SVM Intervalar Linearmente Separável

Adriana Takahashi[*], **Adrião D. Dória Neto**,
Depto de Engenharia de Computação e Automação, DCA, UFRN,
Campus Universitário - Lagoa Nova
59078-900, Natal, RN, Brasil
E-mail: adriana@dca.ufrn.br, adriao@dca.ufrn.br,

Benjamín R. C. Bedregal
Depto de Informática e Matemática Aplicada, DIMAp, UFRN
Campus Universitário - Lagoa Nova
59072-970, Natal, RN, Brasil
E-mail: bedregal@dimap.ufrn.br.

Resumo: *A eficiência das máquinas de vetores-suporte no aprendizado de máquinas tem levado ao desenvolvimento de muitas pesquisas e aplicações associadas, porém, em alguns casos nem sempre é fácil classificar com precisão um determinado padrão entre duas ou mais classes, para problemas de classificação de padrões, e uma vez que, para encontrar o hiperplano de separação ótimo está relacionado diretamente aos dados de entrada aos vetores-suporte, então a teoria intervalar é proposta para casos onde os padrões de entrada não possuam características que modelem alguma classe. Este artigo desenvolve uma nova abordagem para SVM, utilizando SVM associado com a teoria intervalar (SVMI), as máquinas de vetor de suporte intervalares. O objetivo da SVMI é controlar as informações dos padrões de entrada para encontrar vetores-suporte de um hiperplano de separação ótimo quando houver incertezas nos dados ou imprecisões contidas no conjunto de treinamento e controle dos erros computacionais gerados no processo de treinamento da máquina.*

Palavras-chave. SVM, matemática intervalar, espaço linearmente separável.

1 Introdução

As máquinas de vetores-suporte (SVM - Support Vector Machines) tem atraído muita atenção nos últimos anos devido a sua eficiência em aplicações que requerem aprendizado de máquina e por estar bem fundamentado na teoria de aprendizado estatístico [12, 13]. As SVMs foram desenvolvidas inicialmente para resolver problemas de classificação, Burges [2] apresenta um tutorial sobre SVM que trata de problemas de classificação de padrões, e Stitson e autores [4, 13] mostram problemas de regressão, fazendo da SVM uma abordagem abrangente para diversas aplicações que envolvem problemas de modelagem de dados empíricos.

As SVMs foram concebidas através de uma formulação que busca a minimização do risco estrutural e utilizam vetores-suporte para achar hiperplanos ótimos, responsáveis para a separação ótima de classes no espaço de entrada.

Mesmo com a eficiência mostrado em pesquisas relacionadas com SVM existem alguns casos onde a SVM não determina otimamente a separação entre classes distintas através do hiperplano ótimo. Nesses casos, o uso da teoria intervalar pode ser uma alternativa, pois, desenvolve a máquina nos pontos onde não é possível determinar com precisão a pertinência de padrões do conjunto de treinamento.

2 Matemática Intervalar

A representação de um intervalo no conjunto dos números reais $\mathbb{R}$ é denotado por um par ordenado de números reais $X = [\underline{X}, \overline{X}]$, tal que,

[*]Apoio do CNPq

$\underline{X} \leq \overline{X}$. Assim, $X = \{X \in \mathbb{R} \mid \underline{X} \leq X \leq \overline{X}\}$.

Considere as letras maiúsculas como intervalos reais, por exemplo, $X = [\underline{X}, \overline{X}]$ representa um intervalo do conjunto intervalar $\mathbb{IR}$, onde, $\underline{X}$ é denominado de ínfimo de X e $\overline{X}$ denominado de supremo de X.

A representação de um número real exato é dado como $X = [\underline{X}, \overline{X}]$, em que, $\underline{X} = \overline{X}$, ou seja, se $X = 4$, então $X = [4, 4]$. Este tipo de intervalo é chamado de intervalo degenerado.

2.1 Operações Aritméticas Intervalares

Sejam $X = [\underline{X}, \overline{X}]$ e $Y = [\underline{Y}, \overline{Y}]$, onde, X e $Y \in \mathbb{IR}$. As operações aritméticas, tais como, *adição*, *subtração*, *multiplicação* e *divisão* em $\mathbb{IR}$ são definidas sobre os extremos de seus intervalos, [5, 9].

1. Adição intervalar:
   $X + Y = [\underline{X} + \underline{Y}, \overline{X} + \overline{Y}]$
2. Pseudo inverso aditivo intervalar:
   $-X = [-\overline{X}, -\underline{X}]$
3. Subtração intervalar:
   $X - Y = [\underline{X} - \overline{Y}, \overline{X} - \underline{Y}]$
4. Multiplicação intervalar:
   $$\begin{aligned} X \times Y = \; & [min\{\underline{XY}, \overline{X}\underline{Y}, \underline{X}\overline{Y}, \overline{XY}\}, \\ & max\{\underline{XY}, \overline{X}\underline{Y}, \underline{X}\overline{Y}, \overline{XY}\}] \end{aligned}$$
5. Pseudo inverso multiplicativo intervalar, onde $0 \notin X$:
   $X^{-1} = \frac{1}{X} = [\frac{1}{\overline{X}}, \frac{1}{\underline{X}}]$
6. Divisão intervalar: $0 \notin Y$
   $$\begin{aligned} \frac{X}{Y} = \; & [min\{\frac{\underline{X}}{\underline{Y}}, \frac{\underline{X}}{\overline{Y}}, \frac{\overline{X}}{\underline{Y}}, \frac{\overline{X}}{\overline{Y}}\}, \\ & max\{\frac{\underline{X}}{\underline{Y}}, \frac{\underline{X}}{\overline{Y}}, \frac{\overline{X}}{\underline{Y}}, \frac{\overline{X}}{\overline{Y}}\}] \end{aligned}$$
7. Quadrado intervalar:
   $$X^2 = \begin{cases} [\underline{X}^2, \overline{X}^2], \text{ se } \underline{X} \geq 0 \\ [\overline{X}^2, \underline{X}^2], \text{ se } \overline{X} < 0 \\ [0, \max\{\underline{X}^2, \overline{X}^2\}], \text{ senão} \end{cases}$$
8. Raiz quadrada intervalar, onde $\underline{X} \geqslant 0$:
   $\sqrt{X} = \sqrt{[\underline{X}, \overline{X}]} = [\sqrt{\underline{X}}, \sqrt{\overline{X}}]$

2.2 Propriedades Algébricas Intervalares

Sejam X, Y, $Z \in \mathbb{IR}$. As propriedades algébricas para as operações anteriores são, *fechamento*, *comutativa*, *associativa*, *elemento neutro*, *subdistributiva*, e *monotônica*.

1. Fechamento:
   - Se X, $Y \in \mathbb{IR}$, então $X + Y \in \mathbb{IR}$
   - Se X, $Y \in \mathbb{IR}$, então $X \times Y \in \mathbb{IR}$
2. Comutativa:
   - $X + Y = Y + X$
   - $X \times Y = Y \times X$
3. Associativa:
   - $X + (Y + Z) = (X + Y) + Z$
   - $X \times (Y \times Z) = (X \times Y) \times Z$
4. Elemento neutro:
   - $X + [0, 0] = [0, 0] + X = X$
   - $X \times [1, 1] = [1, 1] \times X = X$
5. Subdistributiva:
   $X \times (Y + Z) \subseteq (X \times Y) + (X \times Z)$
6. Inclusão monotônica:
   Sejam X, Y, Z e $W \in \mathbb{IR}$, tais que, $X \subseteq Z$ e $Y \subseteq W$.
   - $X + Y \subseteq Z + W$
   - $-X \subseteq -Z$
   - $X - Y \subseteq Z - W$
   - $X \times Y \subseteq Z \times W$.
   - $\frac{1}{X} \subseteq \frac{1}{Z}$, se $0 \notin Z$
   - $\frac{X}{Y} \subseteq \frac{Z}{W}$, se $0 \notin W$

2.3 Ordem Intervalar

Na literatura encontramos diversas formas de definição de ordens (parciais) para intervalos. As mais conhecidas são, ordem de Moore [9], ordem de Kulisch-Miranker [6, 7], ordem da Informação [1] e ordem da Teoria dos Conjuntos.

A ordem presente neste estudo foi a ordem de Kulisch-Miranker definida por:

1. Ordem de Kulisch-Miranker:
   $X \leq_{KM} Y = [\underline{X}, \overline{X}] \leq [\underline{Y}, \overline{Y}] \Leftrightarrow \underline{X} \leq \underline{Y}$ e $\overline{X} \leq \overline{y}$

2.4 Função Intervalar

Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função real e $F : \mathbb{IR}^n \to \mathbb{IR}$. Se $\forall X \in \mathbb{IR}^n$ e $x \in X$, $f(x) \in F(X)$, então F é uma representação intervalar de f.

A função $F : \mathbb{IR}^n \to \mathbb{IR}$ é representável, se existir uma função $\underline{F} : \mathbb{R}^n \to \mathbb{R}$ e $\overline{F} : \mathbb{R}^n \to \mathbb{R}$, tal que, $\forall X \in \mathbb{IR}^n$, $F(X) = [\underline{F}(\underline{X}), \overline{F}(\overline{X})]$ para F definida monotonicamente com respeito a ordem de Kulisch-Miranker ou $\forall X \in \mathbb{IR}^n$, $F(X) = [\underline{F}(\overline{X}), \overline{F}(\underline{X})]$ para F antimonotônica com respeito a ordem de Kulisch-Miranker [10].

2.5 Matrizes e Vetores Intervalares

Uma matriz $\mathbf{A} = A_{ij}$ é a representação de uma matriz intervalar, matriz com coeficientes intervalares de ordem m por n, se $\mathbf{A}$ for uma matriz com $m \in \mathbb{N}^*$ linhas e $n \in \mathbb{N}^*$ colunas, onde cada elemento $A_{ij} \in \mathbb{IR}$ é representado por um intervalo [8, 11].

2.5.1 Vetor Intervalar

Seja $\mathbf{A} = A_{ij}$ uma matriz intervalar de ordem $m \times n$, se $m = 1$ ou $n = 1$, então $\mathbf{A}$ é denominada de vetor intervalar [11].

2.5.2 Operações Aritméticas Vetoriais Intervalares

Sejam $\mathbf{A} = A_{ij}$ e $\mathbf{B} = B_{ij}$ matrizes intervalares de mesma ordem, $m \times n$, as seguintes operações são definidas [11]:

1. Adição vetorial intervalar:
   $\mathbf{A} + \mathbf{B} = A_{ij} + B_{ij}$

2. Subtração vetorial intervalar:
   $\mathbf{A} - \mathbf{B} = A_{ij} - B_{ij}$

3. Multiplicação de uma matriz intervalar por um intervalo:
   Considerando $I \in \mathbb{IR}$, então,
   $\mathbf{A} \times I = A_{ij} \times I$

4. Multiplicação vetorial intervalar:
   $\mathbf{A} \times \mathbf{B} = \sum_{k=1}^{n} A_{ik} \times B_{kj}$

5. Matriz intervalar transposta:
   Considerando $\mathbf{A} = (A_{ij})_{m \times n}$, sua transposta é dada por, $\mathbf{A}^T = (B_{ji})_{n \times m}$, onde, $B_{ij} = A_{ji}$

6. Distância intervalar:

   (a) Distância intervalar de Moore [10]: a distância entre $X \in \mathbb{IR}$ e $Y \in \mathbb{IR}$ é definida por,

   $$D_M(X, Y) = max(\mid \underline{X} - \underline{Y} \mid, \mid \overline{X} - \overline{Y} \mid)$$

   (b) Distância intervalar de Trindade-Bedregal [14]: a distância entre $X \in \mathbb{IR}$ e $Y \in \mathbb{IR}$ é definida por,
   $D_{TB}(X, Y) = [min\{d(x, y),\ x \in X \text{ e } y \in Y\}, max\{d(x, y),\ x \in X \text{ e } y \in Y\}]$

7. Norma vetorial intervalar:
   Seja a norma de um número real a distância de um ponto com a origem, a norma de um intervalo[1], $X \in \mathbb{IR}$ é definida por,
   $$\parallel X \parallel = \begin{cases} [\underline{X}, \overline{X}], \text{ se } \underline{X} \geq 0 \\ [\mid \overline{X} \mid, \mid \underline{X} \mid], \text{ se } \overline{X} < 0 \\ [0, \max\{\mid \underline{X} \mid, \mid \overline{X} \mid\}], \text{ senão} \end{cases}$$
   A norma intervalar definida de forma geral, baseada na distância intervalar de Trindade-Bedregal, onde, $\mathbf{X} = (X_1, ..., X_n) \in \mathbb{IR}^n$, é,
   $$\begin{aligned} \parallel \mathbf{X} \parallel &= \sqrt{D_{TB}(\mathbf{X}, [0, 0])^2} \\ &= \sqrt{\parallel X_1 \parallel^2 + ... + \parallel X_n \parallel^2} \end{aligned}$$

8. Produto interno intervalar:
   Sejam $\mathbf{A} = (A_1, ..., A_n) \in \mathbb{IR}^n$ e $\mathbf{B} = (B_1, ..., B_n) \in \mathbb{IR}^n$ pontos intervalares, o produto interno de $\mathbf{A}$ e $\mathbf{B}$ é definido por,
   $$< \mathbf{A} \cdot \mathbf{B} > = \begin{array}{l} [\underline{(A_1 \times B_1) + ... + (A_n \times B_n)}, \\ \overline{(A_1 \times B_1) + ... + (A_n \times B_n)}] \end{array}$$

3 Máquinas de Vetores-Suporte Intervalares Linearmente Separáveis

Máquinas de vetores-suporte é uma técnica de aprendizado de máquina baseada na teoria do aprendizado estatístico [3] para aprendizado supervisionado.

A tarefa de classificar padrões é feito através da função:

$$f(\mathbf{x}) = sgn(\sum_{i=1}^{N_{VS}} d_i \alpha_i \mathbf{x_i}^T \mathbf{x_i} + b) \qquad (1)$$

[1]Note que esta definição de norma intervalar é diferente da definição usual de norma intervalar, por ex.: a norma intervalar de Moore [10].

onde, $x_i \in \mathbb{R}^n$ é o vetor de entrada de n-dimensão, $d_i \in \{-1, 1\}$ é a classe que pertence x_i, e α_i e b são parâmetros da função objetivo encontrados durante o treinamento através da resolução de um problema de otimização.

Máquinas de vetores-suporte intervalares (SVMI) é uma extensão intervalar da SVM, que consiste em encontrar um hiperplano intervalar que separe o conjunto de treinamento da forma ótima, utilizando os vetores-suporte intervalares encontrados no treinamento da máquina.

Para uma SVMI, considere uma amostra de treinamento $\{(\mathbf{X}_i, D_i)\}_{i=1}^{N}$, onde, $\mathbf{X}_i$ é um intervalo que representa o padrão de entrada para o i-ésimo exemplo e $D_i = [\underline{D}_i, \overline{D}_i]$ é um intervalo degenerado que representa a saída desejada $D_i \in \{[+1, +1], [-1, -1]\}$ é a resposta desejada do i-ésimo padrão de entrada.

A equação de uma superfície de decisão na forma de hiperplano que realiza a separação entre as classes é definida como:

$$\mathbf{W}^T\mathbf{X} + B = [0, 0] \tag{2}$$

onde, $\mathbf{X}$ é um vetor intervalar de entrada, $\mathbf{W}$ é um vetor intervalar de pesos ajustáveis e B é um intervalo representando o bias.

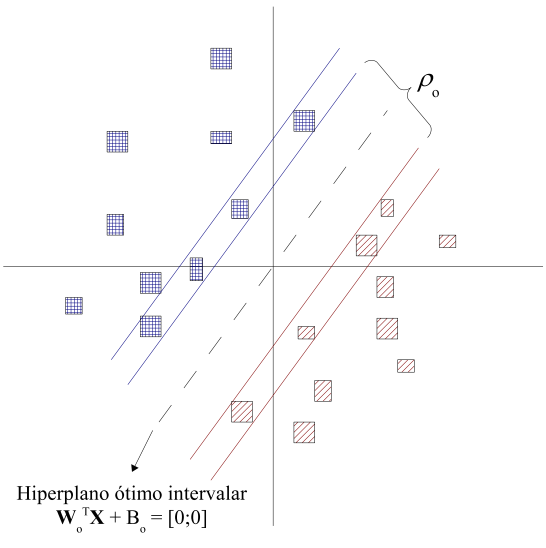

Figura 1: Hiperplano ótimo para padrões intervalares linearmente separáveis

O conjunto de treinamento $\{(\mathbf{X}_i, D_i)\}_{i=1}^{N}$ é dito linearmente separável se existir um vetor intervalar W e um intervalo B que satisfaça:

$$\begin{cases} \mathbf{W}_o^T\mathbf{X}_i + B_o \geq [+1, +1], D_i = [+1, +1] \\ \mathbf{W}_o^T\mathbf{X}_i + B_o \leq [-1, -1], D_i = [-1, -1] \end{cases} \tag{3}$$

que é equivalente a:

$$D_i(\mathbf{W}^T\mathbf{X}_i + B) \geq [1, 1] \tag{4}$$

onde, o par $(\mathbf{W}, B)$ define o hiperplano de separação da equação (2).

Para obter o hiperplano ótimo, ou seja, o hiperplano de máxima margem entre as classes é necessário encontrar a distância de um intervalo $\mathbf{X}_i$ com o hiperplano de separação $(\mathbf{W}, B)$. Considerando a função discriminante:

$$G(\mathbf{X}) = \mathbf{W}_o^T\mathbf{X}_i + B_o \tag{5}$$

onde, $G(\mathbf{X})$ fornece uma medida algébrica da distância de $\mathbf{X}$ até o hiperplano. Dado a distância dos dois hiperplanos[2], a distância entre dois intervalos $D(\mathbf{X}_1, \mathbf{X}_2)$ que satisfaçam a igualdade da equação (3),

$$\begin{aligned} &\mathbf{W}_o(\mathbf{X}_2^{(s)} - \mathbf{X}_1^{(s)}) + B_o - B_o - [2, 2] = [0, 0] \\ &\Rightarrow \mathbf{W}_o(\mathbf{X}_2^{(s)} - \mathbf{X}_1^{(s)}) = [2, 2] \end{aligned} \tag{6}$$

a distância entre dois hiperplanos calculado pela equação (6) é definida como,

$$\frac{\mathbf{W}_o(\mathbf{X}_2^{(s)} - \mathbf{X}_1^{(s)})}{\|\mathbf{W}_o\|} = \frac{[2,2]}{\|\mathbf{W}_o\|} \tag{7}$$

Logo,

$$\rho_o = \frac{[2, 2]}{\| \mathbf{W}_o \|} \tag{8}$$

Para encontrar os parâmetros $\mathbf{W}_o$ e B_o para o hiperplano ótimo dado um conjunto de treinamento, as restrições da equação (4) devem ser satisfeitas. Os pontos intervalares $(\mathbf{X}_i, D_i)$ que satisfazem o sinal de igualdade da equação (4) são chamados de **vetores-suporte** intervalares.

Considerando um vetor de suporte intervalar $\mathbf{X}^{(s)}$ temos:

$$G(\mathbf{X}^{(s)}) = \mathbf{W}_o^T\mathbf{X}^{(s)} + B_o \pm [1, 1], D^{(s)} = \pm[1, 1] \tag{9}$$

Da equação (8) a distância do vetor de suporte intervalar até o hiperplano ótimo:

$$R = \begin{cases} +\frac{[1,1]}{\|\mathbf{W}_o\|} \text{ se } D^{(s)} = +[1, 1] \\ -\frac{[1,1]}{\|\mathbf{W}_o\|} \text{ se } D^{(s)} = -[1, 1] \end{cases} \tag{10}$$

[2]distância entre dois planos no conjunto dos reais, $d(\pi_1, \pi_2) = d(\pi_1, \mathbf{p}_2) = \frac{(\mathbf{p_1} - \mathbf{p2}) \times \mathbf{n_2}}{\|\mathbf{n}\|}$, onde, $\mathbf{p}_1 \in \mathbb{R}$ e $\mathbf{p}_2 \in \mathbb{R}$ são pontos e $\mathbf{n}$ é o vetor normal ao plano

Portanto, considerando que a equação (8) represente o valor intervalar ótimo da margem de separação entre as duas classes do conjunto de treinamento, então:

$$\begin{aligned} \rho = & \ [2,2]R \\ = & \ \frac{[2,2]}{\|\mathbf{W}_o\|} \end{aligned} \quad (11)$$

Da equação (11) temos que, maximizar a margem de separação entre as classes é equivalente a minimizar a norma do vetor intervalar $\mathbf{W}$ com respeito as restrições.

O hiperplano da equação (2) é único e de máxima separação entre as classes.

Para encontrar o hiperplano ótimo utilizando o conjunto de treinamento $\{(\mathbf{X}_i, D_i)\}_{i=1}^N$ e que satisfaça as restrições é necessário encontrar os parâmentros $\mathbf{W}$ e B.

O hiperplano de separação ótimo minimiza a função custo intervalar:

$$\Phi = \frac{1}{2}\mathbf{W}^T\mathbf{W} \quad (12)$$

sujeito as restrições:

$$D_i(\mathbf{W}^T\mathbf{X}_i + B) \geq [1,1]$$

Este é um problema de otimização que pode ser resolvido através do método de multiplicadores de Lagrange adaptado para intervalar:

$$\begin{aligned} J(\mathbf{W}, B, \alpha) = & \ [\tfrac{1}{2}, \tfrac{1}{2}]\mathbf{W}^T\mathbf{W} - \\ & \textstyle\sum_{i=1}^N \alpha_i(D_i(\mathbf{W}_i^T\mathbf{X}_i + B) - \\ & [1,1]) \\ \subseteq & \ [\tfrac{1}{2}, \tfrac{1}{2}]\mathbf{W}^T\mathbf{W} - \\ & \textstyle\sum_{i=1}^N \alpha_i D_i \mathbf{W}_i^T\mathbf{X}_i + \\ & \textstyle\sum_{i=1}^N \alpha_i D_i B - \sum_{i=1}^N \alpha_i \end{aligned} \quad (13)$$

onde, α_i são intervalos representando os multiplicadores de Lagrange.

A função lagrangiana tem que ser minimizada com respeito a $\mathbf{W}$, B e maximizada com respeito a $\alpha_i \geq 0$. Assim, diferenciando $L(\mathbf{W}, B, \alpha)$ em relaçõ a $\mathbf{W}$ B temos as seguintes condições:

$$\begin{aligned} \frac{\partial L(\mathbf{W},B,\alpha)}{\partial B} &= [0,0] \Rightarrow \textstyle\sum_{i=1}^N \alpha_i D_i = [0,0] \\ \frac{\partial L(\mathbf{W},B,\alpha)}{\partial \mathbf{W}} &= [0,0] \Rightarrow \mathbf{W} = \textstyle\sum_{i=1}^N \alpha_i D_i \mathbf{X}_i \end{aligned} \quad (14)$$

Substituindo as condições (14) em (13) temos:

$$\begin{aligned} J(\mathbf{W}, B, \alpha) = & \ [\tfrac{1}{2}, \tfrac{1}{2}]\mathbf{W}^T\mathbf{W} - \\ & \textstyle\sum_{i=1}^N \alpha_i D_i \mathbf{W}_i^T\mathbf{X}_i + \\ & \textstyle\sum_{i=1}^N \alpha_i D_i B - \sum_{i=1}^N \alpha_i \\ = & \ [\tfrac{1}{2}, \tfrac{1}{2}] \textstyle\sum_{i,j=1}^N \alpha_i D_i \mathbf{X}_i \alpha_j D_j \mathbf{X}_j - \\ & \textstyle\sum_{i,j=1}^N \alpha_i D_i \mathbf{X}_i \alpha_j D_j \mathbf{X}_j - \\ & [0,0] + \textstyle\sum_{i=1}^N \alpha_i \\ = & \ -[\tfrac{1}{2}, \tfrac{1}{2}] \textstyle\sum_{i,j=1}^N \alpha_i D_i \mathbf{X}_i \alpha_j D_j \mathbf{X}_j \\ & + \textstyle\sum_{i=1}^n \alpha_i \end{aligned} \quad (15)$$

Fazendo a função objetivo $J(\mathbf{W}, B, \alpha) = Q(\alpha)$ temos:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - [\frac{1}{2}, \frac{1}{2}] \sum_{i,j=1}^N \alpha_i\alpha_j D_i D_j \mathbf{X}_i^T \mathbf{X}_j \quad (16)$$

Logo, o problema dual é dado por:

Maximizar:
$Q(\alpha) = \sum_{i=1}^N \alpha_i - [\frac{1}{2}, \frac{1}{2}] \sum_{i,j=1}^N \alpha_i\alpha_j D_i D_j \mathbf{X}_i^T \mathbf{X}_j$
Sujeito as restrições:
$$\begin{cases} (1) \ \alpha_i \geq [0,0], i = 1, ..., N \\ (2) \ \sum_{i=1}^N \alpha_i D_i = [0,0] \end{cases} \quad (17)$$

Ao encontrar os multiplicadores de Lagrange é possível calcular os pesos intervalares ótimos:

$$\mathbf{W}_o = \sum_{i=1}^N \alpha_{oi} D_i \mathbf{X}_i \quad (18)$$

O valor do bias ótimo B_o é encontrado utilizando os pesos ótimos $\mathbf{W}_o$ encontrados na equação (18) e descrito como:

$$B_o = [1,1] - \mathbf{W}_o^T\mathbf{X}^{(s)} \text{ para } D^{(s)} = [1,1] \quad (19)$$

4 Conclusões

Um formalismo de uma abordagem de SVM associada com a teoria intervalar foi detalhado neste estudo. Foram apresentados os conceitos e as teorias utilizados para o desenvolvimento dessa nova abordagem, e essa extensão intervalar da SVM, que foi denominado de SVMI, representa inicialmente os casos linearmente separáveis.

Considerando a SVMI como um classificador de padrões intervalares, a máquina tratará o conjunto de dados iniciais intervalares e encontrará um hiperplano intervalar de separação entre as classes do conjunto de treinamento, desde que este seja linearmente separável. Com isso, desde o conjunto de treinamento ou conjunto de entrada, terá um tratamento intervalar para uma representação mais precisa, considerando dados faltosos no conjunto de treinamento através dos intervalos, até o processamento da máquina para encontrar o hiperplano intervalar de separação de classes com controle dos possíveis erros gerados computacionalmente.

Conclue-se que, esta nova abordagem, a SVMI, pode tratar de problemas que, ao mesmo tempo em que exijam uma precisão nos resultados obtidos pela máquina, não exijam resultados rígidos, ou seja, através do hiperplano intervalar de separação ótimo é encontrado uma margem de pertinencia de padrões intervalares difícieis de classificação entre as classes, ou seja, um conjunto de padrões imprecisos são encontrados no resultado da classificação para uma posterior análise dependente do problema analisado.

Referências

[1] B. M. Acióly, "Fundamentação Computacional da Matemática Intervalar", Universidade Federal do Rio Grande do Sul, 1991.

[2] C. J. C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, *Data Mining and Knowledge Discovery*, **2** (1998), 121-167.

[3] S. Haykin, "Redes Neurais: Princípios e prática", Bookman, 2001.

[4] M. A. Hearst, Support Vector Machines, *IEEE Inteligent Systems*, **13** (1998), 18-28.

[5] V. Kreinovich and A. Lakeyev J. Rohn and P. Kahl, "Computational Complexity and Feability of Data Processing and Interval Computations", Kluwer Academic Publishers, Canadá, 1998.

[6] U. W. Kulisch and W. L. Miranker, "Computer Arithmetic Theory and Pratice", Academin Press, 1981.

[7] U. W. Kulisch, Computer Arithmetic and Programing Languages, *ACM*, **13** (1982), 176–182.

[8] A. Lyra, "Uma Fundamentação Matemática para o Processamento de Imagens Digitais Intervalares", Universidade Federal do Rio Grande do Norte, Thesis(Ph.D.), Natal, 2003.

[9] R. E. Moore, "Interval Analysis", Pretice Halls, New Jersey, 1966.

[10] R. E. Moore, "Methods and Applications of Interval Analysis", SIAM, Philadelphia, 1979.

[11] P. W. Oliveira and T. A. Diverio and D. M. Claudio, "Fundamentos da Matemática Intervalar", Instituro de Informática da UFRGS: Sagra-Luzzatto, 1997.

[12] M. Pontil and A. Verri, "Proprieties of Support Vector Machines", Massachusetts Institute of Technology, 1997.

[13] M. O. Stitson and J. A. E. Weston and A. Gammerman and V. Vovk and V. Vapnik, "Theory of Support Vector Machines", University of London, 1996.

[14] R. M. P. Trindade and B. R. C. Bedregal and A. D. D. Neto and B. M. Acióly, Uma Métrica Essencialmente Intervalar, *submetido CNMAC*, (2008).