais \ Eixo	Eixo	
	1	2
s_1	0.1481	0
s_2	0.0370	0.1111
s_3	0.0370	0.1111

TABELA 5.4.13

Correlações dos sinais com cada Subespaço

Sinais \ Eixo	1	2
	1	0
s_1	1	0
s_2	0.25	0.75
s_3	0.25	0.75

Da observação destas tabelas vemos que: 1) a energia para o conjunto de sinais em cada dimensão são iguais; 2) o sinal s_1 está totalmente correlacionado com a dimensão 1, o que significa que este ponto está sobre uma reta; 3) os sinais s_2 e s_3 estão correlacionados totalmente no espaço gerado pelos eixos 1 e 2; finalmente, a translação da classe lateral $\mathcal{H} \oplus 2v_1$ de v_1 resulta no código $\mathcal{H}$. Este é o código que utilizaremos para codificar os 3 padrões.

Na Tabela 5.4.14, mostramos os parâmetros dos espaços e códigos para assim encontrar uma relação entre a distância de Hamming e Euclideana Ponderada.

TABELA 5.4.14

Relações entre as d_H e $d_c \chi^2$ para Códigos 3-ários

m	$GF(3^m)$	$M=3^{m-1}$	Espaço de Hamming			Espaço Euclideano Ponderado			
			$n=3^m-1$	$k=m$	d_H	$n=3^m-1$	$k=3^{m-1}-1$	$d_c \chi^2$	$d_s \chi^2$
2	$GF(3^2)$	3	8	2	6	8	2	$\frac{4}{9}$	4
3	$GF(3^3)$	9	26	3	18	26	8	$\frac{16}{27}$	16
4	$GF(3^4)$	27	80	4	54	80	26	$\frac{52}{81}$	52

As relações entre as d_H e d_{χ^2} para códigos 3-ários é dado pela seguinte equação:

$$d_{\chi^2} = d_H - 2$$

5.5. CODIFICAÇÃO DE CLASSES DE PADRÕES

No Capítulo 3, vimos que para detectar um padrão desconhecido, as funções de discriminante devem ser obtidas através dos protótipos de cada classe padrão.

Utilizando a AC e a definição de perfil suplementar, também podemos fazer a codificação de protótipos de cada classe padrão e de classes de padrões. Esta forma de codificação descrevemos a seguir.

No Exemplo 5.4.1 codificamos 4 padrões cujas palavras-código, escolhemos de um código de 8 palavras. Na Tabela 5.4.2, observamos que as 4 palavras-código utilizadas para a codificação, são elementos de 4 classes laterais respectivamente. Para a construção das regiões de decisão no espaço de características ou espaço de sinais e fazer o devido reconhecimento de padrões desconhecidos (Seção 3.5), cada padrão pode ser codificado por duas palavras código, a exceção de um padrão cuja classe lateral associada contém $v_0 = 0$. Deste resultado, concluimos que o número de protótipos de cada classe padrão dependerá do código a ser utilizado.

O número de protótipos a ser associado à um conjunto de palavras-código de um código $\mathcal{C}$ com cardinalidade 2^m pode ser obtido através do teorema de Lagrange. Para isso considere:

- Seja um código com 2^m palavras-código que formam um grupo sob a operação módulo 2 componente-a-componente.
- Seja $\mathcal{H}$ um subcódigo com cardinalidade de 2^{m-1} , ou seja, um subgrupo de $\mathcal{C}$. Então $\mathcal{C}/\mathcal{H}$ possui 2 classes laterais de ordem 2^{m-1} .

Logo, $2^m - 1$ padrões podem ser codificados por duas palavras-código.

Finalmente, no seguinte teorema generalizamos a codificação das classes de padrões quando o número de protótipos de cada classe é potência de um número primo q .

Teorema 5.5.1.: Seja o código $\mathcal{C}$ de cardinalidade q^m e $\mathcal{H}'$ um subcódigo de $\mathcal{C}$ com cardinalidade q^k , com $k < m$, então $\mathcal{C}/\mathcal{H}'$ possui q^{m-k} classes laterais com q^k elementos cada. Logo, q^k padrões podem ser codificados com $(m-k)$ palavras-código.

CAPÍTULO 6

HOMEOMORFISMO ENTRE SUBESPAÇOS MÉTRICOS DE HAMMING, EUCLIDIANO E EUCLIDIANO PONDERADO PARA O CONJUNTO DE SINAIS SIMPLEX

6.1. INTRODUÇÃO

Nos capítulos anteriores apresentamos o casamento entre grupos cíclicos e conjuntos de sinais do tipo Slepian. Além disso, mostramos como codificar os padrões de tal maneira que gerem conjuntos de sinais simplex no espaço de sinais através da AC utilizando para tal código Reed-Muller de primeira ordem.

Sabemos que o melhor arranjo de um conjunto de sinais é o simplex, quando associado um critério de otimalidade do tipo probabilidade de erro. Pela definição de conjunto de sinais do tipo Slepian, os sinais simplex formam uma sub-classe dos sinais tipo Slepian.

Agora, perguntamos: Qual será o grupo casado a um conjunto de sinais simplex? Como rotular os elementos deste conjunto de sinais?

Neste capítulo, mostraremos que os conjuntos de sinais simplex da forma 2^n , $n \in \mathbb{Z}^+$ são casados e rotulados com grupos comutativos não cíclicos. A rotulação é uma consequência direta da obtenção de um isomorfismo entre dois grupos: sendo que um destes grupos contém elementos que darão origem a um código linear (Espaço de Hamming) e o outro contém matrizes ortogonais que determinam transformações ortogonais ou isometrias (Espaço Euclideano). Obtida a rotulação mostramos, então, a existência de um Homeomorfismo entre os subespaços métricos de Hamming, Euclideano e Euclideano ponderado quando utilizamos as suas distâncias, respectivamente.

Na Seção 6.2 mostramos o casamento e rotulação de conjuntos de sinais simplex quando o número de sinais simplex é da forma 2^n , $n \in \mathbb{Z}^+$. A rotulação baseia-se na existência de um isomorfismo entre dois grupos comutativos não cíclicos. Obtida a rotulação dos vetores sinais, na Seção 6.3 mostramos o Homeomorfismo entre os subespaços métricos de Hamming, Euclideano e Euclideano Ponderado.

6.2. CASAMENTO E ROTULAÇÃO DE CONJUNTOS DE SINAIS SIMPLEX A GRUPOS

Para mostrar este casamento e rotulação, utilizaremos as definições e teoremas das Seções 4.4 e 4.5, devido a que o conjunto de sinais simplex é uma subclasse de conjuntos de sinais Slepian. Para o espaço Euclidiano $\mathbb{R}^3$ a distribuição dos elementos de cada um destes conjuntos pode ser vista nas Figuras 4.6 e 5.6. Observe que os 8 elementos do conjunto de sinais Slepian estão distribuídos sobre dois planos paralelos. E os 4 elementos do conjunto de sinais simplex estão distribuídos, de maneira uniforme sobre a esfera em $\mathbb{R}^3$ e além disso, estão sobre dois planos.

Iniciaremos esta seção mostrando a existência de um isomorfismo entre dois subgrupos; um deles associado ao espaço de Hamming e o outro ao espaço Euclidiano.

6.2.1. Isomorfismo entre um Sub-grupo de SO_{2^n-1} e um Subgrupo de $(\mathbb{Z}_2)^{2^n-1}$

Primeiramente, obteremos o isomorfismo para $n=2$, para em seguida, generalizá-lo para todo $n \in \mathbb{Z}^+$.

O conjunto de sinais simplex composto por 4 elementos, e que chamaremos 4-Simplex, é considerado como vértices de um tetraedro. Baseados no tetraedro obteremos a estrutura do grupo que associaremos ao espaço Euclidiano. Primeiramente vejamos quantas simetrias tem o tetraedro (Figura 6.1).

Considerando o grupo de permutações S_4 , e rotulando os vértices por 1, 2, 3 e 4 respectivamente, a ordem do grupo de simetrias de S_4 é dada por $\#S_4 = 4! = 24$.

Se considerarmos somente as simetrias rotacionais teremos um grupo de ordem 12. Em [32] é mostrado que o grupo de simetrias rotacional do tetraedro é isomorfo ao grupo alternante A_4 . Os elementos de A_4 são as permutações pares de S_4 . Uma maneira de "visualizar" este grupo é como mostrado na Figura 6.1.

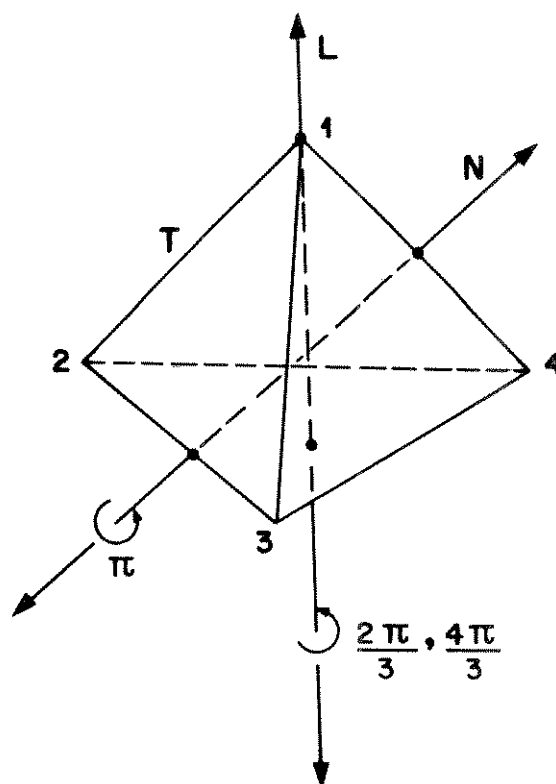


Figura 6.1 - Rotações do tetraedro com respeito aos eixos 4 e N.

Na Figura 6.1, mostramos dois eixos, um eixo rotulado por L que passa pelo vértice 1 e o centróide da face oposta ao vértice. Existem 4 eixos semelhantes a L. A face oposta ao vértice é rotacionada em sentido anti-horário por dois ângulos $2\pi/3$ e $4\pi/3$ radianos. O outro eixo rotulamos por N. Existem 3 eixos desta classe cujas restrições são de π radianos. Portanto, temos $4 \times 2 + 3 = 11$ simetrias mais a simetria identidade resulta em um total de 12 rotações formando assim um grupo de simetrias rotacional de ordem 12. Cada simetria rotacional inclui uma permutação dos vértices e portanto, uma permutação dos primeiros 4 inteiros.

Como o conjunto de sinais 4-simplex possui 4 elementos ~~procuramos~~ de um grupo de simetria rotacional de ordem 4. Quem garante a existência deste grupo é o primeiro teorema Sylow o qual enunciamos a seguir:

Teorema 6.2.1. [33] *Primeiro Teorema de Sylow* : Seja G um grupo finito de ordem $\#G = p^n m$, com $n \geq 1$, onde p não divide m . Então,

1. G contém um subgrupo de ordem p^i , $\forall i, i = 1, \dots, n$;
2. Cada subgrupo H de G de ordem p^i é um subgrupo normal de ordem p^{i+1} , para $i \leq i \leq n$.

Sabemos que a $\#A_4 = 2^2 \cdot 3$, onde 2 não divide 3. Pelo Teorema 6.2.1, A_4 contém um subgrupo de ordem 2 e 4 e cada subgrupo de ordem 2 é um subgrupo normal do subgrupo de ordem 4.

Garantida a existência do subgrupo de ordem 4, procuramos um grupo de matrizes ortogonais de ordem 4. Da Seção 4.5 sabemos que um grupo de matrizes determina um grupo de transformações lineares; portanto, deveremos determinar 4 matrizes ortogonais que gerem um grupo transitivo, isto é, multiplicando cada elemento do grupo de matrizes a um vetor sinal obtêm-se os outros vetores sinais.

Na Figura 6.2a mostramos os 4 vetores sinais $S = (s_1, s_2, s_3, s_4)$ como vértices de um cubo tal que $\sum_{i=1}^4 s_i = 0$ com coordenadas:

$$s_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}; \quad s_2 = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}; \quad s_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{bmatrix}; \quad s_4 = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{bmatrix}$$

e na Figura 6.2b mostramos o cubo no Espaço de Hamming.

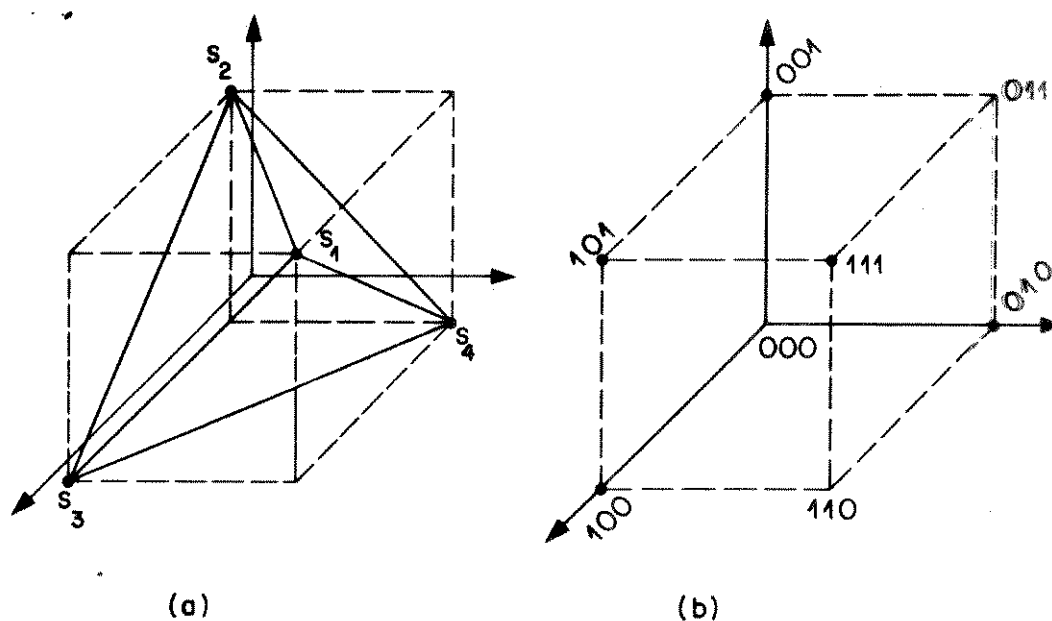


Figura 6.2a - Conjunto de Sinais Simplex, no Espaço de Sinais

b - Cubo no Espaço de Hamming.

O próximo passo, então, será o de determinar o grupo de matrizes ortogonais B_i que formem o grupo transitivo. Para tal, considere a transformação T_{B_i} definida como:

$$\begin{aligned} T_{B_i} : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ s_1 &\longrightarrow B_i s_1 \end{aligned} \quad (6.1)$$

a qual fixa um elemento qualquer do espaço de sinal e transforma este elemento em qualquer outro elemento do espaço de sinal.

$$T_{B_i}(s_1) = B_i s_1$$

$$a) s_1 \longrightarrow s_1$$

$$\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \Rightarrow \underbrace{\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}}_{s_1} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{B_1} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}}_{s_1}$$

$$b) s_1 \longrightarrow s_2$$

$$\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \Rightarrow \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

$s_2 \qquad \qquad \qquad B_2 \qquad \qquad \qquad s_1$

$$c) s_1 \longrightarrow s_3$$

$$\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

$s_3 \qquad \qquad \qquad B_3 \qquad \qquad \qquad s_1$

$$d) s_1 \longrightarrow s_4$$

$$\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \Rightarrow \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

$s_4 \qquad \qquad \qquad B_4 \qquad \qquad \qquad s_1$

O conjunto de matrizes B_1 , B_2 , B_3 e B_4 denotaremos por $\mathcal{B} \subset SO_3$. Na Tabela 6.1 mostramos que os elementos de $\mathcal{B}$ formam um grupo comutativo não cíclico, sob a operação multiplicação de matrizes, as quais determinam o grupo de transformações ortogonais que por sua vez, também formam um grupo transitivo de isometrias. Este grupo associamos ao espaço Euclídeo. Note que este grupo é o grupo 4 de Klein.

Tabela 6.1 - Operação de Multiplicação de Matrizes para
Quaisquer dos Dois Elementos de $\mathcal{B}$.

$\cdot$	B_1	B_2	B_3	B_4
B_1	B_1	B_2	B_3	B_4
B_2	B_2	B_1	B_4	B_3
B_3	B_3	B_4	B_1	B_2
B_4	B_4	B_3	B_2	B_1

Para que exista o isomorfismo entre os grupos associados aos diferentes espaços, o outro grupo associado ao espaço de Hamming e que possui as mesmas características de $\mathcal{B}$, isto é, comutativo e não cíclico é um subgrupo de $(\mathbb{Z}_2)^3$ sobre a operação adição módulo 2. Este grupo denotaremos por Z cujos elementos são: 000, 110, 011, 101. Na Tabela 6.2 mostramos que os elementos de Z formam um grupo.

Tabela 6.2 - Operação de Adição Módulo 2 para Dois
Elementos de Z .

$\oplus$	000	110	011	101
000	000	110	011	101
110	110	000	101	011
011	011	101	000	110
101	101	011	110	000

As Tabelas 6.1 e 6.2, mostram a existência de uma correspondência biunívoca entre seus elementos.

Esta correspondência é dada por:

$$\left. \begin{array}{lcl} z_1 = 000 & \rightarrow & B_1 \\ z_2 = 110 & \rightarrow & B_2 \\ z_3 = 011 & \rightarrow & B_3 \\ z_4 = 101 & \rightarrow & B_4 \end{array} \right\} \quad (6.2)$$

que induzem a existência de um isomorfismo entre os grupos $(Z, \oplus)$ e $(B, \cdot)$

Seja $\varphi : (Z, \oplus) \rightarrow (B, \cdot)$ definida por (6.2) vemos através das tabelas que φ é bijetiva e para todo $z_i, z_j \in Z$, satisfaz:

$$\varphi(z_i \oplus z_j) = \varphi(z_i) \cdot \varphi(z_j)$$

Portanto, Z e B são isomorfos.

Com estes resultados, generalizaremos o isomorfismo entre subgrupos de SO_{2^n-1} e $(Z_2)^{2^n-1}$.

Os 2^n sinais simplex podem ser representados como "polítopos regulares" inscritos em um hipercubo de 2^{n+1} vértices, num espaço Euclidiano de dimensão $2^n - 1$. Considerando o grupo de simetrias teremos um grupo de ordem, $S_{2^n} = (2^n)!$. Dado que a permutação é par, existe um subgrupo de ordem $\frac{(2^n)!}{2}$ que é o grupo alternante A_{2^n} , isomorfo ao grupo de simetrias rotacional.

De forma similar a $n = 2$, observamos que as ordens dos grupos são muito maiores que o número de sinais simplex 2^n . Mas como

$$\begin{aligned} \#A_{2^n} &= \frac{(2^n)!}{2} = \frac{2^n \cdot (2^n - 1) \cdot \dots \cdot 2 \cdot 1}{2} \\ &= 2^{n-1} \cdot (2^n - 1) \cdot \dots \cdot 3 \cdot 1. \end{aligned}$$

Pelo primeiro teorema de Sylow (Teorema 6.2.1) garantimos a existência de um subgrupo normal de simetrias de ordem 2^n .

Garantida a existência de um subgrupo normal de ordem 2^n , o próximo passo passa a ser a determinação das 2^n matrizes ortogonais de ordem $2^{n-1} \times 2^{n-1}$ as quais são originadas por vetores sinais, tal que

$$\sum_{s_i \in S} s_i = 0, \quad s_i \in \mathbb{R}^{2^{n-1}}, \quad \|s_i - s_j\| = \|s_k - s_\ell\| \quad \forall i \neq j \text{ e } k \neq \ell.$$

Os vetores sinais são determinados pela relação que existe entre os vértices dos hipercubos no espaço de Hamming e Euclideano. Esta relação para $n=2$ pode ser vista nas Figuras 6.2a e 6.2b. Em geral, as coordenadas dos vetores sinais obedecem a seguinte regra, quando um elemento do vértice do cubo no espaço de Hamming é "0", a coordenada no espaço Euclideano é $\frac{1}{\sqrt{2^n-1}}$, e se é "1", a coordenada é $-\frac{1}{\sqrt{2^n-1}}$. O número de coordenadas negativas dos vetores sinais sempre será par, porque o número de uns nas coordenadas dos vértices do cubo no espaço de Hamming sempre será par.

Uma vez conhecidos os vetores sinais, as matrizes ortogonais são determinadas de forma similar a (6.1).

As matrizes B_i , $i=1, \dots, 2^n$ formam um grupo de matrizes ortogonais diagonais idempotentes aos quais também denotaremos por, $\mathcal{B}$. O grupo $\mathcal{B} \subset SO_{2^n-1}$ induz um grupo de transformações ortogonais que por sua vez é um grupo transitivo de isometrias.

O grupo associado ao Espaço de Hamming, que possui as mesmas características que as do grupo $\mathcal{B}$ é um subgrupo $\mathcal{Z}$ de $(\mathbb{Z}_2)^{2^n-1}$ de ordem 2^n .

O isomorfismo entre os grupos $\mathcal{Z}$ e $\mathcal{B}$ baseia-se no isomorfismo do grupo aditivo $\mathbb{Z}_2 = \{0, 1\}$ e o grupo multiplicativo $\{1, -1\}$ da seguinte maneira:

- Se o primeiro elemento do $z_1 \in Z$ é "0", o primeiro elemento da diagonal da matriz é "1" e se é "1" o elemento correspondente na diagonal é "-1"; isto é, se $\rho: (0, 1) \rightarrow (1, -1)$ definida por $\rho(0) = 1$ e $\rho(1) = -1$; então,

$$\varphi: Z \subset (\mathbb{Z}_2)^{2^{n-1}} \longrightarrow \mathcal{B} \subset SO_{2^{n-1}} \quad \text{é definido por:}$$

$$\varphi(z_{i1}, z_{i2}, \dots, z_{i2^{n-1}}) = \begin{pmatrix} \rho(z_{i1}) & & & 0 \\ & \rho(z_{i2}) & & \\ & & \ddots & \\ 0 & & & \rho(z_{i2^{n-1}}) \end{pmatrix} = B_i$$

Assim, por exemplo, para

$$z_1 \in (\mathbb{Z}_2)^{2^{n-1}}, \quad z_1 = (0, \dots, 0) \longleftrightarrow I_{2^{n-1} \times 2^{n-1}}$$

I = matriz identidade.

Portanto, os elementos do subgrupo $Z \subset (\mathbb{Z}_2)^{2^{n-1}}$ são isomorfos aos elementos do subgrupo $\mathcal{B} \subset SO_{2^{n-1}}$.

Baseados neste isomorfismo, mostraremos o casamento e rotulação do conjunto de sinais simplex em $\mathbb{R}^{2^{n-1}}$ com os elementos do grupo Z .

6.2.2. Mapeamento Casado e Rotulamento do Conjunto de Sinais Simplex em $\mathbb{R}^{2^{n-1}}$, com Elementos do Grupo $Z \subset (\mathbb{Z}_2)^{2^{n-1}}$

Nesta seção mostraremos o casamento entre o grupo Z e o conjunto de sinais simplex S . Denotaremos por $T_{\mathcal{B}}(S)$ ao grupo de transformações ortogonais determinada pelo grupo $\mathcal{B}$, e que tem a propriedade de ser um grupo transitivo de isometrias. O casamento entre Z e S é uma consequência do isomorfismo existente entre Z e $\mathcal{B}$ e do Teorema 4.5.2 como veremos.

Para $n=2$

Seja $h : Z \rightarrow S$ definida por

$$\begin{aligned} z_1 = 000 &\rightarrow h(000) = T_{B_1}(s_1) = B_1 s_1 = s_1 \\ z_2 = 110 &\rightarrow h(110) = T_{B_2}(s_1) = B_2 s_1 = s_2 \\ z_3 = 011 &\rightarrow h(011) = T_{B_3}(s_1) = B_3 s_1 = s_3 \\ z_4 = 101 &\rightarrow h(101) = T_{B_4}(s_1) = B_4 s_1 = s_4 \end{aligned} \quad (6.3)$$

A distância Euclideana para quaisquer dois elementos de Z satisfaz

$$d(h(z_i), h(z_j)) = d(h(z_i^{-1} \cdot z_j), h(e)) \quad (6.4)$$

Por exemplo, seja z_2 e $z_3 \in Z$ mostraremos que:

$$d(h(z_2), h(z_3)) = d(h(z_2 \oplus z_3), h(z_1))$$

Para tal, considere

$$d(h(z_2), h(z_3)) = d(B_2 s_1, B_3 s_1) = \|B_2 s_1 - B_3 s_1\| = \|s_2 - s_3\|.$$

Dado que o conjunto de sinais é simplex, então

$$\begin{aligned} \|s_2 - s_3\| &= \|s_1 - s_4\| = \|B_4 s_1 - B_1 s_1\| \\ &= d(h(z_4), h(z_1)) = d(h(z_2 \oplus z_3), h(z_1)) \end{aligned}$$

logo,

$$d(h(z_2), h(z_3)) = d(h(z_2 \oplus z_3), h(z_1)) \quad (6.5)$$

Pela Definição 4.4.1 de casamento entre conjunto de sinais e grupos, vemos que h é um mapeamento casado e além disso, é um rotulamento casado, pois h é bijetiva. Desta maneira mostramos o mapeamento casado e o rotulamento entre o grupo Z e o conjunto de sinais em $\mathbb{R}^3$.

Obtida a rotulação dos vértices do tetraedro, através de h , resta nos determinar o peso fase dos elementos de Z . Desse modo, a partir da Definição 4.4.2, temos que o peso de fase do elemento $z_j \in Z$ é dado por:

$$w_p(z_j) = \|h(z_j) - h(z_i)\|^2 \quad (6.6)$$

A seguir, mostraremos que a distância Euclideana quadrática depende da soma dos elementos de Z e do vetor inicial s_i . Sejam os sinais s_i e $s_j \in S$. Pelo casamento entre Z e S temos que:

$$s_i = h(z_i) \quad , \quad s_j = h(z_j) \quad \text{e} \quad \|s_i\|^2 = 1$$

$$\begin{aligned} d_E^2(h(z_i), h(z_j)) &= \|h(z_i) - h(z_j)\|^2 \\ &= \|T_{B_i}(s_i) - T_{B_j}(s_i)\|^2 \\ &= \|B_i s_i - B_j s_i\|^2 \quad ; \quad \|s_i^T B_i\| = 1 \\ &= \|s_i^T B_i B_i s_i - s_i^T B_i B_j s_i\|^2 \\ &= \|1 - s_i^T B_i B_j s_i\|^2 \\ &= \|1 - s_i^T T_{B_i B_j}(s_i)\|^2 \\ &= \|1 - s_i^T h(z_i \otimes z_j)\|^2 \end{aligned}$$

Logo,

$$d_E^2(h(z_i), h(z_j)) = \|1 - s_1^T h(z_i \otimes z_j)\|^2 \quad (6.7)$$

Considerando (6.7), o peso de fase (6.6) de um elemento $z_j \in Z$ é dado por

$$w_p(z_j) = \|1 - s_1^T h(z_j)\|^2 \quad (6.8)$$

De (6.8), a distância de fase entre z_i e $z_j \in Z$ é:

$$d_p(z_i, z_j) = w_p(z_i \otimes z_j) \quad (6.9)$$

esta relação mostra o casamento entre linearidade e distância quando a distância de fase é utilizada.

De (6.7), (6.8) e (6.9) vemos que:

$$d_p(z_i, z_j) = d_E^2(h(z_i), h(z_j)). \quad (6.10)$$

A generalização do casamento e rotulação entre $Z \subset (\mathbb{Z}_2)^{2^n-1}$ e $S \subset \mathbb{R}^{2^n-1}$ é uma consequência do Teorema 4.5.2 e do isomorfismo existente entre Z e $\mathcal{B}$ mostrado anteriormente.

Com o resultado desta rotulação dos vetores sinais e, considerando as relações (6.7), (6.8) e (6.9), generalizaremos a distância Euclideana quadrática, peso de fase e a relação entre distância de fase e peso de fase respectivamente.

Finalizaremos este capítulo, mostrando o Homeomorfismo entre subespaços dos espaços métricos de Hamming, Euclideano e Euclideano Ponderado.

6.3. HOMEOMORFISMO ENTRE SUBESPAÇOS DOS ESPAÇOS MÉTRICOS DE HAMMING, EUCLIDEANO E EUCLIDEANO PONDERADO

O homeomorfismo (ou transformação topológica, constitui de grupos de mapeamentos contínuos que possuem inversos também contínuos) entre os subespaços são mostrados sobre subespaços métricos. Um espaço métrico é definido por um conjunto não vazio e uma métrica. Como os conjuntos de sinais simplex tem a propriedade de que a distância entre os elementos de cada subespaço sempre são iguais, as definiremos como métricas discretas. Os três espaços em consideração são métricos.

6.3.1. Homeomorfismo entre Subespaços de Hamming e Euclideano

Seja $H_{2^{n-1}}$, $n \in \mathbb{Z}^+$, o espaço de Hamming cuja métrica entre os elementos de $Z \subset H_{2^{n-1}}$ é a métrica discreta; isto é, dado $Z \neq \emptyset$, definiremos:

$$d_H: Z \times Z \longrightarrow \mathbb{R}, \text{ para todo } z_i, z_j \in Z, \text{ por}$$

$$d_H(z_i, z_j) = \begin{cases} 2^{n-1} & \text{se } z_i \neq z_j \\ 0 & \text{se } z_i = z_j \end{cases} \quad (6.3.1)$$

O subespaço métrico correspondente, denotaremos por (Z, d_H) .

Agora, $\mathbb{R}^{2^{n-1}}$ o espaço Euclideano cuja métrica entre os elementos de $S \neq \emptyset$, tal que $S \subset \mathbb{R}^{2^{n-1}}$ é a distância Euclideana. Como os elementos de S formam o conjunto de sinais simplex, portanto discreto, terá como distância Euclideana discreta a seguinte métrica definida por:

$$d_E^2(s_i, s_j) = \begin{cases} 2^{n+1}/2^n - 1 & \text{se } s_i \neq s_j \\ 0 & \text{se } s_i = s_j \end{cases} \quad (6.3.2)$$

A este subespaço métrico denotaremos por (S, d_E^2) .

Uma vez que temos as métricas bem definidas tanto no espaço de Hamming como no Euclideo, passaremos, então, à busca da existência do homeomorfismo entre os mesmos. Para tal, consideramos as definições de continuidade e homeomorfismo.

Definição 6.3.1. [16] : Sejam M e N espaços métricos. Uma função $f: M \longrightarrow N$ se diz *continua* no ponto $p \in M$ se, para qualquer $\epsilon > 0$, existe $\delta > 0$ de maneira que $d(x, p) < \delta \implies d(f(x), f(p)) < \epsilon$.

Dizer que f é *contínua* significa que f é continua em todos os pontos de M .

Definição 6.3.2. [16] : Se M e N são espaços métricos uma função $f: M \longrightarrow N$ é chamada *homeomorfismo* se, e somente se

- i) f é bijetiva
- ii) f e sua inversa f^{-1} são contínuas

Neste caso os espaços M e N se dizem *homeomorfos*.

Com estas definições mais a rotulação de cada um dos conjuntos de sinais simplex, mostramos o Teorema 6.3.1.

Teorema 6.3.1. : Sejam S e Z dois espaços métricos. Se $h: Z \longrightarrow S$ é um aplicação de rotulação e as distâncias entre os elementos dos espaços são iguais então os espaços S e Z são homeomorfos.

Prova : Por hipóteses sabemos que h é um aplicação de rotulamento, ou seja, que h é bijetiva. Considerando a definição de continuidade e das distâncias:

Seja $0 < \varepsilon < 2^{n+1}/2^n - 1$ então, existe um $0 < \delta < 2^{n-1}$ tal que

$$d_H(x, z_i) < 2^{n-1} \implies x = z_i \implies d_E^2(h(x), h(z_i)) = 0 < \varepsilon.$$

portanto, h é contínua.

Seja $0 < \varepsilon < 2^{n+1}/2^n - 1$, então, existe um $0 < \delta < 2^{n+1}/2^n - 1$ tal que

$$d(u, h(z_i)) < 2^{n+1}/2^n - 1 \implies u = h(z_i) \implies d(h^{-1}(x), h^{-1}(z_i)) = 0 < \varepsilon$$

portanto, h^{-1} é contínua. Por conseguinte, h é homeomorfismo. Logo Z e S são espaços métricos homeomorfos.

A hipótese de aplicação de rotulação entre Z e S é absolutamente necessária pois Z e S poderiam ter cardinalidades diferentes neste caso a distância em ambos seria um múltiplo de métrica discreta sem que estes fossem homeomorfos.

Das relações (6.3.1), (6.3.2) e a definição de rotulação h (6.3), obtemos que

$$d_H(z_i, z_j) = \frac{(2^n - 1)}{4} d_E^2(h(z_i), h(z_j))$$

e considerando a relação (6.10) temos,

$$d_P(z_i, z_j) = \frac{(2^n - 1)}{4} d_H(z_i, z_j)$$

6.3.2. Homeomorfismo entre Subespaços dos Espaços Métricos Euclideano e Euclidean Ponderado

Com o objetivo de mostrar o homeomorfismo pelo uso da AC obteremos primeiramente, um grupo de matrizes por meio das coordenadas do conjunto de sinais simplex, S' , obtidos através da AC. Para tal, utilizaremos a Tabela 5.4.4 onde estão expostos estes vetores sinais. A obtenção deste grupo é similar ao procedimento utilizado na Seção 6.2.1.

As matrizes obtidas são as seguintes:

$$D_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; D_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}; D_3 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}; D_4 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}.$$

Estas matrizes formam um grupo que denotaremos por $\mathcal{D}$.

Para mostrar o isomorfismo existente entre $\mathcal{D}$ e $\mathcal{B}$, construímos uma tabela com os elementos de $\mathcal{D}$, onde mostramos os resultados das operações entre quaisquer dos dois elementos.

Tabela 6.3 - Resultado das Operações entre Quaisquer dos
Dois Elementos de $\mathcal{D}$.

.	D_1	D_2	D_3	D_4
D_1	D_1	D_2	D_3	D_4
D_2	D_2	D_1	D_4	D_3
D_3	D_3	D_4	D_1	D_2
D_4	D_4	D_3	D_2	D_1

Considerando as Tabelas (6.1) e (6.3), vemos que existe uma correspondência biunívoca entre seus elementos dado da seguinte maneira:

$$\left. \begin{array}{lcl} B_1 & \longrightarrow & D_1 \\ B_2 & \longrightarrow & D_2 \\ B_3 & \longrightarrow & D_3 \\ B_4 & \longrightarrow & D_4 \end{array} \right\} \quad (6.11)$$

esta correspondência, induz o isomorfismo entre $\mathcal{B}$ e $\mathcal{D}$. Logo, $\mathcal{B}$ e $\mathcal{D}$ são isomórfos.

Dado que $\mathcal{B}$ e $\mathcal{D}$ são isomorfos, existe um mapeamento de casamento e rotulação entre o grupo $\mathcal{Z} \subset (\mathbb{Z}_2)^3$ e os sinais $S' = (s'_1, s'_2, s'_3, s'_4)$. Portanto, o peso de fase e a distância de fase são definidos igualmente como em (6.8) e (6.9) respectivamente. O vetor inicial s_i neste caso é $\left(\sqrt{\frac{1}{3}}, \sqrt{\frac{2}{3}}, 0 \right)$.

Finalmente, utilizando o Exemplo 5.4.1, mostraremos que a distância Euclideana e distância χ^2 entre os vetores sinais são iguais para conjuntos de sinais simplex.

Da Seção 2.4 e equação (2.4.1) cada perfil de linha pode ser expresso como:

$$\underline{r}_i = Y \underline{s}'_i \quad \text{onde} \quad Y^T D_c^{-1} Y = I \quad (6.12)$$

Considerando a definição de distância χ^2 entre dois perfis (Seção 2.2.2) temos:

$$\begin{aligned} d_{\chi^2}(\underline{r}_i, \underline{r}_j) &= (Y \underline{s}'_i - Y \underline{s}'_j)^T (Y \underline{s}'_i - Y \underline{s}'_j) \\ &= (\underline{s}'_i - \underline{s}'_j)^T Y^T D_c^{-1} Y (\underline{s}'_i - \underline{s}'_j) \\ &= (\underline{s}'_i - \underline{s}'_j)^T (\underline{s}'_i - \underline{s}'_j) = \|\underline{s}'_i - \underline{s}'_j\|^2 \end{aligned}$$

Portanto,

$$d_{\chi^2}(\underline{r}_i, \underline{r}_j) = d_E^2(s'_i, s'_j), \quad s'_i, s'_j \in \mathbb{R}_p^3 \quad (6.13)$$

Por outro lado, seja $s_i, s_j \in \mathbb{R}^3$

$$\begin{aligned} d_E^2(s_i, s_j) &= \|s_i - s_j\|^2 \\ &= \|B_i s_i - B_j s_i\|^2 \\ &= \|B_i - B_j\|^2 \|s_i\|^2, \text{ dado que } \|B_i - B_j\|^2 = \|D_i - D_j\|^2 \quad \text{e} \\ &\quad \|s_i\|^2 = \|s'_i\|^2 \\ &= \|D_i - D_j\|^2 \|s'_i\|^2 = \|s'_i - s'_j\|^2 \end{aligned}$$

Logo,

$$d_E^2(s_i, s_j) = d_E^2(s'_i, s'_j) \quad (6.14)$$

De (6.13) e (6.14)

$$d_{\chi^2}(\underline{r}_i, \underline{r}_j) = d_E^2(s_i, s_j) = d_E^2(s'_i, s'_j) \quad (6.15)$$

Assim, destes resultados e do Teorema 6.3.1 concluímos que a distância Euclideana entre os vetores sinais coincide com a distância χ^2 . Portanto, existe um homeomorfismo entre os subespaços métricos do espaço Euclideano $\mathbb{R}^3$ e Euclideano Ponderado $\mathbb{R}_p^3$.

A generalização do Homeomorfismo entre dois subespaços pelo uso da AC para $n \geq 3$, passa pela dificuldade de obter o grupo de matrizes ortogonais $\mathcal{D}$, de ordem 2^n . Este problema, ainda em aberto, é de difícil solução. Com o objetivo de ilustrar este fato, iremos a seguir considerar o caso em que $n = 3$. Seja $n = 3$, na Seção 6.2 o grupo de matrizes associado a $\mathbb{R}^7$ é de ordem 8. Estas

matrizes são diagonais idempotentes e possuem a propriedade de comutatividade. Para a obtenção do grupo de matrizes em $\mathbb{R}_p^7$ consideremos o Capítulo 5 da geração de conjuntos de sinais simplex através de classes laterais. Seja, o código

$$\mathcal{C} = \{v_0, v_1, \dots, v_1 \oplus v_2 \oplus v_3 \oplus v_4 \oplus v_5\} \quad \text{e o subcódigo}$$

$$\mathcal{H} = \{v_0, v_1, \dots, v_1 \oplus v_2 \oplus v_3\}$$

as classes laterais são: $\mathcal{H} \oplus v_0$, $\mathcal{H} \oplus v_4$, $\mathcal{H} \oplus v_5$, $\mathcal{H} \oplus v_4 \oplus v_5$.

Nas Tabelas 6.3 e 6.4 apresentamos as coordenadas dos vetores sinais correspondentes às classes laterais $\mathcal{H} \oplus v_4$ e $\mathcal{H} \oplus v_5$, respectivamente.

Observe nas tabelas que as coordenadas dos vetores sinais são diferentes para cada classe lateral. Portanto, as posições dos vetores sinais mudam; mas os vetores sinais correspondentes à mesma classe lateral sempre formam um conjunto de sinais simplex.

TABELA 6.3

Coordenadas dos Vetores Sinais para a Classe Lateral $\mathcal{H} \oplus v_4$

Dimensão Sinal	1	2	3	4	5	6	7
s_1	-0.0000	-0.2535	0.0000	-0.0000	0.0000	0.0000	0.9673
s_2	-0.1238	-0.4151	-0.1794	0.2516	-0.4816	-0.6474	-0.2565
s_3	-0.3364	-0.4151	0.0669	-0.6322	0.4941	0.0209	-0.2565
s_4	-0.4479	0.3747	-0.7044	0.1330	-0.0553	0.3736	-0.0495
s_5	0.4602	-0.4151	0.1125	0.3805	-0.0126	0.6265	-0.2565
s_6	0.4479	0.3747	-0.1163	0.2298	0.6332	-0.4351	-0.0495
s_7	0.4479	0.3747	0.0589	-0.6279	-0.5077	0.0316	-0.0495
s_8	-0.4479	0.3747	0.7618	0.2652	-0.0702	0.0298	-0.0495

TABELA 6.4

Coordenadas dos Vetores Sinais para a Classe Lateral $\mathcal{H}\mathbf{ov}_5$

Dimensão Sinal	1	2	3	4	5	6	7
s_1	0.0000	0.0000	0.8989	0.4383	-0.0000	0.0000	-0.0000
s_2	0.2006	-0.0240	0.0761	-0.4820	-0.5890	-0.4149	-0.4494
s_3	-0.3785	0.3934	0.0761	-0.4820	-0.1766	0.4997	0.4278
s_4	-0.7216	-0.4612	-0.2818	0.2519	-0.0351	0.0930	-0.3373
s_5	0.1778	-0.3695	0.0761	-0.4820	0.7656	-0.0848	0.0216
s_6	-0.1501	0.4612	-0.2818	0.2519	0.1711	-0.6967	0.3274
s_7	0.3879	0.4612	-0.2818	0.2519	0.1955	0.4388	-0.5129
s_8	0.4838	-0.4612	-0.2818	0.2519	-0.3315	0.1649	0.5228

Utilizando a transformação de Householder [34], encontramos parte das matrizes ortogonais. Estas matrizes não são comutativas e, portanto, não é possível generalizar o Homeomorfismo entre os espaços $\mathbb{R}^7$ e $\mathbb{R}_p^7$ pelo uso da AC.

O problema de encontrar o Homeomorfismo entre $\mathbb{R}^7$ e $\mathbb{R}_p^7$ se reduz ao problema em aberto de otimização do valor inicial de sinais do tipo Slepian, na determinação do ponto máximo global.

Para $n = 2$, o homeomorfismo entre os subespaços de Hamming e Euclidiano ponderado é uma consequência dos resultados obtidos na Seção 6.3.1 e 6.3.2.

CAPÍTULO 7

CONCLUSÕES

Neste trabalho desenvolvemos dois itens dentro da área de Comunicações. A primeira trata da Codificação de Padrões Via Análise de Correspondência, a qual está dentro do contexto de Codificação de Fonte. A segunda é sobre a Codificação de Canal quando a Codificação e Modulação, são consideradas como uma só entidade.

Nos Capítulos 2, 3 e 4 podem ser vistos como os principais blocos para a obtenção de uma Codificação Ótima de Padrões e Reconhecimento de Padrão. Este procedimento está fundamentado na AC, com Estruturas Algébricas e Teoria da Decisão, os quais foram abordados em detalhes dentro de cada capítulo; objetivando a melhor Codificação de Padrões de tal maneira que gerem conjuntos de sinais simplex e, conseqüentemente, a probabilidade de erro seja mínima.

Os Capítulos 2, 4 e 5 formam a base para mostrar o homeomorfismo entre o Espaço de Hamming, Euclideano e Euclideano Ponderado, para conjuntos de sinais simplex sobre espaços Euclidianos, $\mathbb{R}^{2^n-1}$, $n \in \mathbb{Z}^+$. Para a obtenção destes homeomorfismos, primeiramente mostramos o isomorfismo entre grupos associados a cada um dos espaços, e como correspondência deste isomorfismo, obtivemos a rotulação de cada um dos vetores sinais por elementos de um grupo comutativo não cíclico.

No Capítulo 2, com o título "Análise de Correspondência", desenvolvemos a teoria da Análise de Correspondência. Iniciamos dando a definição e a construção da tabela de Contingência, a medida utilizada em AC é a distância χ^2 a qual provê um efeito estabilizador ao conjunto de dados. Por ser de nosso interesse obter as coordenadas das categorias de uma variável categórica, desenvolvemos a parte da AC para a matriz que contém os perfis de linha e as respectivas tabelas contendo a distribuição da inércia, contribuição absoluta e relativa para cada um dos vetores e seus respectivos subespaços. Finalmente, fizemos referência ao perfil suplementar que tem a particularidade de não participar da geração dos espaços.

No Capítulo 3 com o título "Reconhecimento de Padrões e Teoria de Decisão" foi descrito o processo do reconhecimento de padrões que envolve os conceitos de espaço padrão, espaço de características e o espaço de classificação. Na parte de seleção e extração de características, foram apresentados duas formas de extração de características; a primeira via transformação diagonal e a segunda via a transformação rotacional com o objetivo de mostrar que a AC engloba estas duas formas de extração de características. A obtenção do vetor de características do padrão através da AC proporciona muito mais informação, porque transforma o espaço padrão em espaço de características considerando em forma simultânea as ponderações de cada dimensão do espaço padrão e de cada padrão em forma individual. E, na última parte do capítulo, descrevemos a classificação de padrões e Teoria de Decisão na sua forma geral e quando a sua classificação baseia-se numa distribuição gaussiana.

No Capítulo 4 com o título "Casamento entre Grupos e Conjuntos de Sinais" mostramos na primeira parte, um breve resumo da Teoria de Grupos e Anéis. Em seguida, fizemos um resumo de Sistemas de Comunicação Digital na qual tratamos de Codificação de Fonte, Codificação de Canal, Modulação e Modulação Codificada, com o objetivo de introduzir o casamento de conjuntos de sinais e grupos. Na parte final do capítulo, apresentamos definições e resultados do casamento entre sinais e grupos bem como o casamento entre sinais M-PSK e o grupo $\mathbb{Z}_M$. Apresentamos também o resultado que mostra que qualquer conjunto de sinais casado a um grupo é do tipo Slepian [24].

No Capítulo 5 sob o título "Codificação e Reconhecimento de Padrões Via Análise de Correspondência". Apresentamos brevemente na primeira parte, conceitos sobre o Anel de Polinômios e Corpos de Galois por serem base para a geração de códigos. Em seguida, foram apresentados os códigos que utilizamos para a codificação de padrões que são os códigos Reed-Muller e Códigos cujos elementos estão em $\mathbb{Z}_p$, p primo. Através de dois exemplos, mostramos o processo

de codificação de padrões de tal maneira, que gerarão conjuntos de sinais simplex.

Este processo de codificação foi uma contribuição no contexto de codificação de padrões.

Dado que os vetores sinais são projeções dos perfis das palavras código então, para interpretar a posição do vetor sinal com respeito aos subespaços, foi utilizado o conceito de contribuição relativa.

No Capítulo 6 com o título "Homeomorfismo entre Subespaços Métricos de Hamming, Euclideano e Euclideano Ponderado para Conjuntos de Sinais Simplex", mostramos na primeira parte um isomorfismo entre os subgrupos $\mathcal{B} \subset SO_{2^n-1}$ e $\mathcal{Z} \subset (\mathbb{Z}_2)^{2^n-1}$. Por meio deste isomorfismo, concluímos que o grupo casado ao conjunto de sinais simplex é um grupo comutativo não cíclico.

Ainda neste capítulo, utilizamos a AC como uma alternativa para rotular os vetores sinais. Este objetivo foi alcançado, uma vez que foi possível rotular sinais em $\mathbb{R}^3$ utilizando o conceito de contribuição relativa, dessa forma mostrando que os sinais que estão correlacionados de igual forma num plano, pertencem a classes laterais distintas.

Uma vez rotulado os sinais, finalmente mostramos que os subespaços métricos de Hamming, Euclideano e Euclideano Ponderado são homeomorfos.

Como sugestões para futuros trabalhos, apresentamos os seguintes tópicos:

- 1) Aplicação da AC como extrator de características;
- 2) Estudo da Análise de Correspondência Generalizada e suas aplicações em Sistemas de Comunicações.

- 3) Com respeito a codificação de padrões, estender este resultado à codificação de classes de padrões via AC, de tal maneira que cada classe padrão pertence a uma classe lateral do código;
- 4) Com respeito a modulação-codificada, a obtenção de códigos tal que através de AC rotulemos conjuntos de vetores sinais que são vértices de diferentes poliedros.
- 5) Relacionar em forma Algébrica e Geométrica, o conceito de classe lateral e geometria dos sinais.
- 6) Casamento de conjunto de sinais com grupos não comutativos.

BIBLIOGRAFIA

- [1] B. Li e C.Y. Suen, "A Knowledge-Based Thinning Algorithm", *Pattern Recognition*, Vol. 24, No. 12, pp. 1211-1221, 1991, Grã Bretanha.
- [2] S.O. Belkasim, S.O. Belkasim, M. Shridhar e M. Ahmadi, "Patter Recognition with Moment Invariants: A Comparative Study and New Results", *Patter Recognition*, Vol. 24, No. 12, pp. 1117-1138, 1991.
- [3] Z-Q. Hong, "Algebraic Feature Extraction of Imagens for Recognition", *Pattern Recognition*, Vol. 24, No. 3, pp. 211-219, 1991.
- [4] Z-Q. Hong e J-Y. Yang, "Optimal Discriminat Plane for a Small Numbber of Samples and Design Method of Classifier on the Plane", *Pattern Recognition*, Vol. 24, No. 4, pp. 317-324, 1991.
- [5] H.O. Hirschfeld, "A Connection Between Correlation and Contingency", *Cambridge Philosophical Soc. Proc.*, Vol. 31, pp. 520-524, 1935.
- [6] M. Richarson e G.F. Kuder, "Making a Rating Scale that Measures", *Personnel J.*, Vol. 2, 36-40, 1933.
- [7] P. Horst, "Measuring Complex Attitudes", *J. Social Psychol.*, Vol. 6, pp. 369-374, 1935.
- [8] J.P. Benzécri, "L'Analyse des Données", Tome 2: *L'Analyze des Correspondences*", Dunod, 2ª edição, 1976.
- [9] J.P. Benzécri, "*Correspondence Analysis Handbook*", Marcel Dekker, New York, 1992.

- [10] L. Lebart, A. Morineu e K. Warwick, *"Multivariate Descriptive Statistical Analysis Correspondence Analysis and Related Techniques for Large Matrices"*, John Wiley & Sons Inc., 1984.
- [11] D.V. Gokhale e S. Kullback, *"The Information in Contingency Tables"*, Marcel Dekker, 1978.
- [12] K. Pearson, "On Linear and Planes of Closest Fit to a System of Points in Space", *Philosophical Magazine and Journal of Science, Series 6*, Vol. 2, pp. 559-572, 1901.
- [13] H. Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components", *Educ. Psychol.*, Vol. 24, pp. 417-441 e 498-520, 1933.
- [14] M.J. Greenacre, *"Theory and Applications of Correspondence Analysis"*, Academic Press, 1984.
- [15] C. Eckart e G. Young, "The Approximation of One Matrix by Another of Lower Rank", *Psychometrika*, Vol. 1, pp. 211-218, 1936.
- [16] H.H. Domingues, *"Espaços Métricos e Introdução à Topologia"*, Editôra da USP, 1982.
- [17] G.S. Sebestyen, *"Decision Making Processes in Pattern Recognition"*, MacMillan, New York, 1962.
- [18] H.C. Andrews, *"Introduction to Mathematical Techniques in Pattern Recognition"*, Wiley-Interscience, 1972.

- [19] K.V. Mardia, J.T. Kent e J.M. Bibby, "*Multivariate Analysis*", Academic Press, 1979.
- [20] J.L. Massey, e T. Mittelholzer, "Convolutional Codes over Rings", in *Proc. 4th Joint Swedish-Soviet Int. Workshop on Inform. Theory*, pp. 14-18, Gotland Sweden, August 27-September 1, 1989.
- [21] J.L. Massey, "A Short Introduction to Coding Theory and Practice", in *Proc. Int. Symp. Signals, Syst., and Electron (ISSSE'89)*, pp. 629-633, Erlangen Germany, September 18-20, 1989.
- [22] J.L. Massey, T. Mittelholzer, T. Riedel e M. Vollenweider, "Ring Convolutional Codes for Phase Modulation", presented at *IEEE Int. Symp. Inform. Theory*, San Diego, CA, January 14-19, 1990.
- [23] J.L. Massey, e T. Mittelholzer, "Systematicity and Rotational Invariance of Convolutional Codes over Rings", in *Proc. 2nd Int. Workshop on Algebraic and Combinat. Coding Theory*, Leningrad, U.S.S.R., pp. 154-158, September 16-22 1990.
- [24] H-A. Loeliger, "Signal Sets Matched to Groups", *IEEE Trans. Inform. Theory*, Vol. 37, pp. 1675-1682, Novembro 1991.
- [25] D. Slepian, "Group Codes for the Gaussian Channel", *Bell Syst. Tech. J.* Vol. 47, pp. 575-602, Abril 1968.
- [26] L. Shu e D.J. Costello Jr., "*Error Control Coding*", Prentice-Hall, 1983.

- [27] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals, *IEEE Trans. on Inform. Theory*, Vol. IT-28, pp. 55-67, 1982.
- [28] H.S.M. Coxeter, "*Regular Polytopes*", New York: Dover, 1973.
- [29] F.J. Macwilllians e N.J.A. Sloane, "*The Theory of Error - Correcting Codes*", 6ª edição, Elsevier Science, 1988.
- [30] R. Palazzo Jr e B.R.M. Sosa, "Optimum Encoding of Patterns Via Correspondence Analysis", *Twenty First Annual 1991 Communication Theory Workshop*, Rhodes, Grécia, 30 de Junho a 6 de Julho de 1991.
- [31] B.R.M. Sosa e R. Palazzo Jr., "On the Optimum Encoding of Patterns Via Correspondence Analysis", *2nd International Symposium on Communication Theory and Applications*, Charlotte Mason College Ambleside, Lake district, UK, Inglaterra, 12-16 de Julho de 1993.
- [32] M.A. Armstrong, "*Groups and Symmetry*", Springer-Verlag, 1988.
- [33] J.B. Fraleigh, "*A First Course in Abstract Algebra*", Addison-Wesley, 1982.
- [34] P.Hr. Petrov, N.D. Christov e M.M. Konstatinov, "*Computational Methods for Linear Control Systems*", Prentice-Hall, 1991.