

Aproximação de funções: polinômios de Bernstein

Medeiros, Heloisa B.*

UFF – IME – GMA

medeiros@mat.uff.br

Menezes, M. Lucia †

UFF-IME – GMA

menezes@mat.uff.br

20 de junho de 2011

Quando usamos alguma máquina para esboçar um gráfico ou determinar um valor como $e^{\sqrt{2}}$, não nos ocorre perguntar como são feitos os cálculos ou quão exatos são. Todavia, um sem número de pesquisas vem sendo desenvolvido para que estas informações sejam mais precisas e obtidas com maior rapidez.

O surgimento dos processadores (em meados do século XX) colocou para a Matemática uma série de questões sobre como representar e calcular valores e funções. Em linhas gerais, sabemos que um processador só é capaz de fazer somas algébricas de modo que todos os cálculos, em última análise, devem se remeter a este tipo de operação. Produtos podem ser efetuados utilizando somas e, conseqüentemente, operações como elevar um valor a um número inteiro podem ser executadas. A possibilidade de calcular x^n torna o uso de polinômios uma ferramenta importantíssima em cálculos realizados por máquinas. Por exemplo, para calcular $\sqrt{3}$ pode ser conveniente usar um procedimento padrão (como método de Newton ou bisseção) para resolver $x^2 - 3 = 0$. Nem sempre é possível reduzir o problema ao cálculo da raiz de um polinômio, mas são muitos os usos dos polinômios nos compiladores, máquinas de calcular e softwares em geral. Um recurso utilizado em ampla escala é a aproximação de funções por polinômios. Eleger o método específico a ser usado depende muito das circunstâncias. A aproximação deve ser feita em um único ponto, ou em um intervalo? Qual o erro máximo que queremos? Qual o processador disponível? Que valor ou função deve ser aproximado? A aproximação deve ter sensibilidade suficiente para captar singularidades isoladas?

Dentre os métodos possíveis, para funções contínuas, os polinômios de Bernstein se destacam por oferecerem uma aproximação uniforme. Como sempre, existe um preço a ser pago: a convergência não é muito rápida, quando comparado a outros métodos de aproximação polinomial. Mesmo assim, são de grande utilidade nos casos em que se necessita aproximar uma função em todo um intervalo como, por exemplo, no esboço de um gráfico. Vale, ainda, notar que os polinômios de Bernstein fornecem uma belíssima demonstração construtiva do Teorema de aproximação de Weierstrass.

Em todo este texto, $f(x)$ será uma função contínua no intervalo $[0, 1]$.

Para definir os polinômios de Bernstein, lembramos da fórmula binomial:

$$(a + b)^n = \sum_{j=0}^n \binom{n}{j} a^j b^{n-j} \quad (1)$$

Escolhemos agora $a = x$, $b = 1 - x$ e definimos $\beta_{nj}(x) := \binom{n}{j} x^j (1 - x)^{n-j}$, $j = 0, 1, \dots, n$. Dividimos o intervalo $[0, 1]$ em n subintervalos, de igual tamanho, $[x_j, x_{j+1}]$, $j = 0, \dots, n - 1$, de modo que $x_j = \frac{j}{n}$. Avaliamos $f(x_j)$ em cada ponto e, com estas constantes, definimos o polinômio de Bernstein

*Rua Mário Santos Braga s/n Valonguinho – 24020-140 Niteroi Rio de Janeiro -RJ

†idem

de grau n da função $f(x)$ como:

$$B_n(f; x) := \sum_{j=0}^n f(x_j) \beta_{nj}(x) \quad (2)$$

Observamos que o conjunto $\{\beta_{nj}(x)\}$, $j = 0, \dots, n$, forma uma base para o espaço vetorial de polinômios de grau menor ou igual à n , e o polinômio de Bernstein é uma combinação linear dos elementos desta base. Para verificar isso, não é difícil ver que cada um dos elementos da base canônica $\{1, x, x^2, \dots, x^n\}$ pode ser escrito como combinação linear dos $\beta_{nj}(x)$. As *curvas de Bézier*, bem familiares aos que usam softwares gráficos, também são formadas como combinação linear de elementos desta base, embora Bézier e Bernstein tenham chegado aos seus resultados de forma independente.

A Figura 1 ilustra algumas características da aproximação obtida pelos polinômios de Bernstein. Nos dois casos, o desenho apresenta o gráfico da função e os gráficos dos polinômios de Bernstein de graus 4, 8, 12 e 16. Primeiro, observamos que, diferentemente de outras aproximações polinomiais, o polinômio de Bernstein de grau n , em geral, não coincide com a função em um número n de pontos; ademais, mesmo quando a função é um polinômio de grau n , o n -ésimo polinômio de Bernstein não é a própria função (como seria, por exemplo, no caso do polinômio de Taylor). Não é difícil verificar que $f(1) = B_n(f; 1)$ e $f(0) = B_n(f; 0)$, o que está ilustrado nos gráficos. Todavia, a propriedade que queremos ressaltar aqui é que a velocidade de convergência será maior, se a variação da função for mais suave. Em ambos os casos, a imagem da função, como conjunto, é essencialmente a mesma e as propriedades gerais da curva também (isto é: é contínua, assume um único ponto de máximo, possui um ponto de inflexão, etc). Mas, no gráfico da esquerda, a variação na vizinhança do ponto de máximo é bem mais brusca. Não é difícil ver que o erro obtido nas aproximações (de mesmo grau) pelos polinômios, na vizinhança do ponto de máximo, é maior no primeiro caso. Este comportamento pode ser entendido observando propriedades das duas funções. Como a teoria está sendo desenvolvida para funções contínuas (classe C^0), grau de diferenciabilidade não é uma hipótese que se queira utilizar, a princípio, para medir “suavidade” de variação.



Figura 1: O gráfico da função é a curva sólida e as outras curvas são os polinômios de Bernstein de grau 4, 8, 12 e 16.

O conceito mais importante, por ora, é o de *módulo de continuidade* de uma função¹. Trata-se, grosso modo, de uma medida do “quão contínua” uma função é. Para fazer esta medida, consideramos subintervalos de $[0, 1]$ com tamanho δ . Em cada um deles, medimos a variação de f e denominamos o maior destes valores de *módulo de continuidade de f em relação à δ* denotando-o por $w(f; \delta)$.

O Teorema 1 fornece uma estimativa da diferença entre f e seu polinômio de Bernstein de grau n . A demonstração do Teorema é bastante técnica e pode ser vista em [4] ou [7].

Teorema 1. *Se $f(x)$ é uma função contínua em $[0, 1]$, para todo $x \in [0, 1]$, $|f(x) - B_n(f; x)| \leq \frac{9}{4} w(f; \frac{1}{\sqrt{n}})$.*

¹Este conceito é muito usado em análise numérica e sua definição mais formal pode ser vista em qualquer texto clássico sobre o assunto como [4] ou [7].

Isto é, a diferença é proporcional à variação máxima de f em subintervalos de $[0, 1]$ com medida $\frac{1}{\sqrt{n}}$. Para f contínua, a variação tende a zero quando n cresce, de modo que B_n tende à f . Assim, a estimativa do Teorema 1 permite tanto mostrar a convergência quanto avaliar o erro da aproximação. Isto é: para n fixo, a diferença entre f e seu polinômio de Bernstein de grau n não ultrapassa $9/4$ do módulo de continuidade $w(f; \frac{1}{\sqrt{n}})$.

Na Figura 1, não é difícil ver que, para um mesmo δ , o módulo de continuidade da função é maior no gráfico da esquerda, onde também o erro é maior. A estimativa do Teorema 1 pode ser melhorada, dependendo das propriedades da função, especialmente se houver algum grau de diferenciabilidade, em que pese não se conseguir estimativas excelentes.

Como estamos falando de aproximações polinomiais que não envolvem interpolações, julgamos conveniente uma rápida comparação com a mais famosa delas (para funções de classe C^k): os polinômios de Taylor. A Figura 2 exhibe o gráfico da função $f(x) = \sin^2(2\pi(x - 1/2))$, de seu polinômio de Taylor de grau 4 em torno de 0,5 e de seu polinômio de Bernstein de grau 4. Do lado esquerdo colocamos o detalhe, restringindo os valores de x ; do lado direito, aparecem os gráficos em todo o intervalo $[0, 1]$. Chama atenção o fato de que a aproximação do polinômio de Taylor, excelente na vizinhança em torno do qual é calculado (0,5), tem um erro muito maior, quando nos afastamos deste ponto. Em geral, pode-se esperar este comportamento, embora existam exceções.

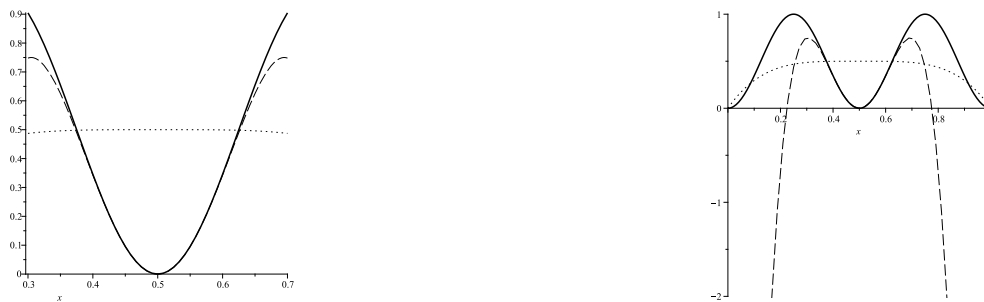


Figura 2: Polinômio de Bernstein (pontilhado) \times polinômio de Taylor (tracejado).

Até agora, dirigimos ao problema um olhar mais próprio da análise numérica: se utilizarmos estes polinômios, qual o tamanho do erro? Como esta aproximação se comporta em relação à outra? Entretanto, nosso esforço para tratar destas perguntas, ainda não indicou qualquer noção do *porquê* os polinômios convergem para f . Este aspecto fica mais claro quando pensamos em teoria das probabilidades, como em [6], e vamos apresentar as linhas gerais do argumento que, por sinal, foi o utilizado pelo próprio Bernstein [1].

Imaginamos um experimento cujo resultado tem apenas as possibilidades A ou B e supomos que A ocorre com probabilidade x , de modo que a probabilidade da ocorrência de B será $(1 - x)$. Em n experimentos, a probabilidade de termos j A 's e $(n - j)$ B 's, em qualquer ordem, é: $\binom{n}{j} x^j (1 - x)^{(n - j)} = \beta_{nj}(x)$. Assim, cada termo do somatório:

$$B_n(f; x) = f(x_0)\beta_{n0}(x) + f(x_1)\beta_{n1}(x) + f(x_2)\beta_{n2}(x) + \cdots + f(x_n)\beta_{nn}(x) \quad (3)$$

é o resultado da multiplicação de $f(x_j)$ pela probabilidade de j ocorrências de A ($\beta_{nj}(x)$). Os termos relevantes no somatório são aqueles em que $(\beta_{nj}(x))$ é grande (ou seja, os que indicam uma probabilidade alta de ocorrência de j A 's.) A pergunta chave é, então, quais valores de j apresentam maior probabilidade? Cabe, assim, analisar como os termos $(\beta_{nj}(x))$ dependem de j . A probabilidade de j ocorrências cresce com j à medida que a diferença $|x_j - x|$ diminui, atingindo um valor máximo quando a diferença é mínima. Por isso, os termos do somatório que realmente importam, são aqueles onde a distância entre x e x_j é pequena. Chamemos de J o índice j tal que $(\beta_{nJ}(x))$ é máximo. Então, os

termos relevantes no somatório 3 são aqueles cujo índice j está próximo de J . Desprezamos os outros termos e aproximamos:

$$B_n(f; x) \approx \sum_{j=J-k}^{J+k} f(x_j) \beta_{nj}(x) \quad (4)$$

para algum valor de k compatível com o número de intervalos da partição (n). Como f é contínua e x está próximo de x_J , para os índices j considerados na aproximação 4, $f(x_j)$ está próximo de $f(x)$, e tornamos a aproximar o somatório por:

$$B_n(f; x) \approx \sum_{j=J-k}^{J+k} f(x) \beta_{nj}(x) = f(x) \sum_{j=J-k}^{J+k} \beta_{nj}(x) \quad (5)$$

Finalmente, notamos que $((x + (1-x))^n = \sum_{j=0}^n \beta_{nj}(x) = 1$. Mais uma vez, desprezamos os valores

de $\beta_{nj}(x)$ se j não for próximo de J para concluir que $\sum_{j=J-k}^{J+k} \beta_{nj}(x) \approx 1$. E, usando este fato em 5,

teremos $B_n(f; x) \approx \sum_{j=J-k}^{J+k} f(x) \beta_{nj}(x) = f(x) \sum_{j=J-k}^{J+k} \beta_{nj}(x) \approx f(x)$. E temos, portanto, uma noção das

ideias subjacentes à construção dos polinômios de Bernstein. Tornar este argumento preciso e formal exige um pouco mais de suor e contas e pode ser visto em [6].

Tanto o Teorema 1 quanto o argumento baseado em teoria das probabilidades permitem provar que os polinômios de Bernstein aproximam uniformemente qualquer função contínua. A existência de uma aproximação polinomial foi mostrada por Weierstrass no final do século XIX, que não a construiu, todavia. Os polinômios de Bernstein possibilitam uma demonstração construtiva dos resultados de Weierstrass.

Aproximações numéricas de funções são um tópico fascinante e são muitos os estudos em desenvolvimento sobre o tema; os polinômios tratados aqui são mais usados no esboço de gráficos. Foram propostos por um matemático ucraniano, Sergei Natanovich Bernstein (falecido em 1968), que contribuiu com diversos resultados importantes para o desenvolvimento da matemática. Intimamente relacionadas aos polinômios de Bernstein são as curvas de Bézier, definidas por um grau n e $(n+1)$

“pontos de controle” P_0, P_1, \dots, P_n , dada por $B_n(t) = \sum_{j=0}^n \binom{n}{j} t^j (1-t)^{n-j} P_j$. Essas curvas foram es-

tudadas por Paul de Casteljaou (físico e matemático da Citroen) que desenvolveu um algoritmo para obtê-las e por Pierre Bézier (um engenheiro e matemático da Renault) que as patenteou e as utilizou para desenhar automóveis, veja [2].

Os gráficos esboçados aqui foram feitos com o software Maple. São inúmeros os trabalhos sobre polinômios de Bernstein e selecionamos alguns dando preferência à facilidade de acesso. No sítio de buscas virtuais de e-books <http://khup.com> é possível encontrar diversos textos em formato pdf sobre os polinômios de Bernstein. Citamos em particular [5] e [6], onde definições e propriedades dos polinômios são bastante explorados. A demonstração do Teorema 1 pode ser vista em [4] ou [7], referências que também apresentam um tratamento clássico e muito bem feito acerca de aproximações numéricas de funções. A demonstração do Teorema de Weierstrass, usando os polinômios de Bernstein, pode ser encontrada em [3] ou em [1].

Referências

- [1] Sergei Bernstein. *Démonstration du théorème du Weierstrass fondé sur le calcul des probabilités*. Reprodução da demonstração original de Bernstein (em francês). Disponível em <http://www.math.technion.ac.il/hat/fpapers/P03.PDF>.
- [2] Bill Casseman. *From Bézier to Bernstein*. Web em 11/2008. Disponível em <http://www.ams.org/samplings/feature-column/fcarc-bezier>.
- [3] Alex Alves Dentamaro e Daniela Mariz Silva Vieira. *Teorema de Aproximação de Weierstrass*. Web em 15/11/2010. Disponível em <http://www.prp.unicamp.br/pibic/congressos/xvicongresso/paineis/041705.pdf>.
- [4] Eugene Isaacson and Herbert B. Keller. *Analysis Of Numerical Methods*. John Wiley & Sons, 1994.
- [5] Kenneth I. Joy. *Bernstein Polynomials*. Web em 10/12/2010. Disponível em http://khup.com/view/1_keyword-bernstein-polynomial/bernstein-polynomials.html.
On-Line Geometric Modeling Notes. Visualization and Graphics Research Group, Department of Computer Science. University of California, Davis.
- [6] George M. Phillips. *Interpolation and Approximation by Polynomials*. Springer, 1st edition, 2003. Capítulo sobre polinômios de Bernstein. Disponível em http://khup.com/view/4_keyword-bernstein-polynomial/bernstein-polynomials.html em dez 2010.
- [7] Theodore J. Rivlin. *An Introduction to the Approximation of Functions*. Dover Publications, 2003.