

CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ
ISSN impresso 1413-9022 / ISSN on-line 2317-4536 - v.41, p.15 - 29, 2016
DOI: 10.12957/cadest.2016.27523

ESTIMAÇÃO DO GRAU DE ASTIGMATISMO PELO MÉTODO *SUPPORT VECTOR REGRESSION* CORRELACIONADO

André Luiz Emidio de Abreu
Centro Universitário Franciscano - FAE
ametodosest@gmail.com

Anselmo Chaves Neto
Universidade Federal do Paraná – Depto de Estatística
anselmo@ufpr.br

Resumo

Este trabalho apresenta a influência do coeficiente de correlação em modelos de regressão. Para o desenvolvimento, utilizou-se o método Support Vector Regression (SVR) como modelo de regressão. O método SVR foi aplicado em exames feitos em pacientes que possuem algum nível de astigmatismo. Para tanto, criou-se uma modificação na fase de ajuste do modelo de regressão, sendo introduzido o coeficiente de correlação linear, avaliando a correlação entre as variáveis preditoras: Ceratometria, subdividida em Eixo Mais Plano e Eixo Mais Curvo; e Refração, subdividido em Esfera e Cilíndrico, sendo o grau de astigmatismo a variável a ser prevista. O novo método proposto, nomeado SVR Correlacionado teve seus resultados comparados com o método convencional SVR, obtendo um desempenho superior, tanto na correlação dos modelos como no valor do erro cometido. Ao todo, utilizaram-se os dados de 26 pacientes com astigmatismo, sendo criadas duas configurações para o ajuste e teste, a primeira sendo composta de 20 observações para ajuste e seis para teste apresentando o erro RMSE = 0,035799, e a segunda composta de 16 observações na fase de ajuste e 10 no teste, gerando RMSE = 0,028518, em ambos os casos, inferiores aos erros gerados pelo método SVR convencional.

Palavras-chave: Regressão, Support Vector Regression, Correlação, Astigmatismo.

1. Introdução

Desenvolver técnicas que possam elaborar previsões ou classificar padrões é de grande utilidade na pesquisa e em aplicações reais, sendo sempre necessário o advento de técnicas robustas e que sejam generalizáveis do ponto de vista técnico. Além da robustez e poder de generalização é desejável que o tempo despendido para a obtenção das previsões seja razoável, uma vez que para uma pessoa com certo tipo de doença o tempo de diagnóstico pode ser crucial para o início do tratamento.

A técnica de aprendizado de máquinas conhecida por SVM (*Support Vector Machine* – Máquina de Vetor de Suporte) vem ganhando grande destaque, sendo utilizada para reconhecimento de padrões e os resultados obtidos são superiores aos alcançados por outras técnicas de aprendizado em várias aplicações (LORENA; CARVALHO, 2003).

A ideia inicial de criar algoritmos para reconhecer padrões surgiu através de Fisher em 1936. Mais tarde, em 1962, Rosenblatt (JONHSON e WICHERN, 1988) sugeriu o algoritmo *Perceptron* de aprendizagem, que é a base das Redes Neurais. Em 1965, Vladimir Vapnik propôs um algoritmo para encontrar um hiperplano separador em uma classificação linear (VAPNIK; CHERVONENKIS, 1974). O objetivo inicial era encontrar um hiperplano separador que tivesse a máxima margem, ou seja, a maior distância entre os vetores suportes usando a metodologia de aprendizagem. Entretanto, em 1992, Bernhard Boser, Isabelle Guyon e Vapnik criaram uma classificação não linear, aplicando as Funções *Kernel* (BOSER, GUYON E VAPNIK, 1992).

Em 1995, Corinna Cortes e Vapnik sugeriram uma modificação no algoritmo, introduzindo variáveis de folgas para obter uma margem suave no hiperplano separador e permitindo a classificação correta dos dados, chamado de modelo de margens flexíveis (CORTES; VAPNIK, 1995).

Baseado no método SVM, na sua variante de regressão SVR, este trabalho apresenta a influência que o coeficiente de correlação linear exerce, melhorando, as previsões quando aplicado ao método SVR. O objetivo do trabalho é demonstrar, de forma empírica, a eficiência da aplicação do método SVR Correlacionado, em comparação com o método SVR convencional.

Astigmatismo é a dificuldade do sistema óptico em formar um ponto focal na retina, devido à diferença na curvatura de uma ou mais superfícies refrativas do globo

ocular. A imagem de um ponto jamais será exatamente um ponto e sim uma linha (MOREIRA, 2001). O astigmatismo total consiste no astigmatismo das faces anterior e posterior da córnea assim como no astigmatismo lenticular ou também denominado de cristalino (MOREIRA, MOREIRA, 2000). O astigmatismo pode ser hereditário sob a forma autossômico dominante, autossômico recessivo ou ligado ao cromossomo X.

O artigo é composto pela introdução (capítulo 1), as revisões bibliográficas sobre o método *Support Vector Machine* (capítulo 2), *Support Vector Regression* (capítulo 3), *Least Squares Support Vector Machine* (capítulo 4), a apresentação do método SVR Correlacionado (capítulo 5), os dados utilizados neste trabalho (capítulo 6), finalizando com os resultados (capítulo 7), conclusão (capítulo 8) e as referências.

2. *Support Vector Machine* – SVM

O SVM é uma metodologia que classifica os padrões em dois conjuntos, podendo ser ampliado para classificações em múltiplas classes. A metodologia utiliza margens de separação para os conjuntos, as quais são deslocamentos com distância igual a um e denominando-se um conjunto de classe -1 e o outro de classe +1. Por se tratar de uma metodologia de aprendizado de máquinas, do tipo aprendizagem supervisionada, introduz-se uma notação para referenciar o conjunto de treinamento e os dados de entrada e saídas. Seja o conjunto de treinamento:

$$S = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\} \subseteq \{X \times Y\}^l \quad (1)$$

onde l é o número de pontos no conjunto de treinamento.

Cada ponto é representado por $\{x_i, y_i\}$, para $i = 1, \dots, l$, onde $x_i \in \mathbb{R}$, $\underline{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$, é o valor de entrada, pertencente ao espaço de entrada, que representa a quantificação das características, e $y_i \in \{-1, +1\}$ é a saída binária correspondente. Por exemplo, para uma classificação linear, pode-se obter a função real $f : X \in \mathbb{R}^n \rightarrow \mathbb{R}$ da seguinte maneira: a entrada $\underline{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$ é considerada de classe +1 se $f(\underline{x}) > 0$ e caso contrário é considerada -1. Assim, a função decisão $f(\underline{x})$ é representada por:

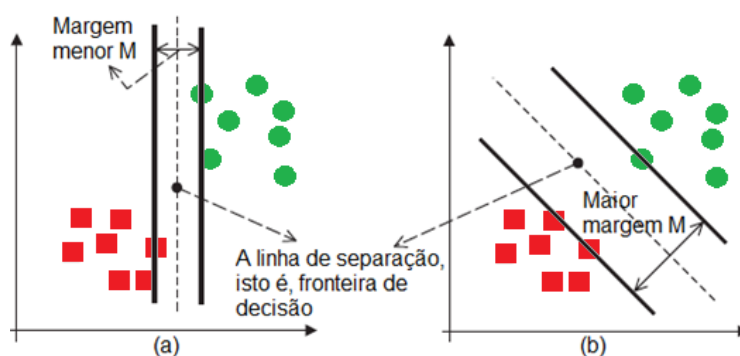
$$f(x) = \underline{w}^t \underline{x} + b, \quad (2)$$

onde $\underline{w} \in \mathbb{R}^n$ é o vetor de pesos, \underline{w}^t é o vetor transposto de \underline{w} e $b \in \mathbb{R}$ é chamado de *bias*.

Caso o espaço de características tenha dimensão dois, o hiperplano separador obtido é uma reta. Caso o espaço de características tenha dimensão n o hiperplano separador obtido é um hiperplano de dimensão $n-1$. O hiperplano separador divide o espaço de características em dois subespaços, sendo um denominado classe -1 e o outro classe $+1$. Um padrão \underline{x}_i é considerado classificado corretamente se ele estiver fora da margem de separação de sua classe, ou seja, se $y_i = +1$ então deve satisfazer $(\underline{w}^t \underline{x} + b \geq +1)$ e se $y_i = -1$ então deve satisfazer $(\underline{w}^t \underline{x} + b \leq -1)$.

O processo de treinamento do SVM consiste na obtenção de valores para os pesos w e do termo *bias* b de forma a maximizar a distância entre as margens. Desta maneira, o SVM se torna robusto a pequenas variações no conjunto de treinamento, possibilitando uma melhor generalização (ABREU, 2016).

Figura 1 – Diferença de margens de separação



Fonte: Abreu, 2016

A Figura 1 apresenta duas separações lineares e também as margens. Para a Figura 2a, a separação não possui ainda a margem máxima passível de ser calculada, ou seja, é considerada não ótima, já a Figura 1b apresenta, para o mesmo conjunto de dados de treinamento, margem máxima possível para esses dados de treinamento, considerada uma separação ótima. Considerando as margens do hiperplano de separação como $\underline{w}^t \underline{x} + b = +1$ e $\underline{w}^t \underline{x} + b = -1$, almeja-se obter a maior distância entre as margens para encontrar a melhor generalização e robustez do modelo do SVM.

O problema de minimização do método SVM normalmente é resolvido por meio de seu modelo dual. Segundo Morales e Simari (2010) apenas alguns trabalhos consideram o resolver diretamente o SVM pelo seu modelo primal.

Assim, considerando o vetor $\underline{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_l)$ com os l multiplicadores de Lagrange e obtém-se então a função a ser maximizada dada por:

$$\begin{aligned} \text{Max } L(\underline{w}, b, \xi, \underline{\alpha}) &= \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle \underline{x}_i, \underline{x}_j \rangle \\ \text{S.a} \quad &\sum_{i=1}^l \alpha_i y_i \geq 0 \\ &0 \leq \alpha_i \leq C \end{aligned} \quad (3)$$

O valor α_i é o Multiplicador de Lagrange correspondente ao padrão i e C é o parâmetro que limita o valor dos Multiplicadores de Lagrange no ajuste do SVM.

3. Support Vector Regression – SVR

Para problemas de regressão, foi criada uma adaptação, ou extensão, do método SVM. Tal extensão se dá pela introdução de uma função de perda, denominada do inglês *loss function*. A formulação do *Support Vector Regression* (SVR) utiliza os mesmos princípios do algoritmo de classificação (SVM), com exceção de alguns detalhes (PARRELLA, 2007).

Ao contrario do SVM, o SVR propõe determinar um hiperplano $f(x)$ ótimo em que, as amostras de treinamento estejam mais próximas possível, não importando qual dos lados da superfície os pontos se localizam, e sim que a distância para a superfície seja a mínima possível, $[y_i - \varepsilon, y_i + \varepsilon]$, onde, desvios são permitidos desde que não ultrapassem a margem especificada. Porém, mesmo com propósitos opostos, ambos buscam estabelecer uma função com máxima capacidade de generalização.

Sapankevych e Sankar (2009) destacam que a aplicação do SVR para a previsão de séries financeiras e de carga elétrica são as aplicações mais estudadas. Nas séries financeiras as análises são, mais especificamente, feitas para prever o índice do mercado financeiro. O objetivo do SVR é encontrar uma função $f(\underline{x})$, com uma margem de erro caracterizada pelo intervalo $[y_i - \varepsilon, y_i + \varepsilon]$, onde, desvios são permitidos desde que não ultrapassem a margem especificada.

O modelo matemático para o SVR trata-se de um problema de programação não linear, dado pela sua formulação dual na equação (3) (WANG e HU, 2005):

$$\begin{aligned}
 \min \quad & -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \langle x_i \cdot x_j \rangle - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) \\
 s.a \quad & \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\
 & \alpha_i, \alpha_i^* \in [0, C], \quad \forall i = 1, \dots, l
 \end{aligned} \tag{4}$$

onde α_i, α_i^* são os multiplicadores de Lagrange, e C é uma constante regularizadora.

4. Least Squares Support Vector Machine – LSSVM

Proposto por Suykens e Vandewalle (1999), LS-SVM é uma variante do método SVM clássico. *Least Squares Support Vector Machine* (máquina de vetores suporte a mínimos quadrados) mantém as mesmas características básicas e a mesma qualidade na solução encontrada que a sua predecessora. Ao contrário do SVM, o LS-SVM considera restrições de igualdade no lugar das desigualdades, com isso, resulta um algoritmo que reduz os problemas ao se aplicar a um conjunto extenso de dados (SANTOS, 2013).

Uma das maiores diferenças, é que, ao contrário do SVM que utiliza a programação quadrática para calcular seus vetores suporte, que demanda um grande tempo computacional e possuem complexidade matemática considerável, o LS-SVM utiliza um sistema de equações lineares e a função de custo por mínimos quadrados (BORIN, 2007; SHAH, 2005). Assim, o modelo primal para o problema é dado pela equação (5) a seguir (TRAFALLIS, SANTOSA, RICHMAN, 2005):

$$\begin{aligned}
 \min \quad & \frac{1}{2} w^t w + \frac{C}{2} \sum_{i=1}^n e_i^2 \\
 s.a \quad & y_i = w^t x_i + b + e_i \\
 & i = 1, 2, \dots, n
 \end{aligned} \tag{5}$$

onde w é o vetor de pesos, C é o parâmetro que penaliza erros altos e é otimizado pelo usuário, e_i são os erros mínimos em relação à reta de regressão.

Assim, tem-se o seguinte sistema de equações, originadas da aplicação do método dos mínimos quadrados (TRAFALLIS, SANTOSA, RICHMAN, 2005):

$$\begin{bmatrix} 0 & 1_n^t \\ 1_n & \Omega + I_n/C \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \tag{6}$$

com $Y = (y_1, y_2, \dots, y_n)^t$, $1_n = (1, 1, \dots, 1)^t$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^t$ e Ω é uma matriz dada pelos elementos $\Omega = \Omega_{ij} = \phi(x_i)\phi(x_j) = K(x_i, x_j)$ com $i, j = 1, \dots, n$, sendo que $K(x_i, x_j)$ é o Kernel utilizado, neste caso, o foi utilizado a função RBF (Gaussiano) para suas abordagens, dada por:

$$K(x, x_i) = \exp\left(-\frac{(x - x_i)(x - x_i)}{2\sigma^2}\right) \quad (7)$$

Optou-se por utilizar o Kernel RBF por ter apresentado os melhores valores para as regressões, pois apresentaram os menores valores para o erro *RMSE* e os maiores valores para o coeficiente de correlação *R*. Vale ressaltar que a obtenção dos valores das constantes C e σ , utilizadas no modelo SVR, e SVR Correlacionado, serão apresentados de forma completa na seção 6 adiante, uma vez que são inseridas pelo pesquisador que está elaborando o ajuste do modelo, sendo parte crucial no processo de otimização dos parâmetros da função de regressão.

E a função de regressão é dada por:

$$y = f(x) = \sum_{i=1}^n \alpha_i K(x, x_i) + b. \quad (8)$$

5. Modelo SVR Correlacionado

O desenvolvimento deste método, ou adaptação do método SVR, se deu no desenvolvimento dos programas para a tese de doutoramento do autor, sendo considerada uma grata surpresa, uma vez que a ideia inicial era verificar se, ao se multiplicar os dados iniciais pelo coeficiente de correlação, haveria algum tipo de mudança no valor da correlação entre as variáveis explicativas e a variável resposta. Após a verificação, optou-se por ajustar estes novos dados ao modelo SVR, e calcular o erro cometido e o ajuste do novo modelo.

Basicamente, antes de efetuar o treinamento, calcula-se a correlação das variáveis independentes e, assim, modifica-se o conjunto de treinamento, destacando as variáveis que possuem uma maior correlação com a variável independente, porém, ainda mantendo-se as variáveis de menor correlação no conjunto de treinamento e teste/validação. É importante ressaltar que tal modificação não altera o valor absoluto da correlação entre as variáveis.

As etapas do desenvolvimento do método seguem os seguintes passos:

- i. Calculam-se as correlações entre as variáveis independentes e a variável dependente, utilizando todos dados existentes;
- ii. Multiplicam-se as variáveis independentes pela sua respectiva correlação em relação a variável dependente;
- iii. Aplicam-se o modelo SVR aos novos dados já multiplicados pela correlação, otimizando os parâmetros C e σ ;
- iv. Apresentam-se os dados da aplicação.

Para o método proposto, usa-se o coeficiente de correlação linear, inserido na fase de ajuste, e é dado por:

$$R = \frac{\sum_{i=1}^n (y_i - \bar{y})(y_i^{(p)} - \bar{y}^{(p)})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (y_i^{(p)} - \bar{y}^{(p)})^2}} \quad (9)$$

onde y_i , \bar{y} são os valores reais e sua média, respectivamente e $y_i^{(p)}$, $\bar{y}^{(p)}$ são as previsões e sua média, respectivamente.

Para o ajuste dos parâmetros C e σ desenvolveu-se um *looping* no programa computacional a fim de obter os seus valores ótimos, assim variou-se os valores de C de 2,0 até 10.000 com σ fixo em 1,2, e ao determinar o valor de C que maximize o coeficiente de correlação e ao mesmo tempo minimize o erro $RMSE$, travou-se a variação do parâmetro C e iniciou-se a variação do parâmetro σ , que varia de 1,2 até 1.000,00, listando os valores que apresentaram menor erro $RMSE$ e maior coeficiente R .

Obtidos os possíveis valores ótimos, inicia-se a fase de ajuste fino dos parâmetros, onde os valores obtidos na fase anterior são “perturbados” gerando variações em torno dos valores obtidos, variando em unidades ou valores decimais, tanto para mais como para menos nos valores obtidos, para os dois parâmetros, a fim de obter erros menores.

Os valores de C e σ obtidos na fase de ajuste do método proposto foram usados também na fase de teste, e os resultados foram comprovadamente melhorados em relação ao método convencional. Vale ressaltar que o tempo computacional dos ajustes foi considerado relativamente baixo, na casa dos 20 segundos para o ajuste automático e não mais de um minuto na fase de ajuste fino.

6. Dados utilizados

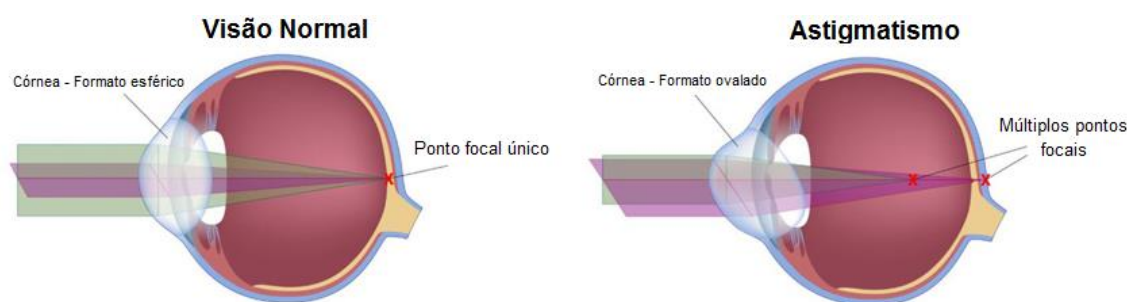
O conjunto de dados disponível corresponde aos resultados dos exames para a avaliação do grau do astigmatismo dos olhos. É composto pelo exame feito em 26 pacientes (MAZUZE, MENDES, 2014), sendo todos com grau de pelo menos 0,25 de astigmatismo, os exames ainda apresentam algumas observações qualitativas, tais como idade, sexo, profissão e olho afetado, que não entraram na modelagem do problema, devido a sua correlação quase nula com o grau de astigmatismo.

Foram determinados dois cenários diferentes para os ajustes do modelo SVR Correlacionado. No primeiro cenário trabalhou-se com 20 observações separadas das 26 disponíveis para a fase de ajuste, deixando as seis restantes para a fase de teste. Já no segundo cenário, optou-se por reservar 16 observações para a fase de ajuste e as 10 restantes foram utilizadas na de teste do modelo. Essas duas configurações diferentes foram pensadas para que fosse possível verificar a robustez do método SVR Correlacionado em diferentes configurações de conjuntos de ajuste, demonstrando que o método não foi desenvolvido apenas para um conjunto específico de dados.

Considerou-se o valor do astigmatismo como sendo a variável dependente a ser prevista com base em quatro variáveis independentes, sendo elas: Ceratometria, subdividida em Eixo Mais Plano e Eixo Mais Curvo; e Refração, subdividido em Esfera e Cilíndrico.

A Figura 2 apresenta a diferença entre a incidência da imagem em um olho sadio, ou seja, sem presença de grau astigmatismo, e um olho com certo grau de astigmatismo, exemplificando como a deformidade na córnea humana influencia na distorção da imagem projetada no fundo do globo ocular.

Figura 2 – Procedimento do método SVR Correlacionado



Fonte: Abreu, 2016

A Tabela 1 apresenta os valores da correlação entre a variável dependente Astigmatismo e as características Ceratometria, dada pelo eixo mais plano e eixo mais curvo, e a Refração, dada pelos valores esféricos e cilíndricos.

Tabela 1 – Procedimento do método SVM Correlacionado

	Ceratometria		Refração	
	Mais plano	Mais curvo	Esfera	Cilíndrico
Astigmatismo	-0,7780	0,4481	0,4953	-0,8066

Fonte: O autor, 2017

Para a avaliação do desempenho do modelo, utilizou-se o coeficiente de correlação R , a Raiz do Erro Quadrático Médio ($RMSE$ – *Root Mean Squared Error*) e a Média Percentual Absoluta do Erro ($MAPE$ – *Mean Absolute Percentage Error*), cujas equações são dadas a seguir:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^{(p)})^2} \quad (10)$$

e

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i^{(p)} - y_i}{y_i} \right) \times 100\% \quad (11)$$

onde y_i , \bar{y} são os valores reais e sua média, respectivamente e $y_i^{(p)}$, $\bar{y}^{(p)}$ são as previsões e sua média, respectivamente.

7. Resultados

7.1 Conjunto 20/6

O primeiro conjunto, onde foram separadas 20 observações para o ajuste e seis para o teste do modelo, a otimização dos parâmetros C e σ retornou os menores valores para o erro $RMSE$. Esses valores foram $C = 81,99$ e $\sigma = 2,87$ e o erro gerado foi igual à $RMSE = 0,030565$, para a aplicação do método SVR Correlacionado. Já para a abordagem convencional, os valores ótimos para os parâmetros C e σ foram $C = 39,01$ e $\sigma = 5,09$, porém o valor do erro cresceu para $RMSE = 0,07350$, se mantendo baixo, mas maior que o valor alcançado pelo SVR Correlacionado. Os valores do erro $MAPE$ são apresentados na Tabela 3 e apresentam a superioridade do método SVR Correlacionado, pois gerou um erro $MAPE$ menor que o método convencional.

A Tabela 2 apresenta os valores para as previsões feitas pelo método SVR, SVR Correlacionado e os valores reais para o grau de astigmatismo dos dados utilizados, bem como as distâncias dos resultados do método SVR aos valores reais (Distância A) e do método SVR Correlacionado (Distância B).

Tabela 2 – Resultados das previsões para o método SVR Correlacionado e comparações

SVR	SVR Correlacionado	Valores reais	Distância A	Distância B
1,6805	1,6867	1,75	0,0695	0,0633
0,6428	0,7247	0,75	0,1072	0,0253
0,9766	0,9993	1,00	0,0234	0,0007
0,5525	0,5272	0,50	0,0525	0,0272
1,0165	1,0141	1,00	0,0165	0,0141
1,3619	1,2552	1,25	0,1119	0,0052

Fonte: O autor, 2017

A Tabela 3 apresenta o resumo dos valores ótimos para os parâmetros e os melhores resultados dos erros obtidos pelo método SVR Correlacionado e SVR convencional, facilitando a comparação dos resultados obtidos para as seis previsões.

Tabela 3 – Resumo dos parâmetros e erros cometidos 20/6

Método	C	σ	$RMSE$	$MAPE$	R
SVR Correlacionado	81,99	2,87	0,030565	2,38%	0,9985
SVR	39,01	5,09	0,073501	6,95%	0,9824

Fonte: O autor, 2017

Assim, a Tabela 3 apresenta a superioridade do método correlacionado em relação ao método convencional, obtendo valores inferiores nos erros $RMSE$ e $MAPE$ e superiores no coeficiente de correlação R .

7.2 Conjunto 16/10

Para a segunda configuração do conjunto de dados o ajuste do modelo foi feito utilizando 16 observações e as 10 restantes foram utilizadas para o teste do modelo. Os valores ótimos obtidos pelo método SVR Correlacionado para os parâmetros foram $C = 3339,75$ e $\sigma = 8,91$, gerando um erro $RMSE = 0,021613$.

O modelo SVR convencional ajustado obteve os valores ótimos dos parâmetros $C = 389,7$ e $\sigma = 18,1$ gerando o erro $RMSE = 0,028518$, sensivelmente maior que o erro cometido pelo método Correlacionado, demonstrando novamente a eficiência do

método proposto. Os valores do erro MAPE para esta abordagem são apresentados na Tabela 5 apresentada adiante.

A Tabela 4 apresenta os valores, a fim de comparação, para as previsões feitas pelo método SVR, SVR Correlacionado e os valores reais para o grau de astigmatismo dos dados utilizados e as distâncias das previsões do método SVR aos valores reais (Distância A) e do método SVR Correlacionado (Distância B).

Tabela 4 – Resultados das previsões para o conjunto 15/11

SVR	SVR Correlacionado	Valores reais	Distância A	Distância B
0,2026	0,2342	0,25	0,0474	0,0158
1,2757	1,2656	1,25	0,0257	0,0156
0,3130	0,3034	0,25	0,0630	0,0534
0,9945	1,0246	1,00	0,0055	0,0246
1,7431	1,7488	1,75	0,0069	0,0012
0,7364	0,7584	0,75	0,0136	0,0084
1,0091	1,0120	1,00	0,0091	0,0120
0,4924	0,5174	0,50	0,0076	0,0174
1,0008	0,9969	1,00	0,0008	0,0031
1,2792	1,2637	1,25	0,0292	0,0137

Fonte: O autor, 2017

Diferente da aplicação para o conjunto 20/6 onde todas as distâncias das previsões SVR Correlacionado foram menores que do convencional, nesta aplicação o método SVR convencional apresentou 4 distâncias menores que as obtidas pelo método correlacionado, porém, ainda assim o erro gerado pelo convencional foi maior que o método correlacionado.

Já Tabela 5 a seguir apresenta um resumo dos parâmetros e erros cometidos pelos métodos SVR Correlacionado e SVR convencional, exemplificando os resultados obtidos na configuração que reserva 16 observações para a fase de ajuste do modelo e 10 observações para a fase de teste do modelo, demonstrando sua eficiência e servindo como um breve resumos dos resultados obtidos.

Tabela 5 – Resumo dos parâmetros e erros cometidos 16/10

Método	C	σ	$RMSE$	$MAPE$	R
SVR Correlacionado	3339,75	8,91	0,021613	6,443762	0,9997
SVR	399,71	18,09	0,028518	8,969937	0,9908

Fonte: O autor, 2017

8. Conclusão

O método *Support Vector Machine* é considerado uma técnica robusta por si só, e é utilizado tanto para classificações, como para regressões, como mostrado nas aplicações apresentadas neste artigo.

Uma das maiores dificuldades, após as implementações dos métodos desenvolvidos foi à calibração dos parâmetros C e σ para os modelos SVR convencional e SVR Correlacionado, uma vez que esta calibração foi feita de maneira automática, porém, o ajuste fino nos valores foi feito pelo autor. Assim, se pode extrair ao máximo o desempenho do método criado.

A partir da Tabela 3 fica clara a eficiência do método SVR Correlacionado frente ao método convencional, quando aplicado a um conjunto 20/6 de observações tanto seu erro *RMSE* como o erro *MAPE* foi menor que a metade do erro gerado pelo método convencional. Já a Tabela 5 apresenta a aplicação ao conjunto 16/10, também retornando valores, em sua maioria, mais próximos aos reais que o método SVR convencional, porém, a diferença entre os erros pelo método Correlacionado e o convencional diminui, mas com o método Correlacionado ainda apresentando valores menores que o método convencional.

Assim, conclui-se que o método SVR Correlacionado pode ser considerado eficiente e robusto, a adaptação da fase de ajuste, com a inclusão do coeficiente de correlação linear como uma constante penalizadora, surtiu efeito positivo, como era desejado no início do trabalho. E ainda assim, o tempo computacional gasto no novo método foi equivalente ao tempo computacional utilizado para rodar os problemas no método convencional.

Este trabalho apresenta a eficácia do método de forma empírica, não retirando a necessidade da demonstração teórica. Assim, o próximo passo é determinar as demonstrações teóricas da generalização do método, ou testa-lo a novos conjuntos de dados. Também se estuda a aplicação do método Correlacionado a problemas de reconhecimento e classificações de padrões, usando o método *Support Vector Machine*.

Referências

- ABREU, A. L. E. **Bootstrap e modelos de Support Vector Machine – SVM** (Tese de doutorado em Métodos Numéricos em Engenharia). Universidade Federal do Paraná, Curitiba, Paraná, 2016.
- BORIN, A. **Aplicações de Máquinas de vetores de suporte por mínimos quadrados (LS-SVM) na quantificação de parâmetros de qualidade de matrizes lácteas**. Tese de Doutorado (Química). Universidade Estadual de Campinas, 2007.
- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. **A Training Algorithm for Optimal Margin Classifiers**. COLT '92 Proceedings of the fifth annual workshop on Computational learning theory, pag. 144-152. ACM New York, NY, USA. 1992.
- CORTES, C. VAPNIK, V. **Support-Vector Networks**. *Machine Learning*, Volume 20, Número 3, Setembro/1995, pag. 273-297, Editora Springer, 1995.
- JONHSON, R. A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. 2nd ed., Prentice-Hall Inc., New Jersey, 1988.
- LORENA, A. C.; CARVALHO, A. C. P. L. F. **Introdução às Máquinas de Vetores Suporte (Support Vector Machines)**. Relatórios técnicos do ICMC, Instituto de Ciências Matemáticas e de Computação, USP, 2003.
- MAZUZE, A. N. F.; MENDES, R. T. **Mudanças refrativas em pós-operatório de pacientes submetidos à cirurgia de Pterígio pela técnica de transplante conjuntival** (Trabalho de conclusão de curso). Curso de Optometria. Universidade do Contestado, Unc. Canoinhas, SC. 2014.
- MORALES, A. K.; SIMARI, G. R. **Advances in Artificial Intelligence**. IBERAMIA 2010: 12th Ibero-American Conference on AI, Bahía Blanca, Argentina, November 1 – 5 , 2010.
- MOREIRA, A. T. R. **Astigmatismo**. Arq. Bras. Oftalmol (online). Vol. 64, n. 3, 2001.
- MOREIRA, S. M.; MOREIRA, L. **Adaptação de lentes de contato em ceratocone**. Ver. Sinopse da Oftalmologia. Ano 2 nº 1 São Paulo, SP 2000.
- PARRELLA, F. **Online Support Vector Regression**. Master Science Thesis. Department of Information Science, University of Genoa, Italy, 2007.
- SANTOS, L. T. **Abordagem da máquina de vetor suporte otimizada por evolução diferencial aplicada à previsão de ventos**. Dissertação de Mestrado (Engenharia Elétrica). Universidade Federal do Paraná, 2013.
- SAPANKEVYCH, N. I.; SANKAR, R. **Time Series Prediction Using Support Vector Machines: A Survey**. IEEE Computational Intelligence Magazine, University of South Florida, USA, 2009.
- SHAH, R. S. **Least Squares Support Vector Machine**. 2005.
- SUYKENS, J. A. K.; VANDEWALLE, J. **Least squares support vector machine classifiers**. Neural Processing Letters, Springer Netherlands, v. 9, p.293 – 300, ISSN 1370-4621, 1999.
- TRAFALIS, T. B.; SANTOSA, B.; RICHMAN, M. B. **Learning networks in rainfall estimation**. *Computational Management Science*. July 2005, Volume 2, Issue 3, pp 229-251.
- VAPNIK, V. N.; CHERVONENKIS, A. Y. **Theory of Pattern Recognition**. Nauka, Moscow, 1974.
- WANG, H.; HU, D. **Comparison of SVM and LS-SVM for Regression**. *Neural Networks and Brain, ICNN&B '05. International Conference on* , vol.1, no., pp.279,283, 13-15 Oct. 2005.

INFLUENCE OF CORRELATION COEFFICIENT ON REGRESSION MODELS VIA SVR

Abstract

The work shows the influence of the correlation coefficient in regression models. For development, we used the method Support Vector Regression (SVR) as the regression model. The SVR method was applied in tests done in patients who have some degree of astigmatism. Therefore, it created a change in the regression model adjusting stage, and introduced the linear correlation coefficient, evaluating the correlation between the predictor variables: Keratometry, subdivided into shaft More Plan and shaft More Curved; and Refraction, subdivided into Sphere and Cylinder, the degree of astigmatism the variable to be predicted. The proposed new method, named SVR Correlated had their results compared to the conventional method SVR, achieving superior performance both in correlation models as the value of the mistake. Altogether, we used the data from 26 patients with astigmatism, Being created two settings for the adjustment and test, the first one being composed of 20 observations for adjustment and 6 for test presenting the error $RMSE = 0.035799$, and the second composed of 15 observations in the adjustment phase and 11 in the test, generating $RMSE = 0,028244$, in both cases, lower than the errors generated by the conventional SVR method.

Key-words: Regression, Support Vector Regression, Correlation, Astigmatism.