Splines por convolução

Praciano-Pereira, T.

Departamento de Matemática
Universidade Estadual Vale do Acaraú
10 de dezembro de 2007
tarcisio@member.ams.org
pré-prints do Curso de Matemática de Sobral
no. 2007.9
Editor Tarcisio Praciano-Pereira
tarcisio@member.ams.org

Resumo

A convolução é uma ferramenta teoricamente muito poderosa que ficou na geladeira até a década de 80 quando foi aquecida pelas possibilidades formais da computação.

Era conhecida até a década de 70 como método de regularização de funções.

O que é particularmente interessante é a observação de que "medida de Dirac", muito conhecida como "função de Dirac" é a identidade relativamente ao "produto por convolução".

Aqui vou usar convolução para construir splines com uma sucessão de pequenos resultados intermediários cujo resultado final será exibir a estrutura de espaço vetorial de dimensão finita que tem os n-splines a suporte compacto e inclusive exibir uma base para este tipo de espaço.

1 A convolução

Neste artigo vou apresentar uma construção relativamente simples de splines, *splines por convolução*. O resultado final do artigo é que podemos descrever de uma forma muito simples e efetiva um certo tipo de espaço de splines, os n-splines a suporte compacto: é um espaço vetorial de dimensão finita.

Iremos mais longe, vamos explicitar uma base para estes espaços, isto será feito na última seção do artigo.

Nesta primeira seção serão apresentados alguns exemplos simples para motivar o leitor a continuar lendo até o final do artigo. Um leitor experiente com convolução pode saltar para a últma seção do artigo diretamente para o resultado final. Vou também apresentar todas as definições necessárias nesta primeira seção.

A ferramenta central deste artigo é o produto de convolução, que existe há muito tempo mas que somente veio a se tornar interessante depois que os meios computacionais se tornaram efetivos. Vou passar a me referir usando apenas a palavra *convolução* em vez de usar a terminologia primitiva *produto de convolução*. Vou começar a apresentar a convolução na sua forma mais geral como sugestão para quem quiser pensar em discussões mais amplas. A convolução é o instrumento para simplificar a expressão das transformadas de Fourier o que permite definir esta transformada dentro da estrutura algébrica que generaliza a reta real, um grupo aditivo, repetindo de forma mais abstrata e geral o que se faz em Cálculo Avançado com esta transformada.

Definição 1 *Dadas duas funções integráveis, f, g podemos definir*

$$f * g(x) = \int_G f(t)g(x-t)d\mu(t) \tag{1}$$

em que G é um grupo aditivo no qual esteja definida um medida.

Por exemplo, $G = \mathbf{R}$, o grupo aditivo dos números reais com a medida de Lebesgue (completação da integral de Riemann), que estamos usando nesta apresentação (observe que não existe nenhuma medida de Riemann...). É a medida de Lebesgue que se usa no Cálculo para *calcular integrais no sentido de Riemann* embora de uma forma restrita (a integral no sentido de Riemann) o que faz com que certas funções não tenham integral. Aqui vou deixar os grupos genéricos e vou me manter no grupo aditivo dos números reais.

Neste caso a expressão acima fica:

$$f * g(x) = \int_{-\infty}^{\infty} f(t)g(x-t)d(t) \tag{2}$$

e obviamente nem sempre está definida. Mas se f, g forem a suporte compacto $f * g$ está bem definida e será de classe de continuidade uma unidade superior a das funções f, g porque, simplesmente, tem derivada contínua bastando que

f, g sejam integráveis. Isto pode ser facilmente provado usando os cálculos que vou fazer mais a frente para demonstrar que a convolução de duas funções características é uma função diferenciável.

Podemos entender mais rapidamente o significado da convolução vendo um gráfico na figura (1) página 2, o efeito da convolução da função característica do

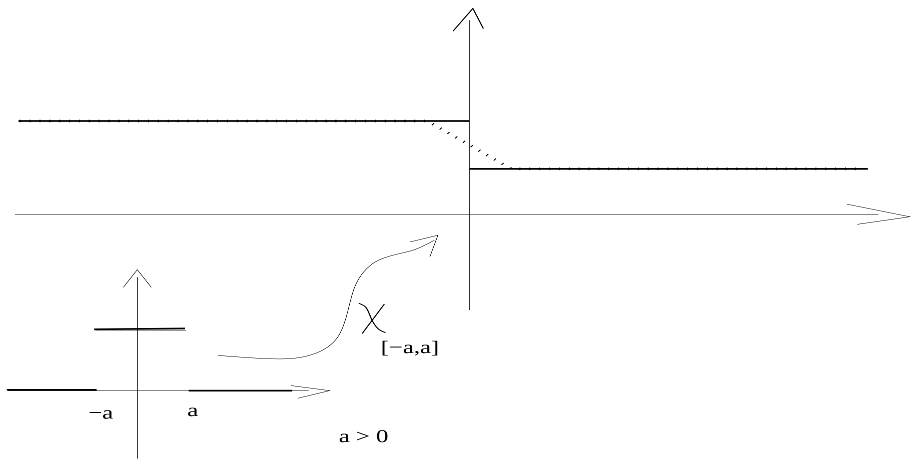

Figura 1: Convolução com a função característica

intervalo $[-a, a]$ com $a = \frac{1}{2}$ com uma função descontínua. Em volta do ponto de descontinuidade a nova função $f * \chi_{[-a,a]}$ corrige a discontinuidade de f com uma interpolação linear. Se repetirmos o produto de convolução, $f * \chi_{[-a,a]} * \chi_{[-a,a]}$ o resultado será o que podemos ver na figura (2) página 2, O gráfico na figura

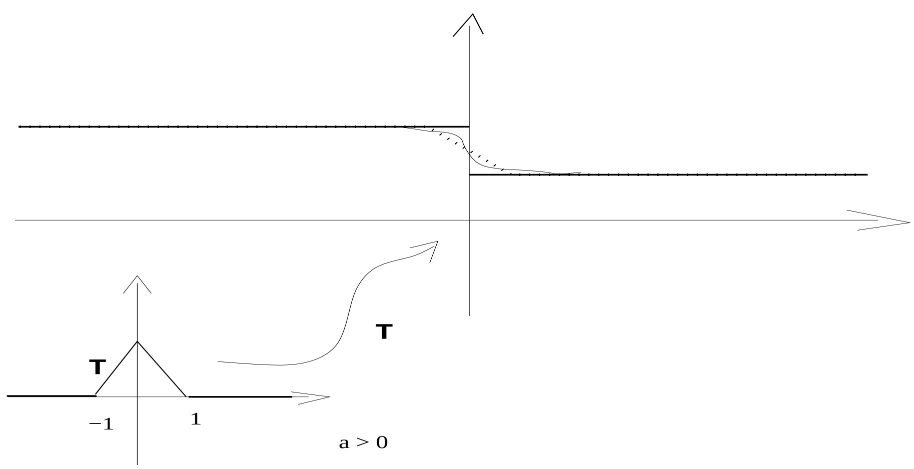

Figura 2: Segunda potência de convolução

(2) foi feito à mão usando `xfig`. Na figura (3) página 3, podemos ver o gráfico

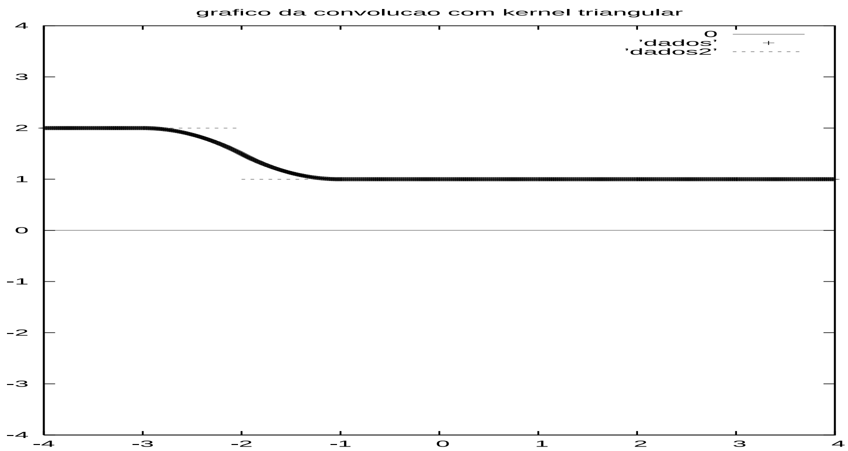

Figura 3: Convolução com kernel triangular T_1

feito com `gnuplot` a partir de um programa escrito em `calc` da convolução da função f com um núcleo triangular T_1.

Definição 2 *O núcleo* T_1

```
Se (x <= -1) T1(x) = 0;
senão se (x <= 0) T1(x) = 1 + x;
senão se (x <= 1) T1(x) = 1 - x;
senão T1(x) = 0;
```

f tem uma discontinuidade de primeira espécie, um salto finito, no ponto $x = -2$, e a convolução "corrige" esta discontinuidade com um "remendo diferenciável" em torno do ponto de discontinuidade. A região de correção depende do suporte do núcleo.

Na figura (4) página 4, temos a convolução com o núcleo T_2 definido por

Definição 3 *O núcleo* T_2

$T_2(x) = 2T_1(2x)$

usando uma dilação com coeficiente 2 que reduz o suporte ao intervalo $[-0.5, 0.5]$ mantendo a integral 1.

Como o suporte tem amplitude (medida) menor, a região de correção é intervalo $[-2.5, -1, 5]$ que tem medida 1. Vemos que o suporte do núcleo é responsável pelo aumento da precisão da correção de forma inversamente proporcional à sua medida: quanto menor a medida do suporte, mais precisa fica a correção.

Podemos generalizar um pouco estas idéias analisando como seria o produto de convolução de duas funções características, ou melhor as funções características de dois intervalos diferentes.

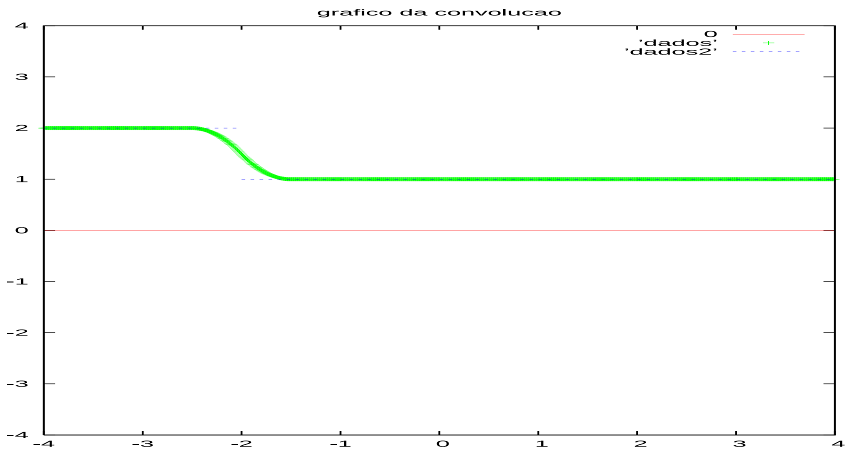

Figura 4: Convolução com o núcleo T_2

Vamos ver que a resposta é que estes gráficos são trapésios, que são a generalização de triângulos. São funções cujos gráficos são trapésios, a suporte compacto como na figura (5) página 4, em que fiz manualmente o gráfico de

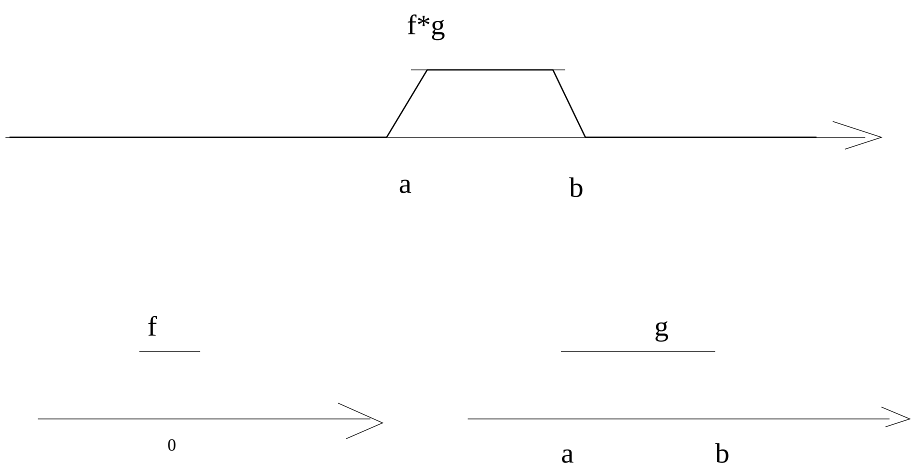

Figura 5: 1-splines por convolução

$f * g$ em que f, g são as duas funções características que aparecem na parte inferior do gráfico. As funções f, g são zero-splines. Logo vou escrever a definição de splines incluindo a definição de *zero-splines* que serão as funções características de intervalos. A convolução delas, das funções características de intervalos, serão 1-splines, uma função polinomial de grau 1 (os pedaços são de grau menor ou igual a 1) e a classe de continuidade é zero, quer dizer é contínua mas a derivada não necessariamente vai sê-lo.

O gráfico do produto de convolução das duas funções características acima, feito com `calc` e `gnuplot` pode ser visto na figura (6) página 5,

Como é um "cálculo numérico", as duas funções f, g tiveram que ser expli-

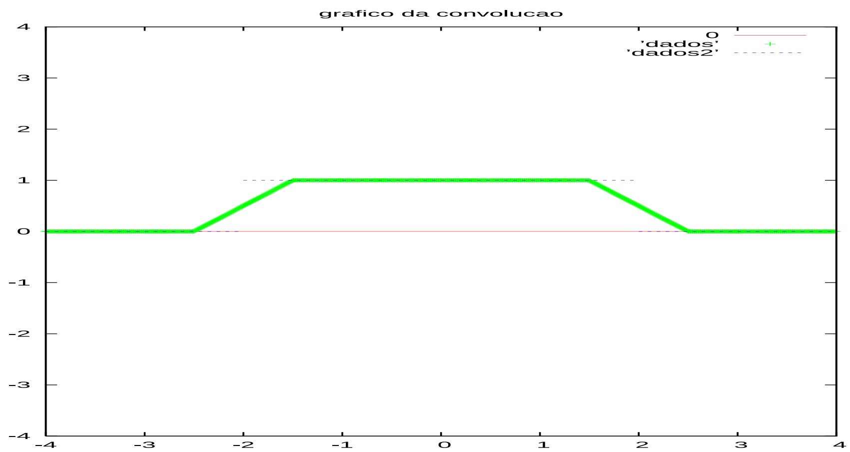

Figura 6: Convolução de duas funções características

citamente definidas,

$$f = \chi_{[-0.5,0.5]}, g = \chi_{[-2,2]}.$$

Os programas usados neste artigo são `convol.calc` e `convol02.calc` que podem ser encontrados em [2, programas]. Os programas mencionados estão escritos em `calc` que é uma linguagem interpretada e você poderá facilmente alterar os programas para fazer outras experiências.

Vamos designar a convolução por

$$h(x) = f * g(x) = \int_{-\infty}^{\infty} f(t)g(x-t)dt$$

Como são duas funções a suporte compacto esta integral pode ser escrita como

$$h(x) = f * g(x) = \int_{-0.5}^{0.5} f(t)g(x-t)dt$$

devido ao suporte de f e como no intervalo $[-0.5, 0.5]$ a função f é a constante 1 podemos escrever

$$h(x) = f * g(x) = \int_{-0.5}^{0.5} g(x-t)dt$$

Foi esta expressão que usamos para escrever os programas `convol.calc, convol02.calc`.

2 Splines

Mostramos nos exemplos que splines são funções polinomiais por pedaços que tem uma certa classe de diferenciabilidade.

Definição 4 *n-Splines*
Um n-splines é uma função polinomial por pedaços, cujos pedaços são polinômios de grau menor ou igual a n e cuja classe de diferenciabilidade é $n-1$. Quer dizer que um n-splines pode ser derivado, contínuamente, $n-1$ vezes.
Um 0-splines é uma combinação linear de funções características de intervalos da reta e pode não ser contínuo.

Esta definição não é fácil de ser usada, constuir um n-splines seguindo este algoritmo é muito trabalhoso. Também não é possível substituir o trabalho difícil por trabalho fácil. Se quisermos construir outro algoritmo precisamos de mais Matemática para obter alguma coisa que seja, por exemplo, computacional e que nos libere do trabalho exaustivo. É isto que vou fazer agora, definindo *splines por convolução.*

2.1 Splines por convolução

Eu incluí entre os splines as funções características como 0-splines, elas são formadas de polinômios de grau zero, funções constantes e não necessariamente contínuas.

Não posso dizer que sejam de classe de diferenciabilidade $0-1$ que não teria sentido. Ou seja 0-splines são polinômios por pedaços de grau zero e eventualmente não contínuos.

Ao incluir as funções características como 0-splines criamos uma facilidade extra na construção: qualquer potência por convolução de uma função característica por outra seria um (n-1)-splines, em que n é a potência.

Vale para uma função característica que é a primeira potência por convolução dela mesma, 0-splines.

A convolução de uma função característica por ela mesma seria a segunda potência por convolução, $n=2$ e teremos um 1-splines. Temos que provar que é um polinômio por pedaços e estes pedaços são de grau menor ou igual a 1. Depois temos que mostrar que é diferenciável uma vez continuamente.

Vamos quebrar este problema em peças mais simples. Vou mostrar que quando $f=\chi_{[0,a]}$ a segunda potência por convolução é um 1-splines, em que $\underline{a}$ é por "convenção" positivo porque ninguém escreve o intervalo $[0,-3]$ mas sim $[-3,0]$.

Por definição, se $f=\chi_{[0,a]}$ então

$$f*f(x)=\int_{-\infty}^{\infty} f(t)f(x-t)dt \tag{3}$$

$$h(x)=f*f(x)=\int_{0}^{a} f(x-t)dt=-\int_{x}^{x-a} f(y)dy=\int_{x-a}^{x} f(y)dy \tag{4}$$

$$\text{tomando } r \in [x-a,x] \; ; \; h(x) = \int_r^x f(y)dy - \int_r^{x-a} f(y)dy \tag{5}$$

$$h'(x) = f(x) - f(x-a) \tag{6}$$

$$\tag{7}$$

Em que na equação (4) usamos o suporte de f e que f é identicamente 1 em cima do suporte e na equação (6) estamos usando o Teorema Fundamental do Cálculo. Na equação (6) temos a diferença de f com a translatada de f por $\underline{a}$.

Se $x < 0$ a equação (6) é zero.

Se $x \in [0,a]$ a equação (6) é 1, o valor de $f(x)$.

Se $x \in [a,2a]$ a equação (6) é -1, o valor de $-f(x-a)$.

Se $x > 2a$ a equação (6) é 0.

Mostramos que $f*f$ é diferenciável (com derivada descontínua) portanto de classe $\mathcal{C}^0$ (é contínua porque tem derivada em todos os pontos exceto em $0, a, 2a$ e podemos obter a equação f usando o Teorema Fundamental do Cálculo).

$$f*f(x) = \begin{cases} \text{se } x \leq 0 & 0 \\ \text{se } x \in [0,a] & x \\ \text{se } x \in [a,2a] & 2a-x \\ \text{se } x \geq 2a & 0 \end{cases} \tag{8}$$

é uma função triângulo de suporte $[0,2a]$ com máximo no ponto $x=a$ portanto uma função polinomial por pedaços, de grau 1 e contínua, um 1-splines.

O caso geral $f = \chi_{[a,b]}$ é uma translação de $f = \chi_{[0,b-a]}$ e deixamos que o leitor faça as contas. Demonstramos assim o teorema

$\boxed{\textbf{Teorema}}$ **1** *Segunda potência de convolução*

A segunda potência de convolução de uma função característica é um 1-splines.

Vamos admitir (hipótese de indução) que isto seja verdadeiro para uma potência n qualquer. Então f, uma função característica do intervalo $[0,a]$ a sua enésima potência por convolução é um (n-1)-splines, com suporte $[0,na]$.

Vamos chamar esta função de g, para simplificar a notação e vamos calcular mais uma convolução com a função f. Um teorema que é fácil de provar usando mudança de variáveis na integral é o lema que vamos precisar agora:

$\boxed{\textit{Lema}}$ **1** *A convolução é comutativa*

Isto nos permite calcular $f*g$ ou $g*f$ conforme seja mais conveniente, neste caso vamos calcular $f*g$

$$h(x) = f*g(x) = \int_{-\infty}^{\infty} f(t)g(x-t)dt = \int_0^a f(t)g(x-t)dt = \tag{9}$$

$$\int_0^a g(x-t)dt = -\int_x^{x-a} g(y)dt = \int_{x-a}^x g(y)dt \tag{10}$$

$$h(x) = \int_r^x g(y)dt - \int_r^{x-a} g(y)dt \tag{11}$$
$$h'(x) = g(x) - g(x-a) \tag{12}$$

valendo as mesmas observações que já fizemos acima. Então h' é uma soma de duas (n-1)-splines portanto qualquer primitiva sua será um n-splines.

Como h' é uma soma de g como uma translação de g então é identicamente nula fora do intervalo $[0,(n+1)a]$. Se re-escrevermos a convolução

$$h(x) = g * f(x) = \int_{-\infty}^{\infty} g(t)f(x-t)dt = \int_0^{na} g(t)f(x-t)dt = \tag{13}$$

vemos que h é nula quando $x \leq 0$ porque é a integral de uma translação por x da função g pela função $t \mapsto f(-t)$ cujo suporte é $[-a,0]$ enquanto que o suporte de g é $[0,na]$.

Qualquer translação de $t \mapsto f(-t)$ por x positivo até $x = na+a$ faz com que os suportes dos termos no integrando tenham uma interseção não vazia e logo a integral é diferente de zero. Mas quando $x > na + a = (n+1)a$ os suportes vão ficar disjuntos e a integral se anula. O suportte de h é $[0,(n+1)a]$ que é a soma dos suportes de f e de g.

Este fato pode ser deduzido em forma geral da expressão da convolução quando se supõe que os fatores por convolução tem suportes compactos.

O caso geral, em que $f = \chi_{[a,b]}$, como no caso inicial, se deduz da análise de que é uma translação de $g = \chi_{[0,b-a]}$.

Demonstramos assim o teorema

Teorema 2 *A $n+1$ potência por convolução de uma função característica é um n-splines a suporte compacto.*

3 O espaço vetorial de dimensão finita Spl-n([a,b])

Estou no ponto para discutir a estrutura geral que têm os n-splines a suporte compacto. A ferramenta básica foi construída na seção anterior que é a potência por convolução de uma função característica que agora vou generalizar para falar de multiplicação por convolução em vez de potência por convolução.

Existe um tipo de função característica que tem efeito importante de aproximação por convolução: as funções características de intervalos de medida 1. Como consequência a integral destas funções características é também 1. As demonstrações que fizemos na seção anterior mostram que uma generalidade muito grande pode ser obtida a partir de certos casos particulares que resta saber escolher. Vou me fixar em funções características equilibradas em volta de zero: quando zero é o ponto médio do intervalo. E como quero que a integral seja 1, tem uma única: $f = \chi_{[-0.5,0.5]}$

As funções positivas cuja integral é 1 recebem o nome de núcleo, então $f = \chi_{[-0.5,0.5]}$ é um núcleo, assim como qualquer de suas translações. Vou me restringir aos núcleos a suporte compacto, quer dizer, neste artigo eu vou adotar a definição

Definição 5 *Núcleo*
Núcleo é uma função positiva, a suporte compacto, cuja integral é 1.

O quadrado por convolução de $f = \chi_{[-0.5,0.5]}$ também tem integral 1. Podemos provar um resultado mais geral:

Teorema **3** *Produto por convolução de núcleos*
Se f for um núcleo, qualquer potência por convolução de f é também um núcleo. **Dem**:
Vou inicialmente supor que o núcleo f tem suporte no intervalo $[-a, a]$. Consequentemente a função $t \mapsto f(-t)$ tem também o seu suporte em $[-a, a]$.

Lema *2* *O suporte de $f * f$ está contido em $[-2a, 2a]$* **Dem**:

$$f * f(x) = \int_{-\infty}^{\infty} f(t)f(x-t)dt = \int_{-a}^{a} f(t)f(x-t)dt = \tag{14}$$

$$se\ x < -a \quad \Rightarrow \quad f * f(x) = 0 \tag{15}$$

$$se\ x > a \quad \Rightarrow \quad f * f(x) = 0 \tag{16}$$

Nas equações (15) e (16), estou usando o fato que uma translação por um valor em módulo superior a $\underline{a}$ faz com que os termos no produto do integrando fiquem com suportes disjuntos *zerando a integral.*

q.e.d .

Vamos ver o valor da integral de $f * f$.

$$I = \int_{-a}^{a} f * f(x)dx = \int_{-a}^{a}\int_{-a}^{a} f(t)f(x-t)dtdx \tag{17}$$

$$I = \int_{-a}^{a} f(t)dt \int_{-a}^{a} f(x-t)dx \tag{18}$$

$$I = \int_{-a}^{a} f(x-t)dx = 1 \tag{19}$$

A equação (18) se justifica porque no integrando temos um produto de termos, uma expressão separável. A passagem da equação (18) para a equação (19) se justifica por f é um núcleo, tem integral 1. Na equação (19) temos uma translação de f o que não altera o valor da integral, (nestes casos usamos a medida de Lebesgue, que é invariante por translação, que está por trás da "integração no sentido de Riemann").
O final da demonstração do teorema segue por indução finita sobre a potência.

q.e.d .

Portanto, se f, g forem núcleos então $f * g$ é também um núcleo. Já demonstramos que para qualquer potência por convolução se f for um núcleo-0-splines então a n-ésima potência por convolução de f é um n-1-splines positivo, a suporte compacto com integral 1. Demonstrei

Teorema **4** *Se f for um 0-splines-núcleo, então $f * f$ é um 1-splines-núcleo a suporte compacto e a n-ésima potência por convolução de f é um n-1-splines núcleo a suporte compacto.*

Considere agora o intervalo $I = [1, n]$ em que $\underline{n}$ é um número inteiro positivo. Considere sobre I a família das funções características dos intervalos $[k, k+1]$, portanto 0-splines, na verdade $n-1$ tranlações do 0-splines $\chi_{[0,1]}$. Ao final você verá que esta observação é importante.

Considere agora um núcleo $n-1$-splines, η com suporte em $[-0.5, 0.5]$.

Pelo teorema (3) o produto de η pelos elementos desta família produz uma família de n-splines a suporte compacto que vamos designar por $\Phi = (\phi_k)_{k=1}^{n-1}$.

$\boxed{\textit{Lema}}$ **3** *Todos os elementos da família Φ são translações de um único n-splines a suporte compacto.* $\boxed{\textbf{Dem}}$*:*

Basta escrever a expressão do elemento genérico da família das funções características multiplicado por η

$$\chi_0 = \chi_{[0,1]} \tag{20}$$

$$\chi_{[k,k+1]}(x) = \chi_0(x-k) \tag{21}$$

$$\phi_k = \eta * \chi_{[k,k+1]} \tag{22}$$

$$\eta * \chi_{[k,k+1]}(x) = \int_{-\infty}^{\infty} \eta(t)\chi_{[k,k+1]}(x-t))dt \tag{23}$$

$$\eta * \chi_{[k,k+1]}(x) = \int_{-\infty}^{\infty} \eta(t)\chi_0(x-k-t))dt \tag{24}$$

$$\eta * \chi_{[k,k+1]}(x) = \int_{-\infty}^{\infty} \eta(t)\chi_k(x-t))dt \tag{25}$$

em que η_k é a k-ésima translação do núcleo η. $\boxed{\textbf{q.e.d .}}$

A família dos 0-splines que usamos de partida não tem a propriedade

$$\sum_{k=1}^{n-1} \chi_k = 1$$

em que a igualdade significa "identicamente". Porque nos extremos dos intervalos a soma é igual a 2 portanto a soma destas funções características é uma função descontínua. Ao multiplicarmos por convolução com η o resultado é uma função contínua portanto a soma

$$\sum_{k=1}^{n-1} \phi_k = 1$$

caracterizando que esta família é uma *partição da unidade* sobre o intervalo $[1, n]$, uma partição da unidade de classe $\mathcal{C}^{n-1}$.

Como são translações são também linearmente independentes e portanto a base de um espaço vetorial que vou chamar $Spl - n([1, n])$.

É um espaço vetorial de dimensão $n-1$ de n-splines a suporte compacto contido no intervalo $[1, n]$ porque somas de polinômios de grau n é um polinômio de grau menor ou igual a n.

Vou agora impor uma restrição sobre o núcleo arbitrário que usei para regularizar a família de 0-splines, estabelecendo que ele seja símetrico em torno de zero e atingindo o ponto máximo em zero. Esta propriedade se espalha para todos os membros da família ϕ_k que terão seus pontos de máximo no ponto central do intervalo $[k, k+1]$. Isto se demontra calculando a derivada e usando a propriedade (fácil de provar)

$$(f * g)' = f' * g = f * g'.$$

Dado um elemento qualquer de $Spl - n([1, n])$, ele é a combinação linear

$$P \in Spl - n([1, n]) \ \ \Rightarrow \ \ P = \sum_{k=1}^{n-1} a_k \phi_k$$

com $P(x_k) = a_k$ em que x_k é ponto médio do intervalo $[k, k+1]$. Portanto o n-splines P interpola os pontos

$$(x_k, a_k)_{k=1}^{n=1}$$

Não há unicidade nesta *solução* como é visível que tudo depende do núcleo que foi usado para regularizar a família.

Se considerarmos uma intervalo $[a, b]$ qualquer, podemos transferir este processo de interpolação de $[1, n]$ para $[a, b]$ com translações e dilações. Neste contexto eu chamo os elementos da família ϕ_k de sinais e vamos obter com translação e dilação uma família de sinais sobre $[a, b]$ que é uma partição da unidade gerando um espaço vetorial $Spl - n([a, b])$, de dimensão $n - 1$ cujos elementos interpolam os pontos

$$(x_k, y_k)$$

em que x_k são os pontos médios dos sub-intervalos da partição uniforme de $[a, b]$ com $n - 1$ sub-intervalos, imagem de $[1, n]$.

O fator de dilação é $\Delta x = \frac{b-a}{n-1}$ e a expressão dos sinais sobre o intervalo $[a, b]$ é

$$\rho_k = \phi_k(\frac{x - a}{\Delta x})$$

Demonstramos assim o teorema

Teorema **5** *Interpolação de pontos numa faixa*

Dada uma faixa do plano perpendicular ao intervalo $[a, b]$ e considerada uma partição uniforme deste intervalo, considere a família de pontos desta faixa do plano

$$(x_k, y_k)$$

em que x_k são os pontos médios dos sub-intervalos da partição uniforme de $[a, b]$ com $n - 1$ sub-intervalos.

Existe um elemento de um espaço vetorial de n-splines que interpola estes pontos, sendo este elemento gerado por uma base ρ_k de n-splines a suporte compacto bem definida a partir de um núcleo $n - 1$-splines com suporte em $[-\Delta x, \Delta x]$.

Referências

[1] `gnuplot` *um programa para fazer grafico e alguns cálculos*
http://www.gnuplot.info

[2] Praciano-Pereira, T.
Programas para Cálculo Numérico
http://www.calculo-numerico.sobralmatematica.org