UNIVERSIDADE FEDERAL DO PARANÁ
Adriano Verdério

SOBRE O USO DE REGRESSÃO POR VETORES SUPORTE PARA A CONSTRUÇÃO DE MODELOS EM UM MÉTODO DE REGIÃO DE CONFIANÇA SEM DERIVADAS

Curitiba, 12 de março de 2015.

UNIVERSIDADE FEDERAL DO PARANÁ
Adriano Verdério

SOBRE O USO DE REGRESSÃO POR VETORES SUPORTE PARA A CONSTRUÇÃO DE MODELOS EM UM MÉTODO DE REGIÃO DE CONFIANÇA SEM DERIVADAS

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Paraná, como requisito parcial à obtenção do Título de Doutor em Matemática Aplicada.

Orientadora: Profa. Dra. Elizabeth Wegner Karas.

Coorientador: Prof. Dr. Lucas Garcia Pedroso.

Curitiba, 12 de março 2015.

UNIVERSIDADE FEDERAL DO PARANÁ

Adriano Verdério

SOBRE O USO DE REGRESSÃO POR VETORES SUPORTE PARA A CONSTRUÇÃO DE MODELOS EM UM MÉTODO DE REGIÃO DE CONFIANÇA SEM DERIVADAS

Profa. Dra. Elizabeth Wegner Karas
Orientadora - Universidade Federal do Paraná - UFPR

Prof. Dr. Lucas Garcia Predroso
Coorientador - Universidade Federal do Paraná - UFPR

Profa. Dra. Sandra Augusta Santos
Universidade Estadual de Campinas - UNICAMP

Profa. Dra. Fernanda Maria Pereira Raupp
Laboratório Nacional de Computação Científica - LNCC

Prof. Dr. Luiz Carlos Matioli
Universidade Federal do Paraná - UFPR

Profa. Dra. Mael Sachine
Universidade Federal do Paraná - UFPR

Curitiba, 12 de março 2015.

Agradecimentos

Ao final de um ciclo sempre começamos outro. Ao colocarmos uma vírgula em nossa trajetória encontramos o momento propício para o reconhecimento do bem feito por alguém. Tantas pessoas, de várias maneiras, contribuem diariamente em nossa vida.

Não posso deixar de começar os agradecimentos à minha família, meu pai Domingos, minha mãe Luiza, meus irmãos Andréia, Alessandro, Alex e Anderson. Sempre me apoiaram e incentivaram, sempre ao meu lado, obrigado.

À Professora Elizabeth e ao Professor Lucas por todas as horas expendidas para a conclusão da tese, por todas as correções, pela visão matemática que agora é mais límpida, obrigado.

Um especial agradecimento à Lehigh University e à Professora Katya Scheinberg por me aceitar como aluno visitante durante um semestre e pela imensa contribuição, obrigado.

Aos professores que demandaram seu tempo para ler e contribuir com o trabalho durante a defesa. Professora Sandra Santos, Professora Fernanda Raupp, Professor Luis Matioli, Professora Mael Sachine, muito obrigado.

Aos professores Arinei e Luiz Eduardo, pela contribuição, obrigado.

À Universidade Federal do Paraná e ao Programa de Pós-graduação em Matemática (em meu coração para sempre PPGMA), pela oportunidade de estar em seus corredores, obrigado. À nossa indispensável Cinthia, desempenhando seu papel e contribuindo com o programa, obrigado. À Cida, à dona Elizia, obrigado.

Aos colegas do PPGMA, que durante muitos anos de convívio me ensinaram a ser companheiro, obrigado.

Aos amigos deixados em Cascavel, especialmente Suzana, Josielli e Claudemir. Não tenho palavras para dizer o quanto vocês fazem parte da minha vida, obrigado. Aos amigos mais que especiais Anderson, Joice, Viviane e Keffy. Vocês são incríveis, palavras são insuficientes para vocês, obrigado.

Aos amigos que encontrei em Curitiba e tive a honra de dividir o mesmo teto. Rafael, Zen, Zé Gui e Alex, por segurarem as pontas e aceitarem minha amizade, obrigado. Ao

irreverente Roberto, que alegra o apartamento 408 e aceita meu mau humor joinvilense, obrigado.

Aos meus amigos e às minhas amigas da Universidade Federal de Santa Catarina do Campus Joinville, apesar de não estarem todos devidamente aqui nomeados em meu coração vocês são titulares, obrigado.

Não posso deixar de agradecer ainda à Lednice e ao Ricardo, obrigado, obrigado e obrigado.

Enfim, são inúmeras as pessoas que contribuíram com um sorriso, com um abraço, com um cinema pré apresentação, com um jantar, com tantos gestos de carinho e amizade, obrigado.

"Não fosse isso
e era menos.
Não fosse tanto
e era quase."

Paulo Leminski

Resumo

Em otimização, os métodos de região de confiança a cada iteração utilizam um modelo que aproxima localmente a função a ser otimizada. Em métodos sem derivadas geralmente os modelos são construídos por interpolação polinomial. Apresentamos a construção de modelos de uma função utilizando vetores suporte, que são uma classe de métodos de aprendizagem de máquinas que podem ser utilizados para a classificação de padrões ou regressão. Apresentamos ainda modificações em um algoritmo de região de confiança livre de derivadas e sua prova de convergência. Mostramos que os modelos construídos por regressão via vetores suporte satisfazem as hipóteses necessárias para a convergência do algoritmo e podem ser utilizados como alternativa à interpolação polinomial. Experimentos numéricos preliminares são apresentados comparando o desempenho do algoritmo com modelos construídos por regressão via vetores suporte e por interpolação polinomial.

Palavras-chave: *Regressão via Vetores Suporte, Região de Confiança, Otimização Sem Derivadas.*

Abstract

In optimization, each iteration of trust-region methods uses a model that locally approximates the function to be minimized. In derivative-free methods, the models generally are built by polynomial interpolation. Alternatively, we present function models built by support vectors, a class of machine learning methods that can be used to pattern classification or regression. We also propose modifications for a derivative-free trust-region algorithm and its global convergence proof. We show that support vector regression models satisfy the assumptions required for the global convergence of the trust-region algorithm. Preliminary numerical experiments are presented to compare the performance of the algorithm using models constructed by support vectors regression and by polynomial interpolation.

Keywords: *Support Vectors Regression, Trust-Region, Derivative-Free Optimization.*

Sumário

Introdução

Métodos de região de confiança correspondem a uma classe de algoritmos para resolver problemas de otimização não linear que se baseiam em modelos para aproximar a função objetivo em uma vizinhança do ponto corrente [8]. Quando as derivadas da função objetivo estão disponíveis, os modelos podem ser construídos por aproximações pelos polinômios de Taylor. Neste caso os métodos de região de confiança podem ser vistos como uma estratégia de globalização para o método de Newton [25].

No entanto, algumas vezes calcular as derivadas de uma função é muito trabalhoso ou computacionalmente inviável. Em outras tal cálculo é impossível, uma vez que nem sempre possuímos a expressão analítica da função, quando ela provém de uma simulação, por exemplo. Mesmo sem conhecer as derivadas, algumas vezes é desejável realizar a otimização, e nesses casos utilizamos métodos sem derivadas.

O desempenho prático de métodos sem derivadas dificilmente supera o de um bom algoritmo baseado em derivadas, especialmente no que tange ao tempo computacional e número de avaliações de função, uma vez que as derivadas carregam informações importantes como inclinação e curvatura de uma função.

Conn, Scheinberg e Vicente em [10] consideram a otimização sem derivadas uma área aberta e desafiadora muito importante na ciência da computação e engenharia, que contempla um enorme potencial prático. A fonte de sua importância é a necessidade cada vez maior em resolver problemas de otimização definidos por funções cujas derivadas são indisponíveis ou disponíveis a um custo proibitivo. O aumento da complexidade em modelagem matemática, a maior sofisticação da computação científica e uma abundância de códigos relacionados são algumas das razões pelas quais a otimização sem derivadas é atualmente uma área de grande demanda.

Os métodos de região de confiança sem derivadas têm como pioneiro Winfield [51] e têm sido estudados por Powell [33, 34, 35], Conn e Toint [12], Scheinberg e Toint [41], Conn, Scheinberg e Vicente [10], Fasano, Morales e Nocedal [17], Gratton, Toint e Tröltzsch [21], entre outros. Provas de convergência para os métodos de região de confiança sem derivadas para problemas irrestritos são discutidas, por exemplo, em [9, 10, 34]. San-

tos, em [39], compila os avanços recentes dos métodos de região de confiança e apresenta ramos de interesse em otimização para os quais métodos de região de confiança podem trazer avanços com a construção de algoritmos robustos e globalmente convergentes.

Conejo et. al. [7] apresentam um algoritmo de região de confiança sem derivadas globalmente convergente para problemas com restrições em que o conjunto viável é convexo e fechado. O algoritmo é bastante geral, pois permite o uso de qualquer técnica para obtenção dos modelos, desde que sejam aproximações locais da função objetivo.

Geralmente, em algoritmos de região de confiança sem derivadas, os modelos são construídos por interpolação polinomial. O objetivo deste trabalho é apresentar uma alternativa à interpolação construindo os modelos por regressão via vetores suporte e provar que os resultados de convergência propostos em [7] permanecem válidos.

As máquinas de vetores suporte são uma classe de algoritmos de aprendizagem de máquinas motivada por resultados da Teoria de Aprendizagem Estatística [50]. A aprendizagem de máquinas é uma área interdisciplinar que visa a construção de métodos computacionais que sejam capazes de *aprender* com dados, no sentido de fazer classificações ou predições relevantes, como por exemplo identificar *spams* entre e-mails, ordenar páginas de busca da internet, traduzir textos automaticamente, entre outras aplicações. A aprendizagem de máquinas conta hoje com vários trabalhos publicados por diversos autores [1, 18, 23, 24, 27, 30, 53].

As máquinas de vetores suporte formam uma pequena parte desse universo, surgindo para classificação de padrões e posteriormente sendo estendidas para a regressão de funções. As máquinas de vetores suporte recentemente atraíram bastante atenção na comunidade de aprendizagem de máquinas e de otimização devido aos excelentes resultados de generalização. Maiores informações sobre máquinas de vetores suporte podem ser encontradas nos trabalhos de Burges [5], de Schölkopf e Smola [42] e de Vapnik [50]. O trabalho de Pontil, Rifkin e Evgeniou [32] traz um estudo relacionando máquinas de vetores suporte para a regressão e para classificação, onde é mostrado que para uma solução do problema de classificação existe uma solução para o problema de regressão que é equivalente para uma certa escolha de parâmetros.

Outra classe de algoritmos de aprendizagem de máquinas que tem chamado atenção são as chamadas máquinas de centro analítico (*analytic center machines*). Raupp e Svaiter em [37] apresentam uma nova formulação baseada no método de pontos interiores para uma máquina de centro analítico. Malyscheff e Trafalis em [26] formulam uma máquina de centro analítico para regressão, o que sugere uma investigação sobre a obtenção de modelos utilizando essa técnica em métodos de região de confiança.

A principal contribuição deste trabalho consiste em mostrar que as condições neces-

sárias apresentadas pelos modelos para garantir a convergência de métodos de região de confiança sem derivadas são asseguradas quando os modelos são construídos usando máquinas de vetores suporte. Apresentamos ainda uma modificação no algoritmo proposto em [7], bem como a análise de convergência revista para este caso.

O texto está organizado da seguinte maneira. No Capítulo 1 são discutidos os conceitos de máquinas de vetores suporte e regressão por vetores suporte. O Capítulo 2 é dedicado à aplicação da técnica de regressão por vetores suporte na construção de modelos lineares e quadráticos de uma função. Provamos que tais modelos, como os modelos obtidos por interpolação polinomial, são boas aproximações locais para a função, no sentido que existem limitantes para a norma da diferença do gradiente da função e dos modelos. No Capítulo 3, discutimos um método de região de confiança para minimização de uma função restrita a um conjunto convexo e fechado, sem fazer uso de suas derivadas. O algoritmo é bastante geral e sua convergência global é provada independente de como os modelos são obtidos, desde que satisfaçam propriedades como as discutidas no Capítulo 2. Experimentos numéricos preliminares são apresentados no Capítulo 4. Por fim, apresentamos algumas conclusões e possibilidades para trabalhos futuros.

Notações

As seguintes notações serão utilizadas durante o trabalho.

x_i: i-ésima componente do vetor x.

$\|\cdot\|$: norma euclidiana, ou seja, $\|\cdot\|_2$.

$\|x\|_\infty$: $\max_{1\leq i\leq n}|x_i|$ onde $x\in\mathbb{R}^n$.

$B(y,\delta)$: $\{x\in\mathbb{R}^n \mid \|x-y\|\leq\delta\}$.

e: vetor cujas componentes são todas iguais a 1, com dimensão dependendo do contexto.

$x\leq a$: $x_i\leq a$ para todo $i=1,\ldots,n$ com $x\in\mathbb{R}^n$ e $a\in\mathbb{R}$.

$f(X)$: vetor cuja i-ésima componente é $f(x^i)$ onde $x^i\in X$.

Capítulo 1

Máquinas de vetores suporte

Neste capítulo, faremos uma breve revisão da literatura sobre aproximações de funções por máquinas de vetores suporte. A primeira parte do capítulo é dedicada à introdução a aprendizagem de máquina, a fim de compor um quadro geral dessa área. Em seguida, falamos sobre máquinas de vetores suporte, uma classe de métodos de aprendizagem que pode ser utilizada para classificação ou regressão. A regressão via vetores suporte apresenta papel fundamental no nosso trabalho.

Os trabalhos de Alpaydin [1], Flach [18], Harrington [23], Mohri, Rostamizadeh e Talwalkar [27], Murphy [30], Sammut e Webb [40] e Winkler, Niranjan e Lawrence [52] podem ser consultados para maiores detalhes sobre a aprendizagem de máquina. Quanto às máquinas de vetores suporte, também podem ser consultados os trabalhos de Cristianini e Shawe-Taylor [13] e de Schölkopf e Smola [42].

1.1 Aprendizagem de máquina

O advento do computador permitiu armazenar e processar uma grande quantidade de informações, bem como acessá-las de praticamente qualquer lugar através de computadores interligados. Com essa grande quantidade de dados surgiu a aprendizagem de máquina (do termo *Machine Learning*, em inglês). Seu objetivo é *aprender* a partir dessa imensa quantidade de dados.

Aprendizagem de máquina não é tentar ter uma conversa com programas de computadores, nem mesmo perguntar a um computador qual é o sentido da vida. Para Harrington [23], a aprendizagem de máquina não busca a criação de seres conscientes, mas sim ter uma visão a partir de um conjunto de dados, de modo que o computador faça predições e encontre padrões a partir desses dados. Flach [18] define a aprendizagem

de máquina como tudo que envolve usar as características certas para construir o modelo certo que alcança o alvo certo.

Em problemas de aprendizagem de máquina, tentamos descobrir uma estrutura nos dados. Um dos requisitos em problemas de aprendizagem é especificar o que exatamente queremos alcançar, minimizar, limitar ou aproximar.

A aprendizagem de máquina é uma área interdisciplinar com intersecção entre Computação, Estatística e Matemática e pode ser aplicada a várias outras áreas, da Política à Geociência, passando pela Engenharia, Medicina, entre outras.

Segundo Murphy [30], um objetivo da aprendizagem de máquina é desenvolver algoritmos e métodos que possam automaticamente reconhecer padrões nos dados, e então usá-los para fornecer informações sobre dados futuros. Aqui podemos salientar a diferença entre "conhecimento prévio" e "aprendizagem com os dados". Por exemplo, se quisermos um método para reconheçer a letra 'A', podemos criar um algoritmo no qual dizemos como a letra 'A' se parece (dois segmentos de reta inclinados com um ponto em comum na parte superior e um segmento de reta horizontal na parte central que liga os dois segmentos inclinados, por exemplo) e então o computador irá classificar uma entrada futura que se enquadra nessa descrição como uma letra 'A'. Por outro lado, podemos tomar 10.000 exemplos de letras 'A' e usar algum método de aprendizagem de máquina para decidir quais são as regras que fazem um 'A' ser um 'A'. Ou seja, a priori um método de aprendizagem de máquina não precisa saber as características relevantes para definir o que é ou não a letra 'A'.

Harrington [23] enumera alguns passos necessários para aplicar um método de aprendizagem de máquina a um problema prático:

1 - Coleta de Dados. É a primeira fase, onde são coletadas informações julgadas relevantes para o problema que se pretende resolver: pode ser coletando informações por meio de observações, pesquisas ou questionários; pode ser utilizando um equipamento que meça a velocidade do vento, a quantidade de glicose no sangue, ou qualquer ente mensurável ou classificável.

2 - Preparação dos Dados. Nesta etapa, os dados coletados são preparados para que o computador possa entendê-los.

3 - Análise dos Dados. Essa é uma análise humana, apenas para verificar se nos dados já se consegue encontrar um padrão ou algum dado totalmente discrepante. Essa etapa pode ser suprimida, quando por exemplo não se tem muito conhecimento sobre o objeto de estudo ou quando a quantidade de dados é humanamente intratável.

4 - Treinamento do Algoritmo. Com posse dos dados computacionalmente compreensíveis, é aplicada alguma das técnicas de aprendizagem a um subconjunto dos dados chamado conjunto de treinamento.

5 - Teste do Algoritmo. Uma vez treinado, o algoritmo deve ser testado em dados para os quais seja possível avaliar o erro cometido. Caso a quantidade de erros não seja adequada, o algoritmo precisa ser treinado novamente com outro conjunto de treinamento ou com outro método de aprendizagem.

6 - Uso do Algoritmo. Com um algoritmo que apresente poucos erros na etapa anterior, pode-se supor que, na prática, também cometerá poucos erros.

Murphy [30] divide a aprendizagem de máquina em duas grandes classes, aprendizagem supervisionada e aprendizagem não supervisionada. Em ambos os casos a ênfase está nos dados. O que as difere é que na primeira, além de fornecer os dados, também é informado o rótulo com o qual cada dado é classificado. Já na segunda, o objetivo é encontrar um padrão apenas com os dados.

Para classificar as letras 'A' no exemplo anterior com aprendizagem supervisionada, são fornecidas várias letras 'A' rotuladas como tal. Já na aprendizagem não supervisionada, apenas são fornecidas letras diversas e o algoritmo deve decidir como rotulá-las.

Entre os métodos de aprendizagem supervisionada, pode-se citar o Algoritmo dos Vizinhos mais Próximos (*K-Nearest Neighbors Algorithm*) [23], aprendizagem por Árvore de Decisão (*Decision Tree Learning*) [23], máquinas de vetores suporte (*Support Vector Machines*) [42], entre outros. Já entre os métodos de aprendizagem não supervisionada, pode-se citar o Algoritmo Agrupamento de K-Médias (*K-Means Clustering Algorithm*) [23] e o Algoritmo Apriori [23].

1.2 Máquinas de vetores suporte

As máquinas de vetores suporte são uma classe de algoritmos de aprendizagem supervisionada motivados por resultados da teoria de aprendizagem estatística [49]. Esses resultados foram usados para a classificação de padrões, em que encontramos um limitante de representação[1] em termos de um subconjunto, geralmente pequeno, do conjunto de amostra, sendo os elementos desse subconjunto os chamados *vetores suporte*, que dão nome à técnica.

[1]No decorrer do trabalho, ainda usaremos as denominações classificador e preditor para nomear a aplicação a ser utilizada para classificar os dados futuros.

Os vetores suporte são assim chamados pois, entre todos os pontos do conjunto amostral, são os que possuem papel relevante, no sentido que caso os demais pontos sejam retirados da amostra, o classificador não muda. Ao representar o classificador por meio de poucos pontos, conseguimos uma certa esparsidade e com isso o trabalho computacional para classificar futuros dados torna-se menor, uma vantagem bastante importante.

O trabalho de Pontil, Rifkin e Evgeniou [32] traz um estudo relacionando máquinas de vetores suporte para a regressão e para classificação, onde é mostrado que para uma solução do problema de classificação existe uma solução para o problema de regressão que é equivalente para uma certa escolha de parâmetros. Uma consequência direta deste resultado é que o caso para classificação pode ser visto como um caso especial do problema para regressão.

Schölkopf, Smola, Williamson e Bartlett em [43] propõem uma nova classe de métodos de máquinas de vetores suporte tanto para a regressão quanto para a classificação que possibilita o controle do número de vetores suporte. Neste trabalho descrevem o algoritmo, apresentam alguns resultados teóricos e também trazem alguns resultados computacionais.

Discutiremos a seguir as técnicas de vetores suporte tanto para classificação quanto para regressão para nos aproximar do nosso objetivo, que é utilizar a regressão por vetores suporte para a construção de modelos em métodos de região de confiança livre de derivadas.

1.2.1 Máquinas de vetores suporte para classificação

Suponha que são conhecidas duas ou mais classes de objetos. Ao encontrarmos um novo objeto, gostaríamos de decidir em qual classe este deve ser classificado. Esse é essencialmente o problema de classificação de padrões.

Para resolver esse problema, os objetos são reduzidos a conceitos abstratos. Sejam $\mathcal{X}$ e $\mathcal{Y}$ subconjuntos de espaços vetoriais normados, usualmente $\mathbb{R}^n$ e $\mathbb{R}$, respectivamente. Suponha que é dado algum conjunto de entradas $X = \{x^1, x^2, \ldots, x^p\} \subset \mathcal{X}$ e seus respectivos rótulos $Y = \{y^1, y^2, \ldots, y^p\}$, onde $y^i \in \mathcal{Y}$ para todo $i = 1, 2, \ldots, p$, de forma que X e Y sejam independentes e identicamente distribuídos[2] de acordo com alguma medida de probabilidade $\mathbb{P}(x, y)$.

Chamaremos X o conjunto de amostra e Y o conjunto de rótulos. Os dados futuros

[2]Uma sequência ou uma coleção de variáveis aleatórias é independente e identicamente distribuída se cada variável aleatória tem a mesma distribuição de probabilidade que as outras e todas são mutuamente independentes.

que desejamos classificar devem pertencer ao conjunto $\mathcal{X}$ e os rótulos futuros devem pertencer a $\mathcal{Y}$. Nosso classificador será uma aplicação h que leva os pontos de $\mathcal{X}$ em $\mathcal{Y}$.

Classificação binária

Na classificação binária, os dados podem ser divididos em apenas duas classes, digamos ± 1. Assim, neste caso, $\mathcal{Y} = \{-1, 1\}$. Ou seja, nosso conjunto de amostra contém p pontos $x^i \in \mathbb{R}^n$ e para cada $i = 1, \ldots, p$ conhecemos $y^i \in \{-1, 1\}$ e queremos classificar futuros pontos do $\mathbb{R}^n$ entre os dois rótulos.

Classificação binária perfeitamente linear

Dentro da classificação binária o caso mais simples é a classificação linear, no qual o preditor h é um hiperplano. O hiperplano será nosso limitante de representação, todos os dados que estão em um lado pertencem a uma classe e todos os outros pontos que estão do outro lado pertencem à outra classe.

Se os dados de amostra são perfeitamente separáveis, precisamos além de encontrar o hiperplano, definir uma região onde nenhum ponto dos dados de amostra estará. Esta região é chamada de margem.

Uma máquina de vetores suporte para classificação binária linear busca encontrar um hiperplano que separe perfeitamente os dados de cada uma das duas classes e cuja margem de separação seja máxima, chamado hiperplano ótimo.

Queremos encontrar $w \in \mathbb{R}^n$ e $b \in \mathbb{R}$ que definem o hiperplano $w^\top x + b = 0$, que por sua vez definirá nosso preditor $h(x) = w^\top x + b$, de modo que os pontos futuros $x \in \mathbb{R}^n$ poderão ser classificados na classe $+1$ caso $h(x) > 0$ ou na classe -1 caso $h(x) < 0$.

Note que é possível encontrar diferentes valores para w e b que definem o mesmo hiperplano. Para garantir a unicidade são impostas algumas condições. Inicialmente, exigimos que:

$$\begin{cases} w^\top x^i + b \geq 1, & \text{para todo } i \text{ tal que } y^i = 1 \\ w^\top x^i + b = 1, & \text{para pelo menos um } x^i \in X \\ w^\top x^i + b \leq -1, & \text{para todo } i \text{ tal que } y^i = -1 \\ w^\top x^i + b = -1, & \text{para pelo menos um } x^i \in X. \end{cases} \tag{1.1}$$

As desigualdades acima podem ser reescritas como $y^i(w^\top x^i + b) \geq 1$, para todo $i \in \{1, \ldots, p\}$.

A distância de um ponto $x^i \in X$ ao hiperplano definido por (w, b) é dada por

$$d_{(w,b)}(x^i) = \frac{|w^\top x^i + b|}{\|w\|} = \frac{y^i(w^\top x^i + b)}{\|w\|} \geq \frac{1}{\|w\|},$$

em que $\|.\| = \|.\|_2$ é a norma euclidiana.

Os pontos que satisfazem as igualdades em (1.1) definirão a margem, portanto a largura da margem será

$$\rho = \frac{2}{\|w\|}.$$

Para um método de reconhecimento de padrões ser eficaz, deve apresentar pelo menos duas características:
- boa capacidade de *generalização*, permitindo que dados semelhantes sejam igualmente classificados;
- boa capacidade de *discriminação*, que assegura a correta separação entre as classes.

Para que os pontos estejam corretamente classificados e a capacidade de classificar futuros dados seja alcançada com menos erros, queremos que a margem seja a maior possível. Para conseguirmos a margem máxima, basta encontrarmos o vetor w com menor norma. Isto juntamente com (1.1) garante a unicidade do hiperplano. Ou seja, precisamos resolver o seguinte problema de otimização,

$$\begin{array}{ll} \underset{(w,b)}{\text{minimizar}} & \|w\| \\ \text{sujeita a} & y^i(w^\top x^i + b) \geq 1, \quad \forall\ i \in \{1, \ldots, p\}. \end{array}$$

Ao minimizar $\|w\|$, certamente as igualdades em (1.1) serão atingidas em pelo menos dois pontos da amostra, um ponto para cada classe. O problema anterior é equivalente ao problema convexo,

$$\begin{array}{ll} \underset{(w,b)}{\text{minimizar}} & \frac{1}{2}\|w\|^2 \\ \text{sujeita a} & y^i(w^\top x^i + b) \geq 1, \quad \forall\ i \in \{1, \ldots, p\}, \end{array}$$

que é mais facilmente tratável.

Os pontos da amostra que satisfazem a igualdade nas restrições do problema acima são os vetores suporte. Se os retirarmos do conjunto de amostra a solução muda, diferentemente dos outros pontos que, caso sejam retirados, o hiperplano separador permanece

o mesmo.

Na Figura 1.1 temos um exemplo em $\mathbb{R}^2$ que mostra duas classes de pontos perfeitamente linearmente separadas, o hiperplano ótimo e a margem de separação. Os vetores suporte são os pontos que estão nos limitantes da margem, mais precisamente os hiperplanos $w^\top x + b = +1$ e $w^\top x + b = -1$.

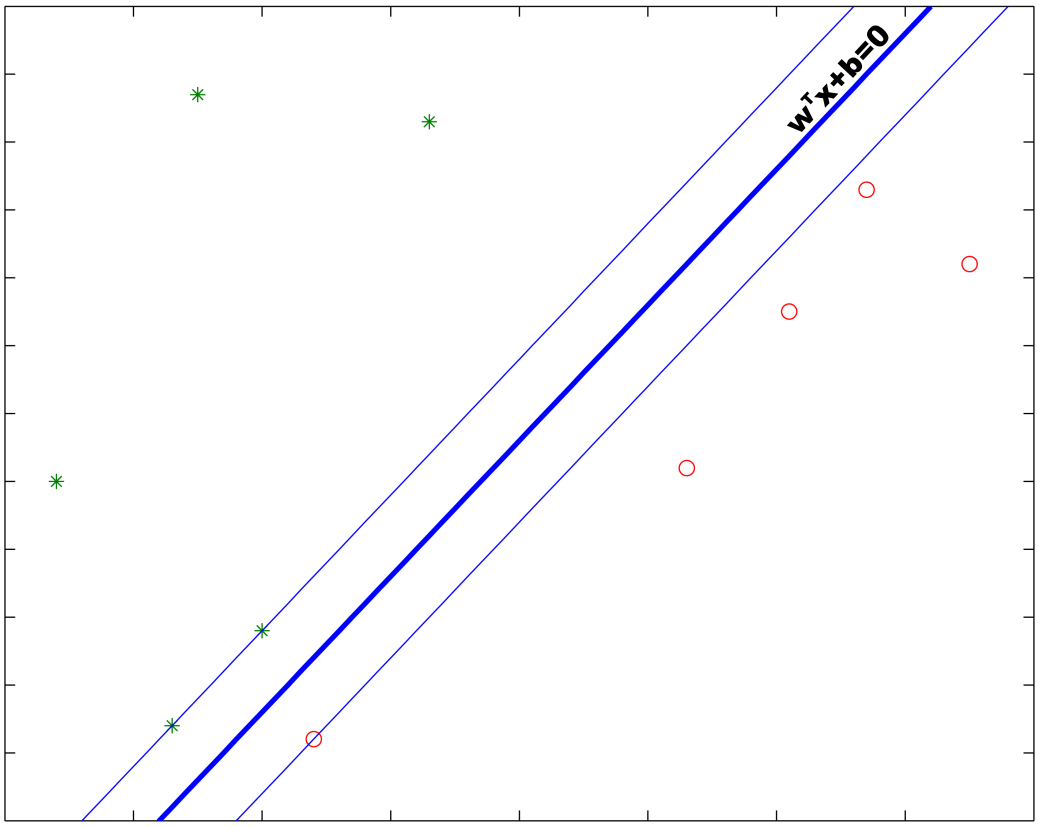

Figura 1.1: Classificação binária linear para um conjunto perfeitamente separável.

Geralmente usamos a formulação dual Lagrangiana para resolver o problema de otimização das máquinas de vetores suporte. Burges em [5] enumera duas razões para isso. A primeira é que as restrições serão trocadas por restrições mais simples. A segunda é que nesta formulação do problema os dados de amostra aparecem apenas em produto interno entre vetores, propriedade de fundamental importância, pois nos permite generalizar o mesmo procedimento para o caso não linear.

O problema dual para o caso em que os dados são perfeitamente linearmente separáveis é

$$\begin{aligned} \underset{\alpha}{\text{maximizar}} \quad & \sum_{i=1}^{p} \alpha_i - \frac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_i \alpha_j y^i y^j (x^i)^\top x^j \\ \text{sujeita a} \quad & 0 \leq \alpha_i, \quad \forall\, i \in \{1, \ldots, p\} \\ & \sum_{i=1}^{p} \alpha_i y^i = 0. \end{aligned}$$

Maiores detalhes sobre a formulação dual do problema serão abordados no caso das máquinas de vetores suporte para regressão. Detalhes sobre dualidade podem ser encontrados nos trabalhos de Bazaraa, Sherali e Shety [3] e de Fletcher [19], e sobre a formulação dual para o problema de classificação por vetores suporte nos trabalhos de Burges [5] e de Schölkopf e Smola [42, 44]. Em [45], Sra, Nowozin e Wrigth fazem uma leitura da otimização aplicada em aprendizagem de máquinas.

Classificação binária linear com margem maleável

Mesmo que os dados de amostra não sejam perfeitamente separáveis, podemos usar a classificação linear se permitirmos que alguns dados do conjunto de amostra sejam classificados de maneira incorreta. Neste caso queremos que os erros de classificação sejam os menores possíveis no conjunto de amostra. Esse caso é conhecido na literatura como *soft margin*, pois permitimos que a margem seja maleável.

Essa maleabilidade é conseguida acrescentando folgas ξ_i em cada uma das restrições do problema original e penalizando essas folgas na função objetivo, ou seja, dado $C > 0$, consideramos o problema

$$\begin{array}{ll} \underset{(w,b,\xi)}{\text{minimizar}} & \dfrac{1}{2}\|w\|^2 + C\displaystyle\sum_{i=1}^{p}\xi_i \\ \text{sujeita a} & y^i(w^\top x^i + b) \geq 1 - \xi_i, \quad \forall\, i \in \{1,\ldots,p\} \\ & \xi_i \geq 0, \quad \forall\, i \in \{1,\ldots,p\}. \end{array}$$

O parâmetro de penalização C faz a ligação entre a maximização da margem e a minimização do erro permitido. Quanto maior o valor de C, menos permitimos erros de classificação nos dados de amostra.

Novamente a solução do problema é encontrada pela formulação dual, que nesse caso é

$$\begin{array}{ll} \underset{\alpha}{\text{maximizar}} & \dfrac{1}{2}\displaystyle\sum_{i=1}^{p}\sum_{j=1}^{p}\alpha_i\alpha_j y^i y^j (x^i)^\top x^j \\ \text{sujeita a} & 0 \leq \alpha_i \leq C, \quad \forall\, i \in \{1,\ldots,p\} \\ & \displaystyle\sum_{i=1}^{p}\alpha_i y^i = 0. \end{array} \tag{1.2}$$

Apesar de não apresentarmos os detalhes da formulação dual, como já dito, precisamos dessa formulação para abordarmos o caso em que os dados são melhores classificados se utilizarmos separadores não lineares.

Classificação binária não linear

Vamos agora resolver o problema no caso em que os dados são separáveis por um classificador não linear. No conjunto de amostra da Figura 1.2 podemos ver que não existe uma reta que separe perfeitamente os dados, mas uma quadrática os separam.

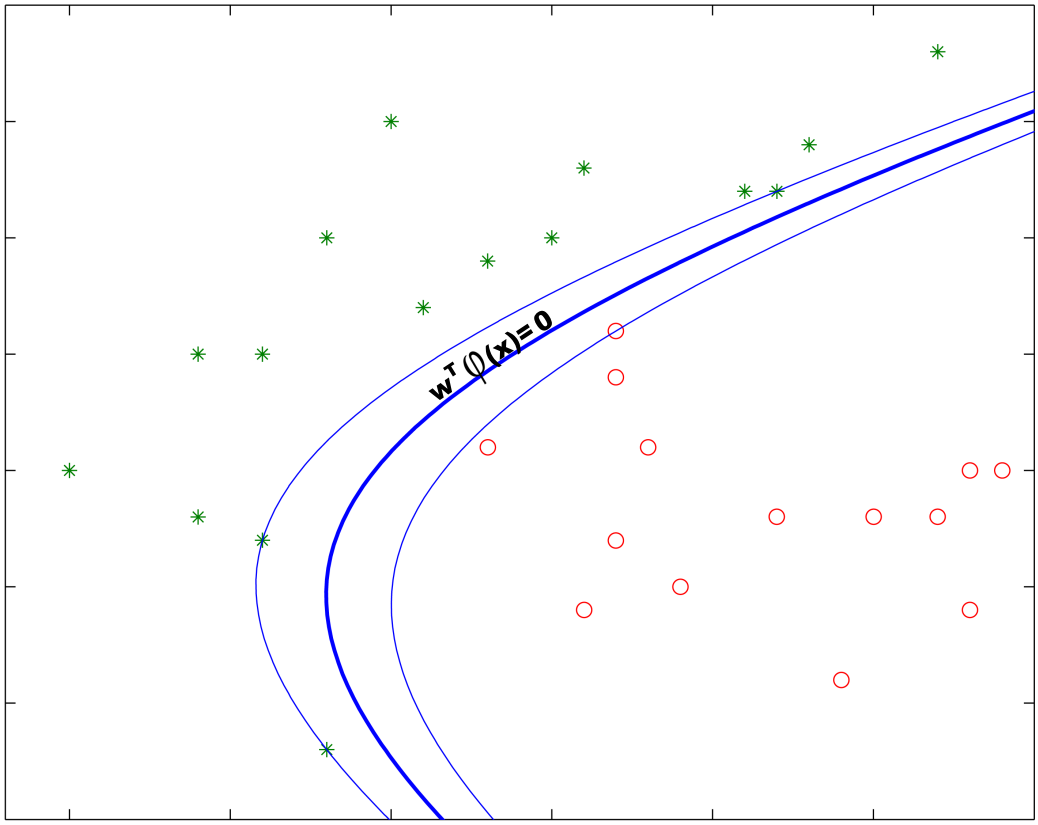

Figura 1.2: Classificação binária não linear para um conjunto perfeitamente separável.

Para resolver o problema nesse caso, levamos os dados a um espaço $\mathcal{H}$ de dimensão maior através de uma aplicação $\varphi : \mathcal{X} \to \mathcal{H}$. Este espaço é conhecido como espaço de características. Uma vantagem é que o espaço de características não precisa ser construído explicitamente.

Observando a formulação dual (1.2), podemos ver que os dados de amostra aparecem apenas na forma de produto interno, e quando mudamos a dimensão dos dados de amostra não efetuamos nenhuma mudança no rótulo. Ou seja, apenas precisamos saber como calcular produtos internos no espaço de características. Para esses cálculos usamos as chamadas funções *kernel*.

Definição 1.1. *Dada uma função $\varphi : \mathcal{X} \to \mathcal{H}$ com $\mathcal{H}$ um espaço com produto interno, definimos como uma classe de kernels a aplicação $\kappa(x, x') = \langle \varphi(x), \varphi(x') \rangle$.*

Como já observamos, a chave está no cálculo do produto interno[3] $(x^i)^\top x^j$. Vamos supor, por exemplo, que nossos dados de amostra estão no $\mathbb{R}^2$ e queremos levar esses dados a um espaço de características de dimensão três por meio da relação

$$\varphi(x^i) = ((x_1^i)^2, (x_2^i)^2, \sqrt{2}x_1^i x_2^i)^\top.$$

Procedemos calculando o produto interno

$$\varphi(x^i)^\top \varphi(x^j) = (x_1^i x_1^j)^2 + (x_2^i x_2^j)^2 + 2x_1^i x_1^j x_2^i x_2^j = ((x^i)^\top x^j)^2,$$

de onde vemos que não precisamos explicitamente construir o espaço de características. Podemos simplesmente definir

$$\kappa(x^i, x^j) = ((x^i)^\top x^j)^2$$

e substituir o produto interno por $\kappa(x^i, x^j)$.

Essas ideias podem ser utilizadas para definirmos *kernels* polinomiais de qualquer dimensão, ou seja,

$$\kappa(x^i, x^j) = ((x^i)^\top x^j)^s, \quad s \in \mathbb{N}.$$

Mais ainda, se incluirmos uma constante

$$\kappa(x^i, x^j) = ((x^i)^\top x^j + 1)^s$$

consideramos os termos de todas as ordens. Por exemplo, seja $\mathcal{X} = \mathbb{R}^2$ e $\mathcal{H} = \mathbb{R}^6$. Podemos construir, nesse caso, φ da seguinte maneira,

$$\varphi(x) = (1, \sqrt{2}x_1, \sqrt{2}x_2, x_1^2, \sqrt{2}x_1x_2, x_2^2)^\top,$$

assim

$$\begin{aligned} \varphi(x)^\top \varphi(y) &= 1 + 2x_1y_1 + 2x_2y_2 + x_1^2y_1^2 + 2x_1x_2y_1y_2 + x_2^2y_2^2 \\ &= (1 + x_1y_1 + x_2y_2)^2 = (1 + x^\top y)^2, \end{aligned}$$

e com isso $\kappa(x, y) = (1 + x^\top y)^2$.

Dessa maneira, a quadrática

$$h(x) = a_0 + a_1x_1 + a_2x_2 + a_3x_1^2 + a_4x_1x_2 + a_5x_2^2$$

[3]Estamos interessados no caso em que $\mathcal{X}$ é um subespaço de $\mathbb{R}^n$, com produto interno $\langle x, y\rangle$ usual denotado por $x^\top y$. Na definição de classe de *kernels*, produtos internos gerais podem ser utilizados.

pode ser construída por um classificador quadrático da forma

$$h(x) = w^\top \varphi(x),$$

com $w = \left(a_0, a_1/\sqrt{2}, a_2/\sqrt{2}, a_3, a_4/\sqrt{2}, a_5\right)^\top$. Como h é linear em w, podemos, através do procedimento descrito, separar os dados de amostra no $\mathbb{R}^2$ por uma quadrática usando as ideias da separação linear.

As funções *kernel* não estão restritas às formas polinomiais. Temos, por exemplo, o *kernel* gaussiano

$$\kappa(x^i, x^j) = \exp\left(\frac{-\|x^i - x^j\|^2}{2\sigma^2}\right),$$

onde o parâmetro σ é conhecido como largura de banda.

Outros exemplos de função *kernel* e também detalhes sobre quais propriedades uma função precisa satisfazer para ser um *kernel* podem ser encontrados no trabalho de Schölkopf e Smola [42].

Classificação não binária

Quando a quantidade de classes é superior a duas, cada ponto do conjunto de amostra pode ser associado a um ou mais entre k possíveis rótulos [18]. Nesse caso, temos duas alternativas principais: a técnica um-contra-todos e a técnica um-contra-um.

Na técnica um-contra-todos, usamos a classificação binária definindo uma classe com o rótulo $+1$ e todas as demais com o rótulo -1. Assim, se tivermos, por exemplo, as classes C_1, C_2, C_3, C_4 e C_5, temos $k = 5$ e precisamos de 4 classificadores:

- C_1 *versus* C_2, C_3, C_4 e C_5;
- C_2 *versus* C_3, C_4 e C_5;
- C_3 *versus* C_4 e C_5;
- C_4 *versus* C_5.

Logo, para k classes precisamos de $k - 1$ preditores.

Na técnica um-contra-um, novamente dividimos o problema em vários problemas de classificação binária, mas não unimos classes. Por exemplo, para $k = 5$, fazemos C_i *versus* C_j para todos $i = 1, \ldots, 5$ e $j = i + 1, \ldots, 5$. Nesse caso precisamos de $k(k-1)/2$ preditores.

Como a ênfase do nosso trabalho reside no uso das máquinas de vetores suporte para regressão, não apresentamos todos os detalhes para a classificação. Maiores informações

sobre esses resultados podem ser encontrados nos trabalhos de Burges [5], de Schölkopf e Smola [42] e de Vapnik [50].

1.2.2 Máquinas de vetores suporte para regressão

As máquinas de vetores suporte para regressão diferem da técnica de classificação no sentido de que, enquanto a segunda apenas busca dividir os dados em diferentes classes e classificar corretamente os dados futuros, a primeira busca encontrar um preditor que aproxime bem os dados. Na classificação, os dados podem estar distantes do preditor, já as máquinas de vetores suporte para regressão são construídas para encontrar uma função $h : \mathcal{X} \longrightarrow \mathcal{Y}$ que aproxima bem os dados de amostra.

A versão de uma máquina de vetores suporte para regressão foi proposta em 1996 por Drucker et al. em [16]. Nesse caso, o método é chamado de regressão por vetores suporte (do inglês, *support vectors regression*).

Diferentemente do caso para classificação, no qual partimos de uma ideia intuitiva, para a regressão vamos partir de alguns conceitos mais formais. Primeiramente, precisamos encontrar um preditor que apresente o menor erro dentre um conjunto de preditores. Para isso precisamos definir o que é o erro cometido por um preditor e uma medida para avaliar seu desempenho. Uma maneira natural de mensurar o erro cometido é através de uma função perda, que mede o erro que um preditor comete no conjunto de amostra.

Definição 1.2. *[42, Definição 3.1] Seja a tripla $(x, y, h(x)) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Y}$ que consiste de um padrão x, um rótulo y e a predição $h(x)$. Então uma aplicação $\ell : \mathcal{X} \times \mathcal{Y} \times \mathcal{Y} \to [0, \infty)$ com a propriedade $\ell(x, y, y) = 0$ para todo $x \in \mathcal{X}$ e $y \in \mathcal{Y}$ é chamada de função perda.*

Um exemplo de função perda é a função perda 0-1, usada para o problema de classificação. Essa é a função perda mais simples. Ela recebe o valor 1 se um dado de amostra é classificado de maneira incorreta ou recebe o valor 0 se um dado de amostra é classificado corretamente, isto é,

$$\ell_{01}(x, y, h(x)) = \begin{cases} 0, & \text{se } y = h(x) \\ 1, & \text{caso contrário.} \end{cases}$$

Para o uso em regressão, a função perda geralmente é do tipo $\ell(x, y, h(x)) = \bar{\ell}(h(x) - y)$, com o intuito de medir a discrepância entre o valor obtido pelo preditor e o valor o qual deveria assumir.

Uma vez definida a função perda a ser utilizada, conseguimos determinar como

os erros são penalizados em cada ponto da amostra. Precisamos agora encontrar uma maneira de combinar essas penalidades locais e avaliar a qualidade de um preditor.

Uma vez que os dados utilizados são independentes e identicamente distribuídos de acordo com alguma medida de probabilidade $\mathbb{P}(x, y)$, o valor esperado da função perda é

$$R[h] = \mathbb{E}[\ell(x, y, h(x))] = \int_{\mathcal{X}\times\mathcal{Y}} \ell(x, y, h(x)) \; d\mathbb{P}(x, y),$$

que é chamado de *risco esperado*, ou simplesmente *risco*. Minimizando o risco de um preditor, encontramos o melhor candidato a representar nossos dados.

No entanto, minimizar o risco esperado de um preditor é impossível, pois não conhecemos a medida de probabilidade $\mathbb{P}(x, y)$. Para resolver este problema, Vapnik em sua Teoria de Aprendizagem Estatística [50] desenvolveu o Princípio Indutivo da Minimização do Risco Empírico, no qual o risco é determinado pelo conjunto de amostra.

O princípio indutivo da minimização do risco empírico pode ser descrito como:

1. Substituir o risco esperado pelo risco empírico

$$R_{emp}[h] = \frac{1}{p}\sum_{i=1}^{p} \ell(x^i, y^i, h(x^i)),$$

   que é a perda média encontrada no conjunto de amostra.

2. Utilizar como preditor a aplicação h que minimiza o risco empírico.

Apesar de em um primeiro momento parecer que o princípio indutivo da minimização do risco empírico resolve o problema, essa técnica sozinha não é suficiente.

Em [5], Burges discorre sobre o equilíbrio para uma dada tarefa de aprendizagem. Com uma quantidade finita de dados de amostra, o melhor desempenho de generalização será alcançado com o balanço entre a precisão alcançada no conjunto de treino e a "capacidade" da máquina, isto é, a capacidade da máquina para aprender com qualquer conjunto de treino sem erro. Uma máquina com demasiada capacidade é como um botânico com uma memória fotográfica que, ao encontrar uma nova árvore, conclui que ela não é uma árvore por ter um número diferente de folhas de qualquer coisa que ele viu antes. Uma máquina com pouca capacidade é como um irmão preguiçoso do botânico, que declara que se é verde, é uma árvore.

Minimizar o risco empírico pode gerar instabilidades numéricas ou ainda não alcançar uma boa generalização [42].

Por exemplo, vamos supor que gostaríamos de resolver o problema de regressão

usando a função perda quadrática

$$\ell(x, y, h(x)) = (y - h(x))^2.$$

Mais que isso, vamos assumir que estamos trabalhando com uma classe de funções do tipo

$$\mathcal{F} = \left\{ h(x) = \sum_{i=1}^{n} \alpha_i h_i(x) \text{ com } \alpha_i \in \mathbb{R} \right\},$$

em que h_i são aplicações do espaço de amostra $\mathcal{X}$ em $\mathbb{R}$.

Usar o Método de Minimização do Risco Empírico consiste em resolver o problema

$$\min_{h \in \mathcal{F}} R_{emp}[h] = \min_{\alpha \in \mathbb{R}^n} \frac{1}{p} \sum_{i=1}^{p} \left(y^i - \sum_{j=1}^{n} \alpha_j h_j(x^i) \right)^2.$$

Calculando a derivada de $R_{emp}[h]$ com respeito a α e definindo $F_{ij} = h_i(x^j)$, obtemos que α^* é solução do problema acima se

$$F^\top F \alpha^* = F^\top y. \tag{1.3}$$

Se $F^\top F$ possuir um número de condição grande, o sistema (1.3) é numericamente difícil de resolver para α. Nesse caso o sistema é dito mal condicionado [46]. Mais que isso, se $n > p$, ou seja, se temos mais funções h_i na base da classe de funções do que dados de amostra x^i, existe uma variedade com dimensão pelo menos $n - p$ que satisfaz (1.3). Tanto do ponto de vista teórico quanto computacional isso é indesejado.

Segundo Schölkopf e Smola [42], uma maneira de evitar esses problemas é restringir a classe de soluções possíveis a um conjunto compacto. Essa técnica foi introduzida por Tykhonov e Arsenin [47] para resolver problemas inversos e tem sido aplicada em problemas de aprendizagem com bastante sucesso, trabalhando com a função risco regularizada.

Suponhamos que o conjunto $\mathcal{F}$ dos minimizadores do risco empírico seja um conjunto compacto. Além disso, assumamos que $R_{emp}[h]$ é contínuo em h. Essa segunda hipótese é facilmente satisfeita para muitos problemas de regressão, como aqueles que usam a função perda quadrática, por exemplo.

Com essas hipóteses, a aplicação inversa do mínimo do risco empírico para seu minimizador $\hat{h}$ é contínua. E com isso existe uma aplicação inversa $h^{-1} : h(X) \to X$ que também é contínua.

Em geral, não especificamos o conjunto compacto $\mathcal{F}$ e sim adicionamos um termo de estabilização $\Gamma[h]$ à função objetivo original, que em nosso caso é o risco empírico

$R_{emp}[h]$. Ou seja, consideramos a seguinte classe de risco regularizado

$$R_{reg}[h] = R_{emp}[h] + \lambda\Gamma[h],$$

em que $\lambda > 0$ é o parâmetro regularizador que especifica o balanço entre a minimização do risco empírico e a simplicidade do nosso preditor, que é alcançado com um $\Gamma[h]$ pequeno.

Outra vantagem de se trabalhar com o risco regularizado é encontrada no Teorema de Representação, enunciado abaixo.

Teorema 1.3. *[42, Teorema 4.2] Sejam $\Gamma : [0, \infty) \to \mathbb{R}$ uma função estritamente crescente, um conjunto $\mathcal{X}$ e uma função perda $\ell : (\mathcal{X} \times \mathbb{R}^2)^p \to \mathbb{R} \cup \{\infty\}$ arbitrária. Então todo minimizador h do risco regularizado admite uma representação na forma*

$$h(x) = \sum_{i=1}^{p} \gamma_i \kappa(x^i, x).$$

Geralmente, escolhemos $\Gamma[h]$ convexo para que um minimizador local do problema de otimização resultante seja um minimizador global.

O problema de classificação de maximização da margem maleável equivale a utilizar como função perda $\ell(x, y, h(x)) = \max\{0, 1 - yh(x)\}$ e como termo regularizador $\frac{1}{2}\|w\|^2$, pois o problema

$$\underset{(w,b)}{\text{minimizar}} \frac{1}{2}\|w\|^2 + \frac{C}{p}\sum_{i=1}^{p} \max\{0, 1 - y^i(w^\top x^i + b)\}$$

é equivalente a

$$\begin{array}{ll} \underset{(w,b,\xi)}{\text{minimizar}} & \frac{1}{2}\|w\|^2 + \frac{C}{p}\sum_{i=1}^{p} \xi_i \\ \text{sujeita a} & y^i(w^\top x^i + b) \geq 1 - \xi_i, \quad \forall\, i \in \{1, \ldots, p\} \\ & \xi_i \geq 0, \quad \forall\, i \in \{1, \ldots, p\}. \end{array}$$

Para manter a natureza esparsa do problema para classificação e estender para regressão as ideias de vetores suporte, Vapnik [50] concebeu a chamada função perda $\varepsilon-$insensível

$$\ell(x, y, h(x)) = |y - h(x)|_\varepsilon = \max\{0, |y - h(x)| - \varepsilon\},$$

a qual não penaliza erros menores que uma tolerância $\varepsilon \geq 0$ escolhida previamente. Seu algoritmo, $\varepsilon-$SVR ($\varepsilon-$*Support Vector Regression*), procura estimar funções por um preditor

$$h(x) = w^\top x + b \quad \text{com} \quad w, x \in \mathbb{R}^n, b \in \mathbb{R} \tag{1.4}$$

baseado nos dados $X = \{x^1, \ldots, x^p\}$ e $Y = \{y^1, \ldots, y^p\}$.

Não penalizar erros menores que uma tolerância $\varepsilon \geq 0$ significa que o preditor escolhido deverá estar numa região de margem ε dos pontos de amostra. Nem sempre isso é possível, e por essa razão queremos minimizar o risco empírico regularizado com essa função perda, ou seja, queremos minimizar

$$\frac{1}{2}\|w\|^2 + \frac{C}{p}\sum_{i=1}^{p} |y^i - h(x^i)|_\varepsilon \tag{1.5}$$

onde $\frac{1}{2}\|w\|^2$ é o termo regularizador e C é o parâmetro de regularização.

Para o problema de regressão, a interpretação geométrica de utilizar o regularizador $\frac{1}{2}\|w\|^2$ é encontrar a função h mais achatada possível com suficiente qualidade na aproximação, além de capturar a ideia principal da Teoria de Aprendizagem Estatística de tentar obter um risco pequeno controlando tanto o erro como a complexidade do modelo [42].

Nosso interesse reside em aproximar uma função $f : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}$ com conhecimento limitado sobre os valores funcionais, ou seja, temos $X = \{x^1, \ldots, x^p\}$ e $Y = f(X) = \{f(x^1), \ldots, f(x^p)\}$ e queremos encontrar um preditor h de modo que $h(x) \approx f(x)$ para todo $x \in \mathcal{X}$.

Desta maneira, iremos resolver o problema

$$\min \frac{1}{2}\|w\|^2 + \frac{\widehat{C}}{p}\sum_{i=1}^{p} \left|f(x^i) - h(x^i)\right|_\varepsilon,$$

que é equivalente a

$$\begin{array}{ll} \underset{(w,b,\xi,\xi')}{\text{minimizar}} & \dfrac{1}{2}\|w\|^2 + C\displaystyle\sum_{i=1}^{p}(\xi_i + \xi_i') \\ \text{sujeita a} & f(x^i) - w^\top x^i - b \leq \varepsilon + \xi_i \qquad \forall\ i = 1, \ldots, p \\ & w^\top x^i + b - f(x^i) \leq \varepsilon + \xi_i' \qquad \forall\ i = 1, \ldots, p \\ & \xi_i, \xi_i' \geq 0,\ \forall\ i = 1, \ldots, p, \end{array} \tag{1.6}$$

que por sua vez é um problema quadrático convexo, onde $C = \widehat{C}/p$.

Ao considerarmos a tolerância $\varepsilon \geq 0$, queremos que a imagem de todos os pontos de amostra pelo preditor estejam a uma distância no máximo ε da função a ser aproximada. Isso determina uma região chamada ε-tubo. No entanto, permitimos que o preditor esteja fora dessa região através das folgas ξ e ξ'. As folgas aparecem também na função objetivo do problema (1.6) para que ao resolvê-lo sejam as menores possíveis. A constante C, escolhida a priori, faz a ligação entre o preditor e o tamanho das folgas.

A função Lagrangiana associada ao problema (1.6) é $L : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$ dada por

$$\begin{aligned} L(w, b, \xi, \xi', \alpha, \gamma, \eta, \eta') = &\frac{1}{2} w^\top w + C \sum_{i=1}^{p} (\xi_i + \xi_i') + \sum_{i=1}^{p} \alpha_i \left(f(x^i) - w^\top x^i - b - \varepsilon - \xi_i \right) \\ &+ \sum_{i=1}^{p} \gamma_i \left(w^\top x^i + b - f(x^i) - \varepsilon - \xi_i' \right) - \sum_{i=1}^{p} \eta_i \xi_i - \sum_{i=1}^{p} \eta_i' \xi_i'. \end{aligned}$$

Note que as restrições do problema (1.6) são lineares, logo satisfazem uma condição de qualificação. Das condições de otimalidade e complementaridade [22], temos que se (w, b, ξ, ξ') é solução do problema (1.6), existem $(\alpha, \gamma, \eta, \eta')$ tais que

(a) $\nabla_w L = w - \sum_{i=1}^{p} \alpha_i x^i + \sum_{i=1}^{p} \gamma_i x^i = 0 \iff w = P^\top(\alpha - \gamma),$

(b) $\nabla_b L = -\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{p} \gamma_i = 0 \iff e^\top \alpha = e^\top \gamma,$

(c) $\nabla_{\xi_i} L = C - \alpha_i - \eta_i = 0$ e $\nabla_{\xi_i'} L = C - \gamma_i - \eta_i' = 0,$ para todo $i = 1, \ldots, p,$

(d) $\alpha_i \geq 0$, $\gamma_i \geq 0$, $\eta_i \geq 0$ e $\eta_i' \geq 0,$ para todo $i = 1, \ldots, p,$

(e) $\alpha_i \left(f(x^i) - w^\top x^i - b - \varepsilon - \xi_i \right) = 0,$ para todo $i = 1, \ldots, p,$

(f) $\gamma_i \left(w^\top x^i + b - f(x^i) - \varepsilon - \xi_i' \right) = 0,$ para todo $i = 1, \ldots, p,$

(g) $\eta_i \xi_i = 0$ e $\eta_i' \xi_i' = 0,$ para todo $i = 1, \ldots, p,$

onde

$$P = \begin{pmatrix} (x^1)^\top \\ (x^2)^\top \\ \vdots \\ (x^p)^\top \end{pmatrix}, \quad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{pmatrix}, \quad \gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_p \end{pmatrix} \quad \text{e} \quad e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Substituindo $w = P^\top(\alpha - \gamma)$ e reescrevendo a Lagrangiana em forma matricial, com $f(P) \in \mathbb{R}^p$ o vetor cuja i-ésima componente é $f(x^i)$, obtemos

$$\begin{aligned} L(w, b, \xi, \xi', \alpha, \gamma, \eta, \eta') &= \tilde{L}(b, \xi, \xi', \alpha, \gamma, \eta, \eta') \\ &= \frac{1}{2}(\alpha - \gamma)^\top P P^\top (\alpha - \gamma) + (Ce - \alpha - \eta)^\top \xi + (Ce - \gamma - \eta')^\top \xi' \\ &\quad + \gamma^\top P P^\top (\alpha - \gamma) - \alpha^\top P P^\top (\alpha - \gamma) + b(\gamma^\top e - \alpha^\top e) \\ &\quad + (f(P) - \varepsilon e)^\top \alpha + (-f(P) - \varepsilon e)^\top \gamma. \end{aligned}$$

Usando (b) e (c),

$$L(w,b,\xi,\xi',\alpha,\gamma,\eta,\eta') = \widehat{L}(\alpha,\gamma) = \frac{1}{2}(\alpha-\gamma)^\top PP^\top(\alpha-\gamma) - (\alpha-\gamma)^\top PP^\top(\alpha-\gamma)$$

$$+(f(P)-\varepsilon e)^\top\alpha + (-f(P)-\varepsilon e)^\top\gamma,$$

ou ainda

$$\widehat{L}(\alpha,\gamma) = -\frac{1}{2}\left(\alpha PP^\top\alpha - \alpha^\top PP^\top\gamma - \gamma^\top PP^\top\alpha + \gamma^\top PP^\top\gamma\right)$$
$$+(f(P)-\varepsilon e)^\top\alpha + (-f(P)-\varepsilon e)^\top\gamma,$$

que pode ser reescrita como

$$\widehat{L}(\alpha,\gamma) = -\frac{1}{2}\begin{pmatrix}\alpha\\ \gamma\end{pmatrix}^\top\begin{pmatrix}PP^\top & -PP^\top\\ -PP^\top & PP^\top\end{pmatrix}\begin{pmatrix}\alpha\\ \gamma\end{pmatrix} - \begin{pmatrix}-f(P)+\varepsilon e\\ f(P)+\varepsilon e\end{pmatrix}^\top\begin{pmatrix}\alpha\\ \gamma\end{pmatrix}.$$

Denotando

$$Q = \begin{pmatrix}PP^\top & -PP^\top\\ -PP^\top & PP^\top\end{pmatrix}, \tag{1.7}$$

$$z = \begin{pmatrix}\alpha\\ \gamma\end{pmatrix}, \quad v = \begin{pmatrix}-f(P)+\varepsilon e\\ f(P)+\varepsilon e\end{pmatrix} \quad \text{e} \quad A^\top = \begin{pmatrix}-e\\ e\end{pmatrix},$$

e utilizando (b), (c) e (d), o problema dual de (1.6) pode ser escrito como

$$\begin{array}{ll}\underset{z}{\text{minimizar}} & \frac{1}{2}z^\top Qz + v^\top z\\ \text{sujeita a} & Az = 0\\ & 0 \le z \le C.\end{array} \tag{1.8}$$

Com posse da solução dual $z = (\alpha,\gamma)$, somos capazes de calcular w e b que definem o preditor h, como veremos a seguir. Por (a) obtemos

$$w = \sum_{i=1}^{p}(\alpha_i - \gamma_i)x^i. \tag{1.9}$$

Usando as condições (c) a (g), segue que para todo $i = 1,\ldots,p$,

$$\alpha_i(\varepsilon + \xi_i - f(x^i) + w^\top x^i + b) = 0, \quad \gamma_i(\varepsilon + \xi_i' + f(x^i) - w^\top x^i - b) = 0,$$

$$(C-\alpha_i)\xi_i = 0 \quad \text{e} \quad (C-\gamma_i)\xi_i' = 0.$$

Através dessas condições e das restrições do problema (1.6), podemos calcular o valor de b. Notemos que, se para algum $i \in \{1,\ldots,p\}$, $0 < \alpha_i < C$, temos que

$$\varepsilon + \xi_i - f(x^i) + w^\top x^i + b = 0 \quad \text{e} \quad \xi_i = 0,$$

o que implica que

$$b = f(x^i) - w^\top x^i - \varepsilon. \tag{1.10}$$

Da mesma forma, se para algum $i \in \{1, \ldots, p\}$, $0 < \gamma_i < C$, temos que

$$b = f(x^i) - w^\top x^i + \varepsilon. \tag{1.11}$$

Substituindo (1.9) em (1.4), temos a expansão por vetores suporte

$$h(x) = \sum_{i=1}^{p} (\alpha_i - \gamma_i)(x^i)^\top x + b,$$

com b dado em (1.10) ou (1.11).

Os vetores suporte são os pontos da amostra associados a α_i ou γ_i não nulos. Sempre que α_i e γ_i são nulos o ponto x^i correspondente não faz nenhuma contribuição para o preditor. Isso ocorre para todos os pontos que estão dentro do ε-tubo, uma vez que se α_i ou γ_i são não nulos temos que $|f(x^i) - w^\top x^i - b| \geq \varepsilon$. Por isso, podemos remover os pontos cuja imagem estão dentro do ε-tubo e obter o mesmo preditor.

Outra consideração a ser feita é referente à existência de b. Como ele não aparece diretamente na solução do problema (1.8) precisamos calculá-lo indiretamente. Se α_i e γ_i são nulos para todo $i = \{1, \ldots, p\}$, então w_i, ξ_i e ξ_i' também são nulos. Com isso podemos escolher $b = f(x^i)$ para qualquer um dos pontos da amostra que obtemos uma solução para os problemas primal e dual. Nesse caso o preditor é simplesmente $h(x) = b$. Isso pode ocorrer quando ε é muito grande.

Por outro lado, se existe pelo menos um vetor suporte, podemos escolher C suficientemente grande de modo que $\alpha_i < C$ ou $\gamma_i < C$, o que garante a existência do b como calculado em (1.10) ou em (1.11).

Para construir um modelo não linear, do mesmo modo que fizemos para classificação, podemos levar nossos dados a um espaço de dimensão maior e construir um modelo linear nesse espaço.

Por exemplo, se queremos construir um modelo quadrático em um espaço de dimensão dois, estamos procurando por (w, b) que definem a função quadrática

$$q(x) = w_1 x_1^2 + w_2 \sqrt{2} x_1 x_2 + w_3 x_2^2 + w_4 x_1 + w_5 x_2 + b.$$

Nesse caso, cada ponto x é mapeado em

$$\varphi(x) = (x_1^2, \sqrt{2} x_1 x_2, x_2^2, x_1, x_2)^\top.$$

Queremos encontrar o modelo quadrático definido por $q(x) = w^\top \varphi(x) + b$ tal que para todo $i = 1, \ldots, p$,

$$|f(x^i) - (w^\top \varphi(x^i) + b)| \leq \varepsilon.$$

Ou seja, precisamos resolver exatamente o mesmo problema que no caso linear, exceto que agora os elementos da matriz Q dada em (1.7) são definidos por $\varphi(x^i)^\top \varphi(x^j)$ em vez de $(x^i)^\top x^j$.

Para o caso em que queremos modelos quadráticos definidos em $\mathbb{R}^n$, fazemos $\varphi(x) = (x_1^2, \sqrt{2}x_1x_2, \sqrt{2}x_1x_3, \ldots, x_2^2, \ldots, \sqrt{2}x_{n-1}x_n, x_n^2, x_1, x_2, \ldots, x_n)^\top$, de modo que $\varphi(x^i)^\top \varphi(x^j) = (x^i)^\top x^j + ((x^i)^\top x^j)^2$, o que facilita a construção da matriz Q do nosso problema de programação quadrática.

Assim, para construirmos um modelo para uma função avaliada em alguns pontos, precisamos resolver um problema do tipo

$$\begin{array}{ll} \underset{z}{\text{minimizar}} & z^\top Q z + v^\top z \\ \text{sujeita a} & Az = 0 \\ & 0 \leq z \leq C \end{array}$$

com Q simétrica e semidefinida positiva.

No próximo capítulo vamos formalizar a construção de modelos lineares e quadráticos de uma função usando a técnica de regressão via vetores suporte. Discutiremos algumas de suas propriedades que serão fundamentais para garantir, no Capítulo 3, a convergência global de um algoritmo de região de confiança livre de derivadas. É neste contexto em que reside a principal contribuição desta tese.

Capítulo 2

Sobre a construção de modelos

Neste capítulo, vamos discutir dois métodos para construção de modelos lineares e quadráticos de uma função diferenciável $f : \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}$ em torno de um ponto $x^0 \in \mathcal{D}$. Supomos que $\mathcal{D}$ é um conjunto aberto e convexo e que conhecemos os valores funcionais em um conjunto de amostra R que contém x_0. Na primeira seção, discutimos propriedades exigidas pela função a ser aproximada. Na segunda seção, discutimos a construção de modelos por interpolação polinomial, método bastante utilizado em otimização livre de derivadas. Na terceira seção, apresentamos nossa contribuição, que reside na construção de modelos utilizando a técnica de regressão por vetores suporte. Na seção seguinte, fazemos uma discussão sobre o controle da geometria do conjunto R dos pontos usados na construção dos modelos. Tal controle se faz necessário para garantirmos que o modelo realmente aproxime a função, no sentido que a norma da diferença entre o gradiente da função e o gradiente do modelo seja limitado numa vizinhança de x^0, o que é mostrado na última seção do capítulo.

Durante todo o trabalho, usamos a nomenclatura *erro dos gradientes* para designar a norma da diferença entre o gradiente do modelo e o gradiente da função. Nesse sentido, o objetivo do presente capítulo é encontrar um limitante para o erro dos gradientes. Conseguimos também encontrar um limitante para o erro cometido nos valores funcionais, ou seja, o módulo da diferença entre o valor do modelo e o valor da função. Quando f é suficientemente suave, é possível também encontrar um limitante para o erro das Hessianas.

2.1 Propriedades da função

Em otimização livre de derivadas, é comum para propósitos teóricos admitir que a função é suficientemente suave, apesar de suas derivadas não serem utilizadas. Em geral, queremos que os modelos aproximem o comportamento da função original e com isso garantir que a utilização dos modelos nos permita encontrar um minimizador da função.

A seguir, apresentamos hipóteses e resultados sobre a suavidade da função, os quais serão usados no decorrer do capítulo para encontrarmos limitantes para a norma da diferença da Hessiana do modelo e da Hessiana da função, para a norma da diferença do gradiente do modelo e do gradiente da função e para o módulo da diferença entre o valor do modelo e o valor da função em todos os pontos de uma vizinhança de um ponto $x^0 \in \mathbb{R}^n$. Tais hipóteses e resultados serão revisitados no capítulo seguinte no contexto dos métodos de região de confiança livre de derivadas.

A1. *A função f é continuamente diferenciável no conjunto aberto e convexo $\mathcal{D}$ e ∇f é Lipschitz com constante $L_g > 0$ em $\mathcal{D}$.*

Com a Hipótese A1 podemos garantir o próximo resultado, adaptado de [14].

Lema 2.1. *[14, Lema 4.1.12] Suponha que a Hipótese A1 seja satisfeita. Então para todo $x \in \mathcal{D}$ e d tal que $x + d \in \mathcal{D}$,*

$$|f(x+d) - f(x) - \nabla f(x)^\top d| \leq \frac{1}{2} L_g \|d\|^2.$$

Demonstração. Como f é continuamente diferenciável, pela forma integral do Teorema do Valor Médio

$$f(x+d) - f(x) = \int_0^1 \nabla f(x+td)^\top d \; dt,$$

sempre que $[x, x+d] \subset \mathcal{D}$.

Somando e subtraindo $\nabla f(x)^\top d$ na igualdade acima, obtemos

$$f(x+d) - f(x) = \nabla f(x)^\top d + \int_0^1 \left(\nabla f(x+td) - \nabla f(x)\right)^\top d \; dt. \tag{2.1}$$

Por outro lado, pela desigualdade de Cauchy-Schwarz e pela Hipótese A1,

$$\begin{aligned} \left| \int_0^1 \left(\nabla f(x+td) - \nabla f(x)\right)^\top d \; dt \right| &\leq \int_0^1 \left| \left(\nabla f(x+td) - \nabla f(x)\right)^\top d \right| \; dt \\ &\leq L_g \|d\|^2 \int_0^1 t \; dt \\ &= \frac{1}{2} L_g \|d\|^2, \end{aligned} \tag{2.2}$$

completando a demonstração. □

No caso em que a função f é de classe $\mathcal{C}^2$, conseguimos um resultado mais forte. Para isso, consideramos a hipótese a seguir.

A2. *A função f é duas vezes continuamente diferenciável no conjunto aberto e convexo $\mathcal{D}$ e $\nabla^2 f$ é Lipschitz com constante $L_h > 0$ em $\mathcal{D}$.*

Com a Hipótese A2 podemos garantir que vale o seguinte resultado.

Lema 2.2. *[14, Lema 4.1.14] Suponha que a Hipótese A2 seja satisfeita. Então para todo $x \in \mathcal{D}$ e d tal que $x + d \in \mathcal{D}$,*

$$\left| f(x+d) - f(x) - \nabla f(x)^\top d - \frac{1}{2} d^\top \nabla^2 f(x) d \right| \leq \frac{1}{6} L_h \|d\|^3.$$

Demonstração. Seja $h : \mathbb{R} \to \mathbb{R}$ uma parametrização de f ao longo do segmento $[x, x+d]$, com $h(t) = f(x + td)$. Seja $r(t) = x + td$. Então pela Regra da Cadeia, para $0 \leq \theta \leq 1$,

$$\frac{dh}{dt}(\theta) = \sum_{i=1}^{n} \frac{\partial f}{\partial r_i}(r(\theta)) \frac{dr_i}{dt}(\theta) = \nabla f(x + \theta d)^\top d,$$

e

$$\begin{aligned} \frac{d^2 h}{dt^2}(\theta) &= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial r_i r_j}(r(\theta)) \left(\frac{dr_i}{dt}(\theta) \right)^2 + \sum_{i=1}^{n} \frac{\partial f}{\partial r_i}(r(\theta)) \frac{dr_i^2}{dt^2}(\theta) \\ &= d^\top \nabla^2 f(x + \theta d) d. \end{aligned}$$

Pelo Teorema Fundamental do Cálculo, dados $a, b \in \mathbb{R}$,

$$\int_a^{a+b} h'(t)\, dt = h(a+b) - h(a).$$

Por outro lado, através de integração por partes, vemos que

$$\int_a^{a+b} h'(t)\, dt = (a+b)h'(a+b) - ah'(a) - \int_a^{a+b} t h''(t)\, dt.$$

Com isso e novamente o Teorema Fundamental do Cálculo,

$$\begin{aligned} h(a+b) - h(a) &= (a+b)h'(a+b) - ah'(a) - \int_a^{a+b} t h''(t)\, dt \\ &= bh'(a) + \int_a^{a+b} (a+b) h''(t)\, dt - \int_a^{a+b} t h''(t)\, dt \\ &= bh'(a) + \int_a^{a+b} (a+b-t) h''(t)\, dt. \end{aligned}$$

Fazendo a mudança de variáveis $t = a + sb$,

$$h(a+b) - h(a) - bh'(a) = \int_0^1 (b - sb)h''(a+sb)b\,ds = \int_0^1 b^2(1-s)h''(a+sb)\,ds.$$

Por outro lado,

$$\frac{1}{2}b^2h''(a) = \int_0^1 b^2(1-s)h''(a)\,ds.$$

Portanto,

$$h(a+b) - h(a) - bh'(a) - \frac{1}{2}b^2h''(a) = \int_0^1 b^2(1-s)[h''(a+sb) - h''(a)]\,ds.$$

Para $a = 0$ e $b = 1$,

$$h(1) - h(0) - h'(0) - \frac{1}{2}h''(0) = \int_0^1 (1-s)[h''(s) - h''(0)]\,ds.$$

Substituindo $h(t) = f(x+td)$, $h'(t) = \nabla f(x+td)^\top d$ e $h''(t) = d^\top \nabla^2 f(x+td)d$ na igualdade acima

$$f(x+d) - f(x) - \nabla f(x)^\top d - \frac{1}{2}d^\top \nabla^2 f(x)d = \int_0^1 (1-s)[d^\top \nabla^2 f(x+sd)d - d^\top \nabla^2 f(x)d]\,ds.$$

Com isso, pela Hipótese A2

$$\begin{aligned}\left|f(x+d) - f(x) - \nabla f(x)^\top d - \frac{1}{2}d^\top \nabla^2 f(x)d\right| &\leq \int_0^1 \left|(1-s)d^\top[\nabla^2 f(x+sd) - \nabla^2 f(x)]d\right| ds \\ &\leq L_h\|d\|^3 \int_0^1 (s - s^2)\,ds = \frac{1}{6}L_h\|d\|^3,\end{aligned}$$

o que completa a demonstração. □

2.2 Interpolação polinomial

Esta seção é dedicada a mostrar que modelos de interpolação aproximam f quando conhecemos seu valor em alguns pontos. Os resultados apresentados são uma revisão do conteúdo apresentado em [10], com pequenas modificações.

Denotamos o espaço dos polinômios definidos em $\mathbb{R}^n$ de grau menor ou igual a a por $\mathcal{P}_n^a$. Particularmente, os polinômios lineares no $\mathbb{R}^n$ são da forma

$$m(x) = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n,$$

nesse caso temos $\dim \mathcal{P}_n^1 = n + 1$. Para os polinômios quadráticos definidos no $\mathbb{R}^n$ temos

$\dim \mathcal{P}_n^2 = (n+1)(n+2)/2$. Em geral,

$$\dim \mathcal{P}_n^a = \frac{(n+a)!}{n!a!}.$$

Considere $s = (n+a)!/(n!a!) - 1$. Uma base $\phi = \{\phi_0(x), \phi_1(x), \ldots, \phi_s(x)\}$ de $\mathcal{P}_n^a$ é um conjunto de $s+1$ polinômios linearmente independentes de grau menor ou igual a a que gera $\mathcal{P}_n^a$. Nesse caso $\dim \mathcal{P}_n^a = s+1$.

Qualquer polinômio $m \in \mathcal{P}_n^a$ pode ser escrito como combinação linear dos elementos de uma base ϕ de $\mathcal{P}_n^a$, ou seja,

$$m(x) = \sum_{j=0}^{s} \mu_j \phi_j(x) = \mu^\top \phi(x),$$

onde $\mu = (\mu_0, \mu_1, \ldots, \mu_s)^\top$ é um vetor de $\mathbb{R}^{s+1}$ e $\phi(x) = (\phi_0(x), \phi_1(x), \ldots, \phi_s(x))^\top$ é o vetor formado com os elementos da base ϕ.

A base $\widehat{\phi}$ formada por monômios de coeficiente 1 é a base canônica de $\mathcal{P}_n^a$. Por exemplo, se $a = 1$, a base canônica é

$$\widehat{\phi} = \{1, x_1, x_2, \ldots, x_n\}.$$

Se $a = 2$, a base canônica é

$$\widehat{\phi} = \{1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2 \ldots, x_1x_n, x_2^2, x_2x_3, \ldots, x_2x_n, x_3^2, x_3x_4, \ldots x_n^2\}.$$

Uma base bastante importante é a chamada base natural. A base natural $\bar{\phi}$ pode ser convenientemente descrita via o uso de índices múltiplos. Seja um vetor $\tau = (\tau_1, \ldots, \tau_n)$ chamado de índice múltiplo, cujas componentes são inteiros não negativos. Para qualquer $x \in \mathbb{R}^n$, seja x^τ definido como

$$x^\tau = \prod_{j=1}^{n} x_j^{\tau_j}.$$

Também, definamos

$$|\tau| = \sum_{j=1}^{n} \tau_j \qquad \text{e} \qquad \tau! = \prod_{j=1}^{n} (\tau_j)!.$$

Considere $\{\tau^0, \tau^1, \ldots, \tau^s\}$ o conjunto de todos os vetores de índice múltiplo, com $|\tau^i| \leq a$ para todo $i = 0, \ldots, s$. Os elementos da base natural de $\mathcal{P}_n^a$ são

$$\bar{\phi}_i(x) = \frac{1}{(\tau^i)!} x^{\tau^i}, \quad i = 0, \ldots, s. \tag{2.3}$$

A base natural pode ser escrita como

$$\bar{\phi} = \{1, x_1, x_2, \ldots, x_n, x_1^2/2, x_1x_2, \ldots, x_1x_n, \ldots, x_{n-1}^{a-1}/(a-1)!, x_n^a/a!\},$$

onde cada termo da base canônica é dividido pelo produtório dos fatoriais de cada expoente. Quando $n = 3$ e $a = 2$, por exemplo, a base natural é

$$\bar{\phi} = \left\{1, x_1, x_2, x_3, \frac{1}{2}x_1^2, x_1x_2, x_1x_3, \frac{1}{2}x_2^2, x_2x_3, \frac{1}{2}x_3^2\right\}.$$

A base natural é assim chamada por Conn, Scheinberg e Vicente [10] pois surge naturalmente da expansão por séries de Taylor de uma função. Vale ressaltar que para os polinômios lineares as bases natural e canônica coincidem.

Dizemos que um polinômio m interpola a função f em um ponto $\bar{x}$ se $m(\bar{x}) = f(\bar{x})$. Se temos um conjunto $R = \{x^0, x^1, \ldots, x^p\} \subset \mathbb{R}^n$ e $m \in \mathcal{P}_n^a$ que interpola f nos pontos de R, então os coeficientes $\mu_0, \ldots, \mu_s$ que definem m em termos de uma base ϕ fixa podem ser determinados pelas condições de interpolação

$$m(x^i) = \sum_{j=0}^{s} \mu_j \phi_j(x^i) = f(x^i), \quad i = 0, \ldots, p. \tag{2.4}$$

Escrevendo (2.4) na forma matricial obtemos

$$M(\phi, R)\mu_\phi = f(R), \tag{2.5}$$

onde

$$M(\phi, R) = \begin{pmatrix} \phi_0(x^0) & \phi_1(x^0) & \cdots & \phi_s(x^0) \\ \phi_0(x^1) & \phi_1(x^1) & \cdots & \phi_s(x^1) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(x^p) & \phi_1(x^p) & \cdots & \phi_s(x^p) \end{pmatrix}, \quad \mu_\phi = \begin{pmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_s \end{pmatrix} \quad \text{e} \quad f(R) = \begin{pmatrix} f(x^0) \\ f(x^1) \\ \vdots \\ f(x^p) \end{pmatrix}.$$

Dizemos que o conjunto R é posicionado para interpolação quando o sistema (2.5) possui solução única, independente de f. Isto ocorre se, e somente se, $p = s$ e $M(\phi, R)$ é não singular. Alguns autores chamam um conjunto posicionado para interpolação polinomial de conjunto unisolvente [36].

Definição 2.3. *[10, Definição 3.1] O conjunto $R = \{x^0, x^1, \ldots, x^p\}$ é posicionado para interpolação polinomial em $\mathbb{R}^n$ se a matriz correspondente $M(\phi, R)$ é não singular para alguma base ϕ em $\mathcal{P}_n^a$.*

Como todas as bases em espaços vetoriais de dimensão finita são equivalentes, a

definição acima vale se $M(\phi, R)$ for não singular, qualquer que seja a base ϕ de $\mathcal{P}_n^a$, como estabelecido pelo lema a seguir.

Lema 2.4. *[10, Lema 3.2] Dada uma função $f : \mathbb{R}^n \to \mathbb{R}$ e um conjunto posicionado para interpolação polinomial $R \subset \mathbb{R}^n$, o polinômio interpolador $m \in \mathcal{P}_n^a$ de f em R existe e é único, independentemente da base de $\mathcal{P}_n^a$ usada.*

Demonstração. Como R é posicionado para interpolação polinomial, m existe e é único para uma base dada ϕ. O que precisamos mostrar é que m não depende da escolha da base. Assim, seja $\psi = B^\top \phi$ uma outra base de $\mathcal{P}_n^a$, onde B é uma matriz $(s+1) \times (s+1)$, não singular e ψ e ϕ estão escritas na forma de vetor. Note que,

$$M(\psi, R) = M(B^\top \phi, R) = M(\phi, R)B.$$

Como $M(\phi, R)$ e B são não singulares, o sistema $M(\psi, R)\mu = f(R)$ admite uma única solução μ_ψ.

Seja μ_ϕ a solução de $M(\phi, R)\mu = f(R)$, assim

$$\begin{aligned} \mu_\phi^\top \phi(x) &= \left(M^{-1}(\phi, R) f(R)\right)^\top \phi(x) \\ &= \left(BB^{-1} M^{-1}(\phi, R) f(R)\right)^\top \phi(x) \\ &= \left(BM^{-1}(\psi, R) f(R)\right)^\top \phi(x) \\ &= \mu_\psi^\top B^\top \phi(x) = \mu_\psi^\top \psi(x). \end{aligned}$$

Portanto, $m(x) = \mu_\psi^\top \psi(x) = \mu_\phi^\top \phi(x)$. □

2.2.1 Interpolação linear

Seja $m_I \in \mathcal{P}_n^1$ um polinômio linear em $\mathbb{R}^n$ que interpola f nos $n+1$ pontos do conjunto $R = \{x^0, \ldots, x^n\}$. Considere $b_I = \mu_0$ e $w_I = (\mu_1, \mu_2, \ldots, \mu_n)^\top \in \mathbb{R}^n$. O polinômio m_I pode ser escrito como

$$m_I(x) = \mu_0 + \mu_1 x_1 + \mu_2 x_2 + \ldots + \mu_n x_n = b_I + w_I^\top x,$$

em relação à base natural $\bar{\phi} = \{1, x_1, x_2, \ldots, x_n\}$.

As condições de interpolação podem ser escritas na forma do sistema linear

$$M(\bar{\phi}, R)\mu = f(R) \quad \Leftrightarrow \quad \begin{pmatrix} 1 & x_1^0 & \cdots & x_n^0 \\ 1 & x_1^1 & \cdots & x_n^1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^n & \cdots & x_n^n \end{pmatrix} \begin{pmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} f(x^0) \\ f(x^1) \\ \vdots \\ f(x^n) \end{pmatrix}.$$

Ao aplicar um passo da eliminação Gaussiana à matriz $M(\bar{\phi}, R)$, obtemos a matriz

$$L = \begin{pmatrix} 1 & x_1^0 & \cdots & x_n^0 \\ 0 & x_1^1 - x_1^0 & \cdots & x_n^1 - x_n^0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & x_1^n - x_1^0 & \cdots & x_n^n - x_n^0 \end{pmatrix},$$

que pode ser escrita por blocos como

$$L = \begin{pmatrix} 1 & (x^0)^\top \\ 0 & L_\ell \end{pmatrix},$$

em que

$$L_\ell = \begin{pmatrix} (x^1 - x^0)^\top \\ \vdots \\ (x^n - x^0)^\top \end{pmatrix} = \begin{pmatrix} x_1^1 - x_1^0 & \cdots & x_n^1 - x_n^0 \\ \vdots & \ddots & \vdots \\ x_1^n - x_1^0 & \cdots & x_n^n - x_n^0 \end{pmatrix}.$$

Lema 2.5. *A matriz L_ℓ é não singular se, e somente se, o conjunto R é posicionado para interpolação linear.*

Demonstração. Basta ver que $L = TM(\bar{\phi}, R)$, com

$$T = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ -1 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

e

$$\det(L_\ell) = \det(L) = \det(T)\det(M(\bar{\phi}, R)) = \det(M(\bar{\phi}, R)).$$

□

Considere $\delta > 0$ tal que $R \subset B(x^0, \delta) = \{x \in \mathbb{R}^n \mid \|x - x^0\| \leq \delta\}$. Para calcularmos o limitante entre o erro da função e do modelo linear por interpolação em $B(x^0, \delta)$,

usamos a matriz com mudança de escala

$$\widehat{L}_\ell = \frac{1}{\delta} L_\ell, \tag{2.6}$$

que corresponde ao conjunto de amostra com mudança de escala

$$\widehat{R} = \{x^0/\delta, x^1/\delta, \cdots, x^n/\delta\} \subset B(x^0/\delta, 1).$$

Lema 2.6. *[10, Teorema 2.11] Suponha que a Hipótese A1 seja satisfeita e o conjunto $R = \{x^0, \ldots, x^n\} \subset \mathcal{D} \cap B(x^0, \delta)$ seja posicionado para interpolação polinomial linear. Então para todo $x \in \mathcal{D} \cap B(x^0, \delta)$ o gradiente ∇m_I do modelo linear por interpolação polinomial satisfaz*

$$\|\nabla f(x) - \nabla m_I(x)\| \leq \kappa_1 \delta,$$

em que $\kappa_1 = L_g + \frac{1}{2} L_g \sqrt{n} \|\widehat{L}_\ell^{-1}\|$.

Demonstração. Como m_I é um polinômio linear, $m_I(x) = w_I^\top x + b_I$, para algum $w_I \in \mathbb{R}^n$ e algum $b_I \in \mathbb{R}$. Das condições de interpolação, para todo $x^i \in R$, vale

$$f(x^i) - f(x^0) = m_I(x^i) - m_I(x^0) = w_I^\top (x^i - x^0). \tag{2.7}$$

Considere $i \in \{1, \ldots, n\}$ arbitrário. Pelo Lema 2.1, com $d = x^i - x^0$,

$$|f(x^i) - f(x^0) - \nabla f(x^0)^\top (x^i - x^0)| \leq \frac{1}{2} L_g \|x^i - x^0\|^2 \leq \frac{1}{2} L_g \delta^2.$$

Usando isto, (2.7), a definição da matriz L_ℓ e o Lema 2.5, obtemos

$$\begin{aligned} \|\nabla f(x^0) - w_I\| &\leq \|L_\ell^{-1}\| \|L_\ell (\nabla f(x^0) - w_I)\| \\ &\leq \sqrt{n} \|L_\ell^{-1}\| \|L_\ell (\nabla f(x^0) - w_I)\|_\infty \\ &\leq \frac{1}{2} L_g \sqrt{n} \|L_\ell^{-1}\| \delta^2. \end{aligned}$$

Usando a definição (2.6) de $\widehat{L}_\ell$, segue que

$$\|\nabla f(x^0) - w_I\| \leq \frac{1}{2} L_g \sqrt{n} \|\widehat{L}_\ell^{-1}\| \delta.$$

Considere agora $x \in \mathcal{D} \cap B(x^0, \delta)$ arbitrário. Usando a desigualdade triangular, a Hipótese A1 e o fato de que $\nabla m_I(x) = w_I$, segue que

$$\begin{aligned} \|\nabla f(x) - \nabla m_I(x)\| &\leq \|\nabla f(x) - \nabla f(x^0)\| + \|\nabla f(x^0) - \nabla m_I(x)\| \\ &\leq \left(L_g + \frac{1}{2} L_g \sqrt{n} \|\widehat{L}^{-1}\| \right) \delta, \end{aligned}$$

concluindo a demonstração. □

O lema anterior estabelece uma limitação para a norma da diferença dos gradientes da função e do modelo obtido por interpolação polinomial na região $B(x^0, \delta)$. O próximo lema estabelece uma limitação para o valor absoluto da diferença dos valores funcionais.

Lema 2.7. *[10, Teorema 2.12] Suponha que a Hipótese A1 seja satisfeita e o conjunto $R = \{x^0, x^1, \ldots, x^n\} \subset \mathcal{D} \cap B(x^0, \delta)$ seja posicionado para interpolação polinomial linear. Então, para todo $x \in \mathcal{D} \cap B(x^0, \delta)$, o modelo linear por interpolação polinomial m_I satisfaz*

$$|f(x) - m_I(x)| \leq \kappa_2 \delta^2,$$

em que $\kappa_2 = \frac{3}{2} L_g + \frac{1}{2} L_g \sqrt{n} \|\widehat{L}^{-1}\|$.

Demonstração. Temos que $m_I(x) = w_I^\top x + b_I$, então $\nabla m_I(x) = w_I$ para todo $x \in B(x^0, \delta)$. Assim,

$$m_I(x) = \nabla m_I(x^0)^\top x + b_I + m_I(x^0) - m_I(x^0) = m_I(x^0) + \nabla m_I(x^0)^\top (x - x^0).$$

Logo,

$$f(x) - m_I(x) = f(x) - m_I(x^0) - \nabla m_I(x^0)^\top (x - x^0).$$

Somando e subtraindo $f(x^0) + \nabla f(x^0)^\top (x - x^0)$ do lado direito obtemos

$$\begin{aligned} f(x) - m_I(x) &= f(x) - f(x^0) - \nabla f(x^0)^\top (x - x^0) + f(x^0) - m_I(x^0) + \\ &\quad + \nabla f(x^0)^\top (x - x^0) - \nabla m_I(x^0)^\top (x - x^0). \end{aligned}$$

Pela desigualdade triangular e a condição de interpolação em x^0, temos que

$$|f(x) - m_I(x)| \leq |f(x) - f(x^0) - \nabla f(x^0)^\top (x - x^0)| + |\left(\nabla f(x^0) - \nabla m_I(x^0)\right)^\top (x - x^0)|.$$

Utilizando o Lema 2.1 e o Lema 2.6, segue que

$$|f(x) - m_I(x)| \leq \frac{1}{2} L_g \|x - x^0\|^2 + \kappa_1 \delta \|x - x^0\| \leq \left(\frac{1}{2} L_g + \kappa_1\right) \delta^2,$$

o que conclui a demonstração. □

2.2.2 Interpolação quadrática

Para a construção de um modelo quadrático único por interpolação, precisamos de $(n + 1)(n + 2)/2$ pontos distintos nos quais conhecemos o valor da função a ser aproximada. Se houvesse menos pontos para interpolação, poderia haver uma infinidade de modelos interpoladores. Nesse caso, uma possibilidade seria a escolha daquele que tem a

Hessiana de norma de Frobenius mínima, como sugerido em [10].

Seja $m_I \in \mathcal{P}_n^2$ um polinômio quadrático em $\mathbb{R}^n$ que interpola f nos pontos do conjunto $R = \{x^0, \ldots, x^q\}$, com $q = (n^2 + 3n)/2$. Utilizando a base natural

$$\bar{\phi} = \{1, x_1, x_2, \ldots, x_n, \frac{1}{2}x_1^2, x_1x_2, x_1x_3, \ldots, x_1x_n, \frac{1}{2}x_2^2, \ldots, x_{n-1}x_n, \frac{1}{2}x_n^2\}, \tag{2.8}$$

podemos expressar

$$m_I(x) = \sum_{j=0}^{q} \mu_j \phi_j(x) = \mu^\top \bar{\phi}(x).$$

As condições de interpolação, nesse caso, podem ser escritas na forma do sistema linear

$$M(\bar{\phi}, R)\mu = f(R) \quad \Leftrightarrow \quad \begin{pmatrix} 1 & \bar{\phi}_1(x^0) & \cdots & \bar{\phi}_q(x^0) \\ 1 & \bar{\phi}_1(x^1) & \cdots & \bar{\phi}_q(x^1) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \bar{\phi}_1(x^q) & \cdots & \bar{\phi}_q(x^q) \end{pmatrix} \begin{pmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_q \end{pmatrix} = \begin{pmatrix} f(x^0) \\ f(x^1) \\ \vdots \\ f(x^q) \end{pmatrix}.$$

Para o próximo lema vamos considerar a matriz

$$L_q = \begin{pmatrix} (\bar{\varphi}(x^1 - x^0))^\top \\ \vdots \\ (\bar{\varphi}(x^q - x^0))^\top \end{pmatrix}, \tag{2.9}$$

em que

$$\bar{\varphi}(x) = \left(x_1, x_2, \ldots, x_n, \frac{1}{2}x_1^2, x_1x_2, x_1x_3, \ldots, x_1x_n, \frac{1}{2}x_2^2, \ldots, x_{n-1}x_n, \frac{1}{2}x_n^2\right)^\top.$$

Essa matriz corresponde à matriz do conjunto de interpolação transladado para a origem $\widehat{R} \subset B(0, \delta)$, excluindo a primeira linha e primeira coluna. Para eliminar a dependência de L_q em δ, vamos considerar a matriz

$$\widehat{L}_q = L_q \begin{pmatrix} D_\delta^{-1} & 0 \\ 0 & D_{\delta^2}^{-1} \end{pmatrix}, \tag{2.10}$$

em que $D_\delta = \delta I_{n\times n}$ e $D_{\delta^2} = \delta^2 I_{(q-n)\times(q-n)}$. Essa matriz corresponde ao conjunto com mudança de escala e transladado para a origem $\widehat{R} = R/\delta \subset B(0, 1)$.

Podemos notar que se considerarmos o conjunto $R \subset B(x^0, \delta)$ posicionado para interpolação quadrática, o conjunto transladado para a origem $\widehat{R} \subset B(0, \delta)$ também será posicionado para interpolação quadrática, consequentemente as matrizes L_q e $\widehat{L}_q$ são não-

singulares.

Cabe ressaltar que o modelo quadrático m_I pode ser escrito na forma

$$m_I(x) = \frac{1}{2}x^\top H_I x + g_I^\top x + b_I,$$

com $b_I = \mu_0 \in \mathbb{R}$, $g_I = (\mu_1, \mu_2, \ldots, \mu_n)^\top \in \mathbb{R}^n$ e $H_I \in \mathbb{R}^{n\times n}$ uma matriz simétrica cujo elemento $(H_I)_{ij} = \mu_{\ell(i,j)}$ com $\ell(i,j) = \frac{i(2n+1-i)+2j}{2}$ para todo $j \geq i$. Assim temos que $\nabla m_I(x) = H_I x + g_I$ e $m_I(x^i) = f(x^i)$ para todo $x^i \in R$, $i = 0, \ldots, q$.

Com essas considerações em mente, mostraremos o seguinte resultado.

Lema 2.8. *[10, Teorema 3.16] Suponha que a Hipótese A2 seja satisfeita e o conjunto $R = \{x^0, \ldots, x^q\} \subset \mathcal{D} \cap B(x^0, \delta)$ seja posicionado para interpolação polinomial quadrática. Então para todo $x \in \mathcal{D} \cap B(x^0, \delta)$ a Hessiana $\nabla^2 m_I$ e o gradiente ∇m_I do modelo quadrático por interpolação polinomial satisfazem*

$$\|\nabla^2 f(x) - \nabla^2 m_I(x)\| \leq \kappa_3 \delta$$

e

$$\|\nabla f(x) - \nabla m_I(x)\| \leq \kappa_4 \delta^2,$$

em que $\kappa_3 = \dfrac{3\sqrt{2}}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\|$ *e* $\kappa_4 = \dfrac{3+3\sqrt{2}}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\|$.

Demonstração. Para todo $i = 0, 1, \ldots, q$ e $x \in \mathcal{D} \cap B(x^0, \delta)$,

$$\frac{1}{2}(x^i - x)^\top H_I (x^i - x) + \nabla m_I(x)^\top (x^i - x) =$$

$$= \frac{1}{2}(x^i)^\top H_I x^i - x^\top H_I x^i + \frac{1}{2} x^\top H_I x + (H_I x + g_I)^\top (x^i - x) =$$

$$= \frac{1}{2}(x^i)^\top H_I x^i - x^\top H_I x^i + \frac{1}{2} x^\top H_I x + x^\top H_I x^i + g_I^\top x^i - x^\top H_I x - g_I^\top x =$$

$$= \frac{1}{2}(x^i)^\top H_I x^i + g_I^\top x^i - \frac{1}{2} x^\top H_I x - g_I^\top x = m_I(x^i) - m_I(x).$$

Ou seja,

$$m_I(x) - m_I(x^i) + \nabla m_I(x)^\top (x^i - x) + \frac{1}{2}(x^i - x)^\top H_I (x^i - x) = 0. \tag{2.11}$$

Considerando a Hipótese A2, temos que a Hessiana da função é Lipschitz e portanto pelo Lema 2.2 segue que

$$|f(x^i) - f(x) - \nabla f(x)^\top (x^i - x) - \frac{1}{2}(x^i - x)^\top \nabla^2 f(x)(x^i - x)| \leq \frac{1}{6} L_h \|x^i - x\|^3$$

para todo $i = 0, 1, \ldots, q$, o que implica que para todo $x \in B(x^0, \delta)$ e para $i = 1, \ldots, q$,

$$|f(x^i) - f(x) - \nabla f(x)^\top (x^i - x) - \frac{1}{2}(x^i - x)^\top \nabla^2 f(x)(x^i - x)| \leq \frac{4}{3} L_h \delta^3,$$

uma vez que $\|x^i - x\| \leq \|x^i - x^0\| + \|x^0 - x\| \leq 2\delta$.

Utilizando (2.11) e as condições de interpolação na desigualdade acima, obtemos para todo $i = 1, \ldots, q$,

$$\left| m_I(x) - f(x) + (\nabla m_I(x) - \nabla f(x))^\top (x^i - x) + \frac{1}{2}(x^i - x)^\top \left(H_I - \nabla^2 f(x)\right)(x^i - x) \right| \leq \frac{4}{3} L_h \delta^3.$$

Subtraindo de todas as desigualdades o caso particular $i = 0$,

$$\left| m_I(x) - f(x) + (\nabla m_I(x) - \nabla f(x))^\top (x^0 - x) + \frac{1}{2}(x^0 - x)^\top \left(H_I - \nabla^2 f(x)\right)(x^0 - x) \right| \leq \frac{1}{6} L_h \delta^3,$$

visto que,

$$\nabla m_S(x)^\top (x^i - x) - \nabla m_S(x)^\top (x^0 - x) = \nabla m_S(x)^\top (x^i - x^0)$$

e

$$\frac{1}{2}(x^i - x)^\top \left(H_I - \nabla^2 f(x)\right)(x^i - x) - \frac{1}{2}(x^0 - x)^\top \left(H_I - \nabla^2 f(x)\right)(x^i - x) =$$

$$\frac{1}{2}(x^i)^\top \left(H_I - \nabla^2 f(x)\right) x^i - x^\top \left(H_I - \nabla^2 f(x)\right) x^i + \frac{1}{2} x^\top \left(H_I - \nabla^2 f(x)\right) x -$$

$$\frac{1}{2}(x^0)^\top \left(H_I - \nabla^2 f(x)\right) x^0 + x^\top \left(H_I - \nabla^2 f(x)\right) x^0 - \frac{1}{2} x^\top \left(H_I - \nabla^2 f(x)\right) x =$$

$$\frac{1}{2}(x^i - x^0)^\top \left(H_I - \nabla^2 f(x)\right)(x^i - x^0) - x^\top \left(H_I - \nabla^2 f(x)\right) x^i +$$

$$(x^i)^\top \left(H_I - \nabla^2 f(x)\right) x^0 - (x^0)^\top \left(H_I - \nabla^2 f(x)\right) x^0 + x^\top \left(H_I - \nabla^2 f(x)\right) x^0 =$$

$$\frac{1}{2}(x^i - x^0)^\top \left(H_I - \nabla^2 f(x)\right)(x^i - x^0) - (x^i - x^0)^\top \left(H_I - \nabla^2 f(x)\right)(x - x^0),$$

segue que para todo $i = 1, \ldots, q$,

$$\left| \left(e^g(x) - E^H(x)(x - x^0)\right)^\top (x^i - x^0) + \frac{1}{2}(x^i - x^0)^\top \left(E^H(x)\right)(x^i - x^0) \right| \leq \frac{3}{2} L_h \delta^3, \quad (2.12)$$

em que

$$e^g(x) = \nabla m_I(x) - \nabla f(x)$$

e

$$E^H(x) = H_I - \nabla^2 f(x).$$

Considerando a matriz L_q dada em (2.9), segue de (2.12) que

$$\left\| L_q \begin{pmatrix} t(x) \\ e^H(x) \end{pmatrix} \right\| \leq \sqrt{q} \left\| L_q \begin{pmatrix} t(x) \\ e^H(x) \end{pmatrix} \right\|_\infty \leq \frac{3}{2} \sqrt{q} L_h \delta^3,$$

em que

$$t(x) = e^g(x) - E^H(x)(x - x^0) \in \mathbb{R}^n \tag{2.13}$$

e $e^H(x)$ é um vetor do $\mathbb{R}^{n(n+1)/2}$ que armazena os elementos $E^H_{kk}(x)$, $k = 1, \ldots, n$ e $E^H_{k\ell}(x)$, $1 \leq \ell < k \leq n$.

A desigualdade acima pode ser reescrita como

$$\left\| L_q \begin{pmatrix} D_\delta^{-1} & 0 \\ 0 & D_{\delta^2}^{-1} \end{pmatrix} \begin{pmatrix} D_\delta t(x) \\ D_{\delta^2} e^H(x) \end{pmatrix} \right\| \leq \frac{3}{2} L_h \sqrt{q} \delta^3.$$

Usando a definição da matriz $\widehat{L}_q$, dada em (2.10), na desigualdade acima, obtemos

$$\left\| \begin{pmatrix} D_\delta t(x) \\ D_{\delta^2} e^H(x) \end{pmatrix} \right\| \leq \frac{3}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\| \delta^3,$$

e com isso

$$\|t(x)\| \leq \|D_\delta^{-1}\| \|D_\delta t(x)\| \leq \frac{3}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\| \|D_\delta^{-1}\| \delta^3 \leq \frac{3}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\| \delta^2, \tag{2.14}$$

e

$$\|e^H(x)\| \leq \|D_{\delta^2}^{-1}\| \|D_{\delta^2} e^H(x)\| \leq \frac{3}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\| \|D_{\delta^2}^{-1}\| \delta^3 \leq \frac{3}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\| \delta. \tag{2.15}$$

O erro nas Hessianas é portanto dado por

$$\left\| H_I - \nabla^2 f(x) \right\| \leq \left\| H_I - \nabla^2 f(x) \right\|_F \leq \sqrt{2} \|e^H(x)\| \leq \frac{3\sqrt{2}}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\| \delta,$$

com isso e (2.13) o erro nos gradientes é

$$\begin{aligned} \|\nabla m_I(x) - \nabla f(x)\| &\leq \|t(x)\| + \left\| H_I - \nabla^2 f(x) \right\| \|x - x^0\| \\ &\leq \frac{3}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\| \delta^2 + \left(\frac{3\sqrt{2}}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\| \delta \right) \delta \\ &\leq \frac{3(1+\sqrt{2})}{2} L_h \sqrt{q} \|\widehat{L}_q^{-1}\| \delta^2, \end{aligned}$$

concluindo a demonstração. □

Lema 2.9. *[10, Teorema 3.16] Suponha que a Hipótese A2 seja satisfeita e o conjunto $R = \{x^0, \ldots, x^q\} \subset \mathcal{D} \cap B(x^0, \delta)$ seja posicionado para interpolação polinomial quadrática. Então para todo $x \in \mathcal{D} \cap B(x^0, \delta)$ o modelo quadrático m_I por interpolação polinomial satisfaz*

$$|f(x) - m_I(x)| \leq \kappa_5 \delta^3$$

em que $\kappa_5 = \frac{1}{6}L_h + \frac{6+9\sqrt{2}}{4}L_h\sqrt{q}\|\widehat{L}_q^{-1}\|$.

Demonstração. Como o modelo é quadrático, pode ser escrito como

$$m_I(x) = m_I(x^0) + \nabla m_I(x^0)^\top (x - x^0) + \frac{1}{2}(x - x^0)^\top H_I (x - x^0),$$

para todo $x \in \mathcal{D} \cap B(x^0, \delta)$. Assim,

$$\begin{aligned} f(x) - m_I(x) &= f(x) - m_I(x^0) - \nabla m_I(x^0)^\top (x - x^0) - \frac{1}{2}(x - x^0)^\top H_I (x - x^0) \\ &= f(x) - f(x^0) - \nabla f(x^0)^\top (x - x^0) - \frac{1}{2}(x - x^0)^\top \nabla^2 f(x^0)(x - x^0) + \\ &\quad f(x^0) - m_I(x^0) + \left(\nabla f(x^0) - \nabla m_I(x^0)\right)^\top (x - x^0) + \\ &\quad \frac{1}{2}(x - x^0)\left(\nabla^2 f(x^0) - H_I\right)(x - x^0), \end{aligned}$$

para todo $x \in \mathcal{D} \cap B(x^0, \delta)$.

Utilizando o Lema 2.2, a condição de interpolação em x^0 e o Lema 2.8 segue que

$$|f(x) - m_I(x)| \leq \frac{1}{6}L_h\delta^3 + \frac{1}{2}\kappa_3\delta^3 + \kappa_4\delta^3,$$

concluindo a demonstração. □

Para o caso em que a quantidade de pontos de interpolação é menor do que a dimensão do espaço $\mathcal{P}_n^a$ dos polinômios de grau menor ou igual a 2 no $\mathbb{R}^n$, a matriz definida pelas condições de interpolação possui mais colunas do que linhas e com isso o polinômio interpolador não é único. Esse é o caso em que temos um modelo de interpolação indeterminado. Uma alternativa nesses casos é construir modelos minimizando a norma de Frobenius de sua Hessiana. Maiores detalhes podem ser encontrados nos trabalhos de Conn, Scheinberg e Vicente [10, 11].

2.3 Regressão por vetores suporte

No Capítulo 1, apresentamos as ideias da regressão por vetores suporte. O objetivo desta seção é encontrar limitantes semelhantes aos apresentados na seção anterior para a norma da diferença do gradiente do modelo e do gradiente da função quando construímos o modelo utilizando regressão por vetores suporte. A primeira parte da seção trata a regressão linear, posteriormente apresentamos a regressão quadrática.

Esta é uma das seções mais importantes da tese. Estes limitantes serão usados,

no Capítulo 3, para provar a convergência global de um algoritmo de região de confiança sem derivadas com os modelos construídos pela técnica de regressão via vetores suporte.

2.3.1 Regressão linear por vetores suporte

Sejam $R = \{x^0, x^1, \ldots, x^p\} \subset \mathcal{D} \subset \mathbb{R}^n$ um conjunto de pontos nos quais conhecemos o valor da função objetivo f, $\varepsilon \geq 0$ a tolerância permitida para o erro entre a função e o modelo que desejamos construir e $C > 0$ o parâmetro de regularização definido na Seção 1.2.2. Considere o problema

$$\begin{array}{ll} \text{minimizar} & \frac{1}{2} z^\top Q_\ell z + v^\top z \\ \text{sujeita a} & Az = 0 \\ & 0 \leq z \leq C, \end{array} \tag{2.16}$$

em que $Q_\ell \in \mathbb{R}^{2(p+1)\times 2(p+1)}$, $v \in \mathbb{R}^{2(p+1)}$ e $A \in \mathbb{R}^{1\times 2(p+1)}$ são dados por

$$Q_\ell = \begin{pmatrix} PP^\top & -PP^\top \\ -PP^\top & PP^\top \end{pmatrix} \text{ onde } P = \begin{pmatrix} (x^0)^\top \\ (x^1)^\top \\ \vdots \\ (x^p)^\top \end{pmatrix},$$

$$v = \begin{pmatrix} -f(R) + \varepsilon e \\ f(R) - \varepsilon e \end{pmatrix} \quad \text{e} \quad A^\top = \begin{pmatrix} -e \\ e \end{pmatrix}. \tag{2.17}$$

Apresentamos a seguir um algoritmo para construção do modelo linear da função objetivo f por regressão via vetores suporte a partir do conjunto de amostra R.

Algoritmo 2.1. *Modelo linear por regressão via vetores suporte*

Dados: $R = \{x^0, x^1, \cdots, x^p\}$, $C > 0$ *e* $\varepsilon \geq 0$.
ENQUANTO b_S *não é calculado.*
  CALCULE $z \in \mathbb{R}^{2(p+1)}$ *como solução do problema* (2.16).
  FAÇA $\alpha = (z_1, z_2, \cdots, z_{p+1})^\top$ *e* $\gamma = (z_{p+2}, z_{p+3}, \cdots, z_{2(p+1)})^\top$.
  DEFINA $w_S = P^\top(\alpha - \gamma)$.
  SE $\alpha \neq 0$ *ou* $\gamma \neq 0$,
    ESCOLHA i *tal que* $0 < \alpha_i < C$ *e calcule* $b_S = f(x^i) - w_S^\top x^i - \varepsilon$.
    SE IMPOSSÍVEL, ESCOLHA i *tal que* $0 < \gamma_i < C$ *e calcule* $b_S = f(x^i) - w_S^\top x^i + \varepsilon$.
    SE IMPOSSÍVEL, AUMENTE C.
  SENÃO
    ESCOLHA $i \in \{0, \ldots, p\}$ *e calcule* $b_S = f(x^i)$.

Para a construção de um modelo linear com regressão por vetores suporte, precisamos de pelo menos um ponto na amostra. O teorema a seguir mostra que se o conjunto de amostra tem pelo menos $n+1$ pontos e se conseguimos controlar o erro entre a função e o modelo nos pontos da amostra, conseguimos limitar o erro entre a função e o modelo e também o erro entre o gradiente da função e o gradiente do modelo em uma vizinhança. Esses limitantes são importantes para garantirmos a convergência de métodos de região de confiança sem derivadas, como será visto no próximo capítulo.

A tolerância $\varepsilon \geq 0$ para construir um modelo por regressão via vetores suporte representa uma região de largura 2ε sobre a função a ser aproximada. Como visto no capítulo anterior, permitimos que o modelo esteja fora desta região através das folgas $\xi, \xi' \geq 0$. Ou seja, pelas restrições do problema (1.6), para todo $i = 1, \ldots, p$,

$$f(x^i) - w^\top x^i - b \leq \varepsilon + \xi_i \qquad \text{e} \qquad w^\top x^i + b - f(x^i) \leq \varepsilon + \xi_i'.$$

Vamos admitir que $\varepsilon \leq c_1\delta^2$ e $\xi, \xi' \leq c_2\delta^2$, com $c_1, c_2 > 0$. Assim,

$$|m_S(x^i) - f(x^i)| \leq (c_1 + c_2)\delta^2, \tag{2.18}$$

para todo $x^i \in R$. Com isso conseguimos controlar o erro em uma vizinhança de um ponto $x^0 \in \mathcal{D}$.

O teorema a seguir estabelece a norma da diferença entre o gradiente da função e o gradiente do modelo linear de regressão por vetores suporte.

Teorema 2.10. *Considere o conjunto de amostra $R = \{x^0, \ldots, x^n\} \subset \mathcal{D} \cap B(x^0, \delta)$ de modo que o conjunto $R^- = \{x^1, \ldots, x^n\}$ seja linearmente independente e suponha que a Hipótese A1 seja satisfeita. Se o modelo linear $m_S(x) = w_S^\top x + b_S$ é construído via regressão por vetores suporte com margem $\varepsilon \leq c_1\delta^2$ e folgas $\xi, \xi' \leq c_2\delta^2$, com $c_1, c_2 > 0$, então para todo $x \in \mathcal{D} \cap B(x^0, \delta)$*

$$\|\nabla f(x) - \nabla m_S(x)\| \leq \kappa_6 \delta,$$

com $\kappa_6 = L_g + \left(\frac{1}{2}L_g + 2(c_1 + c_2)\right)\sqrt{n}\|\widehat{L}_\ell^{-1}\|$.

Demonstração. Considerando a Hipótese A1, temos que o gradiente da função é Lipschitz, portanto pelo Lema 2.1 segue que para $i = 1, \ldots, n$

$$|f(x^i) - f(x^0) - \nabla f(x^0)^\top (x^i - x^0)| \leq \frac{1}{2}L_g\|x^i - x^0\|^2 \leq \frac{1}{2}L_g\delta^2. \tag{2.19}$$

Temos que o modelo linear é $m_S(x) = w_S^\top x + b_S$, e portanto $\nabla m_S(x) = w_S$ para todo $x \in \mathcal{D} \cap B(x^0, \delta)$ e, consequentemente,

$$\left(\nabla f(x^0) - \nabla m_S(x)\right)^\top (x^i - x^0)$$

$$= \nabla f(x^0)^\top (x^i - x^0) + f(x^0) - f(x^i) + f(x^i) - w_S^\top x^i - b_S + w_S^\top x^0 + b_S - f(x^0).$$

Utilizando a desigualdade triangular, (2.19) e o controle do erro entre o valor do modelo e o valor da função nos pontos da amostra (2.18), segue da igualdade acima que para $i = 1, \ldots, n$,

$$\left|\left(\nabla f(x^0) - \nabla m_S(x)\right)^\top (x^i - x^0)\right|$$

$$\leq |\nabla f(x^0)^\top (x^i - x^0) + f(x^0) - f(x^i)| + |f(x^i) - w_S^\top x^i - b_S| + |w_S^\top x^0 + b_S - f(x^0)|$$

$$\leq \frac{1}{2} L_g \delta^2 + (c_1 + c_2)\delta^2 + (c_1 + c_2)\delta^2.$$

Novamente definindo a matriz

$$L_\ell = \begin{pmatrix} (x^1 - x^0)^\top \\ \vdots \\ (x^n - x^0)^\top \end{pmatrix} = \begin{pmatrix} x_1^1 - x_1^0 & \cdots & x_n^1 - x_n^0 \\ \vdots & \vdots & \vdots \\ x_1^n - x_1^0 & \cdots & x_n^n - x_n^0 \end{pmatrix}$$

obtemos

$$\left\|L_\ell \left(\nabla f(x^0) - \nabla m_S(x)\right)\right\| \leq \sqrt{n} \left\|L_\ell \left(\nabla f(x^0) - \nabla m_S(x)\right)\right\|_\infty \leq \left(\frac{1}{2} L_g + 2(c_1 + c_2)\right) \sqrt{n} \delta^2.$$

Como L_ℓ é não singular,

$$\left\|\nabla f(x^0) - \nabla m_S(x)\right\| \leq \|L_\ell^{-1}\| \left\|L_\ell \left(\nabla f(x^0) - \nabla m_S(x)\right)\right\| \leq \left(\frac{1}{2} L_g + 2(c_1 + c_2)\right) \sqrt{n} \|L_\ell^{-1}\| \delta^2.$$

Usando a definição de $\widehat{L}_\ell$, dada em (2.6), para todo $x \in \mathcal{D} \cap B(x^0, \delta)$ segue que

$$\|\nabla f(x^0) - \nabla m_S(x)\| \leq \left(\frac{1}{2} L_g + 2(c_1 + c_2)\right) \sqrt{n} \|\widehat{L}_\ell^{-1}\| \delta. \tag{2.20}$$

Considere agora $x \in \mathcal{D} \cap B(x^0, \delta)$ arbitrário. Assim, usando a desigualdade triangular, a Hipótese A1 e (2.20), segue que

$$\begin{aligned} \|\nabla f(x) - \nabla m_S(x)\| &\leq \|\nabla f(x) - \nabla f(x^0)\| + \|\nabla f(x^0) - \nabla m_S(x)\| \\ &\leq \left(L_g + \left(\frac{1}{2} L_g + 2(c_1 + c_2)\right) \sqrt{n} \|\widehat{L}_\ell^{-1}\|\right) \delta, \end{aligned}$$

concluindo a demonstração. □

O próximo teorema estabelece o erro dos valores funcionais entre o modelo linear de regressão por vetores suporte e a função a ser aproximada em pontos numa vizinhança

que contém o conjunto de amostra.

Teorema 2.11. *Considere o conjunto de amostra* $R = \{x^0, \ldots, x^n\} \subset \mathcal{D} \cap B(x^0, \delta)$ *de modo que o conjunto* $R^- = \{x^1, \ldots, x^n\}$ *seja linearmente independente e suponha que a Hipótese A1 seja satisfeita. Se o modelo linear* $m_S(x) = w_S^\top x + b_S$ *é construído via regressão por vetores suporte com margem* $\varepsilon \leq c_1\delta^2$ *e folgas* $\xi, \xi' \leq c_2\delta^2$*, com* $c_1, c_2 > 0$*, então para todo* $x \in \mathcal{D} \cap B(x^0, \delta)$

$$|f(x) - m_S(x)| \leq \kappa_7 \delta^2,$$

em que $\kappa_7 = \frac{3}{2}L_g + c_1 + c_2 + \left(\frac{1}{2}L_g + 2(c_1 + c_2)\right)\sqrt{n}\|\widehat{L}_\ell^{-1}\|$.

Demonstração. Temos que $m_S(x) = w_S^\top x + b_S$, portanto $\nabla m_S(x) = w_S$ para todo $x \in \mathcal{D} \cap B(x^0, \delta)$, de onde

$$m_S(x) = \nabla m_S(x^0)^\top x + b_S + m_S(x^0) - m_S(x^0) = m_S(x^0) + \nabla m_S(x^0)^\top (x - x^0).$$

Logo,

$$f(x) - m_S(x) = f(x) - m_S(x^0) - \nabla m_S(x^0)^\top (x - x^0).$$

Somando e subtraindo $f(x^0) + \nabla f(x^0)^\top (x - x^0)$ do lado direito obtemos

$$\begin{aligned} f(x) - m_S(x) &= f(x) - f(x^0) - \nabla f(x^0)^\top (x - x^0) + f(x^0) - m_S(x^0) + \\ &\quad \left(\nabla f(x^0) - \nabla m_S(x^0)\right)^\top (x - x^0). \end{aligned}$$

Pela desigualdade triangular, da igualdade anterior segue que

$$\begin{aligned} |f(x) - m_S(x)| &\leq |f(x) - f(x^0) - \nabla f(x^0)^\top (x - x^0)| + |f(x^0) - m_S(x^0)| + \\ &\quad |\left(\nabla f(x^0) - \nabla m_S(x^0)\right)^\top (x - x^0)|. \end{aligned} \tag{2.21}$$

Pelo Lema 2.1, como $x \in \mathcal{D} \cap B(x^0, \delta)$,

$$|f(x) - f(x^0) - \nabla f(x^0)^\top (x - x^0)| \leq \frac{1}{2} L_g \delta^2. \tag{2.22}$$

Por hipótese, o erro nos pontos de amostra é limitado pela soma da margem e da folga (2.18), e com isso

$$|f(x^0) - m_S(x^0)| \leq (c_1 + c_2)\delta^2. \tag{2.23}$$

Usando o Teorema 2.10,

$$|\left(\nabla f(x^0) - \nabla m_S(x^0)\right)^\top (x - x^0)| \leq \|\nabla f(x^0) - \nabla m_S(x^0)\|\|(x - x^0)\| \leq \kappa_6 \delta^2. \tag{2.24}$$

Usando (2.22), (2.23) e (2.24) na desigualdade (2.21), obtemos

$$|f(x) - m_S(x)| \leq \left(\frac{1}{2}L_g + c_1 + c_2 + \kappa_6\right)\delta^2,$$

concluindo a demonstração. □

2.3.2 Regressão quadrática por vetores suporte

Vamos agora considerar o caso em que queremos construir um modelo quadrático. Sejam $R = \{x^0, x^1, \ldots, x^p\} \subset \mathcal{D} \subset \mathbb{R}^n$ pontos nos quais conhecemos o valor da função objetivo f e uma aplicação $\varphi : \mathbb{R}^n \to \mathbb{R}^q$, onde $q = n(n+3)/2$, definida por

$$\varphi(x) = \left(x_1^2, \sqrt{2}x_1x_2, \sqrt{2}x_1x_3, \ldots, x_2^2, \ldots, \sqrt{2}x_{n-1}x_n, x_n^2, x_1, \ldots, x_n\right)^\top. \tag{2.25}$$

A aplicação φ leva os pontos de $\mathbb{R}^n$ a um conjunto de dimensão maior $\mathbb{R}^q$ onde o modelo é linear em $\varphi(x)$, mas quadrático em $x \in \mathbb{R}^n$.

A imagem $\varphi(R)$ do conjunto de amostra sob a aplicação φ é o conjunto

$$\varphi(R) = \{\varphi(x^0), \varphi(x^1), \ldots, \varphi(x^p)\}.$$

Analogamente à seção anterior, considere $\varepsilon \geq 0$ a tolerância permitida para o erro entre a função e o modelo que desejamos construir, $C > 0$ o parâmetro de regularização definido na Seção 1.2.2 e o problema

$$\begin{array}{ll} \text{minimizar} & \frac{1}{2}z^\top Qz + v^\top z \\ \text{sujeita a} & Az = 0 \\ & 0 \leq z \leq C, \end{array} \tag{2.26}$$

onde $Q \in \mathbb{R}^{2(p+1)\times 2(p+1)}$ é dada por

$$Q = \begin{pmatrix} MM^\top & -MM^\top \\ -MM^\top & MM^\top \end{pmatrix} \text{ com } M = \begin{pmatrix} \varphi(x^0)^\top \\ \varphi(x^1)^\top \\ \vdots \\ \varphi(x^p)^\top \end{pmatrix}, \tag{2.27}$$

e $v \in \mathbb{R}^{2(p+1)}$ e $A^\top \in \mathbb{R}^{1\times 2(p+1)}$ estão definidos em (2.17).

Apresentamos a seguir um algoritmo para a construção do modelo quadrático da função objetivo f por regressão via vetores suporte a partir do conjunto de amostra R.

Algoritmo 2.2. *Modelo quadrático por regressão via vetores suporte*

Dados: $R = \{x^0, x^1, \cdots, x^p\}$, $C > 0$ *e* $\varepsilon \geq 0$.
ENQUANTO b_S *não é calculado.*
  CALCULE $z \in \mathbb{R}^{2(p+1)}$ *como solução do problema* (2.26).
  FAÇA $\alpha = (z_1, z_2, \cdots, z_{p+1})^\top$ *e* $\gamma = (z_{p+2}, z_{p+3}, \cdots, z_{2(p+1)})^\top$.
  DEFINA $w_S = M^\top(\alpha - \gamma)$.
  SE $\alpha \neq 0$ *ou* $\gamma \neq 0$,
    ESCOLHA i *tal que* $0 < \alpha_i < C$ *e calcule* $b_S = f(x^i) - w_S^\top \varphi(x^i) - \varepsilon$.
    SE IMPOSSÍVEL, ESCOLHA i *tal que* $0 < \gamma_i < C$ *e calcule* $b_S = f(x^i) - w_S^\top \varphi(x^i) + \varepsilon$.
    SE IMPOSSÍVEL, AUMENTE C.
  SENÃO
    ESCOLHA $i \in \{0, \ldots, p\}$ *e calcule* $b_S = f(x^i)$.
DEFINA $m_S(x) = w_S^\top \varphi(x) + b_S$.

O modelo quadrático $m_S : \mathcal{D} \to \mathbb{R}$ construído por regressão via vetores suporte pode ser escrito como

$$m_S(x) = \frac{1}{2} x^\top H_S x + g_S^\top x + b_S,$$

com $H_S = H_S^\top \in \mathbb{R}^{n \times n}$, $g_S \in \mathbb{R}^n$ e $b_S \in \mathbb{R}$. Assim $\nabla m_S(x) = H_S x + g_S$.

Para o teorema a seguir, analogamente a (2.18), vamos admitir que $\varepsilon \leq c_1\delta^3$ e $\xi, \xi' \leq c_2\delta^3$. Ou seja,

$$|m_S(x^i) - f(x^i)| \leq (c_1 + c_2)\delta^3, \tag{2.28}$$

para todo $x^i \in R$.

Teorema 2.12. *Considere o conjunto de amostra* $R = \{x^0, \ldots, x^q\} \subset \mathcal{D} \cap B(x^0, \delta)$ *de modo que o conjunto* $\varphi(R^-) = \{\varphi(x^1), \ldots, \varphi(x^q)\}$ *seja linearmente independente e suponha que a Hipótese A2 seja satisfeita. Se o modelo quadrático* m_S *é construído via regressão por vetores suporte com margem* $\varepsilon \leq c_1\delta^3$ *e folgas* $\xi, \xi' \leq c_2\delta^3$, *com* $c_1, c_2 > 0$, *então para todo* $x \in \mathcal{D} \cap B(x^0, \delta)$

$$\|\nabla^2 f(x) - \nabla^2 m_S(x)\| \leq \kappa_8 \delta$$

e

$$\|\nabla f(x) - \nabla m_s(x)\| \leq \kappa_9 \delta^2,$$

em que

$$\kappa_8 = \left(\frac{3\sqrt{2}}{2} L_h + 2\sqrt{2}(c_1 + c_2)\right) \sqrt{q} \|\widehat{L}_q^{-1}\| \text{ e } \kappa_9 = \left(\frac{3 + 3\sqrt{2}}{2} L_h + (2 + 2\sqrt{2})(c_1 + c_2)\right) \sqrt{q} \|\widehat{L}_q^{-1}\|.$$

Demonstração. Para todo $i = 0, 1, \ldots, q$ e para todo $x \in \mathcal{D} \cap B(x^0, \delta)$,

$$\frac{1}{2}(x^i - x)^\top H_S(x^i - x) + \nabla m_S(x)^\top (x^i - x) =$$

$$= \frac{1}{2}(x^i)^\top H_S x^i - x^\top H_S x^i + \frac{1}{2}x^\top H_S x + (H_S x + g_S)^\top (x^i - x) =$$

$$= \frac{1}{2}(x^i)^\top H_S x^i - x^\top H_S x^i + \frac{1}{2}x^\top H_S x + x^\top H_S x^i + g_S^\top x^i - x^\top H_S x - g_S^\top x =$$

$$= \frac{1}{2}(x^i)^\top H_S x^i + g_S^\top x^i - \frac{1}{2}x^\top H_S x - g_S^\top x = m_S(x^i) - m_S(x).$$

Ou seja,

$$\nabla m_S(x)^\top (x^i - x) + \frac{1}{2}(x^i - x)^\top H_S(x^i - x) = m_S(x^i) - m_S(x) \tag{2.29}$$

para todo $i = 0, 1, \ldots, q$ e para todo $x \in \mathcal{D} \cap B(x^0, \delta)$.

Considerando a Hipótese A2, temos que a Hessiana da função é Lipschitz e portanto pelo Lema 2.2 segue que

$$\left| f(x^i) - f(x) - \nabla f(x)^\top (x^i - x) - \frac{1}{2}(x^i - x)^\top \nabla^2 f(x)(x^i - x) \right| \leq \frac{1}{6} L_h \|x^i - x\|^3,$$

para todo $i = 0, 1, \ldots, q$. O que implica que, para todo $x \in \mathcal{D} \cap B(x^0, \delta)$ e para $i = 0$,

$$\left| f(x^0) - f(x) - \nabla f(x)^\top (x^0 - x) - \frac{1}{2}(x^0 - x)^\top \nabla^2 f(x)(x^0 - x) \right| \leq \frac{1}{6} L_h \delta^3, \tag{2.30}$$

e para todo $x \in \mathcal{D} \cap B(x^0, \delta)$ e para $i = 1, \ldots, q$,

$$\left| f(x^i) - f(x) - \nabla f(x)^\top (x^i - x) - \frac{1}{2}(x^i - x)^\top \nabla^2 f(x)(x^i - x) \right| \leq \frac{4}{3} L_h \delta^3, \tag{2.31}$$

uma vez que $\|x^i - x\| \leq \|x^i - x^0\| + \|x^0 - x\| \leq 2\delta$.

Subtraindo $\nabla f(x)^\top (x^i - x) + \frac{1}{2}(x^i - x)^\top \nabla^2 f(x)(x^i - x)$ em ambos os lados de (2.29), obtemos que

$$(\nabla m_S(x) - \nabla f(x))^\top (x^i - x) + \frac{1}{2}(x^i - x)^\top \left(H_S - \nabla^2 f(x) \right) (x^i - x) =$$

$$m_S(x^i) - m_S(x) - \nabla f(x)^\top (x^i - x) - \frac{1}{2}(x^i - x)^\top \nabla^2 f(x)(x^i - x),$$

para todo $i = 0, 1, \ldots, q$.

Visto que,

$$\nabla m_S(x)^\top (x^i - x) - \nabla m_S(x)^\top (x^0 - x) = \nabla m_S(x)^\top (x^i - x^0)$$

e

$$\frac{1}{2}(x^i - x)^\top \left(H_S - \nabla^2 f(x)\right)(x^i - x) - \frac{1}{2}(x^0 - x)^\top \left(H_S - \nabla^2 f(x)\right)(x^i - x) =$$

$$\frac{1}{2}(x^i - x^0)^\top \left(H_S - \nabla^2 f(x)\right)(x^i - x^0) - (x^i - x^0)^\top \left(H_S - \nabla^2 f(x)\right)(x - x^0),$$

subtraindo o caso particular $i = 0$ obtemos para todo $i = 1, \ldots, q$,

$$(\nabla m_S(x) - \nabla f(x))^\top (x^i - x^0) + \frac{1}{2}(x^i - x^0)^\top \left(H_S - \nabla^2 f(x)\right)(x^i - x^0) -$$

$$(x^i - x^0)^\top \left(H_S - \nabla^2 f(x)\right)(x - x^0) = m_S(x^i) - m_S(x^0) - \nabla f(x)^\top (x^i - x) -$$

$$\frac{1}{2}(x^i - x)^\top \nabla^2 f(x)(x^i - x) + \nabla f(x)^\top (x^0 - x) + \frac{1}{2}(x^0 - x)^\top \nabla^2 f(x)(x^0 - x) =$$

$$m_S(x^i) - f(x^i) - m_S(x^0) + f(x^0) + f(x^i) - f(x) - \nabla f(x)^\top (x^i - x) - \frac{1}{2}(x^i - x)^\top \nabla^2 f(x)(x^i - x)$$

$$-f(x^0) + f(x) + \nabla f(x)^\top (x^0 - x) + \frac{1}{2}(x^0 - x)^\top \nabla^2 f(x)(x^0 - x).$$

Utilizando a desigualdade triangular, (2.30), (2.31) e a hipótese que o modelo é construído com margem $\varepsilon \leq c_1\delta^3$ e folgas $\xi, \xi' \leq c_2\delta^3$, ou seja vale (2.28), na igualdade acima obtemos

$$\left|\left(\nabla m_S(x) - \nabla f(x) + \frac{1}{2}\left(H_S - \nabla^2 f(x)\right)(x^i - x^0) - \left(H_S - \nabla^2 f(x)\right)(x - x^0)\right)^\top (x^i - x^0)\right|$$

$$\leq \left|m_S(x^i) - f(x^i)\right| + \left|f(x^0) - m_S(x^0)\right| +$$

$$\left|f(x^i) - f(x) - \nabla f(x)^\top (x^i - x) - \frac{1}{2}(x^i - x)^\top \nabla^2 f(x)(x^i - x)\right| +$$

$$+ \left|f(x^0) - f(x) - \nabla f(x)^\top (x^0 - x) - \frac{1}{2}(x^0 - x)^\top \nabla^2 f(x)(x^0 - x)\right|$$

$$\leq (c_1 + c_2)\delta^3 + (c_1 + c_2)\delta^3 + \frac{4}{3}L_h\delta^3 + \frac{1}{6}L_h\delta^3 = \left(\frac{3}{2}L_h + 2(c_1 + c_2)\right)\delta^3.$$

Considerando a matriz

$$L_q = \begin{pmatrix} (\bar{\varphi}(x^1 - x^0))^\top \\ \vdots \\ (\bar{\varphi}(x^q - x^0))^\top \end{pmatrix},$$

em que

$$\bar{\varphi}(x) = \left(x_1, x_2, \ldots, x_n, \frac{1}{2}x_1^2, x_1x_2, x_1x_3, \ldots, x_1x_n, \frac{1}{2}x_2^2, x_2x_3, \ldots, x_{n-1}x_n, \frac{1}{2}x_n^2\right)^\top,$$

segue da desigualdade anterior que

$$\left\| L_q \begin{pmatrix} e_S^g(x) \\ e_S^H(x) \end{pmatrix} \right\| \leq \sqrt{q} \left\| L_q \begin{pmatrix} e_S^g(x) \\ e_S^H(x) \end{pmatrix} \right\|_\infty \leq \left(\frac{3}{2} L_h + 2(c_1 + c_2) \right) \sqrt{q} \delta^3,$$

em que

$$e_S^g(x) = \nabla m_S(x) - \nabla f(x) - \left(H_S - \nabla^2 f(x) \right) (x - x^0) \tag{2.32}$$

e $e_S^H(x)$ é um vetor do $\mathbb{R}^{q-n}$ que armazena os elementos $(H_S - \nabla^2 f(x))_{kk}$, $k = 1, \ldots, n$ e $(H_S - \nabla^2 f(x))_{k\ell}$, $1 \leq \ell < k \leq n$.

Usando a matriz

$$\widehat{L}_q = L_q \begin{pmatrix} D_\delta^{-1} & 0 \\ 0 & D_{\delta^2}^{-1} \end{pmatrix},$$

em que $D_\delta = \delta I_{n \times n}$ e $D_{\delta^2} = \delta^2 I_{(q-n)\times(q-n)}$, na desigualdade acima, obtemos

$$\left\| \widehat{L}_q \begin{pmatrix} D_\delta e_S^g(x) \\ D_{\delta^2} e_S^H(x) \end{pmatrix} \right\| \leq \left(\frac{3}{2} L_h + 2(c_1 + c_2) \right) \sqrt{q} \delta^3.$$

Ou, consequentemente,

$$\left\| \begin{pmatrix} D_\delta e_S^g(x) \\ D_{\delta^2} e_S^H(x) \end{pmatrix} \right\| \leq \left(\frac{3}{2} L_h + 2(c_1 + c_2) \right) \sqrt{q} \|\widehat{L}_q^{-1}\| \delta^3.$$

Assim,

$$\begin{aligned} \|e_S^g(x)\| \leq \|D_\delta^{-1}\| \|D_\delta e_S^g(x)\| &\leq \left(\frac{3}{2} L_h + 2(c_1 + c_2) \right) \sqrt{q} \|\widehat{L}_q^{-1}\| \|D_\delta^{-1}\| \delta^3 \\ &\leq \left(\frac{3}{2} L_h + 2(c_1 + c_2) \right) \sqrt{q} \|\widehat{L}_q^{-1}\| \delta^2, \end{aligned} \tag{2.33}$$

e

$$\begin{aligned} \|e_S^H(x)\| \leq \|D_{\delta^2}^{-1}\| \|D_{\delta^2} e_S^H(x)\| &\leq \left(\frac{3}{2} L_h + 2(c_1 + c_2) \right) \sqrt{q} \|\widehat{L}_q^{-1}\| \|D_{\delta^2}^{-1}\| \delta^3 \\ &\leq \left(\frac{3}{2} L_h + 2(c_1 + c_2) \right) \sqrt{q} \|\widehat{L}_q^{-1}\| \delta. \end{aligned}$$

O erro nas Hessianas é portanto dado por

$$\begin{aligned} \left\| H_S - \nabla^2 f(x) \right\| &\leq \left\| H_S - \nabla^2 f(x) \right\|_F \leq \sqrt{2} \|e_S^H(x)\| \\ &\leq \left(\frac{3\sqrt{2}}{2} L_h + 2\sqrt{2}(c_1 + c_2) \right) \sqrt{q} \|\widehat{L}_q^{-1}\| \delta. \end{aligned}$$

Com isso, de (2.32) e (2.33), o erro nos gradientes é

$$\begin{aligned}\|\nabla m_S(x) - \nabla f(x)\| &\leq \|e_S^g(x)\| + \left\|H_S - \nabla^2 f(x)\right\| \|x - x^0\| \\ &\leq \left(\frac{3+3\sqrt{2}}{2} L_h + (2+2\sqrt{2})(c_1+c_2)\right) \sqrt{q}\|\widehat{L}_q^{-1}\|\delta^2,\end{aligned}$$

concluindo a demonstração. □

Teorema 2.13. *Considere o conjunto $R = \{x^0, \ldots, x^q\} \subset \mathcal{D} \cap B(x^0, \delta)$ de modo que o conjunto $\varphi(R^-) = \{\varphi(x^1), \ldots, \varphi(x^q)\}$ seja linearmente independente e suponha que a Hipótese A2 seja satisfeita. Se o modelo quadrático m_S é construído via regressão por vetores suporte com margem $\varepsilon \leq c_1\delta^3$ e folgas $\xi, \xi' \leq c_2\delta^3$, com $c_1, c_2 > 0$, então para todo $x \in \mathcal{D} \cap B(x^0, \delta)$*

$$|f(x) - m_S(x)| \leq \kappa_{10}\delta^3$$

em que $\kappa_{10} = \frac{1}{6}L_h + c_1 + c_2 + \left(\frac{6+9\sqrt{2}}{4} L_h + (2+3\sqrt{(2)})(c_1+c_2)\right)\sqrt{q}\|\widehat{L}_q^{-1}\|$.

Demonstração. O modelo quadrático pode ser escrito como

$$m_S(x) = m_S(x^0) + \nabla m_S(x^0)^\top (x - x^0) + \frac{1}{2}(x - x^0)^\top H_S (x - x^0),$$

para todo $x \in \mathcal{D} \cap B(x^0, \delta)$, então

$$\begin{aligned}f(x) - m_S(x) &= f(x) - m_S(x^0) - \nabla m_S(x^0)^\top (x - x^0) - \frac{1}{2}(x - x^0)^\top H_S (x - x^0) \\ &= f(x) - f(x^0) - \nabla f(x^0)^\top (x - x^0) - \frac{1}{2}(x - x^0)^\top \nabla^2 f(x^0)(x - x^0) + \\ &\quad f(x^0) - m_S(x^0) + \left(\nabla f(x^0) - \nabla m_S(x^0)\right)^\top (x - x^0) + \\ &\quad \frac{1}{2}(x - x^0)\left(\nabla^2 f(x^0) - H_S\right)(x - x^0).\end{aligned}$$

Utilizando o Lema 2.2, a hipótese que o modelo é construído com margem $\varepsilon \leq c_1\delta^3$ e folgas $\xi, \xi' \leq c_2\delta^3$ nos pontos da amostra e o Teorema 2.12, segue que

$$|f(x) - m_S(x)| \leq \left(\frac{1}{6}L_h + c_1 + c_2\right)\delta^3 + \frac{1}{2}\kappa_8\delta^3 + \kappa_9\delta^3,$$

concluindo a demonstração. □

Vamos agora considerar o caso em que queremos construir um modelo quadrático para uma função f que satisfaz a Hipótese A1, mas não necessariamente satisfaz a Hipótese A2. Vamos considerar $R = \{x^0, x^1, \ldots, x^q\} \subset \mathcal{D} \cap B(x^0, \delta)$ em que $q = n(n+3)/2$.

Novamente vamos admitir que o modelo é construído com margem $\varepsilon \leq c_1\delta^2$ e folgas

$\xi, \xi' \leq c_2\delta^2$, ou seja, para todo x^i no conjunto de amostra

$$|m(x^i) - f(x^i)| \leq (c_1 + c_2)\delta^2. \tag{2.34}$$

Teorema 2.14. *Considere o conjunto de amostra $R = \{x^0, \ldots, x^q\} \subset \mathcal{D} \cap B(x^0, \delta)$ de modo que o conjunto $\varphi(R^-) = \{\varphi(x^1), \ldots, \varphi(x^q)\}$ seja linearmente independente e suponha que a Hipótese A1 seja satisfeita. Se o modelo quadrático m_S é construído por regressão via vetores suporte com margem $\varepsilon \leq c_1\delta^2$ e folgas $\xi, \xi' \leq c_2\delta^2$, com $c_1, c_2 > 0$, então para todo $x \in \mathcal{D} \cap B(x^0, \delta)$*

$$\|\nabla^2 m_S(x)\| \leq \kappa_{11}$$

e

$$\|\nabla f(x) - \nabla m_S(x)\| \leq \kappa_{12}\delta,$$

em que

$$\kappa_{11} = 2\sqrt{2}(c_1 + c_2 + L_g)\sqrt{q}\|\widehat{L}_q^{-1}\| \text{ e } \kappa_{12} = \left((2 + 2\sqrt{2})(c_1 + c_2 + L_g)\right)\sqrt{q}\|\widehat{L}_q^{-1}\|.$$

Demonstração. Como o modelo m_S é quadrático temos que,

$$m_S(x^i) = m_S(x) + \nabla m_S(x)^\top (x^i - x) + \frac{1}{2}(x^i - x)^\top H_S(x^i - x)$$

para todo $i = 0, 1, \ldots, q$ e para todo $x \in \mathcal{D} \cap B(x^0, \delta)$.

Pelo Teorema do Valor Médio, para todo $i = 0, 1, \ldots, q$ existe $y^i \in [x^0, x^i] \subset B(x^0, \delta)$ tal que

$$f(x^i) = f(x^0) + \nabla f(y^i)^\top (x^i - x^0).$$

Com isso temos que

$$\begin{aligned} m_S(x^i) - f(x^i) &= m_S(x) + \nabla m_S(x)^\top (x^i - x) + \frac{1}{2}(x^i - x)^\top H_S(x^i - x) \\ &\quad - f(x^0) - \nabla f(y^i)^\top (x^i - x^0) \end{aligned} \tag{2.35}$$

para todo $i = 0, 1, \ldots, q$.

Subtraindo de (2.35) o caso particular

$$m_S(x^0) - f(x^0) = m_S(x) + \nabla m_S(x)^\top (x^0 - x) + \frac{1}{2}(x^0 - x)^\top H_S(x^i - x) - f(x^0),$$

obtemos que para todo $i = 1, \ldots, q$ vale

$$m_S(x^i) - f(x^i) + f(x^0) - m_S(x^0) + \nabla f(y^i)^\top (x^i - x^0) =$$

$$\nabla m_S(x)^\top (x^i - x) - \nabla m_S(x)^\top (x^0 - x) + \frac{1}{2}(x^i - x)^\top H_S(x^i - x) - \frac{1}{2}(x^0 - x)^\top H_S(x^i - x). \quad (2.36)$$

Por outro lado,

$$\nabla m_S(x)^\top (x^i - x) - \nabla m_S(x)^\top (x^0 - x) = \nabla m_S(x)^\top (x^i - x^0)$$

e

$$\frac{1}{2}(x^i - x)^\top H_S(x^i - x) - \frac{1}{2}(x^0 - x)^\top H_S(x^i - x) =$$

$$\frac{1}{2}(x^i)^\top H_S x^i - x^\top H_S x^i + \frac{1}{2}x^\top H_S x - \frac{1}{2}(x^0)^\top H_S x^0 + x^\top H_S x^0 - \frac{1}{2}x^\top H_S x =$$

$$\frac{1}{2}(x^i - x^0)^\top H_S(x^i - x^0) - x^\top H_S x^i + (x^i)^\top H_S x^0 - (x^0)^\top H_S x^0 + x^\top H_S x^0 =$$

$$\frac{1}{2}(x^i - x^0)^\top H_S(x^i - x^0) - (x^i - x^0)^\top H_S(x - x^0).$$

Substituindo essas igualdades em (2.36) e subtraindo $\nabla f(x)^\top (x^i - x^0)$ de ambos os lados obtemos

$$m_S(x^i) - f(x^i) + f(x^0) - m_S(x^0) + (\nabla f(y^i) - \nabla f(x))^\top (x^i - x^0) =$$

$$(\nabla m_S(x) - \nabla f(x))^\top (x^i - x^0) + \frac{1}{2}(x^i - x^0)^\top H_S(x^i - x^0) - (x^i - x^0)^\top H_S(x - x^0).$$

Ou simplificando,

$$m_S(x^i) - f(x^i) + f(x^0) - m_S(x^0) + (\nabla f(y^i) - \nabla f(x))^\top (x^i - x^0) =$$

$$\left(\nabla m_S(x) - \nabla f(x) - H_S(x - x^0) + \frac{1}{2}H_S(x^i - x^0)\right)^\top (x^i - x^0).$$

Utilizando a desigualdade triangular, a hipótese que o modelo é construído com margem $\varepsilon \leq c_1\delta^2$ e folgas $\xi, \xi' \leq c_2\delta^2$, ou seja, nos pontos da amostra vale (2.34) e a Hipótese A1 obtemos que para todo $x \in \mathcal{D} \cap B(x^0, \delta)$ e para $i = 1, \ldots, q$,

$$\left|\left(\nabla m_S(x) - \nabla f(x) - H_S(x - x^0) + \frac{1}{2}H_S(x^i - x^0)\right)^\top (x^i - x^0)\right| \leq 2(c_1 + c_2 + L_g)\delta^2, \quad (2.37)$$

visto que $\|y^i - x\| \leq 2\delta$.

Considerando a matriz

$$L_q = \begin{pmatrix} (\bar{\varphi}(x^1 - x^0))^\top \\ \vdots \\ (\bar{\varphi}(x^q - x^0))^\top \end{pmatrix},$$

em que

$$\bar{\varphi}(x) = \left(x_1, x_2, \ldots, x_n, \frac{1}{2}x_1^2, x_1x_2, x_1x_3, \ldots, x_1x_n, \frac{1}{2}x_2^2, x_2x_3, \ldots, x_{n-1}x_n, \frac{1}{2}x_n^2\right)^\top,$$

segue da desigualdade anterior que

$$\left\| L_q \begin{pmatrix} r_S^g(x) \\ r_S^H(x) \end{pmatrix} \right\| \leq \sqrt{q} \left\| L_q \begin{pmatrix} r_S^g(x) \\ r_S^H(x) \end{pmatrix} \right\|_\infty \leq 2(c_1 + c_2 + L_g)\sqrt{q}\delta^2,$$

em que

$$r_S^g(x) = \nabla m_S(x) - \nabla f(x) - H_S(x - x^0) \tag{2.38}$$

e $r_S^H(x)$ é um vetor do $\mathbb{R}^{q-n}$ que armazena os elementos $(H_S)_{kk}$, $k = 1, \ldots, n$ e $(H_S)_{k\ell}$, $1 \leq \ell < k \leq n$.

Usando a matriz

$$\widehat{L}_q = L_q \begin{pmatrix} D_\delta^{-1} & 0 \\ 0 & D_{\delta^2}^{-1} \end{pmatrix},$$

em que $D_\delta = \delta I_{n\times n}$ e $D_{\delta^2} = \delta^2 I_{(q-n)\times(q-n)}$, na desigualdade acima, obtemos

$$\left\| \widehat{L}_q \begin{pmatrix} D_\delta r_S^g(x) \\ D_{\delta^2} r_S^H(x) \end{pmatrix} \right\| \leq 2(c_1 + c_2 + L_g)\sqrt{q}\delta^2.$$

Ou, consequentemente,

$$\left\| \begin{pmatrix} D_\delta e_S^g(x) \\ D_{\delta^2} e_S^H(x) \end{pmatrix} \right\| \leq 2(c_1 + c_2 + L_g)\sqrt{q}\|\widehat{L}_q^{-1}\|\delta^2.$$

Assim,

$$\begin{aligned} \|r_S^g(x)\| \leq \|D_\delta^{-1}\|\|D_\delta r_S^g(x)\| &\leq 2(c_1 + c_2 + L_g)\sqrt{q}\|\widehat{L}_q^{-1}\|\|D_\delta^{-1}\|\delta^2 \\ &\leq 2(c_1 + c_2 + L_g)\sqrt{q}\|\widehat{L}_q^{-1}\|\delta, \end{aligned} \tag{2.39}$$

e

$$\begin{aligned} \|r_S^H(x)\| \leq \|D_{\delta^2}^{-1}\|\|D_{\delta^2} r_S^H(x)\| &\leq 2(c_1 + c_2 + L_g)\sqrt{q}\|\widehat{L}_q^{-1}\|\|D_{\delta^2}^{-1}\|\delta^2 \\ &\leq 2(c_1 + c_2 + L_g)\sqrt{q}\|\widehat{L}_q^{-1}\|. \end{aligned}$$

Portanto a Hessiana do modelo é limitada por

$$\begin{aligned}\|H_S\| &\leq \|H_S\|_F \leq \sqrt{2}\|r_S^H(x)\| \\ &\leq 2\sqrt{2}(c_1 + c_2 + L_g)\sqrt{q}\|\widehat{L}_q^{-1}\|.\end{aligned}$$

Com isso, (2.38) e (2.39), o erro nos gradientes é

$$\begin{aligned}\|\nabla m_S(x) - \nabla f(x)\| &\leq \|r_S^g(x)\| + \|H_S\|\,\|x - x^0\| \\ &\leq (2 + 2\sqrt{2})(c_1 + c_2 + L_g)\sqrt{q}\|\widehat{L}_q^{-1}\|\delta,\end{aligned}$$

concluindo a demonstração. □

Teorema 2.15. *Considere o conjunto $R = \{x^0, \ldots, x^q\} \subset \mathcal{D} \cap B(x^0, \delta)$ de modo que o conjunto $\varphi(R^-) = \{\varphi(x^1), \ldots, \varphi(x^q)\}$ seja linearmente independente e suponha que a Hipótese A1 seja satisfeita. Se o modelo quadrático m_S é construído via regressão por vetores suporte com margem $\varepsilon \leq c_1\delta^2$ e folgas $\xi, \xi' \leq c_2\delta^2$, com $c_1, c_2 > 0$, então para todo $x \in \mathcal{D} \cap B(x^0, \delta)$*

$$|f(x) - m_S(x)| \leq \kappa_{13}\delta^2$$

em que $\kappa_{13} = \frac{1}{2}L_g + c_1 + c_2 + (2 + 3\sqrt{2})(L_g + c_1 + c_2)\sqrt{q}\|\widehat{L}_q^{-1}\|$.

Demonstração. O modelo quadrático pode ser escrito como

$$m_S(x) = m_S(x^0) + \nabla m_S(x^0)^\top(x - x^0) + \frac{1}{2}(x - x^0)^\top H_S(x - x^0),$$

para todo $x \in \mathcal{D} \cap B(x^0, \delta)$. Então

$$\begin{aligned}f(x) - m_S(x) &= f(x) - m_S(x^0) - \nabla m_S(x^0)^\top(x - x^0) - \frac{1}{2}(x - x^0)^\top H_S(x - x^0) \\ &= f(x) - f(x^0) - \nabla f(x^0)^\top(x - x^0) + f(x^0) - m_S(x^0) + \\ &\quad \left(\nabla f(x^0) - \nabla m_S(x^0)\right)^\top (x - x^0) - \frac{1}{2}(x - x^0)H_S(x - x^0).\end{aligned}$$

Utilizando o Lema 2.1, a hipótese que o modelo é construído com margem $\varepsilon \leq c_1\delta^2$ e folgas $\xi, \xi' \leq c_2\delta^2$ nos pontos da amostra e o Teorema 2.14 segue que

$$|f(x) - m_S(x)| \leq \left(\frac{1}{2}L_g + c_1 + c_2\right)\delta^2 + \frac{1}{2}\kappa_{11}\delta^2 + \kappa_{12}\delta^2,$$

concluindo a demonstração. □

2.4 Controle da geometria

Considere as matrizes $\widehat{L}_\ell$ e $\widehat{L}_q$ definidas em (2.6) e (2.10), respectivamente. Note que as constantes κ_i, com $i = \{1, \ldots, 13\}$, definidas nas seções anteriores estão relacionadas com $\|\widehat{L}_\ell^{-1}\|$ e $\|\widehat{L}_q^{-1}\|$. Se assumirmos que essas matrizes são limitadas por constantes que não dependem de δ, o erro entre o gradiente da função e o gradiente dos modelos é pelo menos linear em δ e o erro entre a função e os modelos é pelo menos quadrático em δ, dependendo das hipóteses sobre a função. Em geral, essa limitação é alcançada controlando a geometria do conjunto de amostra [11].

Garantir uma boa geometria dos pontos da amostra é uma condição usual em métodos de região de confiança livre de derivadas quando os modelos são construídos por interpolação polinomial. Vamos mostrar que as técnicas usadas para controlar a geometria no contexto de interpolação podem ser empregadas quando os modelos são construídos por máquinas de vetores suporte.

A definição a seguir busca controlar o posicionamento do conjunto R, para garantir limitações em $\|\widehat{L}_\ell^{-1}\|$e em $\|\widehat{L}_q^{-1}\|$.

Definição 2.16. *[10, Definição 3.6] Considere $\phi = \{\phi_0(x), \phi_1(x), \ldots, \phi_q(x)\}$ uma base para $\mathcal{P}_n^a$. Um conjunto $R = \{x^0, \ldots, x^q\} \subset \mathcal{D}$ é Λ-posicionado em $B(x^0, \delta)$ em relação à base ϕ, para uma constante $\Lambda > 0$, se e somente se, para todo $x \in B(x^0, \delta)$, existe $\lambda(x) \in \mathbb{R}^{q+1}$ tal que*

$$\sum_{i=0}^{q} \lambda_i(x)\phi(x^i) = \phi(x) \qquad com \qquad \|\lambda(x)\|_\infty \leq \Lambda.$$

Podemos ver a definição de Λ-posicionamento como um sistema linear em que a norma da solução $\lambda(x) \in \mathbb{R}^{q+1}$ é limitada para todo $x \in B(x^0, \delta)$, ou seja,

$$M(\phi, R)^\top \lambda(x) = \phi(x) \qquad \text{com} \qquad \|\lambda(x)\|_\infty \leq \Lambda, \tag{2.40}$$

em que

$$M(\phi, R) = \begin{pmatrix} \phi_0(x^0) & \phi_1(x^0) & \phi_2(x^0) & \cdots & \phi_q(x^0) \\ \phi_0(x^1) & \phi_1(x^1) & \phi_2(x^1) & \cdots & \phi_q(x^1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_0(x^q) & \phi_1(x^q) & \phi_2(x^q) & \cdots & \phi_q(x^q) \end{pmatrix}.$$

Considere $[I]_A^B$ a matriz mudança de base de uma base ϕ_A para uma base ϕ_B. A

solução $\lambda(x)$ de (2.40) não depende da escolha da base, uma vez que

$$M(\phi_B, R)^\top \lambda(x) = \phi_B(x) \quad \Longrightarrow \quad [I]_A^B M(\phi_A, R)^\top \lambda(x) = [I]_A^B \phi_A(x).$$

Os resultados a seguir mostram que a constante Λ também não é alterada quando mudamos a escala ou efetuamos translações no conjunto de amostra. O primeiro resultado trata da mudança de escala.

Lema 2.17. *[10, Lema 3.8] Sejam $R = \{x^0, x^1, \ldots, x^q\}$ um conjunto de pontos e $\lambda(x) \in \mathbb{R}^{q+1}$ a solução de* (2.40) *para R e uma base ϕ. Então, para qualquer $\delta > 0$, $\lambda(x/\delta)$ é solução de* (2.40) *para $\hat{R}$, em que $\hat{R} = \{x^0/\delta, x^1/\delta, \ldots, x^q/\delta\}$.*

Demonstração. Como a solução $\lambda(x)$ de (2.40) não depende da escolha da base, vamos considerar a base natural $\bar{\phi}$. Temos que $\lambda_i(x)$, $i = 0, \ldots, q$, satisfaz

$$\sum_{i=0}^{q} \lambda_i(x) \bar{\phi}(x^i) = \bar{\phi}(x), \tag{2.41}$$

para todo $x \in B(x^0, \delta)$.

Se multiplicamos cada x^i e x por $1/\delta$, isto corresponde a multiplicar as linhas do sistema (2.41) por diferente escalares, a saber $(1, 1/\delta, 1/\delta^2, \ldots, 1/\delta^a)$ os quais dependem do grau do polinômio da base natural que está em cada linha. E temos que $\lambda(x/\delta)$ satisfaz este novo sistema para $\hat{R}$. □

Segue diretamente do lema anterior o seguinte corolário.

Corolário 2.18. *Se R é Λ-posicionado em $B(x^0, \delta)$, então $\hat{R} = R/\delta$ é Λ-posicionado em $B(x^0, 1)$.*

O próximo resultado mostra que translações do conjunto de amostra não alteram o Λ-posicionamento.

Lema 2.19. *[10, Lema 3.9] Sejam $R = \{x^0, x^1, \ldots, x^q\}$ um conjunto de pontos e $\lambda(x) \in \mathbb{R}^{q+1}$ a solução de* (2.40) *para x dado. Então, para qualquer $a \in \mathbb{R}^n$, $\lambda(x)$ também é solução de* (2.40) *para $R_a = \{x^0 + a, x^1 + a, \ldots, x^q + a\}$ e $x_a = x + a$.*

Demonstração. A solução de (2.40) não depende da escolha da base. Trabalhando com a base natural $\bar{\phi}$ com a notação de índice múltiplo (2.3), temos para $i = 0, \ldots, q$,

$$\bar{\phi}_i(x) = \frac{1}{(\tau^i)!} x^{\tau^i}.$$

Segue que

$$\begin{aligned}\bar{\phi}_i(x+a) &= \frac{1}{(\tau^i)!}(x+a)^{\tau^i} \\ &= \frac{1}{(\tau^i)!}(x^{\tau^i} + x^{\tau^i-1}a + x^{\tau^i-2}a^2 + \ldots + xa^{\tau^i-1} + a^{\tau^i}) \\ &= \bar{\phi}_i(x) + \sum_{k:|\tau^k|<|\tau^i|} \bar{\nu}_k(a)\bar{\phi}_k(x),\end{aligned}$$

onde $\bar{\nu}_k(a)$ são coeficientes que dependem de a mas não de x. Então, existe uma relação biunívoca entre $\bar{\phi}_i(x)$ e $\bar{\phi}_j(x+a)$, logo $\bar{\bar{\phi}} = \{\bar{\phi}_0(x+a), \ldots, \bar{\phi}_q(x+a)\}$ define uma base em $\mathcal{P}_n^a$. Novamente usando o fato que a solução de (2.40) não depende da base, temos

$$\sum_{i=0}^{q} \lambda_i(x)\bar{\phi}(x^i + a) = \bar{\phi}(x+a),$$

concluindo a demonstração. □

Para o caso da construção de modelos lineares, precisamos limitar $\|\widehat{L}_\ell^{-1}\|$ e para isso vamos considerar o seguinte lema.

Lema 2.20. *Considere $R = \{x^0, x^1 \ldots, x^n\} \subset \mathbb{R}^n$ um conjunto Λ-posicionado em $B(x^0, \delta)$ com relação a uma base ϕ de $\mathcal{P}_n^1$. Então para todo $x \in B(x^0, \delta)$, existe $\widehat{\lambda}(x) \in \mathbb{R}^n$ tal que*

$$x - x^0 = \sum_{i=1}^{n} \widehat{\lambda}_i(x)(x^i - x^0) \qquad com \qquad |\widehat{\lambda}_i| \leq \Lambda, \quad i = 1, 2, \ldots, n.$$

Demonstração. Temos que $\phi = \{1, x_1, x_2, \ldots, x_n\}$ é uma base para o espaço $\mathcal{P}_n^1$ dos polinômios lineares no $\mathbb{R}^n$, com isso podemos escrever

$$M(\phi, R) = \begin{pmatrix} 1 & x_1^0 & x_2^0 & \cdots & x_n^0 \\ 1 & x_1^1 & x_2^1 & \cdots & x_n^1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^n & x_2^n & \cdots & x_n^n \end{pmatrix}.$$

Pela Definição 2.16 de Λ-posicionamento temos que para todo $x \in B(x^0, \delta)$,

$$M(\phi, R)^\top \lambda(x) = \phi(x) \qquad \text{com} \qquad \|\lambda(x)\|_\infty \leq \Lambda.$$

Considere $\bar{B}$ uma matriz invertível, assim

$$M(\phi, R)^\top \bar{B}\bar{B}^{-1}\lambda(x) = \phi(x) \qquad \text{com} \qquad \|\lambda(x)\|_\infty \leq \Lambda. \tag{2.42}$$

Em particular para

$$\bar{B} = \begin{pmatrix} 1 & -1 & -1 & \cdots & -1 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix},$$

temos que

$$\bar{B}^{-1} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}.$$

Neste caso

$$M(\phi, R)^\top \bar{B} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ x_1^0 & x_1^1 - x_1^0 & x_1^2 - x_1^0 & \cdots & x_1^n - x_1^0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n^0 & x_n^1 - x_n^0 & x_n^2 - x_n^0 & \cdots & x_n^n - x_n^0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ x^0 & L_\ell^\top \end{pmatrix},$$

e

$$\bar{B}^{-1}\lambda(x) = \begin{pmatrix} \sum_{i=0}^n \lambda_i(x) \\ \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Substituindo isto em (2.42) e observando que $\phi(x) = \begin{pmatrix} 1 \\ x \end{pmatrix}$, temos que

$$x^0 + L_\ell^\top \widehat{\lambda}(x) = x \quad \Rightarrow \quad L_\ell^\top \widehat{\lambda}(x) = x - x^0,$$

onde $\widehat{\lambda}(x) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Usando a definição da matriz L_ℓ segue, para todo $x \in B(x^0, \delta)$, que

$$\sum_{i=1}^n \widehat{\lambda}_i(x)(x^i - x^0) = x - x^0 \quad \text{com} \quad \|\widehat{\lambda}(x)\|_\infty \leq \Lambda,$$

o que completa a demonstração. □

A definição a seguir e o Lema 2.22 serão úteis para conseguirmos uma limitação para $\|\widehat{L}_\ell^{-1}\|$ quando usamos a definição de Λ-posicionamento para o conjunto de amostra, que será feita no Lema 2.23.

Definição 2.21. *Considere $A \in \mathbb{R}^{n\times m}$. Um vetor $v \in \mathbb{R}^m$ é dito vetor singular à direita de A associado a um valor singular σ se existe $u \in \mathbb{R}^n$ tal que $Av = \sigma u$ e $A^\top u = \sigma v$. Consequentemente u é dito vetor singular à esquerda de A.*

Note que $\sigma \geq 0$, uma vez que $A^\top A$ é simétrica e semidefinida positiva e os valores singulares de uma matriz A são raízes quadradas dos autovalores de $A^\top A$.

Lema 2.22. *[10, Lema 3.13] Considere $A \in \mathbb{R}^{n\times n}$ uma matriz não singular e $v \in \mathbb{R}^n$ o vetor unitário singular à direita associado ao maior valor singular de A. Então, para qualquer vetor $r \in \mathbb{R}^n$,*

$$|v^\top r|\|A\| \leq \|Ar\|.$$

Demonstração. Considere σ_1 o maior valor singular de A. Assim

$$|v^\top r|\|A\| = \sigma_1|v^\top r| = |\sigma_1 v^\top r|. \tag{2.43}$$

Como v é vetor singular à direita, existe $u \in \mathbb{R}^n$ tal que $\sigma_1 v^\top = u^\top A$. Consequentemente,

$$|v^\top r|\|A\| = |\sigma_1 v^\top r| = |u^\top Ar| \leq \|u\|\|Ar\|. \tag{2.44}$$

Mas como $\|A\| = \sigma_1$ e v é unitário,

$$\|u\| = \frac{1}{\sigma_1}\|Av\| \leq \frac{1}{\sigma_1}\|A\|\|v\| = 1,$$

o que conclui a demonstração. □

Com esses resultados conseguimos agora limitar $\|\widehat{L}_\ell^{-1}\|$ por uma constante que não dependa do tamanho da região de amostra, desde que o conjunto seja Λ-posicionado em $B(x^0, \delta)$.

Lema 2.23. *Considere $R = \{x^0, x^1, \ldots, x^n\}$ um conjunto Λ-posicionado em $B(x^0, \delta)$ e $\widehat{L}_\ell$ a matriz definida em* (2.6)*. Então*

$$\|\widehat{L}_\ell^{-1}\| \leq \sqrt{n}\Lambda.$$

Demonstração. Considere v um vetor unitário singular à direita correspondente ao maior valor singular σ_1 de $\widehat{L}_\ell^{-1}$. Assim, existe $u \in \mathbb{R}^n$ unitário tal que $\widehat{L}_\ell^{-1}v = \sigma_1 u$. Consequentemente, pela definição de norma euclidiana de matrizes,

$$\|\widehat{L}_\ell^{-1}v\| = \sigma_1\|u\| = \sigma_1 = \|\widehat{L}_\ell^{-1}\|. \tag{2.45}$$

Pelos Lemas 2.17 e 2.19, segue que o conjunto $\widehat{R} = \left\{0, \frac{x^1-x^0}{\delta}, \ldots, \frac{x^n-x^0}{\delta}\right\}$ é Λ-posicionado em $B(0,1)$. Pelo Lema 2.20, existe $\lambda(v) \in \mathbb{R}^n$ com $\|\lambda(v)\|_\infty \leq \Lambda$ tal que

$$\widehat{L}_\ell\lambda(v) = v \qquad \Rightarrow \qquad \lambda(v) = \widehat{L}_\ell^{-1}v.$$

Usando isso, (2.45) e o fato que o conjunto R é Λ-posicionado,

$$\|\widehat{L}_\ell^{-1}\| = \|\lambda(v)\| \leq \sqrt{n}\|\lambda(v)\|_\infty \leq \sqrt{n}\Lambda,$$

concluindo a demonstração. □

Para limitar $\|\widehat{L}_q^{-1}\|$, que aparece quando os modelos são quadráticos, o resultado a seguir será importante.

Lema 2.24. *[10, Lema 6.7] Considere $q(x) = v^\top \bar{\phi}(x)$ um polinômio quadrático, onde $\|v\|_\infty = 1$ e $\bar{\phi}$ é a base natural para $\mathcal{P}_n^2$. Então*

$$\max_{x \in B(0,1)} |v^\top \bar{\phi}(x)| \geq \frac{1}{4}.$$

Demonstração. Temos por hipótese que $\|v\|_\infty = 1$, logo pelo menos uma das componentes de v é 1 ou -1. Então o polinômio $q(x) = v^\top \bar{\phi}(x)$ tem um coeficiente igual a -1, 1, $-\frac{1}{2}$ ou $\frac{1}{2}$. Vamos analisar quando esses coeficientes são positivos. O caso em que são negativos pode ser analisado de modo análogo.

O maior coeficiente em valor absoluto em v corresponde ao termo constante, ou a um termo linear x_i ou ainda a um termo quadrático $x_i^2/2$ ou $x_i x_j$. Vamos mostrar que o máximo valor absoluto do polinômio é pelo menos $\frac{1}{4}$ considerando os 4 casos que correspondem aos maiores coeficientes.

(i) O primeiro caso é quando $q(0) = 1$, em que trivialmente temos $|q(x)| \geq \frac{1}{4}$.

(ii) No segundo caso fazemos $x = e_i$, com i a componente de v que é igual a 1 correspondente a um termo linear de $\bar{\phi}(x)$. Neste caso temos

$$q(e_i) = \alpha/2 + 1 + \beta \quad \text{e} \quad q(-e_i) = \alpha/2 - 1 + \beta,$$

que implica em $\max\{|q(e_i)|, |q(-e_i)|\} \geq 1$.

(iii) No terceiro caso fazemos $x = e_i$, com i a componente de v que é igual a 1 correspondente a um termo quadrático de $\bar{\phi}(x)$ do tipo $x_k^2/2$. Neste caso temos

$$q(e_i) = \frac{1}{2} + \alpha + \beta, \quad q(-e_i) = \frac{1}{2} - \alpha + \beta, \quad q(0) = \beta.$$

Se $|q(e_i)| \geq 1/4$ ou $|q(-e_i)| \geq 1/4$, obtemos o resultado. Por outro lado, se $|q(e_i)| < 1/4$ e $|q(-e_i)| < 1/4$, temos

$$|q(e_i) + q(-e_i)| \leq |q(e_i)| + |q(-e_i)| < \frac{1}{2},$$

logo

$$\left|\frac{1}{2}+\alpha+\beta+\frac{1}{2}-\alpha+\beta\right|<\frac{1}{2},$$

que por sua vez implica em $|1+2\beta|<1/2$, logo $\beta<-1/4$. Como $q(0)=\beta$, temos $|q(0)|>1/4$.

(iv) No quarto caso, consideramos $x=ae_i+be_j$, onde i e j são tais que a componente de v que é igual a 1 correspondente a um termo quadrático de $\bar{\phi}(x)$ do tipo x_kx_l. Neste caso, temos $q(ae_i+be_j)=\alpha a^2/2+\beta b^2/2+ab+\gamma a+\delta b+\epsilon$. Vamos considerar quatro pontos na bola $B(0,1)$:

$$p_1=\frac{\sqrt{2}}{2}e_i+\frac{\sqrt{2}}{2}e_j, p_2=\frac{\sqrt{2}}{2}e_i-\frac{\sqrt{2}}{2}e_j, p_3=\frac{\sqrt{2}}{2}e_j-\frac{\sqrt{2}}{2}e_i \text{ e } p_4=-\frac{\sqrt{2}}{2}e_i-\frac{\sqrt{2}}{2}e_j.$$

Assim,

$$\begin{aligned}
q(p_1)&=\frac{\alpha}{4}+\frac{\beta}{4}+\frac{1}{2}+\frac{\gamma}{\sqrt{2}}+\frac{\delta}{\sqrt{2}}+\epsilon,\\
q(p_2)&=\frac{\alpha}{4}+\frac{\beta}{4}-\frac{1}{2}+\frac{\gamma}{\sqrt{2}}-\frac{\delta}{\sqrt{2}}+\epsilon,\\
q(p_3)&=\frac{\alpha}{4}+\frac{\beta}{4}-\frac{1}{2}-\frac{\gamma}{\sqrt{2}}+\frac{\delta}{\sqrt{2}}+\epsilon,\\
q(p_4)&=\frac{\alpha}{4}+\frac{\beta}{4}+\frac{1}{2}-\frac{\gamma}{\sqrt{2}}-\frac{\delta}{\sqrt{2}}+\epsilon.
\end{aligned}$$

E com isso, tem-se

$$q(p_1)-q(p_2)=1+\delta\sqrt{2} \quad \text{e} \quad q(p_3)-q(p_4)=-1+\delta\sqrt{2}.$$

Se $\delta\geq 0$, então $q(p_1)-q(p_2)\geq 1$, o que implica que se $|q(p_1)|<1/2$, então $q(p_2)\leq -1/2$. O caso $\delta<0$ é análogo, ao analisarmos $q(p_3)-q(p_4)\leq -1$. Logo, existe um ponto $\bar{x}\in B(0,1)$ tal que $|q(\bar{x})|\geq 1/2$.

Considerando os quatro casos, provamos o lema. □

O Lema 2.24 trata de uma estimativa para o caso em que $\|v\|_\infty=1$. Se for dado $\bar{v}\in\mathbb{R}^{q+1}$ com $\|\bar{v}\|=1$, pela equivalência de normas, existe $\beta\in(0,\sqrt{q+1})$ tal que $v=\beta\bar{v}$ satisfaz $\|v\|_\infty=1$. Então,

$$\max_{x\in B(0,1)}|\bar{v}^\top\bar{\phi}(x)|=\max_{x\in B(0,1)}\frac{1}{\beta}|v^\top\bar{\phi}(x)|\geq\frac{1}{\sqrt{q+1}}\max_{x\in B(0,1)}|v^\top\bar{\phi}(x)|\geq\frac{1}{4\sqrt{q+1}}. \tag{2.46}$$

Com isso podemos mostrar o seguinte resultado.

Lema 2.25. *Considere $R=\{x^0,x^1,\ldots,x^q\}$ um conjunto Λ-posicionado em $B(x^0,\delta)$ com*

relação a uma base ϕ de $\mathcal{P}_n^2$ e $\widehat{L}_q$ a matriz definida em (2.10)*. Então*

$$\|\widehat{L}_q^{-1}\| \leq 4\sqrt{(q+1)^3}\Lambda.$$

Demonstração. Considere $\widehat{R} = \left\{0, \frac{x^1-x^0}{\delta}, \ldots, \frac{x^q-x^0}{\delta}\right\}$ e $\bar{\phi}$ a base natural para $\mathcal{P}_n^2$. Pelo Corolário 2.18 e pelo Lema 2.19, $\widehat{R}$ é Λ-posicionado em $B(0,1)$. Considere $\widehat{M} = \widehat{M}(\bar{\phi}, \widehat{R})$ a matriz do respectivo sistema linear (2.40).

Note que, como

$$\widehat{L}_q = \begin{pmatrix} (\bar{\varphi}(x^1 - x^0))^\top \\ \vdots \\ (\bar{\varphi}(x^q - x^0))^\top \end{pmatrix} \begin{pmatrix} D_\delta^{-1} & 0 \\ 0 & D_{\delta^2}^{-1} \end{pmatrix},$$

em que $D_\delta = \delta I_{n\times n}$ e $D_{\delta^2} = \delta^2 I_{(q-n)\times(q-n)}$, então

$$\widehat{M} = \begin{pmatrix} 1 & 0 \\ e & \widehat{L}_q \end{pmatrix}$$

e, consequentemente,

$$\widehat{M}^{-1} = \begin{pmatrix} 1 & 0 \\ -\widehat{L}_q^{-1}e & \widehat{L}_q^{-1} \end{pmatrix}.$$

Utilizando a definição da norma de Frobenius de uma matriz e sua equivalência com a norma euclidiana, seque que

$$\begin{aligned} \|\widehat{L}_q^{-1}\| &\leq \|\widehat{L}_q^{-1}\|_F \leq \|\widehat{M}^{-1}\|_F \\ &\leq \sqrt{q+1}\|\widehat{M}^{-1}\|, \end{aligned} \tag{2.47}$$

uma vez que

$$\widehat{M}^{-\top}\widehat{M}^{-1} = \begin{pmatrix} 1 & -e^\top \widehat{L}_q^{-\top} \\ -\widehat{L}_q^{-1}e & \widehat{L}_q^{-\top}\widehat{L}_q^{-1} \end{pmatrix}.$$

Pelo Λ-posicionamento de $\widehat{R}$, se $\lambda(x) \in \mathbb{R}^{q+1}$ é solução do sistema linear associado para qualquer $x \in B(0,1)$, temos que

$$\Lambda \geq \|\lambda(x)\|_\infty \geq \frac{1}{\sqrt{q+1}}\|\lambda(x)\| = \frac{1}{\sqrt{q+1}}\|\widehat{M}^{-\top}\bar{\phi}(x)\|.$$

Aplicando o Lema 2.22 com $A = \widehat{M}^{-\top}$ e $r = \bar{\phi}(x)$, segue da desigualdade anterior com $x \in B(0,1)$ que maximiza $|\bar{v}^\top\bar{\phi}(x)|$ e de (2.46) que

$$\left(\sqrt{q+1}\right)\Lambda \geq \|\widehat{M}^{-\top}\bar{\phi}(x)\| \geq |\bar{v}^\top\bar{\phi}(x)|\|\widehat{M}^{-\top}\| \geq \frac{1}{4\sqrt{q+1}}\|\widehat{M}^{-\top}\|,$$

portanto

$$\|\widehat{M}^{-1}\| = \|\widehat{M}^{-\top}\| \leq 4(q+1)\Lambda.$$

Com isso e (2.47) temos que

$$\|\widehat{L}_q^{-1}\| \leq \sqrt{q+1}\|\widehat{M}^{-1}\| \leq 4\sqrt{(q+1)^3}\Lambda,$$

concluindo a demonstração. □

Considerando o controle de geometria, conseguimos limitações para $\|\widehat{L}_\ell^{-1}\|$ e $\|\widehat{L}_q^{-1}\|$ que não dependem do raio δ do conjunto de amostra. Consequentemente, junto com os resultados das Seções 2.2 e 2.3 obtemos que os limitantes para o erro entre o gradiente da função e o gradiente do modelo são pelo menos lineares em δ.

2.5 Limitantes para o erro entre modelos e função

Com os resultados das seções anteriores, podemos agora enunciar os resultados que mostram que tanto os modelos construídos por interpolação polinomial quanto os modelos construídos por regressão via vetores suporte aproximam a função e seu gradiente, condição necessária para convergência dos métodos de região de confiança, que será discutida no próximo capítulo.

Interpolação linear

Teorema 2.26. *Considere que o conjunto $R = \{x^0, x^1, \ldots, x^n\} \subset \mathcal{D} \cap B(x^0, \delta)$ seja Λ-posicionado em $B(x^0, \delta)$ com relação a uma base ϕ de $\mathcal{P}_n^1$ e suponha que a Hipótese A1 seja satisfeita. Então existem constantes positivas $\widehat{\kappa}_1$ e $\widehat{\kappa}_2$ tais que para todo $x \in \mathcal{D} \cap B(x^0, \delta)$ valem as seguintes desigualdades*

$$\|\nabla f(x) - \nabla m_I(x)\| \leq \widehat{\kappa}_1 \delta$$

e

$$|f(x) - m_I(x)| \leq \widehat{\kappa}_2 \delta^2.$$

Demonstração. A primeira desigualdade segue dos Lemas 2.6 e 2.23. A segunda desigualdade segue dos Lemas 2.7 e 2.23. □

Interpolação quadrática

Teorema 2.27. *Considere que o conjunto* $R = \{x^0, x^1, \ldots, x^q\} \subset \mathcal{D} \cap B(x^0, \delta)$ *seja* Λ*-posicionado em* $B(x^0, \delta)$ *com relação a uma base* ϕ *de* $\mathcal{P}_n^2$ *e suponha que a Hipótese A2 seja satisfeita. Então existem constantes positivas* $\widehat{\kappa}_3$, $\widehat{\kappa}_4$ *e* $\widehat{\kappa}_5$ *tais que para todo* $x \in \mathcal{D} \cap B(x^0, \delta)$*, valem as seguintes desigualdades*

$$\|\nabla^2 f(x) - \nabla^2 m_I\| \leq \widehat{\kappa}_3 \delta,$$

$$\|\nabla f(x) - \nabla m_I\| \leq \widehat{\kappa}_4 \delta^2$$

e

$$|f(x) - m_I(x)| \leq \widehat{\kappa}_5 \delta^3.$$

Demonstração. As duas primeiras desigualdades seguem dos Lemas 2.8 e 2.25. A terceira desigualdade segue dos Lemas 2.9 e 2.25. □

Regressão linear via vetores suporte

Teorema 2.28. *Considere que o conjunto* $R = \{x^0, x^1, \ldots, x^n\} \subset \mathcal{D} \cap B(x^0, \delta)$ *seja* Λ*-posicionado em* $B(x^0, \delta)$ *com relação a uma base* ϕ *de* $\mathcal{P}_n^1$ *e suponha que a Hipótese A1 seja satisfeita. Se o modelo linear* m_S *é construído via regressão por vetores suporte com margem* $\varepsilon \leq c_1\delta^2$ *e folgas* $\xi, \xi' \leq c_2\delta^2$*, com* $c_1, c_2 > 0$*, então existem constantes positivas* $\widehat{\kappa}_6$ *e* $\widehat{\kappa}_7$ *tais que para todo* $x \in \mathcal{D} \cap B(x^0, \delta)$ *valem as desigualdades*

$$\|\nabla f(x) - \nabla m_S(x)\| \leq \widehat{\kappa}_6 \delta$$

e

$$|f(x) - m_S(x)| \leq \widehat{\kappa}_7 \delta^2.$$

Demonstração. A primeira desigualdade segue do Teorema 2.10 e Lema 2.23. A segunda desigualdade segue do Teorema 2.11 e Lema 2.23. □

Regressão quadrática via vetores suporte

Teorema 2.29. *Considere que o conjunto* $R = \{x^0, x_1, \ldots, x^q\} \subset \mathcal{D} \cap B(x^0, \delta)$ *seja* Λ*-posicionado em* $B(x^0, \delta)$ *com relação a uma base* ϕ *de* $\mathcal{P}_n^2$ *e suponha que a Hipótese A2 seja satisfeita. Se o modelo quadrático* m_S *é construído via regressão por vetores suporte com margem* $\varepsilon \leq c_1\delta^3$ *e folgas* $\xi, \xi' \leq c_2\delta^3$*, com* $c_1, c_2 > 0$*, então existem constantes*

positivas $\widehat{\kappa}_8$, $\widehat{\kappa}_9$ *e* $\widehat{\kappa}_{10}$ *tais que para todo* $x \in \mathcal{D} \cap B(x^0, \delta)$ *valem as desigualdades*

$$\|\nabla^2 f(x) - \nabla^2 m_S(x)\| \leq \widehat{\kappa}_8 \delta,$$

$$\|\nabla f(x) - \nabla m_S(x)\| \leq \widehat{\kappa}_9 \delta^2$$

e

$$|f(x) - m_S(x)| \leq \widehat{\kappa}_{10} \delta^3.$$

Demonstração. As duas primeiras desigualdades seguem do Teorema 2.12 e Lema 2.25. A terceira desigualdade segue do Teorema 2.13 e Lema 2.25.

□

Teorema 2.30. *Considere que o conjunto* $R = \{x^0, \ldots, x^q\} \subset \mathcal{D} \cap B(x^0, \delta)$ *seja* Λ*-posicionado em* $B(x^0, \delta)$ *com relação a uma base* ϕ *de* $\mathcal{P}_n^2$ *e suponha que a Hipótese A1 seja satisfeita. Se o modelo quadrático* m_S *é construído via regressão por vetores suporte com margem* $\varepsilon \leq c_1 \delta^2$ *e folgas* $\xi, \xi' \leq c_2 \delta^2$*, com* $c_1, c_2 > 0$*, então existem constantes positivas* $\widehat{\kappa}_{11}$, $\widehat{\kappa}_{12}$ *e* $\widehat{\kappa}_{13}$ *tais que para todo* $x \in \mathcal{D} \cap B(x^0, \delta)$ *valem as desigualdades*

$$\|\nabla^2 m_S(x)\| \leq \widehat{\kappa}_{11},$$

$$\|\nabla f(x) - \nabla m_S(x)\| \leq \widehat{\kappa}_{12} \delta$$

e

$$|f(x) - m_S(x)| \leq \widehat{\kappa}_{13} \delta^2.$$

Demonstração. As duas primeiras desigualdades seguem do Teorema 2.14 e Lema 2.25. A terceira desigualdade segue do Teorema 2.15 e Lema 2.25. □

Note que as constantes $\widehat{\kappa}_i$, com $i = 1, \ldots, 13$ independem de δ mas dependem do Λ-posicionamento. Faz-se importante notar também que para a construção dos modelos via regressão por vetores suporte precisamos que o erro cometido no valor da função e no valor do modelo nos pontos da amostra seja controlado. O erro ε é escolhido a priori, já os valores de ξ e ξ' precisam ser controlados. Tal controle dependerá do parâmetro de regularização C. Quando o conjunto de amostra é Λ-posicionado, podemos garantir que existe um valor de C grande o suficiente de modo que $\|\xi\|_\infty$ e $\|\xi'\|_\infty$ sejam menores que um limitante pré-estabelecido, uma vez que os modelos construídos por regressão via vetores suporte se aproximariam do modelo construído por interpolação.

Ainda podemos ressaltar que com o controle do erro nos pontos da amostra e com os conjuntos Λ-posicionados os teoremas apresentados nesse capítulo podem ser estendidos

para outras técnicas, além das aqui discutidas de interpolação polinomial e regressão via vetores suporte.

Capítulo 3

Um método de região de confiança sem derivadas

O objetivo deste capítulo é aplicar as técnicas de construção de modelos, discutidas nos capítulos anteriores, na resolução de problemas de otimização sem derivadas por métodos de região de confiança.

Um método de região de confiança, extensamente discutido em [8], define a cada iteração um modelo da função objetivo e uma região em torno do ponto corrente na qual acreditamos que o modelo é confiável, dita região de confiança. Calculamos então um minimizador aproximado do modelo na região de confiança restrito ao conjunto viável. Caso este ponto forneça uma redução razoável no valor da função objetivo, aceitamos o iterando e repetimos o processo. Caso contrário, pode ser que o modelo não represente adequadamente a função. Neste caso, o ponto é recusado e reduzimos o tamanho da região para encontrar um novo minimizador.

Apresentamos um algoritmo de região de confiança sem derivadas, fortemente baseado no trabalho de Conejo et al. [7], para minimização de uma função objetivo em um conjunto convexo e fechado. Embora a função a ser minimizada seja continuamente diferenciável, o interesse reside no caso em que suas derivadas não estejam disponíveis ou que sejam custosas de serem calculadas.

O algoritmo é bastante geral no sentido de que os modelos podem ser obtidos por qualquer técnica desde que sejam satisfeitas algumas hipóteses razoáveis. Além disso, o novo iterando pode ser obtido por qualquer algoritmo interno que forneça um decréscimo suficiente no modelo. O algoritmo apresentado difere do discutido em [7] pela inclusão de um raio δ_k que controla a qualidade do modelo. Em [7], o raio Δ_k da região de confiança desempenha ambas as funções.

Quando as derivadas da função objetivo estão disponíveis, os modelos são baseados em aproximações da série de Taylor da função objetivo. Caso as derivadas não estejam disponíveis, é usual construir os modelos por interpolação polinomial, como em [2, 10, 17, 21, 35, 41]. No capítulo anterior vimos que os modelos podem ser também construídos por regressão via vetores suporte. A convergência global do algoritmo de região de confiança é garantida quando o modelo é obtido por quaisquer dessas técnicas, como veremos neste capítulo. Essa é uma das contribuições da tese.

3.1 O algoritmo

Considere o problema de programação não linear

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeita a} & x \in \Omega, \end{array} \tag{3.1}$$

com $f : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}$ uma função continuamente diferenciável e $\Omega \subset \mathcal{D}$ um conjunto não vazio, convexo e fechado. Estamos particularmente interessados no caso em que as derivadas da função objetivo não estão disponíveis ou estão disponíveis a um custo proibitivo. Além disso, consideramos que seja fácil calcular a projeção de um ponto sobre o conjunto viável Ω.

Como é usual em métodos de região de confiança, em cada iteração $k \in \mathbb{N}$ é considerado o iterando atual $x^k \in \Omega$ e o modelo

$$m_k(x) = b_k + g_k^\top (x - x^k) + \frac{1}{2}(x - x^k)^\top H_k (x - x^k), \tag{3.2}$$

onde $b_k \in \mathbb{R}$, $g_k = \nabla m_k(x^k) \in \mathbb{R}^n$ e $H_k \in \mathbb{R}^{n \times n}$ é uma matriz simétrica. Quando as derivadas da função objetivo estão disponíveis, o modelo (3.2) é baseado na aproximação de Taylor com $b_k = f(x^k)$, $g_k = \nabla f(x^k)$ e H_k uma aproximação da Hessiana ou uma matriz simétrica satisfazendo alguma hipótese de limitação.

Considere a medida de estacionariedade do problema de minimizar o modelo sobre o conjunto convexo e fechado Ω em x^k definida por

$$\pi_k = \|P_\Omega(x^k - g_k) - x^k\|,$$

onde P_Ω denota a projeção ortogonal sobre o conjunto Ω.

Note que o ponto $x^* \in \Omega$ é estacionário para o problema original (3.1) se, e somente

se,

$$\|P_\Omega(x^* - \nabla f(x^*)) - x^*\| = 0.$$

Dado $\Delta_k > 0$, assumimos que as soluções aproximadas, $d^k \in \mathbb{R}^n$, dos subproblemas de região de confiança

$$\begin{array}{ll} \text{minimizar} & m_k(x^k + d) \\ \text{sujeita a} & x^k + d \in \Omega \\ & \|d\| \leq \Delta_k, \end{array} \tag{3.3}$$

satisfazem a condição de decréscimo

$$m_k(x^k) - m_k(x^k + d^k) \geq \theta_1 \pi_k \min\left\{\frac{\pi_k}{1 + \|H_k\|}, \Delta_k, 1\right\}, \tag{3.4}$$

com $\theta_1 > 0$ uma constante independente de k. Supomos que conhecemos um algoritmo que resolve aproximadamente o subproblema (3.3), isto é, que seja capaz de a cada iteração encontrar d_k satisfazendo (3.4).

Condições do tipo (3.4) são bem conhecidas nas abordagens por regiões de confiança e utilizadas por vários autores em diferentes situações. No caso irrestrito, em que $\Omega = \mathbb{R}^n$, a medida de estacionariedade π_k é simplesmente $\|g_k\|$ e o passo clássico de Cauchy d_c^k satisfaz a condição

$$m_k(x^k) - m_k(x^k + d^k) \geq \ \theta_1 \|g_k\| \min\left\{\frac{\|g_k\|}{1 + \|H_k\|}, \Delta_k\right\},$$

como provado por Nocedal e Wright em [31] no caso com derivadas e por Conn, Scheinberg e Vicente em [10] para o caso sem derivadas da função objetivo. Condições do tipo (3.4) também aparecem ao longo do livro de Conn, Gould e Toint [8], em diferentes contextos. Em [20], Gonzaga, Karas e Vanti provam convergência global de um método de filtro para programação não linear, assumindo que as soluções aproximadas dos subproblemas satisfazem uma condição similar a (3.4). Para o caso de otimização não linear e sem derivadas, Tröltzsch [48] também assume esta condição quando trata o problema (3.1) com Ω sendo uma caixa.

Depois de calculada uma solução aproximada do subproblema, analisamos se ela fornece um decréscimo satisfatório na função objetivo. Como usual em métodos de região de confiança, definimos a redução **predita** produzida pelo passo d^k como $pred = m_k(x^k) - m_k(x^k + d^k)$ e a redução **verdadeira** como $ared = f(x^k) - f(x^k + d^k)$ e então, para $pred \neq 0$, calculamos a razão

$$\rho_k = \frac{ared}{pred}. \tag{3.5}$$

O passo d^k será aceito quando a razão ρ_k for maior que uma constante $\eta > 0$ dada.

Neste caso, definimos $x^{k+1} = x^k + d^k$ e repetimos o processo. Caso contrário, recusamos o passo d^k, reduzimos o raio Δ_k e resolvemos o subproblema (3.3) com o novo raio.

A seguir, temos o Algoritmo 3.1 de região de confiança sem derivadas baseado em [7].

Algoritmo 3.1. *Algoritmo de região de confiança sem derivadas*

Dados: $x^0 \in \Omega$, $\beta > 0$, $\delta_0 = \Delta_0 > 0$, $0 < \tau_1 < 1 \leq \tau_2$, $\eta_1 \in (0, 1)$, $0 \leq \eta < \eta_1 \leq \eta_2$.
Faça $k = 0$.
REPITA
  Obtenha o modelo m_k.
  SE $\delta_k > \beta\pi_k$, ENTÃO
    $\delta_{k+1} = \tau_1\delta_k$, *escolha* $\Delta_{k+1} \in [\delta_{k+1}, \Delta_k]$,
    $d^k = 0$ *e* $x^{k+1} = x^k$.
  SENÃO
    Determine uma solução d^k *de* (3.3) *que satisfaça* (3.4).
    SE $\rho_k \geq \eta$, ENTÃO
      $x^{k+1} = x^k + d^k$.
    SENÃO
      $x^{k+1} = x^k$.
    SE $\rho_k < \eta_1$, ENTÃO
      $\delta_{k+1} = \tau_1\delta_k$ *e* $\Delta_{k+1} = \tau_1\Delta_k$.
    SENÃO
      SE $\rho_k > \eta_2$ *e* $\|d^k\| = \Delta_k$, ENTÃO
        $\delta_{k+1} = \tau_2\delta_k$ *e* $\Delta_{k+1} = \tau_2\Delta_k$.
      SENÃO
        $\delta_{k+1} = \delta_k$ *e* $\Delta_{k+1} = \Delta_k$.
  $k = k + 1$.

Quando π_k é pequeno, o iterando está provavelmente perto de uma solução do problema de minimizar o modelo dentro do conjunto viável Ω. Por outro lado, se δ_k é grande, não podemos garantir que o modelo representa adequadamente a função objetivo, como veremos neste capítulo. Então, quando $\delta_k > \beta\pi_k$, o raio δ_k é reduzido, objetivando encontrar modelos mais precisos. Embora possamos tomar $\beta = 1$, este parâmetro deve ser utilizado para balancear a magnitude de π_k e δ_k de acordo com o problema.

Na Seção 3.2 mostraremos que $\delta_k \to 0$ quando $k \to \infty$, o que será fundamental nas provas de convergência. Isto sugere também que, dada uma tolerância $\varepsilon > 0$ e parâmetros $\beta_1, \beta_2 > 0$, a combinação de $\delta_k \leq \beta_1\varepsilon$ e $\pi_k \leq \beta_2\varepsilon$ pode ser utilizada como critério de

parada na implementação do algoritmo.

Pelo Algoritmo 3.1, a razão dada em (3.5) está bem definida, pois, na iteração k em que o algoritmo a calcula, vale $0 < \delta_k \leq \beta\pi_k$. Logo, $\pi_k \neq 0$ e pela condição de decréscimo (3.4), temos que

$$m_k(x^k) - m_k(x^k + d^k) \neq 0.$$

Outra consideração referente ao Algoritmo 3.1 é a utilização de um raio de região de confiança Δ_k e um outro raio δ_k que controla a qualidade do modelo, diferentemente de [7], onde o raio Δ_k desempenha ambas as funções. A inspiração para tal modificação foi o fato que, se do ponto de vista teórico é necessário que o termo que controla a qualidade do modelo convirja a zero, do ponto de vista prático é desejável que o raio de região de confiança seja o maior possível a cada iteração. A inclusão do raio δ_k é uma das contribuições do presente trabalho.

A seguir, discutimos as hipóteses necessárias para a prova de convergência do Algoritmo 3.1.

Hipóteses sobre o problema

Consideraremos a Hipótese A1, já apresentada no capítulo anterior, e reproduzida abaixo.

H1. *A função f é continuamente diferenciável em $\mathcal{D}$ e ∇f é Lipschitz com constante $L_g > 0$ em $\mathcal{D}$.*

Também vamos considerar que a função objetivo é limitada inferiormente no conjunto viável.

H2. *A função f é limitada inferiormente no conjunto Ω.*

Hipótese sobre os modelos

Sobre o modelo, vamos considerar apenas a hipótese a seguir.

H3. *Existe uma constante $\kappa_m > 0$ tal que, para todo $k \in \mathbb{N}$,*

$$\|\nabla f(x) - \nabla m_k(x)\| \leq \kappa_m \delta_k$$

para todo $x \in \mathcal{D} \cap B(x^k, \delta_k)$.

As hipóteses sobre o problema são comuns em análise de convergência para algo-

ritmos de região de confiança com e sem derivadas.

A Hipótese H3 independe da técnica utilizada para a construção dos modelos em cada iteração. A exibição de técnicas de construção de modelos que satisfazem essa hipótese, feita no capítulo anterior, é uma contribuição da tese. O próximo teorema sintetiza condições para que a hipótese seja satisfeita. Tais condições estão relacionadas ao controle da geometria do conjunto de amostra R_k, usado em cada iteração do algoritmo de região de confiança, e do erro cometido em valores funcionais entre o modelo e a função, tanto para modelos lineares quanto para modelos quadráticos.

Teorema 3.1. *Suponha válida a Hipótese H1. Considere ϕ uma base para $\mathcal{P}_n^a$, com $a = 1$ ou $a = 2$, $R_k \subset \mathcal{D} \cap B(x^k, \delta_k)$ um conjunto Λ-posicionado em $B(x^k, \delta_k)$ em relação à base ϕ. Se existe $\kappa_f \geq 0$, independente de k, tal que*

$$|f(y) - m_k(y)| \leq \kappa_f \delta_k^2$$

para todo $y \in R_k$, então existe uma constante κ_m tal que

$$\|\nabla f(x) - \nabla m_k(x)\| \leq \kappa_m \delta_k$$

para todo $x \in \mathcal{D} \cap B(x^k, \delta_k)$.

A demonstração do Teorema 3.1 segue dos resultados apresentados no capítulo anterior, levando em consideração a generalidade requerida, nos moldes do Teorema 2.28 para o caso linear e do Teorema 2.30 para o caso quadrático.

Em [10] há ainda o estudo sobre o caso de modelos por interpolação subdeterminados, ou seja, com uma quantidade inferior de pontos que garantiriam um conjunto posicionado para interpolação polinomial quadrática. Tais modelos também satisfazem a Hipótese H3.

No trabalho de Conejo et al. [7], os autores consideram duas hipóteses para os modelos, uma delas é a exigência de uma limitação para as normas das Hessianas dos modelos, além de uma versão enfraquecida da Hipótese H3, que exige apenas a limitação no ponto corrente.

O próximo resultado garante que se considerarmos as Hipóteses H1 e H3, garantimos que a Hessiana dos modelos são limitadas.

Lema 3.2. *Suponha que as Hipóteses H1 e H3 são satisfeitas. Então existe uma constante $\kappa_h > 0$ tal que*

$$\|H_k\| \leq \kappa_h,$$

para todo $k \in \mathbb{N}$.

Demonstração. Considere $d \in \mathbb{R}^n$ arbitrário com $\|d\| = \delta_k$. Pela definição do modelo (3.2), utilizando a desigualdade triangular e as hipóteses temos que

$$\begin{aligned} \|H_k d\| &= \|\nabla m_k(x^k + d) - \nabla m_k(x^k)\| \\ &\leq \|\nabla m_k(x^k + d) - \nabla f(x^k + d)\| + \|\nabla f(x^k + d) - \nabla f(x^k)\| + \|\nabla f(x^k) - \nabla m_k(x^k)\| \\ &\leq 2\kappa_m \delta_k + L_g \|d\| \\ &= \kappa_h \delta_k, \end{aligned}$$

em que $\kappa_h = 2\kappa_m + L_g$. Consequentemente,

$$\|H_k\| = \max_{\|d\|=\delta_k} \left\| H_k \frac{d}{\|d\|} \right\| = \frac{1}{\delta_k} \max_{\|d\|=\delta_k} \|H_k d\| \leq \kappa_h,$$

o que completa a demonstração. □

Mais do que isso, conseguimos mostrar uma equivalência entre as hipóteses consideradas sobre os modelos em [7] com a Hipótese H3.

Lema 3.3. *Suponha que a Hipótese H1 é satisfeita. Se existem constantes* κ_0 *e* κ_α *tais que para todo* $k \in \mathbb{N}$, $\|\nabla f(x^k) - \nabla m_k(x^k)\| \leq \kappa_0 \delta_k$ *e* $\|H_k\| \leq \kappa_\alpha$, *então para todo* $k \in \mathbb{N}$ *e para todo* $x \in \mathcal{D} \cap B(x^k, \delta_k)$ *vale a desigualdade*

$$\|\nabla f(x) - \nabla m_k(x)\| \leq \kappa_m \delta_k,$$

com $\kappa_m = L_g + \kappa_0 + \kappa_\alpha$.

Demonstração. Pela desigualdade triangular temos que

$$\|\nabla f(x) - \nabla m_k(x)\| \leq \|\nabla f(x) - \nabla f(x^k)\| + \|\nabla f(x^k) - \nabla m_k(x^k)\| + \|\nabla m_k(x^k) - \nabla m_k(x)\|. \quad (3.6)$$

Considerando a Hipótese H1 temos que

$$\|\nabla f(x) - \nabla f(x^k)\| \leq L_g \|x - x^k\|.$$

Por hipótese, temos também que

$$\|\nabla f(x^k) - \nabla m_k(x^k)\| \leq \kappa_0 \delta_k.$$

Pela definição do modelo, dada em (3.2), temos que

$$\|\nabla m_k(x^k) - \nabla m_k(x)\| = \|H_k(x - x^k)\| \leq \|H_k\| \|x - x^k\| \leq \kappa_\alpha \|x - x^k\|.$$

Considerando as três desigualdades acima em (3.6) temos que para todo $k \in \mathbb{N}$ e para todo $x \in \mathcal{D} \cap B(x^k, \delta_k)$

$$\|\nabla f(x) - \nabla m_k(x)\| \leq (L_g + \kappa_0 + \kappa_\alpha)\delta_k,$$

o que conclui a demonstração. □

3.2 Análise de convergência

Esta seção é dedicada à prova de convergência global do Algoritmo 3.1, baseada em [7]. Tanto em [7] como neste trabalho, são consideradas as mesmas hipóteses sobre o problema, ou seja, as Hipóteses H1 e H2. Com respeito aos modelos, considera-se em [7] que existe uma constante $\theta_2 > 0$ tal que, para todo $k \in \mathbb{N}$,

$$\|\nabla f(x^k) - g_k\| \leq \theta_2 \Delta_k,$$

e que a sequência das Hessianas do modelo é limitada. No entanto, estas duas condições são consequências da Hipótese H3. Usando os fatos de que H3 vale em particular em $x = x^k$ e que $\delta_k \leq \Delta_k$ para todo k, a limitação das Hessianas segue do Lema 3.2. O Lema 3.3 estabelece o resultado recíproco, ou seja, considerando as hipóteses apresentadas em [7] conseguimos um resultado semelhante à Hipótese H3, considerando a diferença nos raios.

Assuma, então, que valem as Hipóteses H1, H2 e H3 e que o Algoritmo 3.1 gera uma sequência infinita $\{x^k\} \subset \Omega$. Baseados em [7], provaremos que todo ponto de acumulação da sequência $\{x^k\}$ é estacionário. A prova de convergência independe da técnica utilizada para construção dos modelos da função objetivo, desde que a Hipótese H3 seja satisfeita. Quando as derivadas da função objetivo estão indisponíveis, normalmente os modelos são obtidos por interpolação polinomial [10, 17, 21, 35, 41]. Nossa contribuição nesse trabalho é usar máquinas de vetores suporte na construção dos modelos como alternativa à interpolação polinomial, como discutido no capítulo anterior.

Considere os seguintes conjuntos de índices

$$\mathcal{S} = \{k \in \mathbb{N} \mid \rho_k \geq \eta\} \qquad \text{e} \qquad \overline{\mathcal{S}} = \{k \in \mathbb{N} \mid \rho_k \geq \eta_1\}.$$

O conjunto $\mathcal{S}$ é o conjunto de iterações de *sucesso* e $\overline{\mathcal{S}} \subset \mathcal{S}$.

O lema a seguir garante que se o raio da região de confiança é suficientemente pequeno, então o algoritmo deve executar uma iteração de sucesso. As constantes θ_1,

L_g, κ_h e κ_m são definidas em (3.4), na Hipótese H1, no Lema 3.2 e na Hipótese H3, respectivamente e $\kappa_H = 1 + \kappa_h > 1$.

Lema 3.4. *Suponha válidas as Hipóteses H1 e H3. Considere o conjunto*

$$\mathcal{K} = \left\{ k \in \mathbb{N} \mid \Delta_k \leq \min\left\{ \frac{\pi_k}{\kappa_H}, \frac{(1-\eta_1)\pi_k}{c}, \beta\pi_k, 1 \right\} \right\}, \tag{3.7}$$

em que $c = \dfrac{L_g + \kappa_m + \dfrac{\kappa_H}{2}}{\theta_1}$. *Se* $k \in \mathcal{K}$, *então* $k \in \overline{\mathcal{S}}$.

Demonstração. Considere $k \in \mathcal{K}$ arbitrário. Pelo Teorema do Valor Médio, existe $t_k \in (0,1)$ tal que

$$f(x^k + d^k) = f(x^k) + \nabla f(x^k + t_k d^k)^\top d^k. \tag{3.8}$$

Assim, pela definição de m_k em (3.2) e a igualdade (3.8),

$$\begin{aligned}
|ared - pred| &= \left| f(x^k) - f(x^k + d^k) - m_k(x^k) + m_k(x^k + d^k) \right| \\
&= \left| f(x^k) - f(x^k) - \nabla f(x^k + t_k d^k)^\top d^k + g_k^\top d^k + \frac{1}{2}(d^k)^\top H_k d^k \right| \\
&= \left| -\left(\nabla f(x^k + t_k d^k) - g_k \right)^\top d^k + \frac{1}{2}(d^k)^\top H_k d^k \right|. \\
&= \left| -\left(\nabla f(x^k + t_k d^k) - g_k - \nabla f(x^k) + \nabla f(x^k) \right)^\top d^k + \frac{1}{2}(d^k)^\top H_k d^k \right|.
\end{aligned}$$

Utilizando as desigualdades triangular e de Cauchy-Schwarz, obtemos

$$|ared - pred| \leq \left(\|\nabla f(x^k + t_k d^k) - \nabla f(x^k)\| + \|\nabla f(x^k) - g_k\| \right) \|d^k\| + \frac{1}{2}\|d^k\|^2 \|H_k\|.$$

Pelas Hipóteses H1 e H3 e pelo Lema 3.2, temos

$$|ared - pred| \leq t_k L_g \|d^k\|^2 + \kappa_m \delta_k \|d^k\| + \frac{1}{2}\kappa_h \|d^k\|^2.$$

Visto que $\|d^k\| \leq \Delta_k$, $\delta_k \leq \Delta_k$, $t_k \in (0,1)$ e $\kappa_h = \kappa_H - 1$,

$$|ared - pred| \leq \theta_0 \Delta_k^2, \tag{3.9}$$

em que $\theta_0 = L_g + \kappa_m + \dfrac{1}{2}\kappa_H$.

Pela definição do conjunto $\mathcal{K}$, temos $\Delta_k \leq \beta\pi_k$ e consequentemente $\pi_k > 0$. Pelo Lema 3.2 e de (3.4) segue que

$$pred = m_k(x^k) - m_k(x^k + d^k) \geq \theta_1 \pi_k \min\left\{ \frac{\pi_k}{\kappa_H}, \Delta_k, 1 \right\}. \tag{3.10}$$

Então, segue que $pred \neq 0$. Portanto pela definição de ρ_k, (3.9) e (3.10)

$$\begin{aligned} |\rho_k - 1| &= \left| \frac{ared - pred}{pred} \right| \\ &\leq \frac{\theta_0 \Delta_k^2}{\theta_1 \pi_k \min\left\{ \frac{\pi_k}{\kappa_H}, \Delta_k, 1 \right\}} \\ &= \frac{c\Delta_k^2}{\pi_k \min\left\{ \frac{\pi_k}{\kappa_H}, \Delta_k, 1 \right\}}, \end{aligned}$$

com $c = \dfrac{\theta_0}{\theta_1}$.

Pela definição do conjunto $\mathcal{K}$ em (3.7),

$$\Delta_k \leq \min\left\{ \frac{\pi_k}{\kappa_H}, \Delta_k, 1 \right\} \quad \text{e} \quad \frac{c\Delta_k}{\pi_k} \leq 1 - \eta_1.$$

Logo,

$$|\rho_k - 1| \leq \frac{c\Delta_k^2}{\pi_k \Delta_k} \leq 1 - \eta_1$$

e consequentemente $\rho_k \geq \eta_1$. Portanto $k \in \overline{\mathcal{S}}$, completando a demonstração. □

Com a Hipótese H3 e Lema 3.2 podemos notar que, quanto menor δ_k, melhor o modelo representa localmente a função objetivo. Logo, é razoável que o raio do conjunto de amostra convirja para zero. No lema a seguir mostramos que o algoritmo proposto tem esta propriedade.

Lema 3.5. *Suponha válidas as Hipóteses H1, H2 e H3. Então a sequência $\{\delta_k\}$ converge para zero.*

Demonstração. Se $\overline{\mathcal{S}}$ é finito, então pelo mecanismo de atualização do raio do Algoritmo 3.1, existe $k_0 \in \mathbb{N}$ tal que para todo $k \geq k_0$, $\delta_{k+1} = \tau_1 \delta_k$. Logo, a sequência $\{\delta_k\}$ converge para zero. Por outro lado, se $\overline{\mathcal{S}}$ é infinito, para qualquer $k \in \overline{\mathcal{S}}$, utilizando a definição de ρ_k, a condição (3.4) de decréscimo necessária do modelo e o Lema 3.2, temos

$$f(x^k) - f(x^{k+1}) \geq \eta_1 \left(m_k(x^k) - m_k(x^k + d^k) \right) \geq \eta_1 \theta_1 \pi_k \min\left\{ \frac{\pi_k}{\kappa_H}, \Delta_k, 1 \right\}.$$

Como $k \in \overline{\mathcal{S}}$, calculamos ρ_k, ou seja $\delta_k \leq \beta \pi_k$. Por outro lado, pelo mecanismo do algoritmo $\delta_k \leq \Delta_k$. Assim

$$f(x^k) - f(x^{k+1}) \geq \eta \theta_1 \frac{\delta_k}{\beta} \min\left\{ \frac{\delta_k}{\beta \kappa_H}, \Delta_k, 1 \right\} \geq \eta \theta_1 \frac{\delta_k}{\beta} \min\left\{ \frac{\delta_k}{\beta \kappa_H}, \delta_k, 1 \right\}.$$

Uma vez que $\{f(x^k)\}$ é não crescente e, pela Hipótese H2, limitada inferiormente, o lado

esquerdo da expressão anterior converge para zero. Então,

$$\lim_{k \in \overline{\mathcal{S}}} \delta_k = 0. \tag{3.11}$$

Considere o conjunto

$$\mathcal{U} = \left\{ k \in \mathbb{N} \mid k \notin \overline{\mathcal{S}} \right\}.$$

Se $\mathcal{U}$ é finito, então por (3.11) temos que $\lim_{k \to \infty} \delta_k = 0$.

Agora suponha que $\mathcal{U}$ é infinito. Para $k \in \mathcal{U}$, defina ℓ_k o índice da última iteração em $\overline{\mathcal{S}}$ anterior a k. Pelo mecanismo do Algoritmo 3.1, $\delta_k \leq \tau_2 \delta_{\ell_k}$, o que implica

$$\lim_{k \in \mathcal{U}} \delta_k \leq \tau_2 \lim_{k \in \mathcal{U}} \delta_{\ell_k} = \tau_2 \lim_{\ell_k \in \overline{\mathcal{S}}} \delta_{\ell_k}.$$

Por (3.11), segue que $\lim_{k \in \mathcal{U}} \delta_k = 0$, o que completa a prova. □

O próximo lema garante que a sequência $\{\pi_k\}$ tem uma subsequência que converge para zero.

Lema 3.6. *Suponha que as Hipóteses H1, H2 e H3 sejam válidas. Então* $\liminf_{k \to \infty} \pi_k = 0$.

Demonstração. Suponha por contradição que existem $\varepsilon > 0$ e um inteiro $K > 0$ tais que $\pi_k \geq \varepsilon$ para todo $k \geq K$. Defina

$$\tilde{\Delta} = \min \left\{ \frac{\varepsilon}{\kappa_H}, \frac{(1 - \eta_1)\varepsilon}{c}, \beta\varepsilon, 1 \right\},$$

em que $\kappa_H = \kappa_h + 1$, com κ_h a constante do Lema 3.2, c definida no Lema 3.4, η_1 e $\beta > 0$ parâmetros dados no Algoritmo 3.1.

Considere $k \geq K$. Se $\Delta_k \leq \tilde{\Delta}$, então $k \in \mathcal{K}$, com $\mathcal{K}$ dado na expressão (3.7). Pelo Lema 3.4, $k \in \overline{\mathcal{S}}$ e com isso $\Delta_{k+1} \geq \Delta_k$. Disto segue que o raio somente pode decrescer se $\Delta_k > \tilde{\Delta}$, e neste caso, ou $\delta_k > \beta\pi_k$, e portanto

$$\Delta_{k+1} \geq \delta_{k+1} = \tau_1 \delta_k > \tau_1 \beta \pi_k \geq \tau_1 \beta \varepsilon \geq \tau_1 \tilde{\Delta},$$

ou $\delta_k \leq \beta\pi_k$ e pelo mecanismo do algoritmo

$$\Delta_{k+1} = \tau_1 \Delta_k > \tau_1 \tilde{\Delta}.$$

Em ambas as situações, para todo $k \geq K$,

$$\Delta_k \geq \min \left\{ \tau_1 \tilde{\Delta}, \Delta_K \right\}. \tag{3.12}$$

Considere $k \geq K$ fixo e suponha que $k \in \overline{\mathcal{S}}$. Utilizando a definição de ρ_k dada em (3.5), a condição (3.4), a hipótese de contradição e (3.12) temos que

$$\begin{aligned} f(x^k) - f(x^{k+1}) &\geq \eta_1 \left(m_k(x^k) - m_k(x^k + d^k) \right) \\ &\geq \eta_1 \theta_1 \pi_k \min \left\{ \frac{\pi_k}{\kappa_H}, \Delta_k, 1 \right\} \\ &\geq \eta_1 \theta_1 \varepsilon \min \left\{ \frac{\varepsilon}{\kappa_H}, \Delta_k, 1 \right\} \\ &\geq \eta_1 \theta_1 \varepsilon \min \left\{ \frac{\varepsilon}{\kappa_H}, \min \left\{ \tau_1 \tilde{\Delta}, \Delta_K \right\}, 1 \right\}. \end{aligned}$$

Pela Hipótese H2, a sequência $\{f(x^k)\}$ é limitada inferiormente, e como é monótona não crescente, $f(x^k) - f(x^{k+1}) \to 0$. Como o lado direito da desigualdade acima é constante, o conjunto $\{k \geq K \mid k \in \overline{\mathcal{S}} \,|\}$ é finito. Logo, pelo algoritmo, para todo k suficientemente grande $\delta_k > \beta \pi_k$ ou $\rho_k < \eta_1$. Porém, pelo Lema 3.5, $\delta_k \to 0$ e como $\pi_k \geq \varepsilon$ para todo $k \geq K$ temos que $\rho_k < \eta_1$ para todo k suficientemente grande, o que implica, pelo algoritmo, que $\Delta_{k+1} = \tau_1 \Delta_k$. Consequentemente $\Delta_k \to 0$, contradizendo (3.12). □

Ao assumirmos um decréscimo suficiente na função objetivo definindo $\eta > 0$ no Algoritmo 3.1, podemos provar que não somente existe uma subsequência de $\{\pi_k\}$ convergindo para zero como estabelecido no Lema 3.6, mas que a convergência é em toda a sequência.

Lema 3.7. *Suponha válidas as Hipóteses H1, H2 e H3 e que $\eta > 0$. Então*

$$\lim_{k \to \infty} \pi_k = 0.$$

Demonstração. Suponha por contradição que para algum $\varepsilon > 0$ o conjunto

$$\mathbb{N}' = \{k \in \mathbb{N} \mid \pi_k \geq \varepsilon\} \tag{3.13}$$

é infinito.

Dado $k \in \mathbb{N}'$, considere ℓ_k o primeiro índice tal que $\ell_k > k$ e $\pi_{\ell_k} \leq \varepsilon/2$. A existência de ℓ_k é assegurada pelo Lema 3.6. Assim,

$$\pi_k - \pi_{\ell_k} \geq \frac{\varepsilon}{2}.$$

Utilizando a definição de π_k, a desigualdade triangular e a propriedade de contração das

projeções, temos

$$\begin{aligned}\frac{\varepsilon}{2} &\leq \|P_\Omega(x^k - g_k) - x^k\| - \|P_\Omega(x^{\ell_k} - g_{\ell_k}) - x^{\ell_k}\| \\ &\leq \|P_\Omega(x^k - g_k) - x^k - P_\Omega(x^{\ell_k} - g_{\ell_k}) + x^{\ell_k}\| \\ &\leq 2\|x^k - x^{\ell_k}\| + \|g_k - g_{\ell_k}\|. \end{aligned} \tag{3.14}$$

Por outro lado, como $\delta_k \to 0$ pelo Lema 3.5, existe $k_0 \in \mathbb{N}$ tal que para $k \geq k_0$

$$\delta_k < \frac{\varepsilon}{8\kappa_m}. \tag{3.15}$$

Se $x^k = x^{\ell_k}$, ou seja, se $x^i \notin \mathcal{S}$ para $k \leq i < \ell_k$, temos que

$$\begin{aligned}\frac{\varepsilon}{2} &\leq 2\|x^k - x^{\ell_k}\| + \|g_k - g_{\ell_k}\| \\ &= \|g_k - g_{\ell_k}\| \\ &\leq \|g_k - \nabla f(x^k)\| + \|\nabla f(x^{\ell_k}) - g_{\ell_k}\| \\ &\leq \kappa_m(\delta_k + \delta_{\ell_k}),\end{aligned}$$

onde a última desigualdade segue da Hipótese H3. Disso e (3.15) segue que,

$$\kappa_m(\delta_k + \delta_{\ell_k}) \geq \frac{\varepsilon}{2} \qquad \text{e} \qquad \kappa_m(\delta_k + \delta_{\ell_k}) \leq \frac{\varepsilon}{4},$$

o que é impossível. Portanto, para $k \geq k_0$ e $k \in \mathbb{N}'$ o conjunto

$$C_k = \{j \in \mathcal{S} | k \leq j < \ell_k\}$$

é não vazio.

Agora, somando e subtraindo $\nabla f(x^k)$ e $\nabla f(x^{\ell_k})$ em (3.14) e pela desigualdade triangular, para $k \in \mathbb{N}'$ temos que

$$\begin{aligned}\frac{\varepsilon}{2} &\leq 2\|x^k - x^{\ell_k}\| + \|g_k - \nabla f(x^k) + \nabla f(x^k) - \nabla f(x^{\ell_k}) + \nabla f(x^{\ell_k}) - g_{\ell_k}\| \\ &\leq 2\|x^k - x^{\ell_k}\| + \|g_k - \nabla f(x^k)\| + \|\nabla f(x^k) - \nabla f(x^{\ell_k})\| + \|\nabla f(x^{\ell_k}) - g_{\ell_k}\|.\end{aligned}$$

Utilizando as Hipóteses H1 e H3, tem-se

$$\frac{\varepsilon}{2} \leq (2 + L_g)\|x^k - x^{\ell_k}\| + \kappa_m(\delta_k + \delta_{\ell_k}). \tag{3.16}$$

Usando (3.15), temos que para $k \geq k_0$, $k \in \mathbb{N}'$

$$\frac{\varepsilon}{2} \leq (2 + L_g)\|x^k - x^{\ell_k}\| + \frac{\varepsilon}{4},$$

de onde

$$\|x^k - x^{\ell_k}\| \geq \frac{\varepsilon}{4(2+L_g)}.$$

Temos então que

$$\frac{\varepsilon}{4(2+L_g)} \leq \|x^k - x^{\ell_k}\| \leq \sum_{j \in C_k} \|x^j - x^{j+1}\| \leq \sum_{j \in C_k} \Delta_j. \tag{3.17}$$

Por outro lado, como $C_k \neq \emptyset$ para $k \geq k_0$ e $k \in \mathbb{N}'$ temos pela definição de ρ_k dada em (3.5) e por (3.4) que

$$\begin{aligned} f(x^k) - f(x^{\ell_k}) &= \sum_{j \in C_k} \left(f(x^j) - f(x^{j+1}) \right) \\ &> \sum_{j \in C_k} \eta \left(m_j(x^j) - m_j(x^j + d^j) \right) \\ &\geq \sum_{j \in C_k} \eta \theta_1 \pi_j \min \left\{ \frac{\pi_j}{\kappa_H}, \Delta_j, 1 \right\}. \end{aligned}$$

Pela definição de ℓ_k, temos que $\pi_j > \varepsilon/2$ para todo $j \in C_k$. Dessa forma,

$$f(x^k) - f(x^{\ell_k}) \geq \sum_{j \in C_k} \eta \theta_1 \frac{\varepsilon}{2} \min \left\{ \frac{\varepsilon}{2\kappa_H}, \Delta_j, 1 \right\} \geq \eta \theta_1 \frac{\varepsilon}{2} \min \left\{ \frac{\varepsilon}{2\kappa_H}, \sum_{j \in C_k} \Delta_j, 1 \right\}.$$

Tendo em vista isso, (3.17) e a hipótese de que $\eta > 0$ temos que para $k \geq k_0$ e $k \in \mathbb{N}'$ $f(x^k) - f(x^{\ell_k})$ é uniformemente limitada por uma constante positiva. Por ouro lado, pela Hipótese H2, a sequência $\{f(x^k)\}$ é limitada inferiormente, e pelo algoritmo é monótona não crescente. Consequentemente $f(x^k) - f(x^{\ell_k}) \to 0$, o que é uma contradição, completando a prova. □

Podemos provar agora a convergência global a pontos estacionários de primeira ordem. No teorema a seguir, estabelecemos a relação entre a medida de estacionariedade do problema original e a medida de estacionariedade dada no Lema 3.7, obtendo o resultado de convergência global.

Teorema 3.8. *Suponha que valham as Hipóteses H1, H2 e H3. Então*

(i) Se $\eta = 0$, $\liminf_{k \to \infty} \|P_\Omega(x^k - \nabla f(x^k)) - x^k\| = 0$.

(ii) Se $\eta > 0$, $\lim_{k \to \infty} \|P_\Omega(x^k - \nabla f(x^k)) - x^k\| = 0$.

Demonstração. Pela desigualdade triangular, a propriedade de contração das projeções e

a Hipótese H3, temos que

$$\begin{aligned}
\|P_\Omega(x^k - \nabla f(x^k)) - x^k\| &= \|P_\Omega(x^k - \nabla f(x^k)) - P_\Omega(x^k - g_k) + P_\Omega(x^k - g_k) - x^k\| \\
&\leq \|P_\Omega(x^k - \nabla f(x^k)) - P_\Omega(x^k - g_k)\| + \|P_\Omega(x^k - g_k) - x^k\| \\
&\leq \|\nabla f(x^k) - g_k\| + \|P_\Omega(x^k - g_k) - x^k\| \\
&\leq \kappa_m \delta_k + \pi_k.
\end{aligned}$$

Utilizando os Lemas 3.5, 3.6 e 3.7, completamos a prova. □

Do Teorema 3.8 concluímos que se $\eta > 0$ e o Algoritmo 3.1 gera uma sequência $\{x^k\}$ com algum ponto de acumulação x^*, então o ponto x^* é estacionário de primeira ordem [8, 38]. Uma maneira de garantir a existência de um ponto de acumulação é supor que o conjunto de nível $\{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\}$ é limitado. Note que assim toda a sequência $\{x^k\}$ é limitada e consequentemente possui uma subsequência convergente.

Capítulo 4

Experimentos numéricos

Este capítulo é dedicado a experimentos numéricos a fim de discutir o desempenho do algoritmo de construção de modelos de uma função por regressão via vetores suporte e do algoritmo de região de confiança sem derivadas para minimizar uma função com os modelos construídos por regressão via vetores suporte.

O capítulo é dividido em três seções. Na primeira seção são apresentados modelos de funções definidas em $\mathbb{R}^2$, com o intuito de visualizar quão bem o modelo de regressão via vetores suporte aproxima uma função. A segunda seção é dedicada a comparar o gradiente do modelo por interpolação polinomial e o gradiente do modelo via regressão por vetores suporte. Para os testes utilizamos a coleção de problemas irrestritos organizada por Moré, Garbow e Hillstrom [28]. Por fim, na última seção do capítulo resolvemos os problemas da segunda parte com o algoritmo de região de confiança apresentado no Capítulo 3.

Os algoritmos foram implementados em Matlab® em sua versão R2012a. Os testes foram realizados em um computador portátil com processador Intel® Core™ i5-430M com 3 MB de memória cache, com velocidade do clock de 2.26 GHz e com 4 GB memória RAM, com sistema operacional Windows® 8.1 Pro com arquitetura 64-bits.

4.1 Modelos de regressão via vetores suporte

Nesta seção, apresentamos aproximações de funções por regressão via vetores suporte com o objetivo de visualizar graficamente tais aproximações. Para isso escolhemos as duas primeiras funções da coleção [28], que são funções definidas em $\mathbb{R}^2$.

Apresentamos aproximações com diferentes valores para o raio δ do conjunto dos

pontos da amostra. A tolerância escolhida foi $\varepsilon = \delta \times 10^{-3}$ e o parâmetro de regularização, que balanceia as folgas ξ e ξ', foi fixado para todos os casos em $C = 10^{10}$, escolhido empiricamente. Nas figuras desta seção, no lado esquerdo estão plotados os pontos do conjunto de amostra. Do lado direito temos os valores da função nos pontos do conjunto de amostra, o gráfico da função em azul e do modelo em cinza.

Para escolha dos pontos de amostra, a partir do ponto $x^0 = (1, 1)$ foram tomadas as direções coordenadas e as direções opostas a elas e escolhidos os pontos nessas direções que atingiam a fronteira da bola $B(x^0, \delta)$ respeitando o raio δ em cada caso. Logo, o conjunto de amostra é $\{(1, 1), (1+\delta, 1), (1, 1+\delta), (1-\delta, 1), (1, 1-\delta)\}$. Tal escolha permite verificar que mesmo com menos do que $(n+1)(n+2)/2$ pontos podemos obter modelos razoáveis.

Função de Rosenbrock

A primeira função para a qual construímos o modelo por máquinas de vetores suporte é a função de Rosenbrock $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2.$$

A função de Rosenbrock é a função número 1 da coleção [28].

A Figura 4.1 mostra a aproximação da função de Rosenbrock na vizinhança do ponto ótimo $(1, 1)$ com raio $\delta = 0.5$. O maior erro em termos de valor funcional nos pontos de amostra foi observado no ponto $y = (1, 1)$, para o qual $|f(y) - m(y)| \approx 5.000003 \times 10^{-4}$.

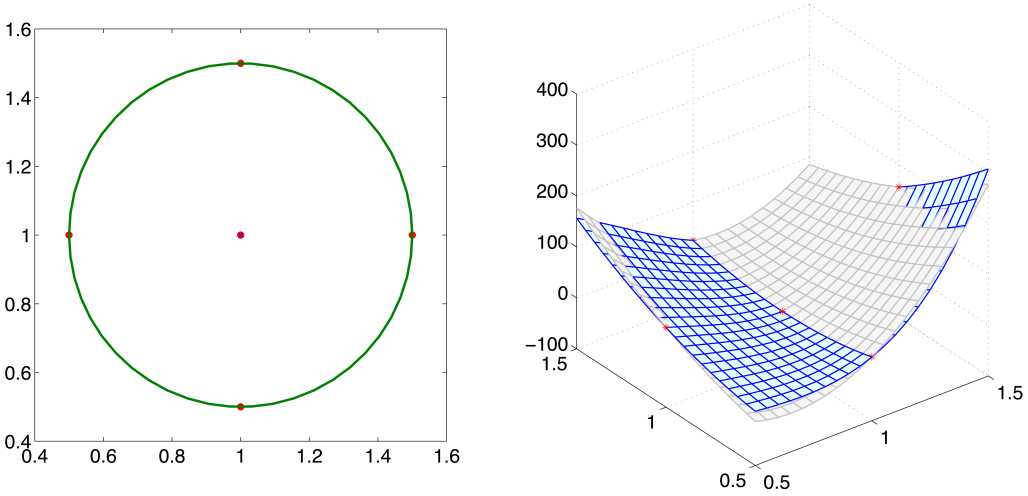

Figura 4.1: Função de Rosenbrock numa vizinhança do ponto $(1, 1)$ com raio $\delta = 0.5$.

A Figura 4.2 mostra a aproximação da função de Rosenbrock próxima ao ponto $(1, 1)$ e com raio $\delta = 0.25$. O maior erro em termos de valor funcional nos pontos de

amostra foi observado no ponto $y = (1, 1)$, para o qual $|f(y) - m(y)| \approx 2.500014 \times 10^{-4}$.

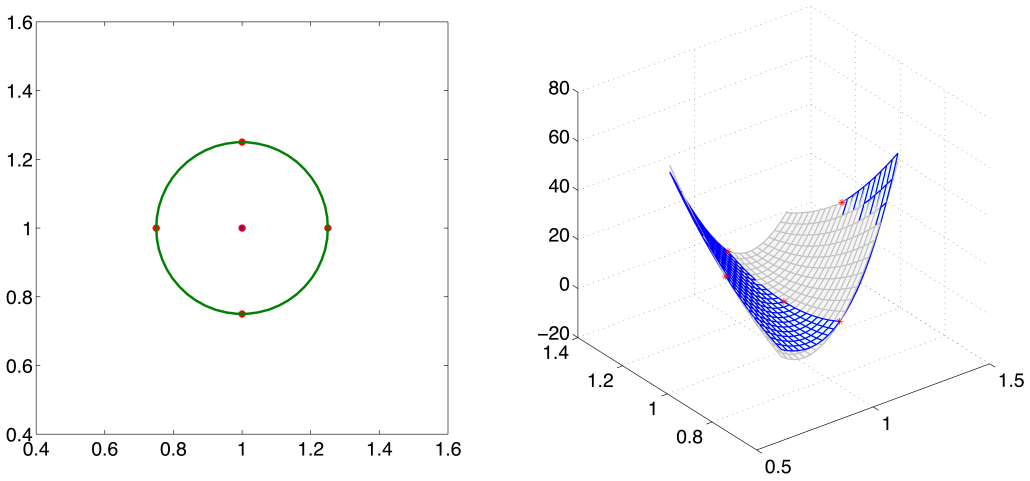

Figura 4.2: Função de Rosenbrock numa vizinhança do ponto $(1, 1)$ com raio $\delta = 0.25$.

A Figura 4.3 mostra a aproximação da função de Rosenbrock na vizinhança do ponto $(1, 1)$ com raio $\delta = 0.1$. O maior erro em termos de valor funcional nos pontos de amostra foi observado no ponto $y = (1, 1)$, para o qual $|f(y) - m(y)| \approx 1.000090 \times 10^{-4}$.

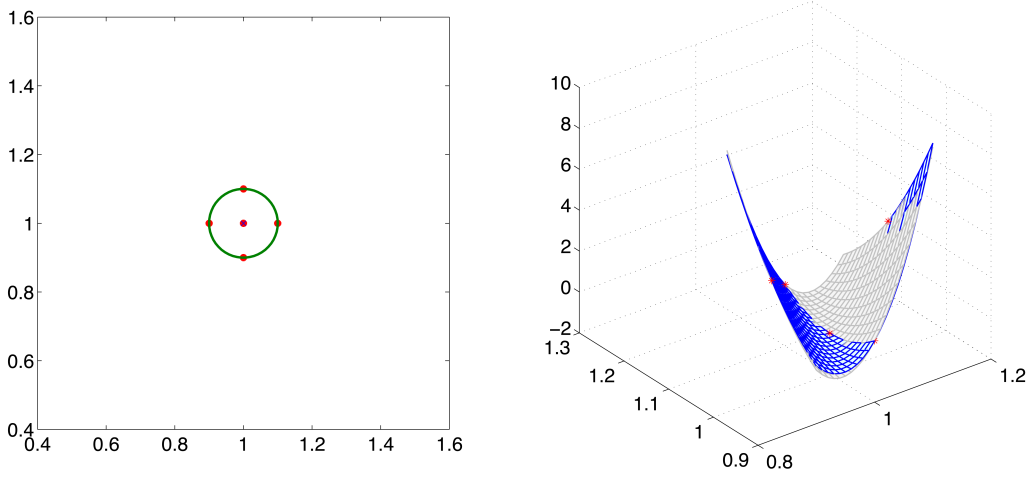

Figura 4.3: Função de Rosenbrock numa vizinhança do ponto $(1, 1)$ com raio $\delta = 0.1$.

Função Freudenstein e Roth

A função de Freudenstein e Roth $f : \mathbb{R}^2 \to \mathbb{R}$, definida por

$$f(x) = (-13 + x_1 + ((5 - x_2)x_2 - 2)x_2)^2 + (-29 + x_1 + ((x_2 + 1)x_2 - 14)x_2)^2,$$

é a função número 2 da coleção [28].

A Figura 4.4 mostra a aproximação da função de Freudenstein e Roth próxima ao ponto $(1, 1)$ e com raio $\delta = 0.5$. O maior erro em termos de valor funcional nos pontos de amostra foi observado no ponto $y = (1, 0.5)$, para o qual $|f(y) - m(y)| \approx 5.000004 \times 10^{-4}$.

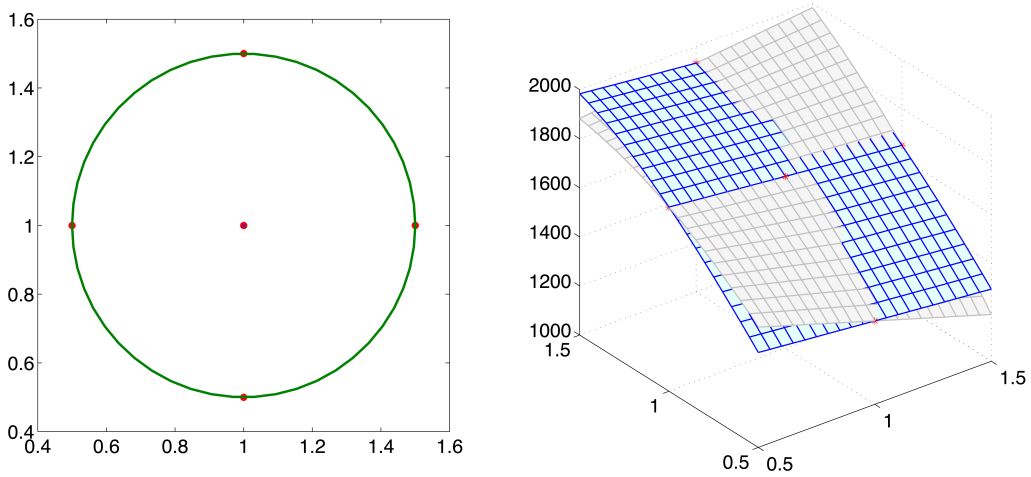

Figura 4.4: Função de Freudenstein e Roth numa vizinhança do ponto $(1, 1)$ com raio $\delta = 0.5$.

A Figura 4.5 mostra a aproximação da função de Freudenstein e Roth próxima ao ponto $(1, 1)$ e com raio $\delta = 0.25$. O maior erro em termos de valor funcional nos pontos de amostra foi observado no ponto $y = (1, 0.75)$, para o qual $|f(y) - m(y)| \approx 2.500015 \times 10^{-4}$.

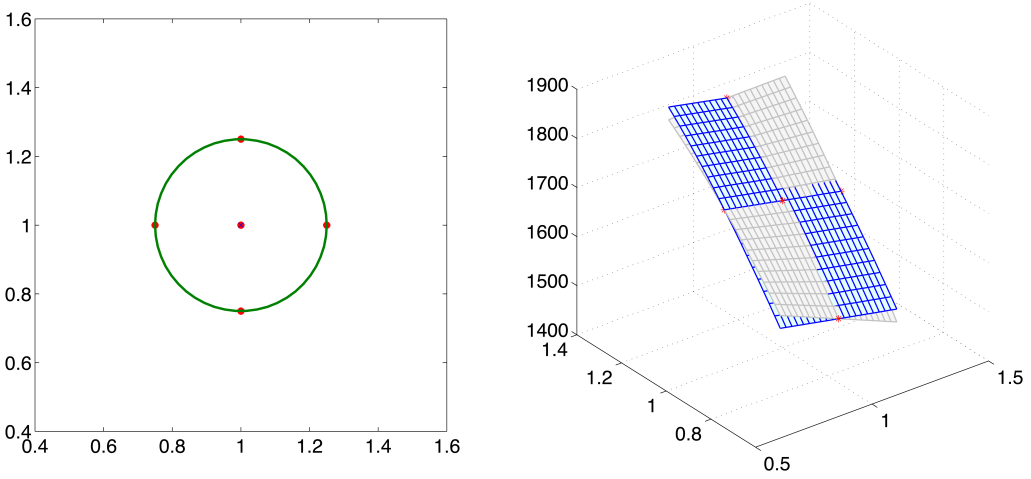

Figura 4.5: Função de Freudenstein e Roth numa vizinhança do ponto $(1, 1)$ com raio $\delta = 0.25$.

A Figura 4.6 mostra a aproximação da função de Freudenstein e Roth próximo ao ponto $(1, 1)$ e com raio $\delta = 0.1$. O maior erro em termos de valor funcional nos pontos de amostra foi observado no ponto $y = (1, 0.9)$, para o qual $|f(y) - m(y)| \approx 1.000090 \times 10^{-4}$.

As figuras apresentadas nesta seção mostram a aproximação de duas funções definidas em $\mathbb{R}^2$ em que conseguimos visualizar a melhora da qualidade da aproximação quando

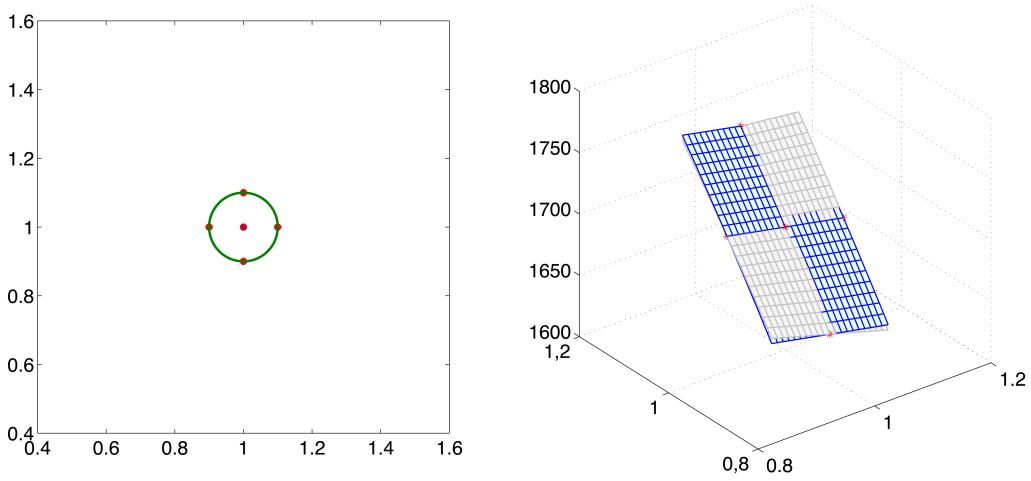

Figura 4.6: Função de Freudenstein e Roth numa vizinhança do ponto $(1, 1)$ com raio $\delta = 0.1$.

diminuímos o raio δ do conjunto de amostra. Isso sugere que para modelos mais precisos, precisamos considerar raios cada vez menores. Esse comportamento é semelhante aos modelos construídos por interpolação polinomial.

4.2 Comparação dos modelos

Nesta seção, faremos comparações entre os modelos quadráticos construídos por regressão via vetores suporte e os modelos quadráticos construídos por interpolação polinomial.

Os testes aos quais o método foi submetido constituem todos os problemas da coletânea organizada por Moré, Garbow e Hillstrom [28]. Trata-se de um conjunto de 35 problemas diferenciáveis de minimização irrestrita, onde as funções objetivo são somas de quadrados. Isto significa que cada função é da forma

$$f(x) = \sum_{i=1}^{m} (f_i(x))^2, \tag{4.1}$$

onde $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, são funções dadas. Para algumas funções a dimensão é fixada, e em outras pode ser escolhida pelo usuário. Implementações em Matlab e em Fortran deste banco de funções estão disponíveis em

`http://www.mat.univie.ac.at/~neum/glopt/test.html`.

Para a construção dos modelos de regressão via vetores suporte são considerados dois valores para o parâmetro C, $C = 10^8$ e $C = 10^{12}$. A tolerância para a construção dos modelos em ambos os casos foi $\varepsilon = 5 \times 10^{-6}$, o que corresponde a $\varepsilon = 0.05\delta^2$ uma vez que

foi usado $\delta = 0.01$. O modelo por interpolação foi construído através dos polinômios de Lagrange, utilizando para isso o Algoritmo 6.2 de [10]. Os três modelos são construídos com o mesmo conjunto de amostra para cada um das 35 funções dadas em [28]. A partir do ponto inicial x^0, fornecido com a coleção, são tomadas as direções coordenadas e as opostas a elas de modo que os pontos escolhidos estejam na fronteira da bola $B(x^0, \delta)$, conseguindo assim $2n+1$ pontos, em que n é a dimensão do problema. A partir desse conjunto inicial, escolhemos novos pontos na bola $B(x^0, \delta)$ para formarmos um conjunto com $(n+1)(n+2)/2$ pontos com o melhor Λ-posicionamento possível utilizando o Algoritmo 6.2 de [10].

Consideramos o erro relativo entre o gradiente do modelo e o gradiente da função no ponto inicial x^0 fornecido para cada problema, calculado pela fórmula

$$E_r = \frac{\|\nabla f(x^0) - \nabla m(x^0)\|}{\|\nabla f(x^0)\|}.$$

A Figura 4.7 apresenta a comparação entre o erro relativo E_r para cada um dos problemas. Vemos que o erro relativo E_r para a construção dos modelos por regressão via vetores suporte com $C = 10^8$ é menor do que o erro relativo dos modelos construídos por interpolação polinomial em um número maior de problemas.

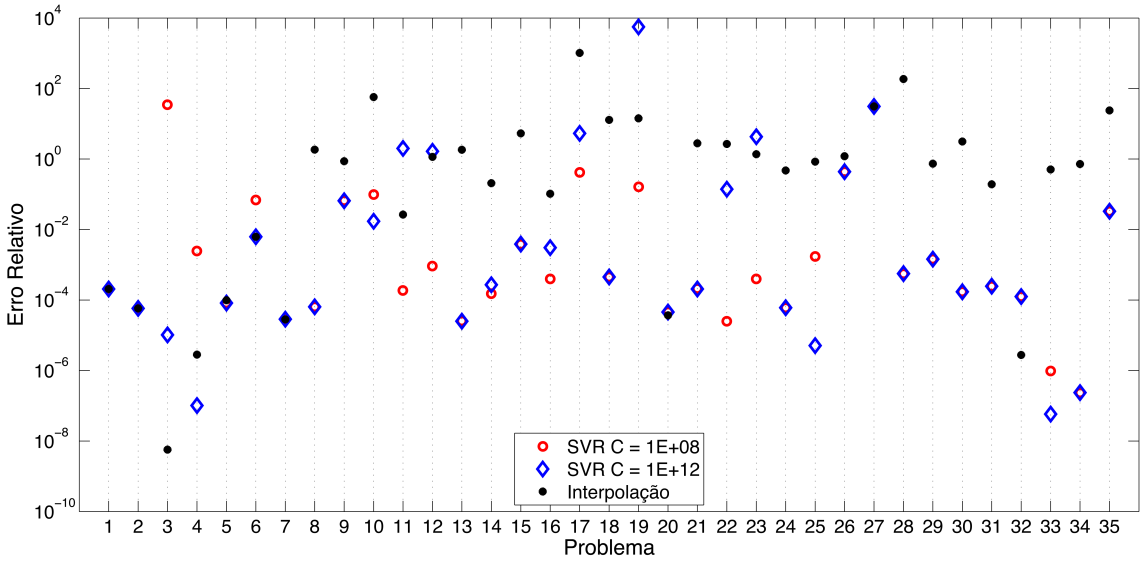

Figura 4.7: Erro relativo na norma do gradiente do modelo.

A Figura 4.8 apresenta a quantidade de problemas divididos em 6 intervalos para o erro E_r. Vemos que a quantidade de problemas com um erro relativo E_r nos intervalos mais baixos é maior quando os modelos são construídos por regressão via vetores suporte. Isso sugere que o gradiente dos modelos assim construídos estão aproximando melhor o gradiente da função no ponto inicial x^0.

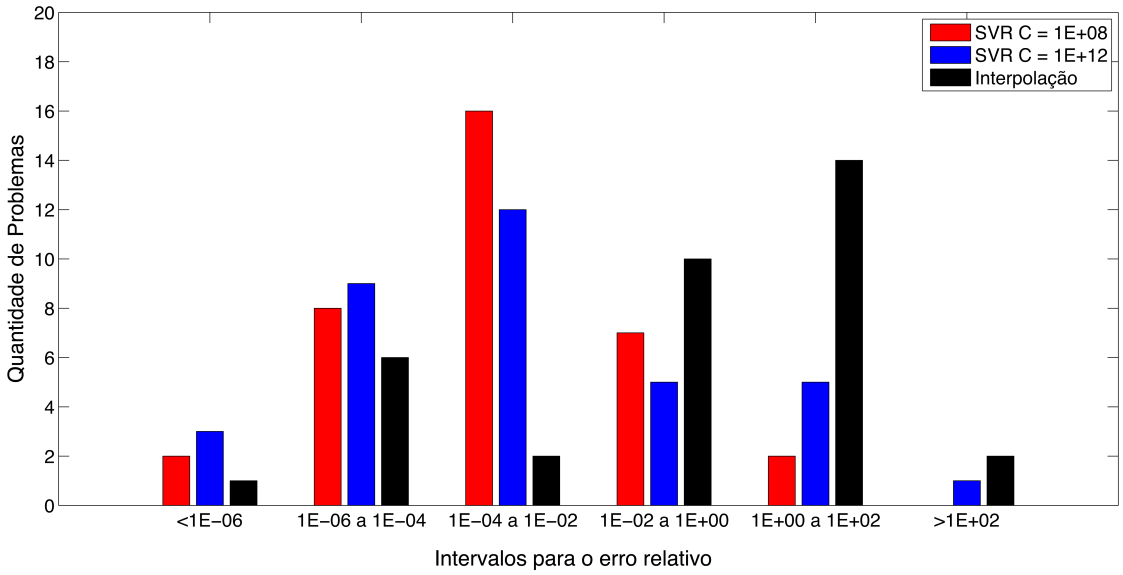

Figura 4.8: Erro relativo na norma do gradiente do modelo por intervalos.

4.3 O método de região de confiança

O objetivo desta seção é comparar diferentes estratégias para a construção do modelo quadrático m_k no Algoritmo 3.1 na resolução dos 35 problemas da coleção [28]. Utilizamos em todas as iterações do algoritmo $\delta_k = \Delta_k$. Além disso, foram adotados os seguintes parâmetros:

$$\delta_0 = 1, \quad \beta = 1, \quad \tau_1 = 0.5, \quad \tau_2 = 2.2, \quad \eta = 0.1, \quad \eta_1 = 0.25 \quad \text{e} \quad \eta_2 = 0.75.$$

O critério de parada adotado, como sugerido em [7], foi o tamanho do raio, ou seja, o algoritmo para numa iteração k quando

$$\|\delta_k\| \leq 10^{-8}. \tag{4.2}$$

Os subproblemas quadráticos (3.3) foram resolvidos pela rotina `trust` do Matlab. Duas técnicas foram utilizadas na construção dos modelos, a saber, o Algoritmo 2.2, que constrói os modelos por máquinas de vetores suporte, e interpolação polinomial.

Independentemente da técnica utilizada, os conjuntos de amostra possuem $(n+1)(n+2)/2$ pontos, onde n é a dimensão do problema. O primeiro conjunto é construído da seguinte maneira. A partir do ponto inicial x^0, são tomados passos de tamanho δ_0 nas direções coordenadas e opostas a elas, obtendo assim $2n+1$ pontos. Os pontos restantes são obtidos pelo Algoritmo 6.2 apresentado em [10]. Quanto à atualização do conjunto de amostra, a cada tentativa de um novo iterando do método de região de confiança, há duas possibilidades: o ponto ser aceito ou recusado. Se o ponto é aceito, ele é incluído

no conjunto de amostra substituindo o ponto mais distante dele. Caso seja recusado, verificamos se ele está mais próximo do iterando atual que o ponto mais distante. Em caso afirmativo, trocamos o ponto mais distante pelo ponto tentativo. Caso contrário, o conjunto de amostra permanece inalterado.

Na construção do modelo m_k pelo Algoritmo 2.2 foi adotado $\varepsilon = 0.05\delta_k^2$. Na resolução do problema (2.26) foi utilizada a rotina `quadprog` do Matlab. Diferentes valores para o parâmetro C foram adotados, a saber, $C = 10^i$, $i = 5, ..., 12$.

Quando a técnica considerada é a interpolação polinomial, os modelos são obtidos através do Algoritmo 6.2 de [10]. Como de uma iteração para outra do Algoritmo 3.1 apenas um ponto interpolador é alterado, uma única iteração do Algoritmo 6.2 é necessária para atualizar o modelo.

Em suma, o Algoritmo 3.1 foi testado com as seguintes estratégias para a construção dos modelos:

- C_5: Algoritmo 2.2 com $C = 10^5$.
- C_6: Algoritmo 2.2 com $C = 10^6$.
- C_7: Algoritmo 2.2 com $C = 10^7$.
- C_8: Algoritmo 2.2 com $C = 10^8$.
- C_9: Algoritmo 2.2 com $C = 10^9$.
- C_{10}: Algoritmo 2.2 com $C = 10^{10}$.
- C_{11}: Algoritmo 2.2 com $C = 10^{11}$.
- C_{12}: Algoritmo 2.2 com $C = 10^{12}$.
- Int: Interpolação polinomial [10, Algoritmo 6.2].

Seja $f(\bar{x})$ o valor da função objetivo encontrado pelo Algoritmo 3.1 ao resolver um problema utilizando a estratégia E. Similarmente ao apresentado por Bueno et al. em [4], consideramos que o problema foi resolvido pela estratégia E se

$$\frac{f(\bar{x}) - f_{min}}{\max\{1, |f(\bar{x})|, |f_{min}|\}} \leq 0.1, \tag{4.3}$$

em que f_{min} é o menor valor da função objetivo entre todas as estratégias.

Outro critério de solução, semelhante ao adotado em [6], usa a solução f_{MGH} apresentada em [28], no lugar de f_{min}, ou seja,

$$\frac{f(\bar{x}) - f_{MGH}}{\max\{1, |f(\bar{x})|, |f_{MGH}|\}} \leq 0.1. \tag{4.4}$$

4.3.1 Análise de desempenho das nove estratégias

Nas Tabelas 4.3-4.9 são apresentados os resultados obtidos mostrando o número do problema (cf. [28]), o número de variáveis, a dimensão m que aparece em (4.1), o tempo em segundos gasto para resolver o problema, a solução encontrada e a solução apresentada em [28]. O símbolo $\diamond$ indica que a solução encontrada não satisfez (4.3). O símbolo $\ddagger$ indica que a solução encontrada não satisfez (4.4).

A Tabela 4.1 mostra o número de problemas considerados resolvidos pelo Algoritmo 3.1 com cada uma das estratégias, considerando os critérios (4.3) e (4.4).

Tabela 4.1: *Número de problemas resolvidos do total de 35 problemas.*

Estratégia / **Critério**	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	Int
(4.3)	25	28	29	31	30	27	25	24	34
(4.4)	23	26	27	29	27	25	23	22	30

Utilizamos o conceito de perfil de desempenho (*performance profile*) [15], uma ferramenta para comparar a performance de um conjunto de métodos quando aplicados para resolver uma coletânea de problemas. Adotamos o número de avaliações de função como medida de desempenho na comparação das estratégias.

A Figura 4.9 apresenta o perfil de desempenho para o método de região de confiança com modelos construídos por interpolação polinomial e modelos construídos via regressão por vetores suporte com diferentes valores para o parâmetro de regularização C. Neste caso é considerado (4.3) como critério de solução. Com base na figura da esquerda e nas informações da Tabela 4.1, vemos que as três estratégias mais robustas foram: Int, C_8 e C_9. Pela figura da direita, vemos que a estratégia Int foi a mais eficiente, com 40% dos problemas resolvidos com o menor número de avaliações de função. No entanto, as estratégias C_8 e C_9 resolvem esta quantidade de problemas gastando não mais que 1.1 vezes o número de avaliações de função da melhor estratégia. Gastando não mais que o dobro do número de avaliações de função da melhor estratégia, C_8 e C_9 resolveram 80% dos problemas, enquanto Int resolve 63%.

Se considerarmos (4.4) como critério de solução, a robustez de todas as estratégias diminui, o que pode ser inferido também pela Tabela 4.1. A Figura 4.10 apresenta o perfil de desempenho relativo ao número de avaliações de função considerando (4.4) como critério de solução. Novamente as estratégias Int, C_8 e C_9 estão entre as mais robustas e as mais eficientes. Pela figura da direita vemos que a estratégia mais eficiente foi a construção de modelos por interpolação polinomial, com 37% dos problemas resolvidos com o menor número de avaliações de função. A estratégia C_8 levou aproximadamente

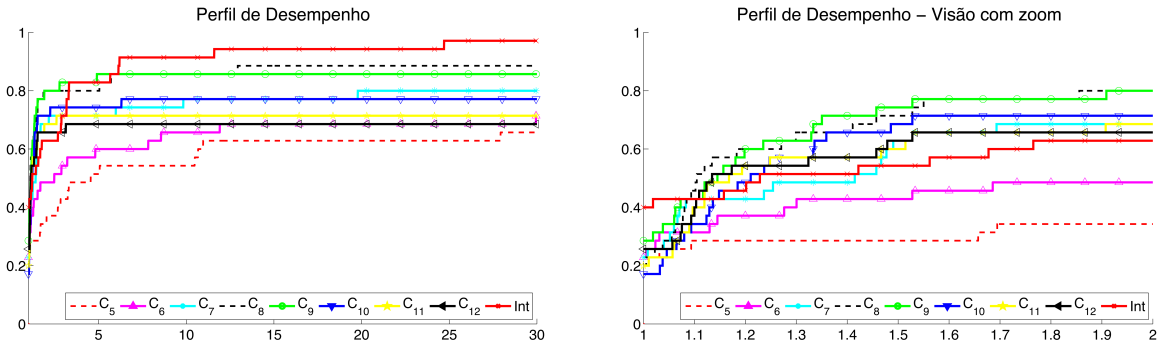

Figura 4.9: Perfil de desempenho relativo ao número de avaliações de função usando (4.3).

1.1 vezes o número de avaliações de função da melhor estratégia para atingir essa marca. A estratégia C_9 levou pouco menos do que 1.2 vezes o número de avaliações de função da melhor estratégia para resolver 37% dos problemas. Com até 2 vezes o número de avaliações de função da melhor estratégia, C_8 resolveu 74% dos problemas.

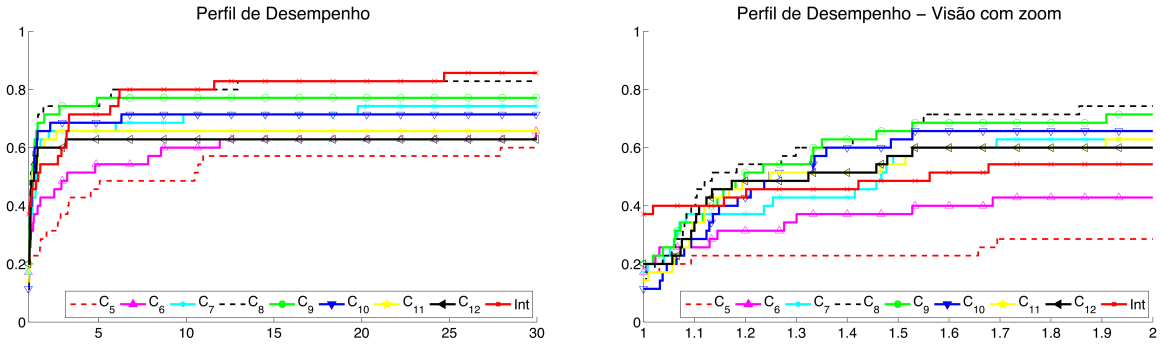

Figura 4.10: Perfil de desempenho relativo ao número de avaliações de função usando (4.4).

A Figura 4.11 apresenta o gráfico de perfil de dados [29], considerando (4.4) como critério de solução. A construção de modelos por regressão via vetores suporte com o parâmetro de regularização $C = 10^8$ se mostrou competitiva, pois resolveu 80% dos 35 problemas com pouco mais de 700 avaliações de função, enquanto o método de região de confiança com modelos construídos por interpolação polinomial gastou quase 1400 avaliações de função para resolver esta mesma quantidade de problemas.

A Figura 4.12 apresenta o gráfico de perfil de desempenho relativo ao tempo computacional gasto pelas estratégias para a resolução dos problemas. A figura da esquerda é o perfil de desempenho considerando (4.3) como critério de solução, enquanto na figura da direita foi adotado o critério (4.4). Os resultados obtidos são semelhantes ao desempenho considerando o número de avaliações de função. Com isso, podemos inferir que o tempo para a construção de um modelo por interpolação polinomial ou por regressão via vetores suporte é similar.

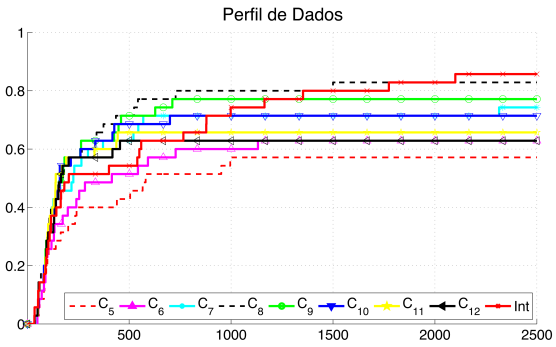

Figura 4.11: Perfil de dados relativo ao número de avaliações de função usando (4.4).

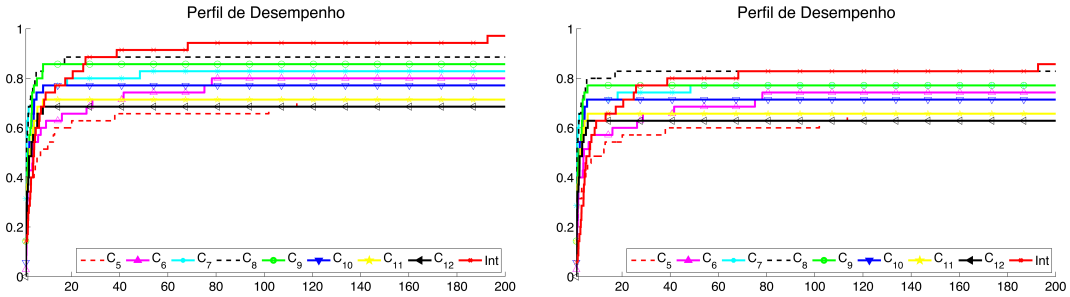

Figura 4.12: Perfil de desempenho relativo ao tempo computacional usando critérios (4.3) e (4.4).

Outra investigação realizada foi a análise da robustez das estratégias quando melhores soluções são exigidas. Ou seja, analisamos a robustez de cada estratégia para diferentes valores de $\varepsilon_{sol} > 0$, considerando que a estratégia encontrou uma solução se

$$\frac{f(\bar{x}) - f_{MGH}}{\max\{1, |f(\bar{x})|, |f_{MGH}|\}} \leq \varepsilon_{sol}. \tag{4.5}$$

A Figura 4.13 apresenta a robustez de cada estratégia para três valores de ε_{sol}. Quando levamos em conta uma melhor qualidade na solução, a robustez de todas as estratégias diminui, mas as estratégias Int, C_8 e C_9 continuam entre as mais robustas.

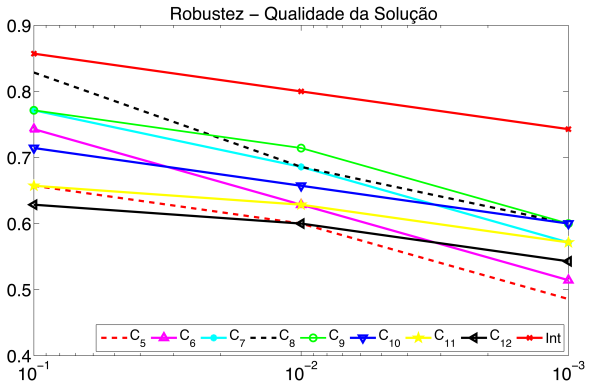

Figura 4.13: Robustez das estratégias considerando diferentes valores de ε_{sol} em (4.5).

4.3.2 Análise de desempenho das três melhores estratégias

Nesta seção procuramos analisar o desempenho das três melhores estratégias entre as discutidas na seção anterior, ou seja, C_8 e C_9, que foram a mais robusta e mais eficiente, respectivamente, entre as que utilizam o Algoritmo 2.2, e Int, que foi a mais robusta e eficiente de todas as estratégias.

As Figuras 4.14 e 4.15 mostram o decréscimo da função objetivo ao longo das iterações do algoritmo considerando cada uma das três estratégias para minimização de diferentes funções da coleção [28]. Infere-se destas figuras que o algoritmo de região de confiança com as três estratégias parece ter o mesmo comportamento em termos da variação dos valores da função objetivo, mas tem dificuldade em obter progresso à medida que se aproxima da solução ou em perceber que é momento de parar.

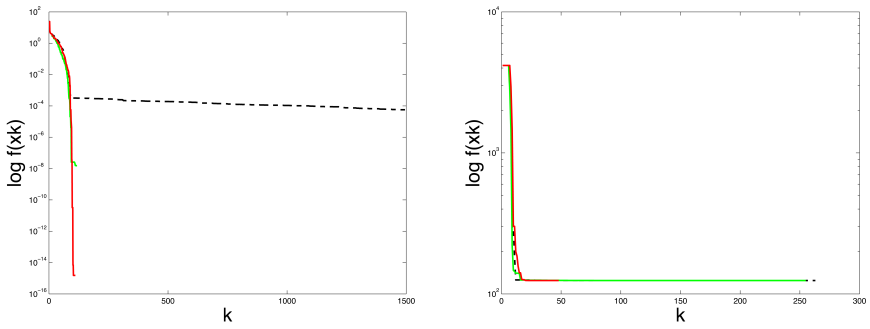

Figura 4.14: Variação da função ao longo das iterações na minimização das funções 1 (esquerda) e 6 (direita).

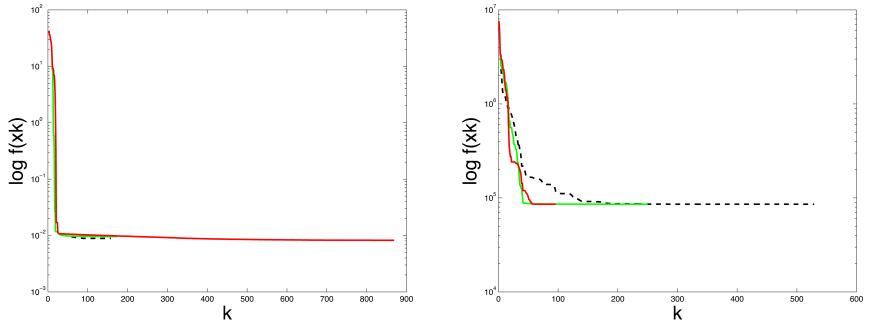

Figura 4.15: Variação da função ao longo das iterações na minimização das funções 8 (esquerda) e 16 (direita).

Como o valor ótimo f_{MGH} de cada uma das funções é conhecido, os testes numéricos com estas três estratégias foram refeitos, incluindo um critério de parada adicional a (4.2). O algoritmo de região de confiança foi parado prematuramente quando, em alguma iteração k,

$$\frac{f(x^k) - f_{MGH}}{\max\{1, |f(x^k)|, |f_{MGH}|\}} \leq 10^{-3}. \tag{4.6}$$

Acrescentar esse novo critério de parada evita que o algoritmo fique rodando quando já está próximo da solução, porém não impede que o algoritmo pare longe da solução pelo critério de parada (4.2).

A Figura 4.16 apresenta o gráfico de perfil de desempenho em relação ao número de avaliações de função com (4.3) como critério de solução.

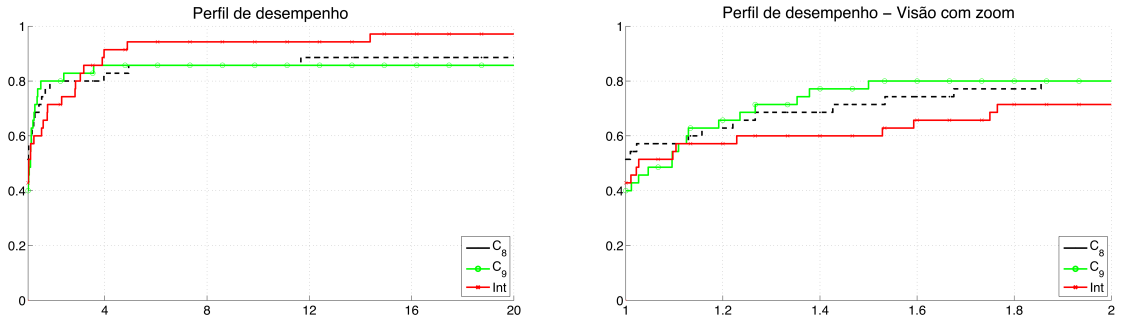

Figura 4.16: Perfil de desempenho em relação ao número de avaliações de função usando (4.3) e critério de parada (4.6).

O método de região de confiança com a estratégia de interpolação polinomial foi o mais robusto, resolvendo 97% dos problemas, a exceção foi o problema 21. Por outro lado, o método com a estratégia C_8 deixou de resolver 5 problemas (3, 4, 10, 22 e 25) e com a estratégia C_9 deixou de resolver 6 problemas (3, 4, 10, 11, 18 e 25).

A figura da direita é uma visão ampliada da figura da esquerda, construída para melhor identificarmos a eficiência das estratégias. O método de região de confiança com a estratégia de interpolação polinomial foi o mais eficiente em 57.1% dos problemas, seguido do uso das estratégias C_8, mais eficiente em 48.6%, e C_9, mais eficiente em 34.3% dos problemas.

A Figura 4.17 apresenta o gráfico de perfil de desempenho em relação ao número de avaliações de função com (4.4) para análise da solução. O método com os modelos construídos por interpolação resolveu 30 dos 35 problemas. Com a estratégia C_8 foram resolvidos 29 problemas e com a estratégia C_9, 27 problemas. Observe que são os mesmos resultados obtidos com o critério de parada original, utilizado na seção anterior. A figura da direita, com a visão ampliada, mostra que interpolação polinomial foi mais eficiente em 45.7% dos problemas. As estratégias C_8 e C_9 foram mais eficiente em 40% e 31.4% dos problemas, respectivamente.

A Tabela 4.2 apresenta o número de problemas em que a estratégia parou de acordo com um dos dois critérios de parada: critério usual (4.2) ou parada prematura, por causa de (4.6).

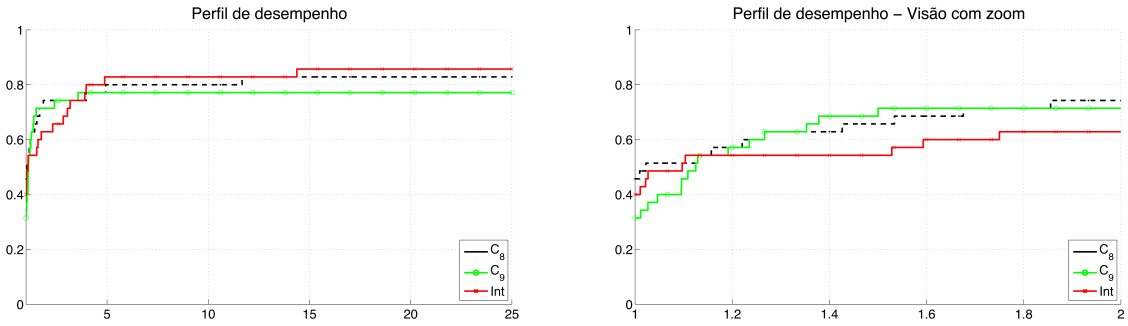

Figura 4.17: Perfil de desempenho em relaçao ao número de avaliações de função usando (4.4) e critério de parada (4.6).

Tabela 4.2: *Número de problemas por critério de parada*

Estratégia / **Critério de parada**	C_8	C_9	Int
(4.2)	21	19	12
(4.6)	14	16	23

O método de região de confiança com modelos construídos por interpolação polinomial foi o mais influenciado pela mudança no critério de parada, seguido pela construção com estratégia C_9. A construção dos modelos com a estratégia C_8 foi menos influenciado pela mudança no critério de parada.

Com essa influência da escolha do critério de parada, modificamos a precisão do critério de parada original e refizemos os testes para as três estratégias, a fim de identificarmos a influência do tamanho do raio ao resolvermos os problemas. Com o novo critério de parada considerado, o método para se

$$\delta_k \leq 10^{-5}. \tag{4.7}$$

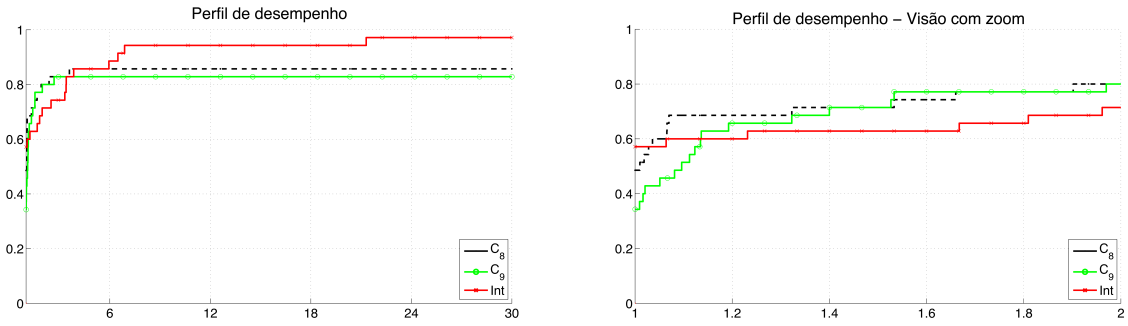

Figura 4.18: Perfil de desempenho em relação ao número de avaliações de função, usando (4.3) e critério de parada (4.7).

A Figura 4.18 apresenta o gráfico de perfil de desempenho em relação ao número de avaliações de função, usando (4.3) como critério de solução e (4.7) como critério de

parada. Os resultados apresentados foram semelhantes aos apresentados com a parada prematura quando próximo da solução ótima f_{MGH}. Quando o critério de solução é (4.4), os resultados são também semelhantes, uma vez que o critério de parada (4.6) engloba o critério de solução (4.4).

No geral, os testes mostram que a construção de modelos por regressão via vetores suporte para métodos de região de confiança sem derivadas são uma alternativa razoável à interpolação polinomial. Maiores investigações, no entanto, são necessárias para incorporar melhorias na técnica e com isso melhorar seu desempenho.

Tabela 4.3: *Resultados Numéricos com* (4.2) *como critério de parada*

P	**n**	**m**	**estratégia**	**tempo(s)**	**#f**	**sol f^***	**falha**	**f_{MGH}**
1	2	2	C_5	108.00	6533	2.2067E−05		0.0000E+00
			C_6	79.70	5716	1.0222E−04		
			C_7	51.30	5333	2.7177E−05		
			C_8	18.10	1500	5.7195E−05		
			C_9	1.06	123	1.5328E−08		
			C_{10}	2.44	261	1.6910E−04		
			C_{11}	1.09	123	2.2877E−06		
			C_{12}	1.16	128	4.6964E−05		
			Int	9.13	116	1.5270E−15		
2	2	2	C_5	79.80	5567	4.8984E+01		4.8984E+01
			C_6	55.00	3643	4.8987E+01		
			C_7	0.74	88	4.8984E+01		
			C_8	0.92	93	4.8984E+01		
			C_9	0.79	80	4.8984E+01		
			C_{10}	1.02	80	4.8985E+01		
			C_{11}	0.70	88	4.8990E+01		
			C_{12}	0.47	57	5.5034E+01	◇ ‡	
			Int	4.75	60	4.8984E+01		
3	2	2	C_5	0.80	59	1.3515E−01	◇ ‡	0.0000E+00
			C_6	0.88	63	1.3518E−01	◇ ‡	
			C_7	0.52	51	1.3518E−01	◇ ‡	
			C_8	0.57	52	1.3520E−01	◇ ‡	
			C_9	0.53	52	1.3520E−01	◇ ‡	
			C_{10}	0.50	49	1.3519E−01	◇ ‡	
			C_{11}	0.47	47	1.3520E−01	◇ ‡	
			C_{12}	0.45	44	1.3519E−01	◇ ‡	
			Int	44.00	559	2.7110E−11		
4	2	3	C_5	1.14	73	9.8439E+11	◇ ‡	0.0000E+00
			C_6	1.00	60	9.8085E+11	◇ ‡	
			C_7	0.56	48	9.7999E+11	◇ ‡	
			C_8	0.52	43	9.7350E+11	◇ ‡	
			C_9	0.59	55	9.7529E+11	◇ ‡	
			C_{10}	0.64	47	9.7822E+11	◇ ‡	
			C_{11}	0.67	69	9.8569E+11	◇ ‡	
			C_{12}	0.53	68	9.9767E+11	◇ ‡	
			Int	108.00	1353	3.2492E−06		
5	2	3	C_5	7.52	581	6.5977E−09		0.0000E+00
			C_6	3.55	255	1.9858E−08		
			C_7	1.27	115	3.7989E−07		
			C_8	0.63	55	8.9304E−08		
			C_9	0.58	54	1.9681E−08		
			C_{10}	0.63	55	5.8612E−08		
			C_{11}	0.64	53	5.3492E−09		
			C_{12}	0.66	57	1.7356E−09		
			Int	4.23	54	1.2893E−16		

Tabela 4.4: *Resultados Numéricos com* (4.2) *como critério de parada*

P	**n**	**m**	**estratégia**	**tempo(s)**	**#f**	**sol f^***	**falha**	**f_{MGH}**
6	2	10	C_5	8.19	566	1.2436E+02		1.2436E+02
			C_6	6.00	415	1.2436E+02		
			C_7	5.50	521	1.2436E+02		
			C_8	2.84	268	1.2436E+02		
			C_9	2.79	260	1.2436E+02		
			C_{10}	3.53	334	1.2436E+02		
			C_{11}	1.44	138	1.2436E+02		
			C_{12}	1.77	166	1.2436E+02		
			Int	4.16	53	1.2436E+02		
7	3	3	C_5	53.90	3267	8.2214E−04		0.0000E+00
			C_6	59.10	3504	4.2409E−04		
			C_7	25.80	2315	2.7585E−05		
			C_8	1.62	152	9.1626E−07		
			C_9	1.63	158	1.0523E−06		
			C_{10}	1.67	159	5.7110E−07		
			C_{11}	1.42	131	5.0316E−07		
			C_{12}	2.05	173	3.1319E−05		
			Int	9.22	117	5.8418E−15		
8	3	15	C_5	6.84	440	8.3009E−03		8.2149E−03
			C_6	98.00	5022	8.3047E−03		
			C_7	4.78	231	9.2063E−03		
			C_8	3.64	167	8.8898E−03		
			C_9	6.70	183	9.5954E−03		
			C_{10}	3.67	160	8.8955E−03		
			C_{11}	3.42	170	8.9058E−03		
			C_{12}	3.61	155	8.8982E−03		
			Int	70.1	878	8.2332E−03		
9	3	15	C_5	2.38	164	1.1404E−08		1.1279E−08
			C_6	1.30	89	1.2038E−08		
			C_7	0.38	39	2.8642E−08		
			C_8	0.36	38	2.0573E−08		
			C_9	0.33	36	2.3534E−08		
			C_{10}	0.38	36	2.2353E−08		
			C_{11}	0.39	36	2.1417E−08		
			C_{12}	0.41	36	1.6278E−08		
			Int	8.30	104	4.0724E−08		
10	3	16	C_5	1.49	85	6.9743E+06	◇ ‡	8.7946E+01
			C_6	2.14	129	6.9756E+06	◇ ‡	
			C_7	1.33	107	6.9747E+06	◇ ‡	
			C_8	1.79	127	6.9980E+06	◇ ‡	
			C_9	0.94	88	6.9797E+06	◇ ‡	
			C_{10}	0.78	70	6.9780E+06	◇ ‡	
			C_{11}	0.88	94	6.9834E+06	◇ ‡	
			C_{12}	0.45	55	6.8697E+06	◇ ‡	
			Int	3.39	43	6.0151E+06	‡	

Tabela 4.5: *Resultados Numéricos com* (4.2) *como critério de parada*

P	**n**	**m**	**estratégia**	**tempo(s)**	**#f**	**sol f^***	**falha**	**f_{MGH}**
11	3	20	C_5	3.13	201	3.7404E−02		0.0000E+00
			C_6	1.20	71	3.7413E−02		
			C_7	0.64	62	3.7475E−02		
			C_8	0.63	67	3.6908E−02		
			C_9	0.40	44	2.8634E−01	◇ ‡	
			C_{10}	0.36	39	2.8651E−01	◇ ‡	
			C_{11}	0.31	39	2.8651E−01	◇ ‡	
			C_{12}	0.31	39	2.8651E−01	◇ ‡	
			Int	16.30	204	3.6520E−02		
12	3	20	C_5	1.27	85	3.8250E−03		0.0000E+00
			C_6	23.00	728	5.1324E−03		
			C_7	0.88	91	3.4117E−03		
			C_8	1.20	108	2.9216E−04		
			C_9	1.29	113	6.0807E−03		
			C_{10}	1.00	96	1.2234E−02		
			C_{11}	0.97	106	1.0769E−01	◇ ‡	
			C_{12}	0.33	41	9.0741E+02	◇ ‡	
			Int	169.00	2100	4.2190E−05		
13	4	4	C_5	9.41	505	8.8502E−03		0.0000E+00
			C_6	3.47	199	1.3032E−02		
			C_7	2.81	226	1.6198E−04		
			C_8	2.45	181	3.8508E−04		
			C_9	2.25	177	3.3927E−04		
			C_{10}	2.11	160	3.4930E−04		
			C_{11}	2.92	191	2.3253E−04		
			C_{12}	2.33	153	1.7167E−04		
			Int	144.00	1774	1.6475E−08		
14	4	6	C_5	10.10	553	7.8770E+00	◇ ‡	0.0000E+00
			C_6	14.90	569	7.8303E+00	◇ ‡	
			C_7	226.00	4880	2.4668E−01	◇ ‡	
			C_8	10.30	435	8.6536E−02		
			C_9	10.90	448	1.3712E−04		
			C_{10}	10.10	427	8.6133E−06		
			C_{11}	11.20	437	1.9455E−05		
			C_{12}	11.60	455	4.2611E−06		
			Int	32.20	401	1.8896E−14		
15	4	11	C_5	3.59	235	5.3906E−04		3.0751E−04
			C_6	4.48	283	4.2876E−04		
			C_7	1.77	149	4.3487E−04		
			C_8	1.19	97	4.0979E−04		
			C_9	1.13	88	4.0695E−04		
			C_{10}	1.81	100	4.0925E−04		
			C_{11}	1.56	99	4.0466E−04		
			C_{12}	1.49	93	4.0782E−04		
			Int	43.80	540	3.6651E−04		

Tabela 4.6: *Resultados Numéricos com* (4.2) *como critério de parada*

P	**n**	**m**	**estratégia**	**tempo(s)**	**#f**	**sol f^***	**falha**	**f_{MGH}**
16	4	20	C_5	21.80	996	8.5822E+04		8.5822E+04
			C_6	28.00	1132	8.5822E+04		
			C_7	9.71	569	8.5824E+04		
			C_8	9.31	543	8.5823E+04		
			C_9	5.87	264	8.5827E+04		
			C_{10}	2.03	115	8.5827E+04		
			C_{11}	1.75	96	8.5823E+04		
			C_{12}	1.87	95	8.5829E+04		
			Int	9.02	110	8.5822E+04		
17	5	33	C_5	49.20	21	8.7903E−01	‡	5.4649E−05
			C_6	45.70	21	8.7903E−01	‡	
			C_7	39.60	21	8.7903E−01	‡	
			C_8	39.70	21	8.7903E−01	‡	
			C_9	39.80	21	8.7903E−01	‡	
			C_{10}	39.90	21	8.7903E−01	‡	
			C_{11}	39.50	21	8.7903E−01	‡	
			C_{12}	39.40	21	8.7903E−01	‡	
			Int	4.86	60	8.7868E−01	‡	
18	6	13	C_5	154.00	952	2.8335E−03		5.65565E−03
			C_6	55.40	543	1.1855E−02		
			C_7	8.45	266	5.2474E−02		
			C_8	7.67	188	8.1439E−02		
			C_9	5.15	78	2.7028E−01	◇ ‡	
			C_{10}	2.11	55	2.7150E−01	◇ ‡	
			C_{11}	3.17	55	2.7150E−01	◇ ‡	
			C_{12}	2.84	55	2.7150E−01	◇ ‡	
			Int	101.00	1164	1.9598E−03		
19	11	65	C_5	8.42	105	2.0934E+00	‡	4.0138E−02
			C_6	9.05	105	2.0934E+00	‡	
			C_7	6.08	105	2.0934E+00	‡	
			C_8	6.10	105	2.0934E+00	‡	
			C_9	15.00	105	2.0934E+00	‡	
			C_{10}	26.20	105	2.0934E+00	‡	
			C_{11}	35.50	105	2.0934E+00	‡	
			C_{12}	35.60	105	2.0934E+00	‡	
			Int	12.20	129	2.0482E+00	‡	
20	6	31	C_5	3.17	107	1.9459E−01	◇ ‡	2.2877E−03
			C_6	6.68	174	1.6276E−02		
			C_7	6.49	217	1.0345E−02		
			C_8	11.00	321	7.9470E−03		
			C_9	5.02	173	9.5107E−03		
			C_{10}	7.02	214	1.1567E−02		
			C_{11}	13.30	330	8.8306E−03		
			C_{12}	7.27	203	1.0465E−02		
			Int	46.40	550	4.8301E−03		

Tabela 4.7: *Resultados Numéricos com* (4.2) *como critério de parada*

P	**n**	**m**	**estratégia**	**tempo(s)**	**#f**	**sol f^***	**falha**	**f_{MGH}**
21	8	8	C_5	13.80	165	1.6333E+01	◇ ‡	0.0000E+00
			C_6	31.30	130	1.6429E+01	◇ ‡	
			C_7	62.50	611	2.8888E−01	◇ ‡	
			C_8	165.00	728	7.8389E−02		
			C_9	219.00	712	8.2023E−02		
			C_{10}	76.00	151	1.6453E+01	◇ ‡	
			C_{11}	60.90	132	1.8457E+01	◇ ‡	
			C_{12}	19.50	79	3.0821E+01	◇ ‡	
			Int	19.10	220	1.6572E+01	◇ ‡	
22	8	8	C_5	26.90	309	2.4470E−01	◇ ‡	0.0000E+00
			C_6	18.90	241	7.2595E−03		
			C_7	15.80	298	4.0267E−03		
			C_8	26.50	340	1.0655E−03		
			C_9	46.20	460	2.8886E−04		
			C_{10}	103.00	222	1.6549E+01	◇ ‡	
			C_{11}	6.97	74	7.1570E+01	◇ ‡	
			C_{12}	6.92	76	6.6444E+01	◇ ‡	
			Int	70.00	766	4.2657E−07		
23	10	11	C_5	6.94	104	8.0631E+03	◇ ‡	7.0877E−05
			C_6	215.00	592	8.6097E−05		
			C_7	58.70	545	7.8488E−05		
			C_8	121.00	524	2.0125E−04		
			C_9	209.00	628	7.4371E−05		
			C_{10}	264.00	700	7.6234E−05		
			C_{11}	48.20	150	3.9509E+04	◇ ‡	
			C_{12}	14.80	107	5.2254E+04	◇ ‡	
			Int	94.20	879	7.9822E−05		
24	10	20	C_5	3.63	94	7.2624E+00	◇ ‡	2.9366E−04
			C_6	3.78	94	8.9322E+00	◇ ‡	
			C_7	48.50	372	3.4087E−04		
			C_8	39.30	374	3.1561E−04		
			C_9	97.40	426	2.9921E−04		
			C_{10}	96.50	418	2.9868E−04		
			C_{11}	88.70	444	3.0087E−04		
			C_{12}	90.00	418	3.2646E−04		
			Int	106.00	998	3.1217E−04		
25	10	12	C_5	6.00	94	1.1901E+05	◇ ‡	0.0000E+00
			C_6	5.13	94	6.3565E+04	◇ ‡	
			C_7	14.40	96	1.8791E+05	◇ ‡	
			C_8	82.80	174	1.6856E+03	◇ ‡	
			C_9	731.00	490	2.9886E+00	‡	
			C_{10}	579.00	389	3.6772E+00	◇ ‡	
			C_{11}	539.00	365	1.2968E+01	◇ ‡	
			C_{12}	517.00	347	9.9505E+01	◇ ‡	
			Int	92.40	865	2.8720E+00	‡	

Tabela 4.8: *Resultados Numéricos com* (4.2) *como critério de parada*

P	**n**	**m**	**estratégia**	**tempo(s)**	$\#f$	**sol** f^*	**falha**	f_{MGH}
26	10	10	C_5	2.86	93	7.0758E−03		0.0000E+00
			C_6	3.64	93	7.0758E−03		
			C_7	2.04	93	7.0758E−03		
			C_8	2.22	93	7.0758E−03		
			C_9	2.28	93	7.0758E−03		
			C_{10}	2.83	93	7.0758E−03		
			C_{11}	3.01	93	7.0758E−03		
			C_{12}	3.04	93	7.0758E−03		
			Int	8.13	93	7.0758E−03		
27	10	10	C_5	4.69	99	4.9783E−10		0.0000E+00
			C_6	6.16	99	5.4302E−10		
			C_7	7.00	101	5.8673E−09		
			C_8	14.30	102	1.7373E−09		
			C_9	17.10	103	3.6262E−07		
			C_{10}	18.00	105	1.7458E−07		
			C_{11}	20.70	103	1.4652E−10		
			C_{12}	23.50	103	5.8539E−09		
			Int	8.73	96	6.4725E−22		
28	10	10	C_5	3.60	90	7.8852E−04		0.0000E+00
			C_6	2.72	90	7.8852E−04		
			C_7	1.98	90	7.8852E−04		
			C_8	2.55	90	7.8852E−04		
			C_9	3.33	90	7.8852E−04		
			C_{10}	4.67	90	7.8852E−04		
			C_{11}	4.42	90	7.8852E−04		
			C_{12}	4.56	90	7.8852E−04		
			Int	7.99	90	7.8852E−04		
29	10	10	C_5	7.31	140	8.9209E−07		0.0000E+00
			C_6	6.26	131	2.5783E−06		
			C_7	3.90	128	3.2493E−07		
			C_8	4.43	140	7.0304E−07		
			C_9	4.68	136	2.4964E−07		
			C_{10}	6.49	147	1.9985E−07		
			C_{11}	6.91	140	3.7594E−07		
			C_{12}	7.99	140	1.1849E−07		
			Int	17.60	182	1.5848E−04		
30	6	6	C_5	3.30	105	2.1873E−03		0.0000E+00
			C_6	6.41	134	9.9939E−05		
			C_7	6.38	112	4.1618E−06		
			C_8	12.90	116	1.2160E−06		
			C_9	10.70	109	1.4049E−06		
			C_{10}	14.00	131	5.2545E−06		
			C_{11}	12.20	111	4.1688E−06		
			C_{12}	13.30	139	5.2906E−06		
			Int	13.70	164	4.0117E−11		

Tabela 4.9: *Resultados Numéricos com* (4.2) *como critério de parada*

P	**n**	**m**	**estratégia**	**tempo(s)**	**#f**	**sol f^***	**falha**	**f_{MGH}**
31	5	5	C_5	7.81	242	2.9699E−06		0.0000E+00
			C_6	2.91	119	1.5657E−06		
			C_7	2.39	120	1.2691E−05		
			C_8	3.15	135	1.1096E−06		
			C_9	3.50	147	7.8358E−07		
			C_{10}	3.73	141	9.0180E−07		
			C_{11}	3.57	139	2.1380E−06		
			C_{12}	4.47	132	2.4532E−06		
			Int	11.60	143	8.7727E−14		
32	6	6	C_5	1.03	50	5.4580E−10		0.0000E+00
			C_6	1.02	50	2.1098E−10		
			C_7	0.79	53	6.3070E−10		
			C_8	0.95	56	4.1272E−09		
			C_9	1.40	56	8.3774E−09		
			C_{10}	1.87	54	7.4388E−09		
			C_{11}	1.92	56	2.0826E−09		
			C_{12}	1.84	55	1.3241E−08		
			Int	13.70	166	3.6722E−12		
33	6	6	C_5	1.33	58	1.1538E+00		1.1538E+00
			C_6	1.83	59	1.1538E+00		
			C_7	0.91	51	1.1538E+00		
			C_8	1.29	51	1.1538E+00		
			C_9	3.98	51	1.1538E+00		
			C_{10}	4.05	52	1.1538E+00		
			C_{11}	3.68	53	1.1538E+00		
			C_{12}	3.15	51	1.1538E+00		
			Int	2.86	35	1.1538E+00		
34	6	6	C_5	1.33	61	2.6667E+00		2.6667E+00
			C_6	1.28	55	2.6667E+00		
			C_7	1.41	55	2.6667E+00		
			C_8	3.70	55	2.6667E+00		
			C_9	4.57	55	2.6667E+00		
			C_{10}	5.17	55	2.6667E+00		
			C_{11}	4.33	55	2.6667E+00		
			C_{12}	4.36	55	2.6667E+00		
			Int	2.94	36	2.6667E+00		
35	9	9	C_5	6.19	82	2.8883E−02		0.0000E+00
			C_6	5.90	82	2.8883E−02		
			C_7	3.58	82	2.8883E−02		
			C_8	17.10	82	2.8883E−02		
			C_9	20.20	82	2.8883E−02		
			C_{10}	19.50	82	2.8883E−02		
			C_{11}	20.50	82	2.8883E−02		
			C_{12}	20.60	82	2.8883E−02		
			Int	6.89	82	2.8883E−02		

Conclusões

Em métodos de região de confiança sem derivadas, geralmente os modelos são construídos por interpolação polinomial. Nosso interesse foi estudar uma maneira alternativa para a construção de tais modelos. Para garantir a convergência de métodos de região de confiança, com ou sem derivadas, precisamos garantir a qualidade do modelo, no sentido que aproxime bem a função a ser otimizada. Neste trabalho, estudamos a possibilidade de construção dos modelos por regressão via vetores suporte.

Apresentamos uma revisão sobre aprendizagem de máquinas e máquinas de vetores suporte a fim de familiarizar o leitor com tais conceitos. As máquinas de vetores suporte são uma classe de algoritmos de aprendizagem supervisionada e podem ser utilizadas para a classificação de padrões ou para regressão. Apresentamos também uma revisão sobre a construção dos modelos por interpolação polinomial. Uma das contribuições da tese consiste nos resultados que mostram que os modelos construídos por regressão via vetores suporte aproximam bem a função e satisfazem as hipóteses necessárias para a convergência de um método de região de confiança livre de derivadas.

No Capítulo 3, apresentamos um algoritmo de região de confiança sem derivadas, baseado no trabalho de Conejo et al. [7], para a minimização de uma função objetivo em um conjunto convexo e fechado. Os modelos podem ser obtidos por qualquer técnica desde que sejam satisfeitas algumas hipóteses razoáveis. O algoritmo apresentado difere do discutido em [7] pela inclusão de um raio δ_k que controla a qualidade do modelo e mantém um raio Δ_k para a região de confiança.

O resultados obtidos no Capítulo 2 permitem o uso dos modelos construídos via regressão por vetores suporte em outros algoritmos de região de confiança, como por exemplo nos Algoritmos 10.1 e 10.3 de [10]. Tal uso é possível uma vez que com os resultados apresentados, os modelos construídos via regressão por vetores suporte são plenamente lineares considerando a Hipótese A1 para a função objetivo e são plenamente quadráticos considerando a Hipótese A2 para a função objetivo.

Por fim, no Capítulo 4, são compilados experimentos numéricos para ilustrar os

capítulos anteriores. Os testes mostram que modelos construídos por regressão via vetores suporte são boas aproximações para funções. Também comparamos o uso de modelos por interpolação polinomial com modelos por regressão via vetores suporte com diferentes escolhas do parâmetro de regularização. Nos testes preliminares, o método de região de confiança apresentado apresenta desempenho ligeiramente superior quando os modelos são construídos por interpolação polinomial. A técnica de regressão por vetores suporte apresentou um desempenho similar, o que nos motiva a continuar trabalhando para aprimorá-la em trabalhos futuros. Mais testes são necessários para verificar se existe uma classe de problemas em que a regressão via vetores suporte se mostre uma alternativa robusta e eficiente.

O trabalho realizado nos trouxe ainda outros pontos que podem ser respondidos em trabalhos futuros. Entre esses, estão: análise da influência da quantidade de pontos no conjunto de amostra para a qualidade do modelo, com a possibilidade de podermos trabalhar com problemas de dimensão alta; discussão sobre a resolução de problemas em que o valor da função objetivo é fornecido com ruído, em que modelos de regressão são menos afetados do que os modelos de interpolação polinomial; investigação referente ao uso de valores diferentes para o raio da região de confiança e do conjunto de pontos de amostra usados na construção dos modelos, no algoritmo implementado; discussão sobre o uso de máquinas de centro analítico para regressão como técnica de construção dos modelos.

Podemos concluir que a técnica de regressão via vetores suporte para a construção de modelos para um método de região de confiança se mostrou uma alternativa razoável, com garantia de boas propriedades teóricas mas que exige uma maior investigação para ser usada em problemas práticos.

Referências Bibliográficas

[1] E. Alpaydin. *Introduction to Machine Learning*. The MIT Press, Cambridge, 2004.

[2] A. S. Bandeira, K. Scheinberg e L. N. Vicente. Computation of sparse low degree interpolating polynomials and their application to derivative-free optimization. *Mathematical Programming*, 134(1):223–257, 2012.

[3] M. S. Bazaraa, H. D. Sherali e C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*. John Wiley & Sons, New York, 1993.

[4] L. F. Bueno, A. Friedlander, J. M. Martínez e F. N. C. Sobral. Inexact restoration method for derivative-free optimization with smooth constraints. *SIAM Journal on Optimization*, 23(2): 1189–1213, 2013.

[5] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1998.

[6] P. D. Conejo, E. W. Karas e L. G. Pedroso. A trust-region derivative-free algorithm for constrained optimization. *Optimization Methods and Software*, (no prelo), 2015.

[7] P. D. Conejo, E. W. Karas, L. G. Pedroso, A. A. Ribeiro e M. Sachine. Global convergence of trust-region algorithms for convex minimization without derivatives. *Applied Mathematics and Computation*, 220:324–330, 2013.

[8] A. R. Conn, N. I. Gould e Ph. L. Toint. *Trust-Region Methods*. Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[9] A. R. Conn, K. Scheinberg e Ph. L. Toint. On the convergence of derivative-free methods for unconstrained optimization In: M. D. Buhmann e A. Iserles (editores), *Approximation theory and optimization: Tributes to M. J. D. Powell*, pag 83 – 108. Cambridge University Press, Cambridge, 1997.

[10] A. R. Conn, K. Scheinberg e L. N. Vicente. *Introduction to derivative-free optimiza-*

tion. Society for Industrial and Applied Mathematics, Philadelphia, 2009.

[11] A. R. Conn, K. Scheinberg e L. N. Vicente. Geometry of sample sets in derivative-free optimization: polynomial regression and underdetermined interpolation. *IMA Journal of Numerical Analysis* , 28:721–748, 2008.

[12] A. R. Conn e Ph. L. Toint. An algorithm using quadratic interpolation for unconstrained derivative-free optimization. In: G. Di Pillo e F. Giannessi (editores), *Nonlinear Optimization and Applications*, pag 27–47. Springer-Verlag, New York, 1996.

[13] N. Cristianini e J. Shawe-Taylor. *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, New York, 2000.

[14] J. E. Dennis Jr e R. B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*. Society for Industrial and Applied Mathematics, Philadelphia, 1996.

[15] E. D. Dolan e J. J. Moré, Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91:201–213, 2002.

[16] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola e V. N. Vapnik. Support vector regression machines. *Advances in Neural Information Processing Systems*, 9:155–161, 1996.

[17] G. Fasano, J. L. Morales e J. Nocedal. On the geometry phase in model-based algorithms for derivative-free optimization. *Optimization Methods and Software*, 24:145–154, 2009.

[18] P. Flach. *Machine learning: The art and science of algorithms that make sense of data*. Cambridge University Press, New York, 2012.

[19] R. Fletcher. *Practical methods of optimization*. John Wiley & Sons, Chichester, 1987.

[20] C. C. Gonzaga, E. W. Karas e M. Vanti. A globally convergent filter method for nonlinear programming. *SIAM Journal on Optimization*, 14:646–669, 2003.

[21] S. Gratton, Ph. L. Toint e A. Tröltzsch. An active set trust-region method for derivative-free nonlinear bound-constrained optimization. *Optimization Methods and Software*, 26:873–894, 2011.

[22] H. W. Kuhn e A. W. Tucker. Nonlinear programming. In *2nd Berkeley Symposium on Mathematical Statistics and Probabilistics*, pag 481–492, Berkeley, 1951.

[23] P. Harrington. *Machine learning in action*. Manning, Shelter Island, 2012.

[24] W. W. Hsieh. *Machine learning methods in the environmental sciences*. Cambridge Press, New York, 2009.

[25] A. Izmailov e M. Solodov. *Otimização. Volume 2: Métodos computacionais*. Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 2007.

[26] A. M. Malyscheff e T. B. Trafalis. An analytic center machine for regression. *Machine Learning*, 46(1-3):203–223, 2002.

[27] M. Mohri, A. Rostamizadeh e A. Talwalkar. *Foundation of machine learning*. The MIT Press, Cambridge, 2012.

[28] J. J. Moré, B. S. Garbow e K. E. Hillstrom. Testing unconstrained optimization software. *ACM Transactions on Mathematical Software*, 7(1):17–41, 1981.

[29] J. J. Moré e S. M. Wild. Benchmarking derivative-free optimization algorithms. *SIAM Journal on Optimization*, 20(1):172–191, 2009.

[30] K. P. Murphy. *Machine Learning: A probabilistic perspective*. The MIT Press, Cambridge, 2012.

[31] J. Nocedal e S. J. Wright. *Numerical optimization*. Springer Series in Operations Research. Springer-Verlag, New York, 2006.

[32] M. Pontil, R. M. Rifkin e T. Evgeniou. From regression to classification in support vector machines. In *7th European Symposium on Artificial Neural Networks*, pag 225–230, Bruges, 1999.

[33] M. J. D. Powell. UOBYQA: Unconstrained optimization by quadratic approximation. *Mathematical Programming*, 92(3):555–582, 2002.

[34] M. J. D. Powell. The NEWUOA software for unconstrained optimization without derivatives. In: G. Di Pillo e M. Roma (editores), *Large-Scale Nonlinear Optimization*, pag 255–297. Springer-Verlag, New York, 2006.

[35] M. J. D. Powell. On the convergence of trust region algorithms for unconstrained minimization without derivatives. *Computational Optimization and Applications*, 53(2):527–555, 2012.

[36] A. Quarteroni, R. Sacco e F. Saleri. *Numerical Mathematics*. Texts in Applied Mathe-

matics. Springer, New York, 2007.

[37] F. M. P. Raupp e B. F. Svaiter Analytic center of spherical shells and its application to analytic center machine. *Computational Optimization and Applications*, 43:329–352, 2007.

[38] A. A. Ribeiro e E. W. Karas. *Otimização contínua: Aspectos teóricos e computacionais*. Cengage Learning, São Paulo, 2013.

[39] S. A. Santos. Trust-region-based methods for nonlinear programming: Recent advances and perspectives. *Pesquisa Operacional*, 34(3):447–462, 2014.

[40] C. Sammut e G. I. Webb. *Encyclopedia of machine learning*. Springer, New York, 2010.

[41] K. Scheinberg e Ph. L. Toint. Self-correcting geometry in model-based algorithms for derivative-free unconstrained optimization. *SIAM Journal on Optimization*, 20:3512–3532, 2010.

[42] B. Schölkopf e A. J. Smola. *Learning with kernels: Support vector machines, regularization, optimization and beyond*. The MIT Press, Cambridge, 2002.

[43] B. Schölkopf, A. J. Smola, R. C. Williamson e P. L. Barlett. New support vector algorithms. *Neural Computation*, 12:1207–1245, 2000.

[44] A. J. Smola e B. Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14:199–222, 2004.

[45] S. Sra, S. Nowozin e S. J. Wright *Optimization for machine learning*. The MIT Press, Cambridge, 2012.

[46] J. Stoer e R. Bulirsch. *Introduction to numerical analysis*. Springer-Verlag, New York, 1980.

[47] A. N. Tikhonov e V. Y. Arsenin. *Solution of ill-posed problems*. V. H. Winston & Sons, Washington, 1977.

[48] A. Tröltzsch. *An active-set trust-region method for bound-constrained nonlinear optimization without derivatives applied to noisy aerodynamic design problems*. Tese de doutorado, Université de Toulouse, 2011.

[49] V. N. Vapnik. *Statistical learning theory*. John Wiley & Sons, New York, 1998.

[50] V. N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag, New York, 2000.

[51] D. Winfield. Function minimization by interpolation in a data table. *Journal of the Institute of Mathematics and its Applications*, 12:339–347, 1973.

[52] J. Winkler, M. Niranjan e N. Lawrence. *Deterministic and statistical methods in machine learning*. Springer, New York, 2004.

[53] Y. Zhang. *Application of machine learning*. In-Teh, Vukovar - Croácia, 2010.