Entendendo cálculo variacional

Tarcisio Praciano-Pereira
outros receiosos de se identificarem

20 de fevereiro de 2009- 4ªEdição

Resumo

Neste artigo estou estudando o chamado Cálculo Variacional e como esta demonominação somente tem sentido para os envolvidos, começo discutindo o que significa este "Cálculo".

Aos leitores deste artigo, primeiro que tudo, não se exasperem com a aparente dificuldade do mesmo, é trabalho em andamento e portanto o texto ainda está muito verde. Aguardo até mesmo colaboradores para que ele melhore. O trabalho está se desenvolvendo em colaboração com Diego Frota, mas os erros que houverem por aqui, por enquanto, são meus.

O objetivo é a solução aproximada de equações diferenciais e vou começar com a equações ordinárias, apesar de que neste caso, das *equações diferenciais ordinárias*, a forma variacional (ou fraca) e a sua forma forte, são equivalentes porque não há soluções fracas que não sejam também soluções fortes portanto *não vamos descobrir novas soluções* como acontece com as *equações diferenciais parciais*. Ainda assim tem sentido discutir esta metodologia para as equações diferenciais ordinárias como um método de aproximação e @ leitor logo vai ver a sua aplicabilidade na segunda seção.

Na quarta seção apresento um exemplo usando uma equação diferencial linear, depois vou mostrar que este exemplo pode ser generalizado para finalmente chegar a expressão da forma variacional de uma equação diferencial ordinária, apresentando sua definição.

Finalmente este resumo está velho! eu não estou com tempo para atualizar o resumo.

* UeVA - Matemática - tarcisio@member.ams.org

1 A forma variacional

Encontrar soluções aproximadas é a razão deste trabalho, porque a forma variacional (ou fraca) se presta mais facilmente para a construção de soluções aproximadas, como a própria apresentação vai mostrar, quer dizer, que a forma variacional se transforma num método para encontrar soluções aproximadas de equações diferenciais ordinárias. É ainda mais interessante apresentar este método porque, quando se passa ao estudo das equações diferenciais parciais, não somente há soluções fracas autênticas, como também o método que vou apresentar aqui é muito útil na busca de soluções aproximadas das EDP, enfim este é o objetivo, envolvendo uma técnica denominado *elementos finitos*.

O objetivo do Cálculo Variacional são os funcionais lineares, em vez de fazer Cálculo Diferencial e Integral em funções reais de variáveis reais, êle tem por objetivo derivar e integrar funcionais, usando outra linguagem, operamos no espaço dual.

A formulação *variacional*, ainda chamada *fraca*, de um problema se constitue de apresentá-lo sob a forma de um funcional que se aplica a um conjunto de funções (ou curvas). O exemplo que tradicionalmente aparece nos livros introdutórios é o problema da braquistocrona (no mínimo tempo), *qual é a curva, caminho, para o mais rápido acesso entre dois pontos.* Se não houver restrições, é um segmento de reta, infelizmente o meio a ser percorrido em geral oferece restrições portanto a curva pode não ser um segmento de reta. Esta simples forma de apresentar o problema nos diz que o operador "caminho mínimo entre dois pontos" atinge seu mínimo numa família de curvas e portanto a sua formulação seria algo do tipo

$$\|\Gamma(\omega)\| \leq C \tag{1}$$

querendo dizer que estamos procurando as curvas ω que fazem com o operador Γ, que calcula o comprimento de uma curva γ, atinja o mínimo, naturalmente, neste caso $\underline{C}$ deve ser o comprimento do segmento de reta que une dois pontos, queremos procurar uma família[1] de curvas que minizam o tempo gasto entre dois pontos.

Uma situação geométrica comum é aquela em que os dois pontos estejam numa superficie que não seja plana, pode haver várias curvas que ofereçam o menor caminho e possivelmente um segmento de reta seja um caminho impossível.

Esta situação é comum e representa a situação generalizada que o Cálculo Diferencial e Integral sobre funções reais de uma variável real, não consegue mostrar uma vez que, neste contexto restrito, é muito comum que o ponto de extremo seja único, entretanto os funcionais atingem o seu máximo em famílias de curvas[2] esta observação sozinha já justifica investirmos em uma linguagem mais sofisticada que é o *Cálculo Variacional.*

[1] Que @ leitor não tenha um grande estress para compreender cada fórmula, algumas serão vagas, como a que se encontra na equação (1)

[2] Que podem ser famílias de trajetórias de distribuição de produtos

1.1 Generalizando funções com funcionais lineares

Parece importante conhecer o caso das *distribuições* que vou agora apresentar.

Vou definir um espaço vetorial de funções-teste e mostrar como podemos generalizar funções através de funcionais lineares, com um exemplo. Para isto vou definir um espaço vetorial de funções, uma classe de funções altamente diferenciáveis (quero mesmo dizer de classe $\mathcal{C}^{\infty}(R)$) e que se anulem fora de um intervalo fechado. Esta classe de funções é um espaço vetorial, a soma de duas tais funções é uma função da classe e o produto de uma função desta classe por um número real, ainda é uma função da classe, esta é uma forma resumida (e incompleta) de estabelecer que este espaço de funções é um espaço vetorial. O espaço destas funções é chamado de "espaço de funções teste", algumas vezes, na literatura se usa a expressão "funções admissíveis", e vou, rapidamente, mostrar porque precisamos de tais funções.

Primeiro um gráfico para ajudar na concretização das idéias, a figura (1) página 2, mostra três elementos ϕ, γ, deste espaço. dois tem mesmo suporte,

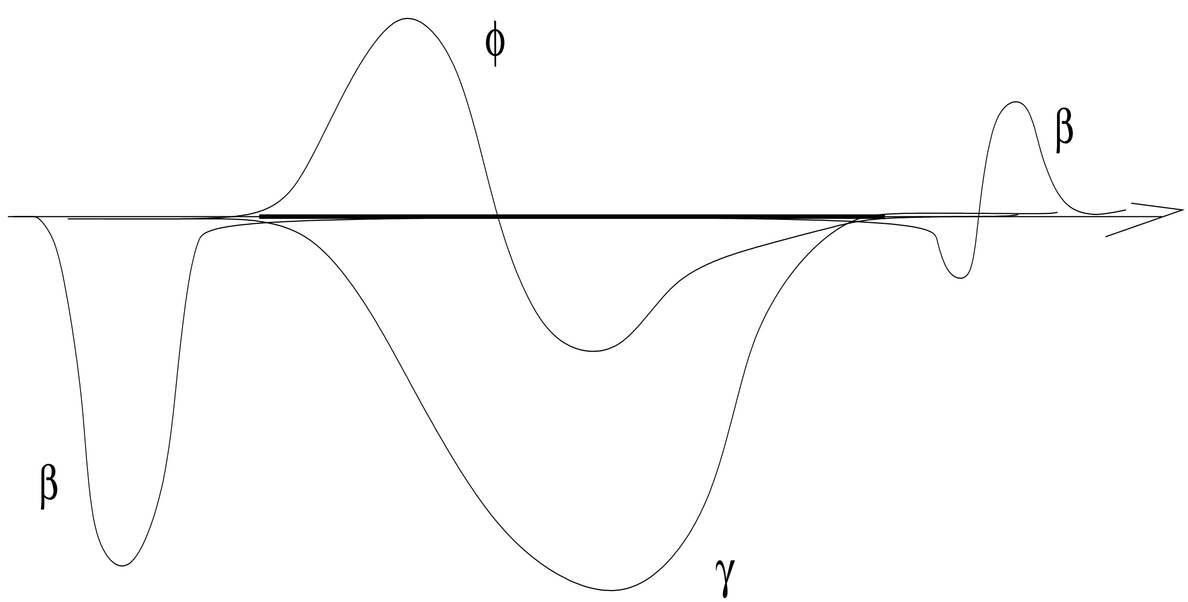

Figura 1: Espaço de funções teste

γ, ϕ e um terceiro elemento, β tem suporte disjunto com os outros dois. Inclui β depois para deixar claro que os elementos não precisam ter todos o mesmo suporte, e qualquer intervalo compacto da reta pode ser suporte de algum elemento de $\mathcal{D}_c(R)$.

Este tipo de funções têm diversos usos e nomes, *alguma vezes são chamadas de sinais, outra vezes são chamadas de janelas, ou ainda, com frequência, pulsos de energia unitária, quando a integral delas for 1 e elas forem positivas*[3]. Este *sistema* de funções foi, e é, muito utilizado em diversas situações,

[3]Este é um caso particularmente importante, estas funções são *pesos* e servem para calcular valores médios.

como ferramentas auxiliares, por exemplo, com os polinômios trigonométricos em comunicações com o objetivo de localizar um sinal, fazer com que o sinal se concentrasse em um intervalo do tempo, porque elas se anulam fora de um intervalo compacto. Como elas se anulam fora de um intervalo, elas são muito interessantes para usar em conexão com integração por partes, e aqui vou repetir um cálculo habitual em livros introdutórios de *teoria das distribuições* para definir uma generalização da derivada.

Alguma notação é importante para ajudar também a fixar as idéias.

Vou designar por $\mathcal{D}_c(R)$ o espaço das funções[4] teste e vou usar as letras $\omega, \phi \in \mathcal{D}_c(R)$ para representar elementos genéricos do espaço de funções teste.

Considere agora uma função real diferenciável f, então

$$\int_{-\infty}^{\infty} f'(x)\phi(x)dx = f(x)\phi(x)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(x)\phi'(x)dx = -\int_{-\infty}^{\infty} f(x)\phi'(x)dx \quad (2)$$

$$\int_{-\infty}^{\infty} f'(x)\phi(x)dx = -\int_{-\infty}^{\infty} f(x)\phi'(x)dx \quad (3)$$

porque os limites $\phi(-\infty) = \phi(\infty) = 0$. Se a função f não for diferenciavel, a expressão à direita na equação (3) serve de definição para a expressão à esquerda, e neste caso, a integral deve ser substituida por um símbolo indicando que o objeto f' foi aplicado à função teste ϕ usando a equação à direita como definição deste objeto:

$$< f', \phi >= -\int_{-\infty}^{\infty} f(x)\phi'(x)dx \quad (4)$$

Posso logo usar este cálculo para mostrar uma forma simples de entender a chamada função "*delta* de Dirac", que não é uma função, e sim um funcional linear. A *delta* de Dirac é a derivada[5], da função de Heaviside[6] definida pelo sistema de equações

$$\begin{cases} x < 0 & H(x) = 0 \\ x \geq 0 & H(x) = 1 \end{cases} \quad (5)$$

Este é um bom exemplo de função que não é diferenciável que podemos colocar na equação. Chamando a função de Heaviside de H, como é habitual, temos,

[4] Justificando o índice "c" na notação, primeiro porque foi assim que Schwarz designou, depois ele o fez para caracterizar que as funções se anulam fora de um intervalo compacto.

[5] Que não existe como função.

[6] Tirado da wikipedia, Oliver Heaviside (18 May 1850 – 3 February 1925) foi um autodidata, engenheiro elétrico, matemático e físico.

aplicando a equação (5)

$$< H', \phi >= - \int_{-\infty}^{\infty} H(x)\phi'(x)dx \tag{6}$$

$$< H', \phi >= - \int_{0}^{\infty} H(x)\phi'(x)dx \tag{7}$$

$$< H', \phi >= - \int_{0}^{\infty} \phi'(x)dx = -\phi(\infty) + \phi(0) = \phi(0) \tag{8}$$

$$< H', \phi >=< \delta_0, \phi >= \phi(0) \tag{9}$$

- Na equação (6) estou usando a definição "generalizada de derivada", que vou explicar melhor em seguida,
- na equação (7) estou usando a definição de H que é nula na semireta negativa o que reduz a integral à semireta positiva,
- na equação (8) estou usando a definição de H que é identicamente 1 na semireta positiva e finalmente estou calculando a integral de ϕ' usando o teorema fundamental do Cálculo Integral em que estou escrevendo a expressão $\phi(\infty) = 0$ em lugar do limite usual.

Dirac identificou como a integral da "função de Dirac" multiplicada pela função ϕ' e definiu a sua "função de Dirac" como uma função que seria nula na reta inteira, tendo um salto para o infinito no ponto zero, mas de tal modo que sua integral fosse 1, ou seja, uma função que teria toda a sua área concentrada num único ponto, no ponto zero, e valendo 1.

Uma função deste tipo, se existisse, serveria para calcular médias (por que tem integral positiva igual a 1, e como é concentrada num ponto, então calcula, exatamente, o valor da função contra quem estiver sendo integrada, naquele ponto de concentração, é o que vemos acima, $\phi(0)$.

Então a derivada da Heaviside é a Dirac, apenas não é uma função, e sim um funcional linear, quer dizer uma integral - as integrais são funcionais lineares, um exemplo de tais funcionais lineares é a integral de Riemann ou melhor, a sua completação[7], a integral de Lebesgue

A Dirac é o funcional linear que corresponde a derivada da Heaviside e podemos deduzir da equação (9)

$$< H', \phi >=< \delta_0, \phi > \tag{10}$$

Observe que o significado da igualdade $H' = \delta_0$ está descrito na equação (10) em que estamos dizendo como calculamos a expressão à esquerda, com a expressão à direita, que foi obtida do sistemas de equações equação (6)-(9)

Este cálculo já serviu para mostrar que estamos fazendo alguma coisa interessante, uma vez que justificamos a famosa "função de Dirac", isto foi feito

[7] É uma forma simples de explicar a integral de Lebesgue, ela completa a integral de Riemann como os números reais completam o conjunto dos números racioanais.

em 1945, por L. Schwarz, e de forma simultânea, pelo matemático português, Sebastião e Silva, este apenas obfuscado pela importância de Schwarz, francês, membro do grupo Bourbaki, professor da chamada *Sorbonne*, enquanto que Sebastião e Silva[8] era apenas professor de Matemática da Escola de Agronômia de Lisboa.

Mostrei, na verdade repeti o que Schwarz e Sebastião e Silva fizeram em 1945, salvando a monstruosidade que Dirac criara no começo do século 20 e levou 40 desafiando os matemáticos com a observação simples "*funciona*", e realmente funciona, com simples exemplos computacionais[9] como vou apresentar aqui.

1.2 No sentido generalizado, uma função é infinitamente diferenciável

Já que defini a derivada da função H, também posso definir a sua segunda derivada. Vou partir de uma função f que tenha duas derivadas:

$$< f'', \phi >= - \int_{-\infty}^{\infty} f'(x)\phi'(x)dx = \int_{-\infty}^{\infty} f(x)\phi''(x)dx =< f, \phi'' > \tag{11}$$

$$< H'', \phi >= \phi''(0) \tag{12}$$

Usei uma função duas vezes diferenciável para descobrir a fórmula para a derivada segunda de uma função qualquer (integrável), e calculei a segunda derivada da função de Heaviside. O resultado é que as derivada sucessivas se apresentam com uma alternância de sinal da forma $(-1)^n$ em que n é a ordem derivação:

$$< f^{(n)}, \phi >= (-1)^n < f, \phi^{(n)} > \tag{13}$$

em que no segundo membro da equação (13) se encontra uma integral comum se f for uma função integrável, caso contrário é simplesmente a expressão do funcional linear f, por exemplo podemos usar esta expressão para calcular a derivada de ordem 2 da Dirac ou de qualquer ordem.

1.3 Derivada no sentido das distribuições

A primeira pergunta que @ leit@r pode se estar fazendo é sobre a "interpretação geométrica" desta nova forma de derivar. Alguns ajustes têm que ser feitos antes de responder a esta pergunta que é inteiramente lógica e legal. A primeira reação que eu tive quando me encontrei com a "derivada generalizada" foi a de que se tratava de uma definição abstrata, que funcionava, mas que não me pareceu ter sentido concreto, e tem!

[8]da Wikipedia, José Sebastião e Silva (Mértola, 12 de Dezembro de 1914—Lisboa, 25 de Maio de 1972) foi um matemático português.Em 1960 foi nomeado, por convite, professor catedrático Instituto Superior de Agronomia da Universidade Técnica de Lisboa.

[9]Na versão definitiva deste trabalho, como certeza.

O primeiro passo nesta direção deve vir numa visão mais ampla da integral que não aparece nos cursos de graduaçao e deveria. Se apresenta a integral com uma visão distorcida, restrita, o fruto natural de uma primeira visão, necessária, da integral como área, mas seria importante logo mostrar que a integral é uma expressão para definir um operador linear numa classe de funções e que esta classe de funções é um espaço vetorial de funções. Até porque a integral, como é apresentada, se torna meio absurda uma vez que os chamados *métodos de integração*, rapidamente são levados ao descrédito, como métodos, ao se encontrarem funções, que são integráveis, mas às quais não sabemos aplicar o *fundamental* teorema do Cálculo. Aqui vimos a importância da *integração por partes* para generalizar a derivação e igual importância têm os *outros métodos de integração* em desenvolvimento teóricos que é o que justifica o estudo deles.

Alterando-se esta visão abririamos uma espectro de aplicação muito mais amplo para integral. Por exemplo, e um caso simples, do dia-a-dia, o preço dos terrenos, numa cidade qualquer, é uma integral do tipo

$$\int\int_Q P(x,y)dxdy \tag{14}$$

Na equação (14) se encontra uma função definida pela corrupção imobiliária que desnivela os valores dos terrenos nas áreas urbanas (e mesmo nas áreas não urbanas) em função de questões pseudo econômicas ou fatores geográficos dos quais tira proveito ajudando a criar *sem terras* ou *sem tetos*. Deixando esta corrupção de lado, mas continuando a usar os seus efeitos, a função P que se fosse 1, simplesmente estabeleceria que o preço de um terreno seria diretamente proporcional à sua área, como não é 1 e sim uma função que tem picos notáveis na proximidades de alguns pontos interessantes, como rios, lagos, centro da cidade, é uma função que se assemelha a um chapeu com várias copas, dependendo do tamanho da cidade e dos seus pontos "turísticos" interessantes.

Este exemplo, embora elemento maligno do capitalismo, serve como aplicação da integral à Economia, e mostra que nas integrais existe uma função peso e portanto a forma natural de definir integral é a *integral de Riemann-Stieltjes*:

Definição 1 *Integral de Riemann-Stieltjes*

Seja p uma função integrável na reta inteira.

A integral de Riemann-Stieltjes da função f, no intervalo $[a,b]$ é integral de Riemann

$$< p, f >= \int_a^b f(x)p(x)dx \tag{15}$$

a função p se chama uma função peso.

Exemplos comuns é a integral que define o preço dos terrenos, com a função P da corrupção imobiliária, ou as integrais da Estatística em que aparecem uma probabilidade no lugar da função p, um caso muito comum é aparecer a *gaussiana.*

Em geral queremos que a função peso seja algo estável que aplicamos em todas as funções de uma classe que chamamos admissivel o que nos conduziria a construir um espaço vetorial de funções admissíveis. Foi isto que levou Schwarz e Sebastião Silva, entre muitos outros, mas estes dois descreveram com perfeição a saída para o problema que Dirac criara 40 anos antes.

Para terminar esta seção, a função p que aparece na integral de Riemann-Stieltjes é de fato uma função, ela assume o papel que as funções deveriam sempre ocupar, a de distribuição no espaço em que ela estão definidas, que é como os estatísticos entendem funções, pelo menos aqueles que entendem. Foi este o sentido que levou Schawarz a criar a *teoria das distribuições* - a teoria de novos objetos que vem a substituir as funções e que incluem estas como casos particulares. A função de Dirac é uma distribuição que não é uma função.

A partir de agora quando dissermos que f é uma função, entendemos que existe um espaço de funçoes-teste ϕ e nos interessa analisar qual é o efeito de f aplicada no conjunto das funções teste

$$< f, \phi >= \int_{-\infty}^{\infty} \phi(x) f(x) dx \tag{16}$$

e quando falarmos "função" somente nos interessa pelas que forem localmente[10] integráveis na reta inteira, as distribuições do espaço $\mathbf{R}$.

E estaremos vendo sendo as funções como funcionais lineares, dentro de integrais, digamos. É esta a interpretação geométrica que damos as funçoes, a figura (2) página 8, mostra o significado geométrico da derivada de uma probabilidade.

2 Transformadas integrais

Na seção anterior descrevi um processo "estático" voltado para calcular o valor aproximado $f(a)$ ou um conjunto de valores discreto de um funcional. Vou apresentar nesta seção e na próxima uma forma de transformar um espaço n'outro podendo assim transformar

$$f \in \mathcal{E} \Rightarrow \tilde{f} \in \mathcal{F}$$

em que $\mathcal{E}, \mathcal{F}$ são dois *espaços de função* definidos no momento certo.
A convolução define uma transformação entre espaços e vou descrevê-la aqui como um método que pode ser utilizado no Cálculo Variacional.Uma pequena generalização do exemplo "convolução" produz uma outra transformação linear que é um paradigma mais genêrico que também vai ser apresentado aqui, *transformadas integrais*.
A linguagem vai ser um pouco vaga uma vez que não é assunto novo e pode ser encontrado em diversos textos com todo rigor, se espera que o leitor conheça o assunto.

[10] Integráveis em qualquer intervalo fechado da reta.

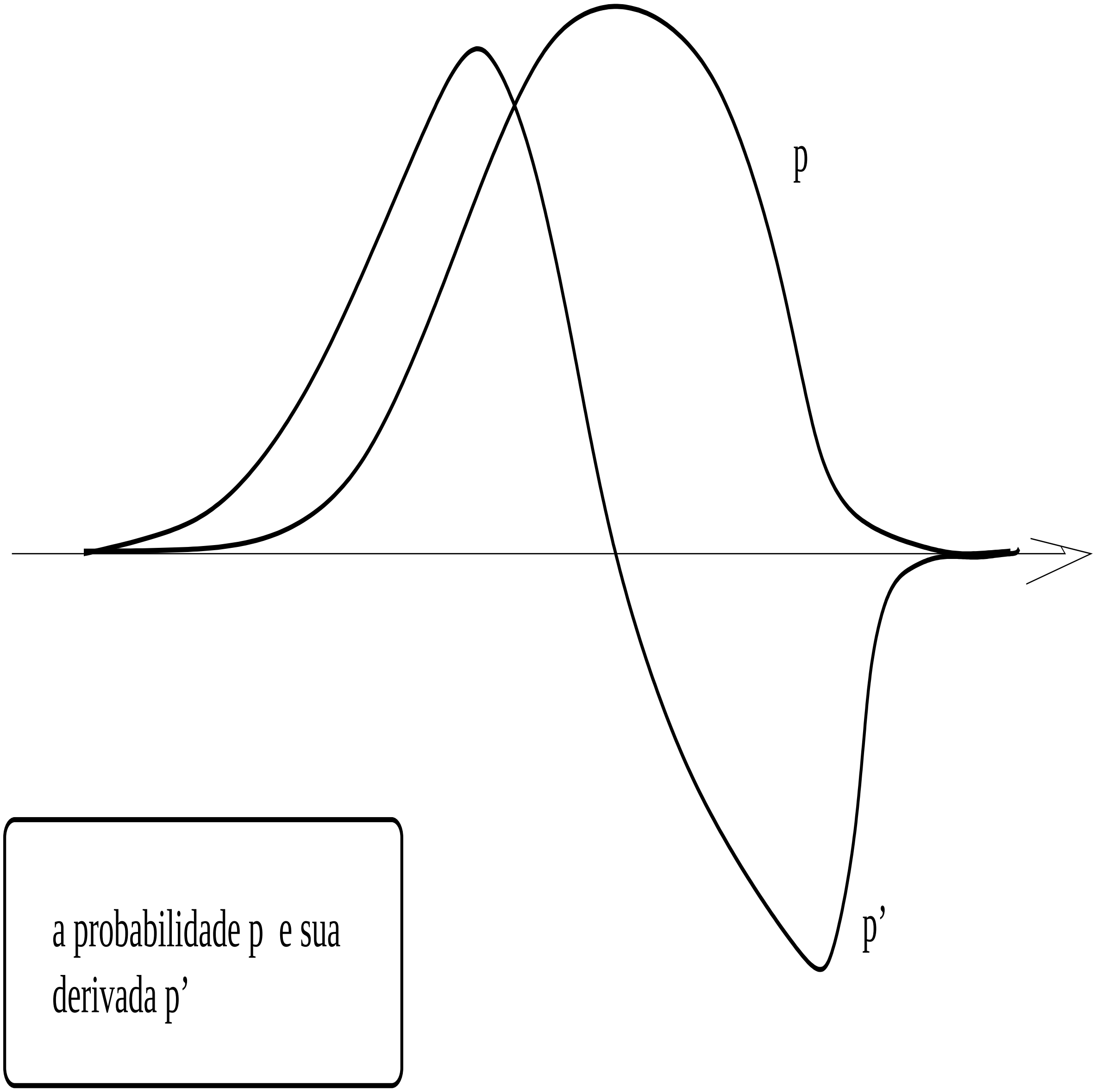

Figura 2: A derivada de uma probabilidade

2.1 O produto de convolucao

Considere duas funções integráveis[11] f, g então o produto delas também é integrável e podemos definir a integral *dependendo de um parâmetro*

$$t \mapsto \int_{-\infty}^{\infty} f(x)g(t-x)dx \tag{17}$$

Observação 1 (exemplo de hiperplano) *Um hiperplano do espaço.*

Vou dar um exemplo para mostrar que estes cálculos acompanham o estudante de Matemática desde a sua infância.

[11] Como desde o início desta artigo, estamos sempre pensando em integrais no espaço inteiro.

Se pensarmos em f como uma função fixa, e g como um elemento genérico de um espaço de funções, teremos definido com a equação (17) uma transformação deste espaço de funções em outro espaço de funções. Este "processo" é muito comum e @ leit@r já se deparou com ele em diversas instâncias, por exemplo, a equação de um plano, em Geometria Analítica é um exemplo disto quando ela se escreve

$$< (A, B, C), (x - a, y - b, z - c) >= 0$$

em que foi eleito um vetor fixo do espaço, como vetor perpendicular ao plano que passa no ponto (a, b, c), um plano é um hiperplano do $\mathbf{R}^3$ e o produto escalar é um modelo de operação que pode ser realizado com uma integral e nós assim poderiamos definir um hiperplano de um determinado espaço de funções se escrevessemos

$$< f, g(x - x_0) >= \int_{-\infty}^{\infty} f(x)g(x - x_0)dx = 0$$

em que f seria um elemento do dual do espaço de funções considerado. Esta expressão define o conjunto de todas as funções que sejam perpendiculares à função f.

A operação definida na equação (17) se chama *produto de convolução* e a razão é a de que ela tem as mesmas propriedades que o produto de números tem relativamente à adição de números, agora se pensando nestas operações aplicadas às funções elementos de um certo espaço de tal forma que se pode obter uma álgebra de funções e o que é mais interessante, a Dirac aparece aqui como o elemento neutro deste produto quer dizer que se pode definir uma álgebra com unidade, eu já @ adverti que a linguagem está sendo mantida vaga... Apenas para excitar o interesse d@ leit@r, os polinômio trigonométricos podem ser definidos como a convolução de uma determinada função por uma função particular:

$$c_n = f * e_n = \int_{S^1} f(x)e^{inx}dx$$

é o coeficiente de Fourier (complexo) de f e o caso real se deduz do caso complexo, [**?**, RudinRCV]

Retornando ao "caso contínuo" a transformação integral definida na equação (17) define uma nova função, da variável t e foi poristo que eu descrevi aquela integral como "dependendo de um parâmetro".

2.2 Transformadas integrais

Eu vou agora tomar o exemplo da convolução para definir uma transformada integral um pouco mais geral que possivelmente servirá melhor aos nosso propósitos futuros. Vou substituir na equação (17) a função f, que sugeri como fixa, por uma função bivariada, o resultado é a expressão

$$t \mapsto \int_{-\infty}^{\infty} f(x,t)g(x)dx \tag{18}$$

Este exemplo é exatamente o mesmo que encontramos em Álgebra Linear quando se multiplica uma matriz retangular $n \times m$, portanto uma função bivariada, por um vetor. Na multiplicação matricial, como na integral da equação (18), a soma "consome" uma das variáveis e o resultado é uma função da outra variável, e o resultado, na Álgebra Linear de dimensão finita, é que transformamos um vetor de uma espaço de dimensão n n'outro vetor de um espaço de dimensão m, aqui, na equaçao (18), cuidadosamente evitando de falar de dimensão, transformamos um vetor de um espaço n'outro vetor de outro espaço. Neste contexto a função bivariada poderia ser chamada simplesmente "matriz da transformação linear" mas o hábito levou a que fosse chamada de "núcleo da transformaçao integral", para que @ leit@r veja claramente a comparação, se trabalharmos com uma aproximação (discretização) da integral na equação (18), vamos ter vetores em espaços de dimensão finita, representando g, e se usarmos somas de Riemann para fazer aproximação da integral, vamos ver aparecer uma matriz $n \times m$ representando f.

Desenvolvemos assim a linguagem necessária para a nossa próxima etapa em que vamos formular fracamente (variacionalmente) uma equação diferencial.

3 Otimizando um funcional

Nesta seção vou dar exemplo de um funcional e calcular uma otimização do mesmo.
Vou usar o método abstrato que é tipico da Álgebra em que suponho que uma solução exista e a represento com variáveis às quais passo a aplicar condições e, no presente caso, operações do Cálculo Diferencial e Integral.

3.1 O método variacional

Esta seção ainda está incompleta, é trabalho em andamento, *estou tentando entender* como se pode usar o método variacional usando uma técnica que domino bem, splines a suporte compacto.

A idéia é a de que a solução de um problema, em vez de ser procurada diretamente, venha através da minimização de um funcional (integral) sobre uma região. Há vários exemplos, a *brachistocrona, a melhor curva* trafegando-se dentro de um determinado meio, a luz por exemplo, ou uma superfície mínima, uma *melhor superfície*, ou em cima de uma determinada superfície a distância mínima entre dois pontos, trafegando-se sobre uma determinada superfície (condicionamento) *geodésicas*.

Os ingredientes do método estão associados

- a um operador, uma integral, ou mais geralmente uma transformação linear;
- uma classe de variedades (curvas, superfícies) admissíveis;
- uma equação diferencial.

Para entender como o método funciona, vou analisar a equação de Laplace, associada a uma região **R** do plano,

$$\Delta u = u_{xx} + u_{yy} = 0 \tag{19}$$

e lhe aplicar o operador

$$I(u) = \int\int_{\mathbf{R}} (\omega_{xx} + \omega_{yy}) v(x,y) dxdy \tag{20}$$

em que ω é um elemento de uma *3-splines-partição da unidade* subordinada a uma cobertura de **R**[12], [6], Isto nos permite escrever a integral na equação (20) como uma soma finita em que ω é cada um dos elementos da partição da unidade, a suporte compacto. Esta família pode ser construída de tal forma que a medida dos suportes diminuam com o número de elementos da família de modo que cada "elemento" na soma se reduz ao Laplaciano aplicado a um destes elementos na integral. Uma forma de definir esta família é por variáveis separáveis, considere

$$\eta_{n,m}(x,y) = \rho_n(x)\rho_m(y)$$

em que ρ_k é um splines por convolução (univariado), veja na figura (3) página 12, temos o gráfico da segunda derivada de um elemento de $\mathcal{D}_c(R)$, os elementos da base de $\mathcal{D}_c(R^2)$ podem ser funções à *variáveis separáveis*[13], o gráfico aqui é o de um dos fatores do produto, ou seja o caso univariado[14].

O gráfico foi feito com um programa em `C++` que pode ser disponibilizado a pedido, mas é um gráfico apenas para efeito ilustrativo, uma vez que usei $exp(-1/x^2)$ para construir uma função diferenciável a suporte compacto, isto pode ser feito com a equação

$$exp(x) = e^{-1/x^2};\ \text{o prato de classe } \mathcal{C}^\infty \tag{21}$$

$$Exp(x) = \begin{cases} x < 0 & 0 \\ x \geq 0 & exp(x) \end{cases} \quad \text{prato diferencialmene quebrado} \tag{22}$$

$$\rho(x) = \frac{1}{A} Exp(x+1) * Exp(1-x) \tag{23}$$

[12]Uma família de 3-splines a suporte compacto, positivos, cuja soma, ponto a ponto, é 1.

[13]Ninguém usa esta caracterização, $\phi(x,y) = \phi_1(x)\phi_2(y)$, chamando-a de "variáveis separáveis", mas a idéia é esta.

[14]Vou fazer o gráfico no caso multivariado, mas não para incluir no trabalho, vou deixá-lo num programa que o usuário rode, usando `gnuplot`, porque estes gráficos ficam claros apenas quando você os manipula.

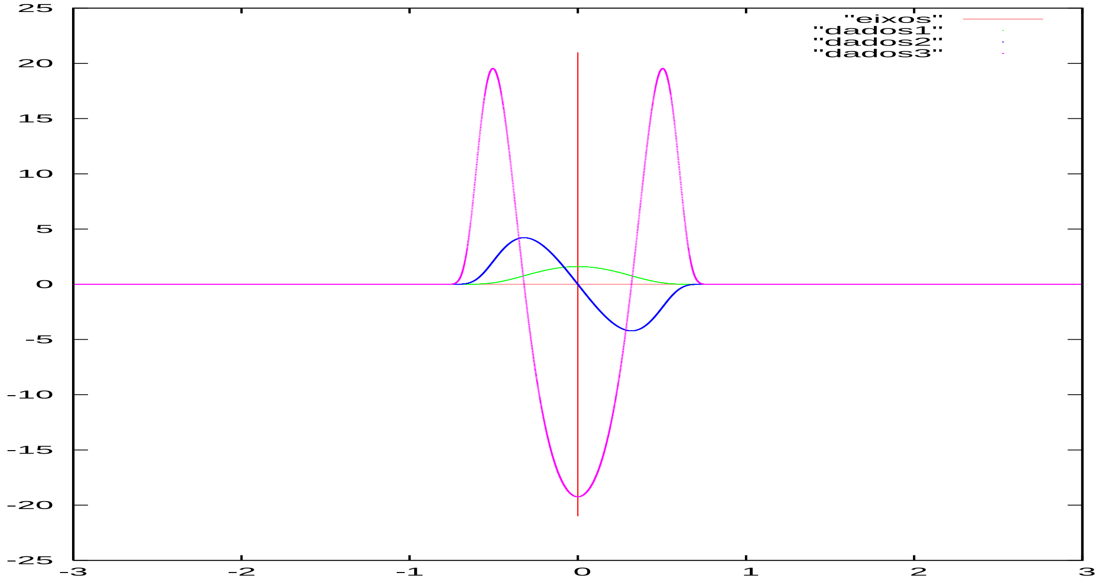

Figura 3: ρ, ρ', ρ''

Na equação (21) está o "prato" de classe $\mathcal{C}^\infty$, na equação (22) está definido o "prato diferencialmente quebrado" usando a função *exp* na equação para reais positivos, continuada de forma infinitamente diferenciável pela constante zero, para reais negativos. Na equação (23) defini uma função usando duas versões do "prato quebrado", uma translatada de -1 que é zero para $x < -1$ e a outra translatada de 1 mas trocando o sinal do parâmetro para usar um "prato quebrado" que é zero para valores de x maiores que 1 de modo que ao fazer o produto o resultado é uma função de classe $\mathcal{C}^\infty$ e a suporte compacto.

Rode em `gnuplot`, o script abaixo, para reproduzir este gráfico

```
 %% variacional02.gnuplot
pow(x,n) = x**n;
Exp(x) = exp(-1/pow(x,2));
a = 1; ## use a diferente de 1 para experimentar
qu(x) = (x<=0)?0:a*Exp(a*x);
rho(x) = qu(x+1)*qu(1-x);
set xrange [-2:2];
plot rho(x), 0;
pause -2; delta = 0.00000001;
drho(x) = (rho(x+delta)-rho(x))/delta;
epsilon = 0.0001;  ## para minimizar o erro na segunda derivada...
ddrho(x) = (drho(x+epsilon)-drho(x))/epsilon;
plot rho(x), drho(x), ddrho(x),0;
pause -2;
```

Experimente usar $epsilon = 0.00000001$ e rode novamente o comando `plot` para observar o efeito do erro na segunda derivada, isto mostra que ao calcularmos a segunda derivada, usando quocientes de diferenças, se pode obter um melhor resultado usando menor precisão no passo.

Observação 2 *Aqui havia um erro na descrição*

Eu troquei maior por menor na descrição da equação (22) na versão anterior. Espero que tenham observado o erro. Melhorei agora a descrição.

Esta função eu vi pela primeira vez numa aula que assisti de B. Malgrange, em Grenoble, em Setembro de 1970, saindo da graduação em Fortaleza, uma função infinitamente diferenciável que se anula na semireta negativa... *bem elementar, mas na época um choque para mim.*

No programa eu calculei, (usando soma de Riemann), a integral do produto destas duas translatadas e assim obtive o número A que aparece na definição de ρ na equação (23).

Redefini ρ usando o inverso da integral como coeficiente, de modo que a integral desta função, ρ, seja 1. Isto não é muito importante para este exemplo, mas é essencial saber como construir funções pesos, e aqui esta um algoritmo para construí-las na classe $\mathcal{C}^\infty$.

O resultado disto está na figura (3), feito com `gnuplot` a partir do meu programa em `C++`.

Com translações de funções do tipo

$$\psi(x,y) = \rho(x)\rho(y) \tag{24}$$

é possível construir uma família que chamamos de *partição da unidade*, porque a soma sobre a família é 1 e esta soma se anula fora de um compacto (um conjunto fechado e limitado do $\mathbf{R}^2$.

Vou supor que eu tenha uma *partição da unidade*[15], uma família de funções, no presente caso, finita, ψ_k, obtidas como na equação (24), mas sendo splines polinomiais [16], posso re-escrever a equação (20) - e vou repetí-la como primeira equação para facilitar a comparação, evitar que você tenha que "virar a página":

$$I(u) = \int\int_{\mathbf{R}} (\omega_{xx} + \omega_{yy})v(x,y)dxdy \tag{25}$$

$$I(u) = \sum_{k=0}^{n-1} \int\int_{\mathbf{R}} (\psi_{k,xx} + \psi_{k,yy})v(x,y)dxdy \tag{26}$$

$$\int\int_{\mathbf{R}} (\psi_{k,xx} + \omega_{k,yy})v(x,y)dxdy = \int\int_{\mathbf{R}} (\psi_k v_{xx}(x,y) + \psi_k v_{yy}(x,y)dxdy \tag{27}$$

Na equação (26) estou usando uma soma, não se esqueça da hipótese de que tenho uma *partição da unidade*, a família de funções ψ_k cuja soma é 1 e entra em lugar de ω da equação (20). Na equação (27) usei derivação no sentido das

[15] Isto torna o texto difícil, @ leitor não tem nenhuma razão para crer que isto seja possível, peço que aceite, e garanto que vou construir isto detalhadamente, agora eu estou apenas pensando, ou melhor, escrevendo para aprender.

[16] Acreditem, é fácil e elementar de construir. E por que splines polinomiais? Porque é fácil de calcular-lhes as integrais!

distribuições. Se o suporte[17] for suficientemente pequeno posso ter $v(x_k, y_K)$ em que (x_k, y_j) são os nós da malha considerada na região $\mathbf{R}$.

Aqui parece que tem um erro, vou ter que examinar melhor isto antes que alguém ache o erro e eu fique desmoralizado.

A questão com minha reputação me preocupa muito! eu não sei o que é interessante aqui neste ponto, se é ter uma partição da unidade *ou ter uma família de funções cuja integral seja 1 - tenho a impressão de que preciso da segunda hipótese, os cálculos vão me mostrar isto.*

Vou chamar R_k os suportes das funções das funções ϕ_k e vou reescrever a equação (27)

$$\int\int_{\mathbf{R}}(\psi_{k,xx} + \omega_{k,yy})v(x,y)dxdy = \int\int_{\mathbf{R}}(\psi_k v_{xx}(x,y) + \psi_k v_{yy}(x,y)dxdy = \tag{28}$$

$$= \sum_{k=1}^{N} \int\int_{\mathbf{R}_k} (\psi_{k,xx} + \omega_{k,yy})v(x,y)dxdy = \tag{29}$$

$$\sum_{k=1}^{N} \int\int_{\mathbf{R}}(\psi_k v_{xx}(x,y) + \psi_k v_{yy}(x,y)dxdy = \tag{30}$$

$$= \sum_{k=1}^{N} VM(v_{xx}) + VM(v_{yy}) \tag{31}$$

em que VM é a notação para "valor médio integral" e que eu deveria indicar que está sendo calculado no suporte de ψ_k.

E aqui acho que cheguei em algo interessante, porque eu tenho a equação diferencial, é uma expressão para a qual posso calcular os valores médios o que dá sentido à equação (28).

Tenho que revisar estas cálculos, e vou passar para todo mundo para ver ser alguém corrige algum erro existente.

3.2 Um primeiro exemplo

Na figura (4) página 15, tenho um domínio retangular *Omega* onde está parametrizada uma família de curvas

$$\Gamma = \{(x(t), y(t)); t \in [0,1]; x(0) = P_1; ; x(1) = P_2 \tag{32}$$

para as quais escolhi um único intervalo de parametrização, comum a todas as curvas, o intervalo $[0,1]$ de medida 1, sem nenhuma razão especial (poderia ser um intervalo qualquer, o de parametrização). Na figura podemos ver, a título de ilustração três exemplos de curva.

Considere agora um funcional

$$\omega = F(x,y) \in \mathbf{R} \tag{33}$$

[17] Suporte é o conjunto em que uma função é essencialmente diferente de zero, na figura (3) página 12, o suporte é $[-1,1]$.

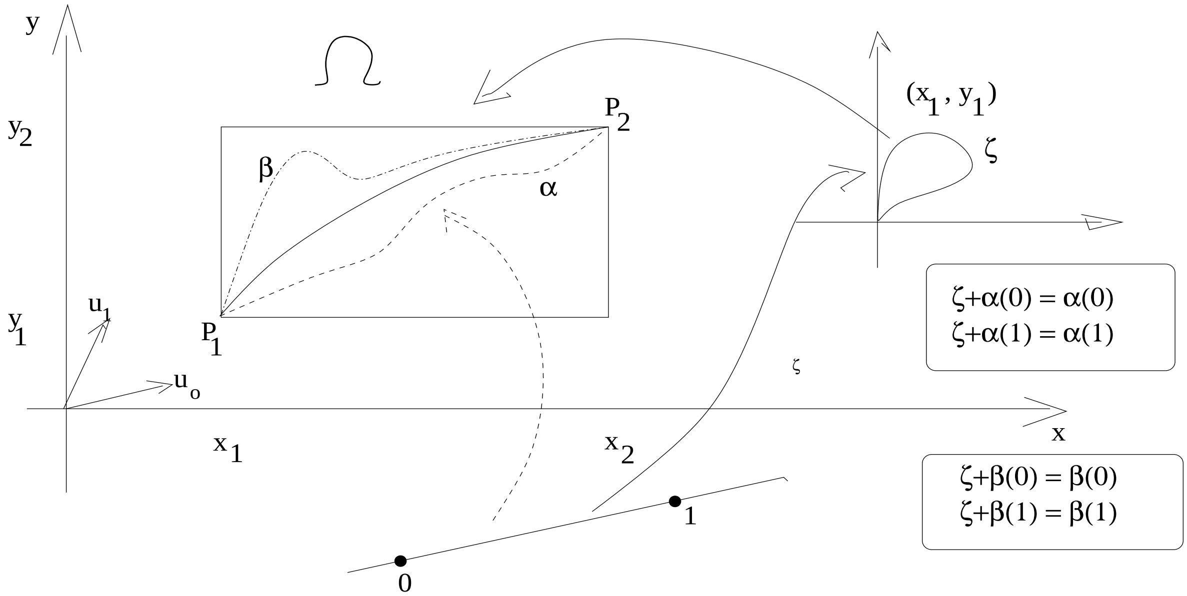

Figura 4: Maximização de rotas

em que (x, y) representa uma curva na região Ω. As curvas que nos interessam podem representar diversos fenômenos, por exemplo percursos de um objeto indo do ponto $P_1 = (x_1, y_1)$ ao ponto $P_2 = (x_2, y_2)$ ou poderia ser o formato de uma viga submetida a uma determinada deflexão (carga), com suas extremidades presas aos pontos P_1, P_2 e, neste segundo caso, talvez algumas das curvas que apareçam na figura (4) não sejam uma boa representação para vigas.

Esta segunda hipótese acrescenta ao nosso exemplo uma linguagem que já usamos na segunda seção ao nos referirmos às "funções admissíveis" e aqui falariamos em "curvas admissíveis". Observe que não precisamos nos preocupar, de partida, com quais são as "curvas admissíveis", o método deve, automaticamente selecionar as curvas que interessa, a partir das condições que se acrescentarem ao problema, esta é uma das características mais notáveis do método abstrato em que designamos de forma vaga a solução do problema e depois traduzimos por equações suas propriedades e poderemos, posteriormente, com manipulação algébrica obter a solução[18].

Neste caso posso incluir algumas condições de contorno que vou inicialmente expressar com palavras e logo traduzir por equações. Observe que vou pensar ao mesmo tempo nas duas possibilidades que descrevi acima, de que as curvas representem rotas de de um objeto se deslocando entre os pontos P_1, P_2 ou uma viga com extremidades presas neste ponto submetida a uma determinada carga. Portanto não cabe pensar em curvas que não se originem no ponto P_1 e findem no ponto P_2. Aqui estamos também falando de orientação, que pode ou não ser importante para um determinado problema, para vigas não seria, mas talvez o seja para rotas de objetos trafegando neste domínio. Para

[18] A tecnologia de programa orientada a objeto vem usando com grande sucesso esta forma abstrata de pensar, nesta tecnologia se manipulam os objetos com operações lógicas que incluem as manipulações algébricas como caso particular.

as rotas talvez não interesse as condições sobre as derivadas. Estes são dados particulares que podem ser excluidos da formulação, simplificando o problema com um determinado objetivo.

Como desejamos calcular uma otimização, vamos partir da hipótese de que temos a curva que otimiza o funcional e incluir uma nova curva no processo que vai servir de variável auxiliar, e vou descrever isto com um sistema de condições:

- $t \mapsto (x(t), y(t))$; $(x(0), y(0)) = P_1$; $(x(1), y(1)) = P_2$
- $t \mapsto (x_1(t), y_1(t))$; $(x_1(0), y_1(0)) = (0, 0)$; $(x_1(1), y_1(1)) = (0, 0)$

Estas condições descrevem duas curvas, uma que satisfaz às condições de fronteira que me interessam, partindo do ponto P_1 e chegando ao ponto P_2 e outra que sai da origem e retorna para origem, de modo que se eu somar as duas curvas, vou obter uma curva satisfazendo as condições de fronteira que me interessam.

Agora vou incluir um parâmetro de deformação na segunda curva, de modo que a soma represente um pequeno afastamento da primeira curva

$$t \mapsto (x(t), y(t)) \; ; \; (x(0), y(0)) = P_1 \; ; \; (x(1), y(1)) = P_2 \tag{34}$$
$$t \mapsto \lambda(x_1(t), y_1(t)) \; ; \; (x_1(0), y_1(0)) = (0, 0); \; (x_1(1), y_1(1)) = (0, 0) \tag{35}$$
$$t \mapsto (x(t) + \lambda x_1(t), y(t) + \lambda y_1(t)) \tag{36}$$

de modo que as condições de adição que descrevi acima ainda continuam verdadeiras, mas agora tenho uma curva variável de acordo com o parâmetro λ.

$$\omega(\lambda) = F(x + \lambda x_1, y + \lambda y_1) \tag{37}$$

Calculando a derivada de F, relativamente à λ vamos encontrar

$$\frac{d\omega}{d\lambda} = \frac{\partial F}{\partial x}\frac{dx}{d\lambda} + \frac{\partial F}{\partial y}\frac{dy}{d\lambda} = 0 \tag{38}$$

que deverá ser zero, quando $\lambda = 0$ porque parti da hipótese de que já tinha um caminho ótimo para o funcional, um ponto de extremo.

Como esta diferencial é exata[19] ela é o modelo de uma variedade linear tangente ao objeto $\omega = F(x, y, x', y')$ e se conhecermos uma solução, um ponto de tangência, (a, b, c, d)[20] podemos escrever a equação da variedade linear tangente

$$\omega - \omega_0 = \frac{\partial F}{\partial x}(x - a) + \frac{\partial F}{\partial y}(y - b) + \frac{\partial F}{\partial x'}(x' - c) + \frac{\partial F}{\partial y'}(y' - d) = 0 \tag{39}$$

$$d\omega = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy + \frac{\partial F}{\partial x'}dx' + \frac{\partial F}{\partial y'}dy' = 0 \tag{40}$$
$$\frac{\partial F}{\partial y}dy = -\frac{\partial F}{\partial x}dx - \frac{\partial F}{\partial x'}dx' - \frac{\partial F}{\partial y'}dy' \tag{41}$$

[19] A obtivemos por derivação de F

[20] (a,b) é um ponto sobre uma curva e (c,d) é paralelo ao vetor tangente a este ponto.

3.3 Outro exemplo

Usando a notação e muitos dos cálculos na seção anterior vamos nos especializar num resultado. Vamos supor que o funcional F é o menor valor (menor rota) entre os pontos P_1, P_2. Agora temos uma expressão para o funcional

$$\omega = F(x, y, x', y') = \int_0^1 \sqrt{1 + x'^2 + y'^2 + x''^2 + y''^2} \tag{42}$$

a fórmula do comprimento de arco de uma curva no espaço.

4 Formulação variacional de uma edo

$\square$

4.1 Transformando uma equação numa integral

Uma equação diferencial linear homogênea é uma expressão que pode ser apresentada (formulada) com auxílio de um polinômio P

$$P(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \tag{43}$$

$$P(D)(y) = 0 \tag{44}$$

$$\tag{45}$$

em que D representa o operador derivada.

Para interpretar variacionalmente a equação diferencial, o que vamos fazer é perguntar como interpretariamos a expressão da equação diferencial quando ela estivesse aplicada a um espaço de funções admissíveis (é a linguagem comum para os físicos ou engenheiros) ou um espaço de funções-teste. O espaço das funções teste é forma de funções infinitamente diferenciáveis e que se anulem fora de alguma intervalo fechado, são todas do tipo que aparecem nas figuras (1) ou (2). Como eu já disse, elas tiveram papeis preponderantas em diversas ocasiões, na tecnologia das telecomunicações, por exemplo onde criaram o caminho para um novo tipo de tranformada, as transformadas *wavelets* que vieram substituir (generalizando)[21], as *transformadas de Fourier*.

Quer dizer, se f for uma função, então

$$< f, \phi >= \int_{-\infty}^{\infty} f(x)\phi(x)dx \tag{46}$$

é a formulação "variacional" das funções.

[21] Esta é um história longa, consulte [**?**].

Se ϕ estiver concentrada na vizinhança do ponto $a \in \mathbf{R}$, é positiva, com integral 1, e suporte com medida pequena, então a equação (46) nos fornece um valor aproximado para $f(a)$, uma função-teste ϕ com estas propriedades é chamada de "sinal" ou "um pulso de energia 1", elas servem para *enjanelar* um fenômeno em volta de um ponto.

Com este método eu posso enjanelar a função f na vizinhança de um ponto qualquer, usando uma função ϕ apropriada, para encontrar um valor aproximado para $f(a)$.

Aqui @ leit@r poderia me perguntar, *e para que vou* enjanelar *uma função que conheço, para calcular o seu valor médio num ponto se eu sei o valor da função no ponto?*

E o objetivo não seria este, estou considerando uma função f, escrevi a sua integral, equação (46) o que supõe que conheço bem f e então seria inútil calcular o valor médio $f(a)$.

Mas se f for um experimento a integral na equação (46) vai ser discretizada, aproximada por uma soma de Riemann ou uma aproximação polinomial, ou possivelmente vamos usar elementos finitos para calcular (46) e é aqui está o ponto de interesse.

Fique bem claro: eu estou procurando entender o assunto, e escrevendo para aprender, também estou estudando isto.

Para não deixá-los inseguros, lembro que eu sei um bocado de coisa disto que estou falando, o que ainda não sei é como é que vou expressar variacionalmente as equações diferenciais. É um assunto novo para mim.

Acho que parte do problema está encaminhado, a equação (46) representa aproximadamente o valor $f(a)$ de uma função que conheço, consequentemente (46) pode expressar com precisão os resultados de um experimento no ponto $\underline{a}$.

Ora, se f for uma solução da equação diferencial

$$P(D)y = 0 \tag{47}$$

então

$$< P(D)f, \phi >=< f, P(\tilde{D})\phi >= \int f(x)P(\tilde{D})\phi(x)dx \tag{48}$$

em que $\tilde{D}$ é a derivada generalizada, com as trocas de sinal definidas. Na equação (48), transferimos para as funções-teste a expressão da equação diferencial com um ganho imediato: elas, as funções teste, são infinitamente diferenciáveis, portanto, qualquer derivada delas existe e assim a integral que está escrita na equação (48) existe, ponto essencial para que possamos aplicar métodos numéricos em seu cálculo.

4.2 Expressão variacional de uma e.d.o.

Vou agora tentar resolver, eu estou aprendendo agora a escrever equações diferenciais variacionamente, o problema que Diego me propos (ou é propôs? como ficam as novas regras gramáticais ?).

Qual a expressão variacional para

$$\frac{d^2 f(x)}{dx^2} + u(x) = 0 \tag{49}$$

Eu suponho que o problema na verdade é

Qual a expressão variacional para

$$\frac{\partial^2 f(x)}{\partial x^2} + u(x) = 0 \tag{50}$$

Nesta equação diferencial procuramos uma função f sobre a qual aplicamos um operador diferencial linear e o resultado é $-u$ e depois teremos que discutir questões de contorno e condições iniciais. Qjuer dizer que o problema é

Qual a expressão variacional para

$$\frac{\partial^2 f(x)}{\partial x^2} = -u(x) \tag{51}$$

cuja tradução variacional seria “para qualquer que seja a função admissível (função-teste) ϕ a imagem de

$$\int \int_T f(x,y) \frac{\partial^2 \phi(x)}{\partial x^2} dxdy = -u \tag{52}$$

Referências

[1] Arfken, G. *Mathematical Methods for Physicists*
Academic Press, INC. 1985

[2] Buck, R. C. and Buck E. F. *Advanced Calculus* McGraw-Hill - 1965

[3] Bo Thidé *A course in Electro Magnetism*

http://www.plasma.uu.se/CED/Book

[4] Praciano-Pereira, T. Cálculo numérico computacional

http://www.calculo-numerico.sobralmatematica.org/textos

[5] Praciano-Pereira, T. *Splines por convolução*

http://www.sobralmatematica.org/preprints

[6] Praciano-Pereira, T. *Splines por convolução bivariados* expansão de *Splines por convolução* a ser escrito em breve - caso multivarido.

http://www.sobralmatematica.org/preprints

[7] Simmons, G.F.
Differential Equations with App. and Hist. Notes.
McGraw-Hill - Book Company - 1978

[8] Hirsch, e Smale S. *Linear Algebra, differential equations and dynamical systems* - Academic Press

[9] Boyce, William E e Diprima, R. C. *Equações diferenciais elementares e problemas de valores de contorno* Editora: LTC ISBN-13: 9788521614999 8a Edição - 2006 - 450 pág.

[10] Claus I. Doering e Artur O. Lopes *Equações Diferenciais Ordinárias* Primeira Edição Coleção Matemática Universitária - IMPA

[11] *Wikipedia, the free enciclopedia in the Internet*

http://en.wikipedia.org