

Aproximação Adaptativa de Funções com bases de elementos tensoriais compactos

Gilcélia Regiane de Souza,

UFSJ - Departamento de Física e Matemática, Campus Alto do Paraopeba,
36420-000, Ouro Branco, MG
E-mail: gilcelia@ufsj.edu.br,

Jorge Stolfi

Unicamp - Departamento de Teoria da Computação, IC,
13083-852, Campinas, SP
E-mail: stolfi@ic.unicamp.br.

Resumo: *O presente trabalho se enquadra na área de Análise Numérica, com enfoque em técnicas modernas de aproximação para funções em domínios multidimensionais que requerem resolução espacial adaptativa.*

Especificamente, o objetivo é encontrar algoritmos eficientes para aproximação de funções que apresentam detalhes importantes mas de pequena escala (alta frequência espacial) em regiões relativamente pequenas do domínio. Vamos supor que a função a aproximar (função objetivo) é amostrada em um número finito de pontos com posições arbitrárias, cuja densidade também pode variar bastante de uma região para outra. Nestas situações é desejável que a aproximação também seja adaptada à função objetivo, com maior parâmetros nas regiões onde há mais detalhes e/ou pontos.

A aproximação de uma função pode ser desejável por diversos motivos, incluindo eficiência computacional, considerações teóricas sobre a grandeza física representada pela função, ou pelo fato dela ser conhecida apenas parcialmente. Espera-se que as funções aproximadoras pertençam uma classe mais simples e sejam mais fáceis de calcular e manipular, por derivadas, integrais, etc.

Uma abordagem para aproximação adaptativa é utilizar combinações lineares de uma base de *elementos radiais* [1], com um elemento centrado em cada ponto dado. De modo geral, esta escolha dos centros dos elementos melhora a estabilidade dos algoritmos de aproximação [3].

Entretanto, quando o número de pontos é muito grande, mas a função amostrada é relativamente suave, podemos obter uma aproximação adequada com bases muito menores. Nesses casos, uma vez que não há um elemento da base para cada ponto, não há mais razão para que os elementos sejam centrados em pontos de amostragem.

Por isso, optamos por usar uma base em que os centros dos elementos são um subconjunto de uma *grade regular de centros*, independente dos pontos de amostragem.

Uma vez que optamos por uma grade regular de centros, o uso de elementos radiais (isotrópicos) também fica sem justificativa, pois o espaçamento entre os elementos depende da direção. Por isso, optamos pelo uso de *elementos tensoriais*, que são produto de elementos univariados cada qual dependendo de uma única coordenada do domínio. Os elementos tensoriais que usamos são derivados de um pequeno número de funções-mãe com suporte limitado (ou efetivamente limitado), por exemplo, escolhemos duas funções-mãe, propostas por Wendland:

$$\Phi_{3,2}^P(r) = (1-r)_+^4(4r+1) \quad \text{e} \quad \Phi_{5,4}^P(r) = (1-r)_+^7(16r^2+7r+1).$$

A função-mãe $\Phi^G(r)$, que denominamos de *Gaussiana*, definida pela fórmula

$$\Phi^G(r) = e^{-\frac{1}{2}r^2},$$

para todo $r \in \mathbb{R}$. Embora tenha suporte infinito, a função Φ^G diminui muito rapidamente quando o argumento $|r|$ aumenta. Os elementos B-splines, ou seja, a função-mãe genérica Φ_g^S , destes elementos é indexada pelo grau $g \in \mathbb{N}$ e consiste de $g+1$ polinômios de grau g definidos nos intervalos $[k, k+1]$, para k variando de $-(g+1)/2$ até $(g-1)/2$ inclusive, de 1 em 1. A função-mãe Φ_g^S tem continuidade $g-1$ nos extremos desses intervalos, e é zero para $|r| \geq (g+1)/2$, um exemplo,

$$\Phi_2^S(r) = \begin{cases} 1 - \frac{4}{3}r^2 & \text{se } |r| \leq \frac{1}{2} \\ \frac{2}{3}\left(\frac{3}{2} - |r|\right)^2 & \text{se } \frac{1}{2} < |r| < \frac{3}{2} \\ 0 & |r| > \frac{3}{2} \end{cases}$$

Algumas destas funções-mãe foram originalmente propostas para uso em bases radiais, cujos centros c_j eram irregularmente distribuídos. Na abordagem mais básica (monoescala), todos os elementos são idênticos exceto pela sua posição. Este modelo de aproximação é adequado quando se espera que a função objetivo seja igualmente complicada em todas partes do domínio.

Muitas vezes, a função a aproximar é praticamente nula em grande parte do domínio. Nestes casos, a base de aproximação não precisa cobrir uniformemente todo o domínio; podemos usar apenas os elementos da base que afetam as regiões onde a função é significativamente diferente de zero, com economia de espaço e tempo de processamento. Chamaremos as estratégias para escolher esses elementos de *esquemas adaptativos de aproximação*. (Ao contrário de muitos autores, nosso uso do termo ‘adaptativo’ não implica o uso de elementos de escala diferentes).

Nas ciências naturais e nas engenharias, é comum encontrar funções em que detalhes de tamanhos diferentes têm causas, efeitos ou relevâncias diferentes. Por exemplo, a forma da superfície da Terra tem detalhes de dezenas de quilômetros (montanhas) devidos ao movimento de placas tectônicas, e detalhes de centenas ou dezenas de metros devidos principalmente à erosão. Usualmente, a modelagem computacional de tais fenômenos requer soma de várias aproximações, onde cada aproximação (nível) captura detalhes significativos de um certo tamanho.

Para esta finalidade, podem ser usadas bases com fatores de escala espacial diferentes para cada nível. Vamos usar *bases multiescala* cujos fatores de escala diminuem de nível para nível em progressão geométrica. Esta abordagem é semelhante à análise wavelet, mas fornece um modelo analítico explícito para a função aproximadora e pode ser usada mesmo quando os pontos de amostragem não formam uma grade regular. Além de possíveis justificativas físicas, aproximações multiescala são a chave para algoritmos numéricos eficientes, para filtragem, integração de equações diferenciais e outros problemas.

Uma dificuldade recorrente em problemas de aproximação é o *efeito das bordas*: quando \mathbb{D} é um subconjunto limitado de \mathbb{R}^d , o erro de aproximação perto das bordas do domínio \mathbb{D} é geralmente diferente do erro na parte central. Por exemplo se os centros estão distribuídos uniformemente em \mathbb{D} , os pontos próximos à borda de \mathbb{D} serão cobertos por menos elementos da base do que os pontos na região central. Esta variação dificulta a análise comparativa de diferentes bases e diferentes métodos de aproximação.

Para evitar esta complicação, vamos supor que \mathbb{D} é um *toro d dimensional*; ou seja, uma caixa alinhada com os eixos, $\mathbb{D} = [0, L_1] \times [0, L_2] \times \dots \times [0, L_d]$, sendo que cada face de \mathbb{D} é identificada conceitualmente com a face oposta. Desta forma, todos os pontos ficam equivalentes em relação ao domínio e todos os elementos de uma base uniforme ficam equivalentes.

A suposição de domínio toroidal efetivamente elimina as bordas do domínio, e os problemas decorrentes delas. Esta suposição também permite reduzir drasticamente o custo computacional quando a base e a norma $\|\cdot\|_{\mathcal{E}}$ tem estrutura uniforme sobre a caixa \mathbb{D} .

Uma desvantagem de adotar esta hipótese é que ela pressupõe que a função a aproximar f tem valores e derivadas iguais nas faces opostas de \mathbb{D} . Se isto não é verdade, tratar o domínio \mathbb{D} como toroidal torna a função f descontínua na fronteira da caixa \mathbb{D} .

Por outro lado, ainda podemos usar esta abordagem quando a função f não é periódica; basta mapear o domínio real \mathbb{D}' do problema para uma região suficientemente pequena no interior da caixa \mathbb{D} , e atribuir valores adequados à função f na região $\mathbb{D} \setminus \mathbb{D}'$ de forma a torná-la periódica, sem introduzir descontinuidades excessivas nessa região.

Observamos que tanto as funções objetivo quanto as aproximantes são definidas por seus valores em uma lista finita $p = (p_1, p_2, \dots, p_N)$ de N pontos do domínio. Neste caso, as funções podem ser vistas como vetores de \mathbb{R}^N ; e tanto as normas quanto os produtos escalares sobre \mathcal{F} (espaço das funções objetivo), \mathcal{A} (espaço das funções aproximadoras) e \mathcal{E} (espaço dos resíduos) passam a ser normas e produtos internos sobre \mathbb{R}^N .

Podemos então especificar um problema de *aproximação* como consistindo de

- Um subespaço \mathcal{F} de \mathbb{R}^N , o espaço das *funções objetivo*;
- Um subespaço \mathcal{A} de \mathbb{R}^N , o espaço das *funções aproximadoras*;
- Um *critério de aproximação*, um predicado sobre $\mathcal{F} \times \mathcal{A}$.

As coordenadas dos vetores de \mathcal{A} e \mathcal{F} são chamadas *amostras*. Vários dos métodos e resultados podem ser estendidos para situações onde as amostras são valores de quaisquer outros N funcionais lineares independentes.

Uma *instância* de tal problema de aproximação é um vetor $f \in \mathcal{F}$. Uma *solução* para essa instância é um vetor $s \in \mathcal{A}$ tal que o par (f, s) satisfaz o critério de aproximação. Na escolha de s vamos considerar apenas o critério de aproximação ótima com norma hilbertiana, que resultam em operadores de aproximação $\mathcal{O} : \mathcal{F} \rightarrow \mathcal{A}$ e de resíduo $\mathcal{R} : \mathcal{F} \rightarrow \mathcal{E}$ que são lineares e idempotentes.

Vamos supor que os subespaços \mathcal{F} e \mathcal{A} são definidos por bases finitas de dimensões m e n , respectivamente. Os elementos dessas bases serão representados pelas colunas de duas matrizes reais F e A , de dimensões $N \times m$ e $N \times n$ respectivamente. Para que o problema seja bem posto é necessário que o número de pontos de amostragem N seja maior ou igual à dimensão n do espaço \mathcal{A} . (Em muitos casos tem-se $N \gg n$). O operador de coeficientes \mathcal{C} é portanto um operador linear de \mathcal{F} para \mathbb{R}^N .

Vamos analisar experimentalmente a qualidade das aproximações obtidas com bases tensoriais *uniformes* — cujos centros são dispostos em uma grade ortogonal regular sobre o domínio \mathbb{D} , e cujos os elementos têm o mesmo fator de escala ρ em todos os eixos.

Esquema de redução de base

Em muitas aplicações a função a aproximar é diferente de zero apenas em uma parte pequena do domínio. Nestas aplicações podemos economizar muito espaço e tempo de processamento, usando um esquema de *redução de base*. Um esquema deste tipo para um domínio \mathbb{D} consiste de

- Uma sequência finita de funções $\phi_1, \phi_2, \dots, \phi_m$ linearmente independentes, que chamaremos de *pré-base*, sendo cada ϕ_j uma função de \mathbb{D} para \mathbb{R} .
- Um *algoritmo de redução* que determina uma subsequência da pré-base ϕ ,

$$\hat{\phi} = \hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_n.$$

Uma aproximação dentro deste esquema é uma combinação linear

$$\hat{s} = \sum_{k=1}^n \hat{\alpha}_k \hat{\phi}_k, \tag{1}$$

sendo cada $\widehat{\alpha}_k$ um coeficiente real. O problema computacional inclui portanto a escolha da base $\widehat{\phi}$, além da determinação dos coeficientes $\widehat{\alpha}_1, \dots, \widehat{\alpha}_n$.

A principal vantagem de uma base menor é que o cálculo da função \widehat{s} pela fórmula (1) fica mais eficiente. Além disso, o custo de armazenar a base reduzida $\widehat{\phi}$ e os coeficientes $\widehat{\alpha}$ é proporcional ao tamanho n da base reduzida.

Algoritmos de redução de base são indicados quando se espera que a base $\widehat{\phi}$ seja muito menor que a pré-base ϕ . Tipicamente, isso ocorre quando a função f é praticamente nula numa região grande \mathbb{D}_0 do domínio e os elementos da pré-base dual [4] ϕ^* tem suporte essencialmente limitado. Nestas condições, qualquer elemento ϕ_j tal que o suporte de ϕ_j^* está inteiramente contido em \mathbb{D}_0 pode ser eliminado sem afetar muito o erro de aproximação. Em geral, a redução da base é ineficaz quando os elementos da base dual ϕ^* não tem suporte limitado.

Nos esquemas que consideramos, a pré-base ϕ é fixada independentemente da função f a aproximar e deve ser linearmente independente em relação aos pontos de amostragem p_1, p_2, \dots, p_N usados. Na versão discreta do problema, isto significa que a matriz da pré-base A , com $A_{ij} = \phi_j(p_i)$, $i \in \{1, \dots, N\}$ e $j \in \{1, \dots, m\}$, precisa ter posto m . Como observamos, o número m de elementos da pré-base ϕ não pode exceder o número N de pontos de amostragem.

Algoritmo Reduz($\phi, \alpha, e, \epsilon_{\max}$)

1. Calcule os valores de $\delta_j = |\alpha_j| \|\phi_j\|_{\infty}$ para $j = 1, \dots, m$
2. Para cada índice j em $1, \dots, m$, em ordem de δ_j crescente, faça:
 3. Se $\delta_j > 2\epsilon_{\max}$ vá para o passo 6.
 4. Se para todo i temos $|e_i| \leq \epsilon_{\max}$ e $|e_i + \alpha_j \phi_j(p_i)| \leq \epsilon_{\max}$, ou $|e_i| > \epsilon_{\max}$ e $|e_i + \alpha_j \phi_j(p_i)| \leq |e_i|$ então faça
 5. $e \leftarrow e + \alpha_j \phi_j$; $\alpha_j \leftarrow 0$.
6. Devolva a base $\widehat{\phi}$ que consiste dos elementos ϕ_j com $\alpha_j \neq 0$, e os respectivos coeficientes $\widehat{\alpha}$.

Para analisar o desempenho do algoritmo *Reduz* nas pré-bases escolhidas, vamos utilizar quatro funções objetivo denotadas por f_O, f_C, f_G e f_F , definidas a seguir. Em cada caso, mapeamos o domínio $\mathbb{D} = [0, L]^2$ para o domínio natural das funções objetivo $[-1/2, 1/2] \times [-1/2, 1/2]$ pelas fórmulas $X = (x/L) - 1/2$ e $Y = (y/L) - 1/2$, sendo (X, Y) os argumentos naturais da função objetivo e (x, y) o ponto de \mathbb{D} .

As funções f_O, f_C e f_F não são nulas na fronteira de \mathbb{D} , e portanto apresentam descontinuidades de valor e/ou derivada nessa fronteira quando consideramos \mathbb{D} com topologia toroidal. Entretanto a descontinuidade é muito pequena e não afeta os testes.

As funções:

- $f_O(X, Y) = \exp(-\frac{1}{0.015}(X^2 + Y^2)) \cos((10X + 5Y)\pi)$, esta função é contínua e diferenciável em todas as ordens exceto na fronteira de \mathbb{D} , onde seu valor absoluto é menor que 6×10^{-8} .

- $f_C(X, Y) = H_0 f_0(x, y) + H_1 f_1(x, y)$, sendo

$$f_0 = \begin{cases} 0 & \text{se } X^2 + Y^2 \geq S_0^2 \\ (1-v)^2 & \text{caso contrário} \end{cases} \quad \text{e } f_1 = \begin{cases} 0 & \text{se } X^2 + Y^2 \leq S_1 \text{ ou } x^2 + y^2 \geq S_2 \\ 4u(1-u) & \text{caso contrário} \end{cases}$$

em que $v = \frac{x^2+y^2}{S_0^2}$ e $u = \frac{x^2+y^2-S_1^2}{S_2^2-S_1^2}$. Os coeficientes são $S_0 = 0,1196$, $S_1 = 0,2208$, $S_2 = 0,4232$, $H_0 = 1,0$ e $H_1 = 0,5$. Esta função é identicamente nula perto da borda de \mathbb{D} , e tem descontinuidade nas derivadas de segunda ordem nos círculos de raios S_0, S_1 e S_2 .

- f_F , é a função de Franke [2], é um exemplo popular em estudo de aproximação.

$$\bullet f_G = g(X, Y) \sin(t) \text{ sendo } g(X, Y) = \begin{cases} \frac{v}{(1-u)^2 \sqrt{1+v^2}} & \text{se } u < 1 \\ 0 & \text{caso contrário,} \end{cases}$$

em que

$u = \frac{X^2+Y^2}{R_1^2}$ $v = \frac{S^2}{R_0}$ $t = \angle(y, x) + R_2 \frac{R_0}{r}$ $S = \sqrt{R_0^2 + X^2 + Y^2}$. Os parâmetros são $R_0 = 0,05$, $R_1 = 0,45$ e $R_2 = 15$. Esta função é identicamente nula perto da borda de \mathbb{D} , e tem descontinuidade nas derivadas de segunda ordem no círculo de raio R_1 .

A Tabela 1 mostra o resultado obtido com o algoritmo de aproximação adaptativa *Reduz*, usando as funções objetivo dadas acima e pré-bases uniformes baseadas nas funções-mãe Φ^G (Gaussiana), $\Phi_{3,2}^P$, $\Phi_{5,4}^P$ (pulsos de Wendland), Φ_2^S , Φ_3^S , Φ_4^S e Φ_5^S (pulsos B-spline).

Tabela 1: Tamanho n da base adaptativa e o erro final $\|e\|_\infty$ obtida para cada uma das funções de teste, com cada função-mãe para $\epsilon_{\max} = 2,5 \times 10^{-3}$.

Base	f_O		f_C		f_F		f_G	
	n	$\ e\ _\infty$	n	$\ e\ _\infty$	n	$\ e\ _\infty$	n	$\ e\ _\infty$
Φ^G	208	$1,2 \times 10^{-4}$	1600	$1,8 \times 10^{-3}$	1499	$1,7 \times 10^{-3}$	1572	$1,3 \times 10^{-3}$
$\Phi_{3,2}^P$	216	$6,5 \times 10^{-4}$	1260	$2,1 \times 10^{-3}$	332	$2,4 \times 10^{-3}$	982	$1,9 \times 10^{-3}$
$\Phi_{5,4}^P$	204	$3,3 \times 10^{-4}$	1492	$1,9 \times 10^{-3}$	442	$2,0 \times 10^{-3}$	1040	$1,5 \times 10^{-3}$
Φ_2^S	256	$1,0 \times 10^{-3}$	868	$2,3 \times 10^{-3}$	357	$2,7 \times 10^{-3}$	896	$2,3 \times 10^{-3}$
Φ_3^S	250	$4,2 \times 10^{-4}$	976	$2,0 \times 10^{-3}$	354	$2,3 \times 10^{-3}$	870	$1,8 \times 10^{-3}$
Φ_4^S	240	$3,8 \times 10^{-4}$	1060	$1,9 \times 10^{-3}$	350	$2,1 \times 10^{-3}$	858	$1,5 \times 10^{-3}$
Φ_5^S	238	$2,2 \times 10^{-4}$	1188	$1,9 \times 10^{-3}$	391	$1,9 \times 10^{-3}$	850	$1,4 \times 10^{-3}$

Referências

- [1] M. D. BUHMANN, *Radial basis functions*. Acta Numerica 9, 1-38.
- [2] R. FRANKE, *Scattered data interpolation: Tests of some methods*. Mathematics of Computation, 38(157), 181-200, 1982.
- [3] A. ISKE, *Multiresolution Methods in Scattered Data Modelling*. Springer 2000.
- [4] G. SOUZA, *Aproximação de Funções Irregularmente Amostradas com Bases Hierárquicas Adaptativas de Elementos Tensoriais Compactos*. Master's thesis, IMECC – Unicamp, Campinas, São Paulo, Outubro 2013.