Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza
Programa de Pós Graduação em Matemática

Renato Soares de Oliveira

Cálculo Variacional com Aplicação ao Problema de Sturm-Liouville

Recife

2016

Renato Soares de Oliveira

Cálculo Variacional com Aplicação ao Problema de Sturm-Liouville

Dissertação apresentada ao Programa de Pós-graduação em Matemática da Universidade Federal de Pernambuco como requisito parcial para obtenção do título de Mestre em Matemática.

Orientador: Eduardo Shirlippe Góes Leandro

Recife

2016

Catalogação na fonte
Bibliotecária Elaine Cristina de Freitas, CRB4-1790

O48c Oliveira, Renato Soares de
Cálculo Variacional com Aplicação ao Problema de Sturm-Liouville / Renato Soares de Oliveira . – 2016.
53f.

Orientador: Eduardo Shirlippe Góes Leandro.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Matemática, Recife, 2016.
Inclui referências.

1. Análise Matemática. 2. Cálculo Variacional. I. Leandro, Eduardo Shirlippe Góes (orientador). II. Título.

515 CDD (22. ed.) UFPE-MEI 2017-207

Renato Soares de Oliveira

Cálculo Variacional com Aplicação ao Problema de Sturm-Liouville

Dissertação apresentada ao Programa de Pós-graduação em Matemática da Universidade Federal de Pernambuco como requisito parcial para obtenção do título de Mestre em Matemática.

Trabalho aprovado. Recife, 29 de Julho de 2016:

Eduardo Shirlippe Góes Leandro
Orientador

Airton Temístocles Gonçalves de Castro
Convidado 1

Carlos Alberto Batista da Silva Filho
Convidado 2

Recife
2016

Resumo

O cálculo de variações é um problema matemático que consiste em buscar máximos e mínimos (ou, mais geralmente, extremos relativos) de funções contínuas definidas sobre algum espaço funcional. Constituem uma generalização do cálculo elementar de máximos e mínimos de funções reais de uma variável. Ao contrário deste, o cálculo das variações lida com os funcionais, enquanto o cálculo ordinário trata de funções. Funcionais podem, por exemplo, ser formados por integrais envolvendo uma função incógnita e suas derivadas. O interesse está em funções extremas - aquelas que fazem o funcional atingir um valor máximo ou mínimo - ou de funções fixas - aquelas onde a taxa de variação do funcional é precisamente zero. Talvez o exemplo mais simples seja o de encontrar a curva com o menor comprimento possível ligando dois pontos. Se não houver restrições, a solução é (obviamente) uma linha reta ligando estes pontos. No entanto, se as possibilidades para esta curva estiverem restritas a uma determinada superfície no espaço, então a solução é menos óbvia e, possivelmente, muitas soluções podem existir. Tais soluções são conhecidas como geodésicas. Um problema relacionado a este é representado pelo princípio de Fermat: a luz segue o caminho de menor comprimento óptico ligando dois pontos, onde o comprimento óptico depende do material de que é composto o meio. Um conceito correspondente em mecânica é o princípio da mínima ação. Nesta dissertação faremos uma exposição dos principais conceitos do cálculo variacional com ênfase na equação de Euler-Lagrange, que trata-se de uma condição necessária para extremos locais de uma determinada classe de funcionais. Nosso objetivo principal é estudar os problemas de extremidades fixas com e sem vínculos para tratar o problema de Sturm-Liouville por meio de uma abordagem variacional. Veremos que cada autovalor do problema de Sturm-Liouville é obtido pela resolução de um problema variacional de minimização para depois, através desse fato, conseguirmos estimativas para esses autovalores.

Palavras-chave: Equação de Euler-Lagrange. Cálculo Variacional. Sturm-Liouville. Quociente de Rayleigh.

Abstract

Calculus of variations is a field of mathematical analysis that deals with maximizing or minimizing functionals, which are mappings from a set of functions to the real numbers. Functionals are often expressed as definite integrals involving functions and their derivatives. The interest is in extremal functions that make the functional attain a maximum or minimum value – or stationary functions – those where the rate of change of the functional is zero. A simple example of such a problem is to find the curve of shortest length connecting two points. If there are no constraints, the solution is obviously a straight line between the points. However, if the curve is constrained to lie on a surface in space, then the solution is less obvious, and possibly many solutions may exist. Such solutions are known as geodesics. A related problem is posed by Fermat's principle: light follows the path of shortest optical length connecting two points, where the optical length depends upon the material of the medium. One corresponding concept in mechanics is the principle of least action. In this dissertation we present the main concepts of the variational calculus emphasizing the Euler-Lagrange equation, that is a necessary condition for local extrema of a particular class of functionals. Our main aim is to study the problems of fixed ends with and without constraints to address the Sturm-Liouville problem through a variational approach. We will see that each eigenvalue of the Sturm-Liouville problem is obtained by solving a variational minimization problem and then, by this fact, we get estimates for these eigenvalues.

Keywords: Euler-Lagrange equation. Variational Calculus. Sturm-Liouville. Rayleigh Quotient

Sumário

1 Introdução

Um tópico importante no cálculo diferencial são os chamados problemas de otimização, que visam maximizar ou minimizar funções reais em um subconjunto qualquer do $\mathbb{R}^n$. Similarmente, podemos lidar com problemas mais gerais caracterizados por funções definidas em espaços vetoriais não necessariamente de dimensão finita. A teoria matemática que estuda as ferramentas para tratar deste tipo de problema é chamada cálculo variacional.

Este campo, desenvolvido principalmente nos séculos XVIII e XIX, foi aplicado em diversos problemas físicos e matemáticos desde a sua criação. A variedade de aplicações práticas da teoria é bastante surpreendente e seu desenvolvimento serviu para impulsionar diversos ramos da matemática, como a análise e as equações diferenciais parciais. Na física, uma das consequências mais importantes do cálculo variacional foi a reformulação da mecânica clássica de Isaac Newton feita por Lagrange em seu grande tratado, o Mécanique Analytique. Nesta obra, Lagrange revolucionou a mecânica formulando os princípios variacionais através de funcionais que dependem de um conjunto de parâmetros, chamado coordenadas generalizadas, em vez das coordenadas cartesianas que caracterizam a posição de um sistema. Deste modo, o tratamento variacional da mecânica tornou-se mais vantajoso em relação à mecânica vetorial.

A dissertação está organizada em dois capítulos. O primeiro capítulo tem por finalidade discutir os principais aspectos teóricos do cálculo variacional. Começaremos apresentando, na primeira seção, conceitos importantes como a variação de Gâteaux e extremos locais, nela veremos também que a variação de Gâteux necessariamente se anula nesses extremos. Na segunda seção faremos um breve estudo sobre a teoria de diferenciabilidade em espaços normados quaisquer . A terceira seção trata essecialmente do teorema dos multiplicadores de Lagrange, que caracteriza os extremos locais de funcionais sujeitos a restrições. Nas últimas seções aplicaremos o teorema dos multiplicadores de Lagrange no problema de extremidades fixas para obter a equação de Euler-Lagrange e posteriormente trataremos do problema de extremidades fixas com vínculos. As referências utilizadas nesse capítulo foram (TROUTMAN, 2012), (LI, 2012) , (FALSET, 2007), (SAGAN, 2012), (NACHBIN, 1976) e (BAGGETT, 1992).

Iniciamos o capítulo 2 fazendo na primeira seção uma breve exposição do problema de Sturm-Liouville cujas principais referências utilizadas foram (SOTOMAYOR, 1979) e (CAVALHEIRO, 2012), faremos um tratamento sucinto com o objetivo de entender seu principais aspectos para posteriormente aplicarmos o cálculo variacional dentro da teoria. Nessa aplicação veremos que todo problema de Sturm-Liouville possui um funcional associado, chamado quociente de Rayleigh, o qual desempenha papel fundamental para deduzir informações a respeito dos autovalores. É através do quociente de Rayleigh que o problema de Sturm-Liouville pode ser tratado de modo variacional, achar um autovalor para o operador de Sturm-Liouville equivale a minimizar o quociente de Rayleigh em um espaço com extremidades fixas acrescido de vínculos. Finalmente, vamos obter o princípio minimax de Courant para comparar os autovalores de dois problemas de Sturm-Liouville distintos e com isto conseguir estimativas para os autovalores.

2 Cálculo variacional

Os problemas que envolvem máximos e mínimos de funções são conhecidos na literatura por problemas variacionais. Um exemplo típico é dado como se segue:

Dado dois pontos (t_0, y_0) e (t_1, y_1) no plano, determinar a curva de menor comprimento que os une. Suponhamos que as curvas em questão possuem parametrizações do tipo

$$y : [t_0, t_1] \longrightarrow \mathbb{R};\ y(t_0) = y_0,\ y(t_1) = y_1,$$

onde y é função continuamente diferenciável. Obtemos seu comprimento pela fórmula

$$J(y) = \int_a^b \sqrt{1 + y'(t)^2} dt.$$

Logo, o problema pode ser formulado como:

$$\begin{cases} \text{minimizar } J(y) = \int_a^b \sqrt{1 + y'(t)^2} dt \\ \text{sujeito a} \\ y \in \{C^1[t_0, t_1] / \ y(t_0) = y_0,\ y(t_1) = y_1\}. \end{cases} \tag{2.0.1}$$

Uma característica comum em vários problemas variacionais, assim como no exemplo acima, é que a função a ser otimizada associa a cada elemento de uma classe de funções um certo número real. Neste ponto, podemos perceber que o cálculo variacional extrapola as fronteiras do cálculo usual.

2.1 Variação de Gâteaux e Extremos Locais

Sendo asssim, de acordo com as considerações iniciais, parece razoável começarmos com a seguinte definição:

Definição 2.1.1. Um **funcional** consiste de um mapa $J : D(J) \longrightarrow \mathbb{R}$ cujo domínio $D(J)$ é um subconjunto de um espaço vetorial H.

Exemplo 2.1.1. O mapa $J(y) = \int_a^b \operatorname{sen}^3(x) + y(x)^2 dx$ é um funcional definido no espaço $C[a, b]$ das funções contínuas em $[a, b]$.

Definição 2.1.2. Um funcional J é dito **linear** se $D(J)$ é todo o espaço H e J satisfaz a relação

$$J(ax + by) = aJ(x) + bJ(y)$$

para todo $a, b \in \mathbb{R}$ e todo $x, y \in H$.

Exemplo 2.1.2. A aplicação $J(\phi) = \phi'(0)$ define um funcional linear em $C^1[-1, 1]$ o espaço das funções continuamente diferenciáveis em $[-1, 1]$.

A seguir, apresentaremos uma definição que estende naturalmente a noção de derivada direcional do cálculo diferencial.

Definição 2.1.3. Um funcional J definido em um subconjunto aberto D de um espaço vetorial normado H é dito ter **variação de Gâteaux em um vetor** $x \in D$ se o limite

$$\delta J(x, v) = \lim_{\epsilon \to 0} \frac{J(x + \epsilon v) - J(x)}{\epsilon}$$

existe para toda direção v de H. Neste caso podemos definir o funcional $\delta J(x) : H \longrightarrow \mathbb{R}$ dado por $\delta J(x)(v) = \delta J(x, v)$, o qual é chamado **variação de Gâteaux de** J **em** x.

Observação 2.1.1. Notemos que a noção de limite aqui faz sentido uma vez que existe uma vizinhança V de 0 na qual a função $\epsilon \mapsto \frac{J(x+\epsilon v) - J(x)}{\epsilon}$ está bem definida em $V - \{0\}$. Com efeito, como D é aberto existe $\delta > 0$ tal que $\|y - x\| < \delta \implies y \in D$. Logo a vizinhança em questão é $V = (-\delta/\|v\|, \delta/\|v\|)$.

Observação 2.1.2. A variação de Gâteaux de J em x na direção v é a derivada de $J(x + \epsilon v)$ em relação à variável ϵ e avaliada em $\epsilon = 0$, isto é, $\delta J(x, v) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} J(x + \epsilon v)$.

Exemplo 2.1.3. Seja H um espaço normado e $J : H \longrightarrow \mathbb{R}$ um funcional definido por $J(x) = \|x\|^2$. Se $h \in H$ e $x = 0$ tem-se

$$\delta J(0, h) = \lim_{t \to 0} \frac{t^2 \|h\|^2}{t} = 0.$$

Portanto J tem variação de Gâteaux em 0 e a variação de Gâteaux de J em 0 é o funcional nulo.

Exemplo 2.1.4. Seja H um espaço normado e $J : H \longrightarrow \mathbb{R}$ um funcional definido por $J(x) = \|x\|$. Se $h \in H$ é não-nulo e $x = 0$ tem-se

$$\delta J(0, h) = \lim_{t \to 0} \frac{|t| \|h\|}{t}.$$

Como o limite anterior não existe, J não tem variação de Gâteaux em 0.

Para calcularmos a variação de Gâteuax, as propriedades abaixo são importantes:

Propocição 2.1.1. *As seguintes afirmações abaixo são verdadeiras:*

a) A variação de Gâteaux é linear nos funcionais, isto é, dados J,K funcionais definidos em um aberto D do espaço vetorial normado H e $a \in H$ tem-se:

$$\delta(J + aK)(x, v) = \delta J(x, v) + a\delta K(x, v).$$

b) (***Regra do produto***) *Sejam* K,L *funcionais definidos em um aberto* D *do espaço vetorial normado* H *e* $J(x) = K(x)L(x)$ *para todo* $x \in D$. *Se* K *e* L *possuem variações de Gâteaux em* x_0 *então* J *possui variação de Gâteaux em* x_0 *e é dada por*

$$\delta J(x_0, v) = \delta K(x_0, v)L(x_0) + \delta L(x_0, v)K(x_0).$$

c) (***Regra do quociente***) *Sejam* K L *funcionais definidos em um aberto* D *do espaço vetorial normado* H *e* $J(x) = \dfrac{K(x)}{L(x)}$ *para todo* $x \in D$. *Suponha que* L *não se anula em* D. *Se* K *e* L *possuem variações de Gâteuax em* $x_0 \in D$ *então* J *possui variação de Gâteuax e é dada por*

$$\delta J(x_0, v) = \frac{\delta K(x_0, v)L(x_0) - \delta L(x_0, v)K(x_0)}{L(x_0)^2}.$$

Demonstração. a) De fato, da observação (2.1.2) segue que

$$\begin{aligned} \delta(J + aK)(x, v) &= \frac{d}{d\epsilon}\Big|_{\epsilon=0}(J + aK)(x + \epsilon v) = \frac{d}{d\epsilon}\Big|_{\epsilon=0}(J(x + \epsilon v) + aK(x + \epsilon v)) \\ &= \frac{d}{d\epsilon}\Big|_{\epsilon=0} J(x + \epsilon v) + a\frac{d}{d\epsilon}\Big|_{\epsilon=0} K(x + \epsilon v) = \delta J(x, v) + a\delta K(x, v). \end{aligned}$$

b) Aplicando a regra do produto para funções reais de uma váriavel real, temos que

$$\begin{aligned} \delta J(x_0, v) &= \frac{d}{d\epsilon}\Big|_{\epsilon=0}(K(x_0 + \epsilon v)L(x_0 + \epsilon v)) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} K(x_0 + \epsilon v)L(x_0 + \epsilon v)\Big|_{\epsilon=0} \\ &+K(x_0 + \epsilon v)\Big|_{\epsilon=0}\frac{d}{d\epsilon}\Big|_{\epsilon=0} L(x_0 + \epsilon v) = \delta K(x_0, v)L(x_0) + K(x_0)\delta L(x_0, v). \end{aligned}$$

c) Analogamente, basta calcular a derivada

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0}\frac{K(x_0 + \epsilon v)}{L(x_0 + \epsilon v)}$$

através da regra do quociente para funções reais de uma variável real.

□

Definição 2.1.4. Um funcional J é **contínuo ao longo de cada direção em** x se

$$\lim_{\epsilon \to 0} J(x + \epsilon v) = J(x),$$

para todo $v \in H$.

Exemplo 2.1.5. O funcional $J : C[0, 1] \longrightarrow \mathbb{R}$, dado por $J(\phi) = \phi(0)$, tem variação de Gâteaux em cada $\phi \in C[0, 1]$. A variação pode ser calculada facilmente e é dada por $\delta J(\phi, \Delta\phi) = \Delta\phi(0)$. Se considerarmos $C[0, 1]$ com a norma L^2: $\|f\|_2 = \sqrt{\int_0^1 f(x)^2 dx}$, concluimos que J é descontínua em todo seu domínio. Com efeito, fixado $\phi \in C[0,1]$, construamos a sequência

$$x_n(t) = \begin{cases} 1, & \text{se} \quad 0 \leq t \leq \frac{1}{n} \\ \sqrt{2-nt}, & \text{se} \quad \frac{1}{n} \leq t \leq \frac{2}{n} \\ 0, & \text{se} \quad \frac{2}{n} \leq t \leq 1 \end{cases}$$

Note que $\|x_n\| = \frac{3}{2n}$. Daí $\psi_n = \phi + x_n \to \phi$, mas $J(\psi_n) \nrightarrow J(\phi)$ pois $J(\psi_n) - J(\phi) = 1$ para todo $n \in \mathbb{N}$.

No entanto, J é contínuo ao longo de cada direção em todo $x \in C[0,1]$. Isto acontece devido à observação seguinte:

Observação 2.1.3. Se J tem variação de Gâteaux em x então J é contínuo ao longo de cada direção em x. De fato, a função $J(x+\epsilon v)$ é diferenciável em $\epsilon = 0$ logo contínua em $\epsilon = 0$.

Definição 2.1.5. Seja $J : D \subseteq H \longrightarrow \mathbb{R}$ um funcional definido em um subconjunto D do espaço vetorial normado H. O vetor $x^* \in D$ é dito um **vetor máximo (mínimo) em D para** J se $J(x) \leq J(x^*)$ $(J(x) \geq J(x^*))$ para todo $x \in D$. Quando isto acontece dizemos que x^* é um **vetor extremo em D para** J. O vetor $x^* \in D$ é dito **um vetor máximo local (mínimo local) em D para** J se existe $\rho > 0$ tal que $J(x) \leq J(x^*)$ $(J(x) \geq J(x^*))$ para todo $x \in D \cap B(x^*, \rho)$, onde $B(x^*, \rho) = \{x \in H : \|x - x^*\| < \rho\}$ é a bola centrada em x^* de raio ρ.

Nós dizemos que $x^* \in D$ é um **vetor extremo local em D para** J se x^* é um vetor máximo local ou mínimo local em D para J. Neste caso, dizemos que J tem um **extremo local em** x^* e $J(x^*)$ é um **valor extremo local para J em** D.

A variação de Gâteaux tem papel fundamental na proposição a seguir que nos dá uma condição necessária para um extremo local.

Propocição 2.1.2. *Seja $J : D \subseteq H \longrightarrow \mathbb{R}$ um funcional definido no aberto D do espaço vetorial normado H. Se x^* é um extremo local para J em D e J tem variação de Gâteaux em x^* então a variação de Gâteaux de J em x^* se anula, isto é, $\delta J(x^*, v) = 0$ para todo $v \in H$.*

Demonstração. Sem perda de generalidade, podemos assumir que x^* é um mínimo local em D para J. Dado $v \in H$, temos

$$\frac{J(x^* + \epsilon v) - J(x^*)}{\epsilon} \geq 0,$$

para todo $\epsilon > 0$ suficientemente pequeno. Como J tem variação de Gâteaux em x^* segue que

$$\lim_{\epsilon \to 0^+} \frac{J(x^* + \epsilon v) - J(x^*)}{\epsilon} \geq 0.$$

Analogamente,

$$\lim_{\epsilon \to 0^-} \frac{J(x^* + \epsilon v) - J(x^*)}{\epsilon} \leq 0.$$

Logo $\delta J(x^*, v) = \lim_{\epsilon \to 0} \dfrac{J(x^* + \epsilon v) - J(x^*)}{\epsilon} = 0$ e portanto a variação de Gâteuax em x^* é nula. $\square$

Exemplo 2.1.6. Considere o funcional

$$L(\phi) = \int_0^{\pi/2} \{2\phi(x)^3 + 9\text{sen}(x)\phi(x)^2 + 12\text{sen}^2(x)\phi(x) - \cos(x)\}dx,$$

definido em $C[0, \pi/2]$. Como

$$\begin{aligned} L(\phi + \epsilon\psi) &= \int_0^{\pi/2} \{2(\phi(x) + \epsilon\psi(x))^3 + 9\text{sen}(x)(\phi + \epsilon\psi(x))^2 \\ &\quad +12\text{sen}^2(x)(\phi(x) + \epsilon\psi(x)) - \cos(x)\}dx \\ &= L(\phi) + \int_0^{\pi/2} \{6\epsilon\phi(x)^2\psi(x) + 6\epsilon^2\phi(x)\psi(x)^2 + 2\epsilon^3\psi(x)^3 \\ &\quad +18\epsilon\text{sen}(x)\phi(x)\psi(x) + 9\epsilon^2\text{sen}(x)\psi(x)^2 + 12\epsilon\text{sen}^2(x)\psi(x)\}dx, \end{aligned}$$

tem-se

$$\begin{aligned} \frac{L(\phi + \epsilon\psi) - L(\phi)}{\epsilon} &= \int_0^{\pi/2} \{6\phi(x)^2\psi(x) + 18\text{sen}(x)\phi(x)\psi(x) + 12\text{sen}^2(x)\psi(x)\}dx \\ &\quad + \int_0^{\pi/2} \{6\epsilon\phi(x)\psi(x)^2 + 2\epsilon^2\psi(x)^3 + 9\epsilon\text{sen}(x)\psi(x)^2\}dx \end{aligned}$$

e consequentemente

$$\delta J(\phi, \psi) = \int_0^{\pi/2} \{6\phi(x)^2\psi(x) + 18\text{sen}(x)\phi(x)\psi(x) + 12\text{sen}^2(x)\psi(x)\}dx.$$

Se ϕ^* é um extremo local para L em $C[0, \pi/2]$, então pela proposição (2.1.2) concluimos que

$$6\phi^*(x)^2 + 18\text{sen}(x)\phi^*(x) + 12\text{sen}^2(x) = 0$$

e portanto

$$\phi^*(x) = -\text{sen}(x) \text{ ou } \phi^*(x) = -2\text{sen}(x).$$

Assim, temos dois candidatos a extremo local, os quais não são extremos globais. De fato,

$$\begin{aligned} L(\lambda) &= \int_0^{\pi/2} \{2\lambda^3 + 9\text{sen}(x)\lambda^2 + 12\text{sen}^2(x)\lambda - \cos(x)\}dx \\ &= \pi\lambda^3 + \left(9\int_0^{\pi/2} \text{sen}(x)dx\right)\lambda^2 + \left(12\int_0^{\pi/2} \text{sen}^2(x)dx\right)\lambda - \int_0^{\pi/2} \cos(x)dx \end{aligned}$$

Como

$$\lim_{\lambda \to \infty} L(\lambda) = \infty \text{ e } \lim_{\lambda \to -\infty} L(\lambda) = -\infty$$

segue que L não admite máximo absoluto nem mínimo absoluto.

2.2 Diferencial de Fréchet

Exclusivamente nesta secção vamos estender o ambiente dos funcionais para estudarmos as aplicações entre espaços vetoriais normados. Sejam E, F espaços vetoriais normados, e A um subconjunto aberto de E. Considere uma função $f : A \subset E \longrightarrow F$. Motivados pelo cálculo diferencial em dimensão finita, poderíamos nos perguntar se esta função pode ser aproximada localmente por uma função afim. Nesse contexto, surge o conceito de funções Fréchet diferenciáveis introduzido pelo matemático M. Fréchet.

Definição 2.2.1. A função $f : A \longrightarrow F$ é dita **Fréchet diferenciável** em x_0 se existe uma aplicação linear e contínua $L : E \longrightarrow F$ tal que

$$\lim_{h \to 0} \frac{\|R(x_0, h)\|}{\|h\|} = 0,$$

onde $R(x_0, h) = f(x_0 + h) - f(x_0) - L(h)$.

Propocição 2.2.1. *A aplicação linear e contínua L da definição anterior é unicamente determinada.*

Demonstração. Suponhamos que L e T são duas aplicações nas condições da definição anterior. Seja $x_0 + h \in A$, $h \neq 0$. Temos que

$$\frac{L(h) - T(h)}{\|h\|} = \frac{f(x_0 + h) - f(x_0) - T(h)}{\|h\|} - \frac{f(x_0 + h) - f(x_0) - L(h)}{\|h\|}.$$

Portanto

$$\lim_{h \to 0} \frac{L(h) - T(h)}{\|h\|} = 0.$$

Seja $S = L - T$. Então

$$\lim_{h \to 0} \frac{\|S(h)\|}{\|h\|} = 0.$$

Logo, para todo $\epsilon > 0$, existe $\delta > 0$ tal que $h \neq 0, \|h\| < \delta$ implicam

$$\|S(h)\| < \epsilon \|h\|.$$

Seja $x \in E$, $x \neq 0$. Façamos

$$h = \frac{\delta x}{2\|x\|}.$$

Então $\|h\| = \dfrac{\delta}{2}$, logo $\dfrac{\|S(h)\|}{\|h\|} < \epsilon$, o que implica $\|S(x)\| < \epsilon \|x\|$. Assim, $\|S\| < \epsilon$ para todo $\epsilon > 0$, onde

$$\|S\| = \ \sup\{\|S(x)\| \ : \ x \in E, \ \|x\| \leq 1\}.$$

Portanto, $\|S\| = 0$. Daí $S = 0$, isto é, $L = T$. □

A unicidade obtida pela proposição anterior nos permite nomear a aplicação L.

Definição 2.2.2. A aplicação linear e contínua L da definição (2.2.1) é chamada de **diferencial de Fréchet** da função f em x_0, a qual será representada por $Df(x_0)$.

Definição 2.2.3. Sejam E e F espaços vetoriais normados. Uma aplicação $T : E \longrightarrow F$ é dita **afim** quando existem $b \in F$ e uma aplicação linear contínua $L : E \longrightarrow F$ tais que, para todo $x \in E$,

$$T(x) = b + L(x).$$

Definição 2.2.4. Sendo E e F espaços vetoriais normados, $A \subseteq E$ aberto e $f : A \longrightarrow F$ Fréchet diferenciável em $x_0 \in A$, diz-se que a aplicação afim contínua dada por

$$x \in E \mapsto f(x_0) + Df(x_0)(x - x_0) \in F$$

é **tangente** a f em x_0.

Como é de se esperar , Fréchet diferenciabilidade implica continuidade:

Propocição 2.2.2. *Se $f : A \subset E \longrightarrow F$ é Fréchet diferenciável em x_0 então f é contínua em x_0.*

Demonstração. Pela definição (2.2.1),

$$f(x_0 + h) - f(x_0) = Df(x_0)(h) + \frac{f(x_0 + h) - f(x_0) - Df(x_0)(h)}{\|h\|}\|h\|.$$

Como

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0) - Df(x_0)(h)}{\|h\|} = 0$$

e $Df(x_0)$ é contínua segue que

$$\lim_{h\to 0}\{f(x_0 + h) - f(x_0)\} = 0.$$

□

Agora provaremos algumas regras básicas de diferenciabilidade e em seguida a regra da cadeia para a composição de duas funções Fréchet diferenciáveis.

Teoreoma 2.2.1. *Sejam E e F espaços vetoriais normados e A um subconjunto aberto de E.*

a) Seja $f : A \longrightarrow F$ uma função constante, isto é, $f(x) = b$ para algum $b \in F$. Então ,para cada $x \in A$, f é Fréchet diferenciável em x e $Df(x)$ é a aplicação nula.

b) Seja $f : E \longrightarrow F$ uma aplicação linear e contínua. Então, para cada $x \in E$, f é Fréchet diferenciável em x e $Df(x) = f$.

c) *Suponha que $f : A \longrightarrow F$ e $g : A \longrightarrow F$ são funções Fréchet diferenciáveis em x. Então a função $f + g$ é diferenciável em x e $D(f + g)(x) = Df(x) + Dg(x)$.*

d) *Se $f : A \longrightarrow F$ é Fréchet diferenciável em x e c é um escalar então $g = cf$ é Fréchet diferenciável em x e $Dg(x) = cDf(x)$.*

Demonstração. a) Fixe $x \in A$ e considere a função nula $o : A \longrightarrow F$. Faça

$$\mathcal{E}(x, h) = \frac{f(x + h) - f(x) - o(h)}{\|h\|},$$

e em seguida note que $\mathcal{E}(x, h)$ é nula para qualquer h, portanto $\lim_{h \to 0} \mathcal{E}(x, h) = 0$. Assim concluimos que f é Fréchet diferenciável em x e, por unicidade, $Df(x) = 0$.

b) Fixe $x \in A$. Seja

$$\mathcal{E}(x, h) = \frac{f(x + h) - f(x) - f(h)}{\|h\|}.$$

Então, pela linearidade de f, temos $\mathcal{E}(x, h) = 0$ para todo $h \in E$. Isto nos diz que:

$$\lim_{h \to 0} \mathcal{E}(x, h) = 0.$$

Como f é linear e contínua, segue que f é Fréchet diferenciável em x e $Df(x) = f$.

c) Façamos

$$\mathcal{E}_f(x, h) = \frac{f(x + h) - f(x) - Df(x)(h)}{\|h\|}, \; \mathcal{E}_g(x, h) = \frac{g(x + h) - g(x) - Dg(x)(h)}{\|h\|}.$$

Então temos $\mathcal{E}_{f+g}(x, h) = \mathcal{E}_f(x, h) + \mathcal{E}_g(x, h)$, onde

$$\mathcal{E}_{f+g}(x, h) = \frac{(f + g)(x + h) - (f + g)(x) - (Df(x) + Dg(x))(h)}{\|h\|}.$$

Note que $\lim_{h \to 0} \mathcal{E}_f(x, h) = \lim_{h \to 0} \mathcal{E}_g(x, h) = 0$ implica em $\lim_{h \to 0} \mathcal{E}_{f+g}(x, h) = 0$. Como $Df(x) + Dg(x)$ é uma aplicação linear contínua temos que $f + g$ é Fréchet diferenciável em x e $D(f + g)(x) = Df(x) + Dg(x)$.

d) Façamos

$$\mathcal{E}_{cf}(x, h) = \frac{(cf)(x + h) - (cf)(x) - (cDf(x))(h)}{\|h\|}.$$

Como $\mathcal{E}_{cf}(x, h) = c\mathcal{E}_f(x, h)$, temos $\lim_{h \to 0} \mathcal{E}_{cf}(x, h) = 0$. Portanto cf é Fréchet diferenciável com diferencial de Fréchet $D(cf)(x) = cDf(x)$.

□

Teoreoma 2.2.2. ***(Regra da cadeia)*** *Sejam E, F e G espaços vetoriais normados tais que A e B são abertos de E e F, respectivamente. Suponhamos que $g(A) \subseteq B$. Se $g : A \longrightarrow F$ é Fréchet diferenciável em x_0 e $f : B \longrightarrow G$ é Fréchet diferenciável em $g(x_0)$ então $f \circ g$ é Fréchet diferenciável em x_0 e $D(f \circ g)(x_0) = Df(g(x_0)) \circ Dg(x_0)$.*

Demonstração. Por definição, podemos escrever

$$\begin{aligned} f(g(x_0+h)) &= f(g(x_0)+Dg(x_0)(h)+R_g(x_0,h)) \\ &= f(g(x_0))+Df(g(x_0))(Dg(x_0)(h)+R_g(x_0,h)) \\ &\quad +R_f(g(x_0),Dg(x_0)(h)+R_g(x_0,h)). \end{aligned}$$

Disto seque que

$$f(g(x_0+h)) = f(g(x_0))+Df(g(x_0))(Dg(x_0)(h))+R(x_0,h),$$

onde

$$R(x_0,h) = Df(g(x_0))(R_g(x_0,h))+R_f(g(x_0),Dg(x_0)(h)+R_g(x_0,h)).$$

Precisamos mostrar que $\lim_{h\to 0}\dfrac{\|R(x_0,h)\|}{\|h\|}=0$. Observemos que

$$\lim_{h\to 0}\frac{\|R_g(x_0,h)\|}{\|h\|}=0 \tag{2.2.1}$$

e a continuidade de $Df(g(x_0))$ implicam

$$\lim_{h\to 0}\frac{\|Df(g(x_0))(R_g(x_0,h))\|}{\|h\|}=0.$$

Além disso,

$$\frac{\|R_f(g(x_0),Dg(x_0)(h)+R_g(x_0,h))\|}{\|h\|} =$$
$$\frac{\|R_f(g(x_0),Dg(x_0)(h)+R_g(x_0,h))\|}{\|Dg(x_0)(h)+R_g(x_0,h))\|}\frac{\|Dg(x_0)(h)+R_g(x_0,h))\|}{\|h\|} \quad \longrightarrow 0$$

quando $h\to 0$. De fato, consideremos a desigualdade

$$\begin{aligned} \frac{\|Dg(x_0)(h)+R_g(x_0,h)\|}{\|h\|} &\leq \frac{\|Dg(x_0)(h)\|}{\|h\|}+\frac{\|R_g(x_0,h)\|}{\|h\|} \\ &\leq \|Dg(x_0)\|+\frac{\|R_g(x_0,h)\|}{\|h\|}. \end{aligned} \tag{2.2.2}$$

A afirmação (2.2.1) nos permite limitar o lado direito de (2.2.2) numa vizinhança de 0. Logo

$$\frac{\|Dg(x_0)(h)+R_g(x_0,h)\|}{\|h\|}$$

é limitada numa vizinhança de 0. Consequentemente,

$$\lim_{h\to 0}\frac{\|R_f(g(x_0),Dg(x_0)(h)+R_g(x_0,h))\|}{\|Dg(x_0)(h)+R_g(x_0,h))\|}=0$$

implica no resultado. Portanto $\lim_{h\to 0}\dfrac{\|R(x_0,h)\|}{\|h\|}=0$.

□

Teoreoma 2.2.3. *Se um funcional $J : D \longrightarrow \mathbb{R}$ é Fréchet diferenciável em x_0 então J possui variação de Gâteaux em x_0 e é dada pela diferencial de Fréchet em x_0, isto é,*

$$\delta J(x_0, v) = DJ(x_0)(v).$$

Demonstração. Dado $v \in H$, temos

$$\frac{J(x_0 + \epsilon v) - J(x_0)}{\epsilon} = \frac{DJ(x_0)(\epsilon v) + R(x_0, \epsilon v)\|\epsilon v\|}{\epsilon}$$

para todo $\epsilon \neq 0$. A linearidade de $DJ(x_0)$ implica em

$$\frac{J(x_0 + \epsilon v) - J(x_0)}{\epsilon} = DJ(x_0)(v) + R(x_0, \epsilon v)\|v\|\frac{|\epsilon|}{\epsilon}.$$

Como $\lim_{\epsilon \to 0} R(x_0, \epsilon v) = 0$ temos que

$$|R(x_0, \epsilon v)\|v\|\frac{|\epsilon|}{\epsilon}| = |R(x_0, \epsilon v)|\|v\| \longrightarrow 0$$

quando $\epsilon \to 0$. Portanto $\lim_{\epsilon \to 0} \dfrac{J(x_0 + \epsilon v) - J(x_0)}{\epsilon} = Df(x_0)(v)$, isto é,

$$\delta J(x_0, v) = DJ(x_0)(v).$$

□

Nem todo funcional com variação de Gâteaux num determinado ponto é Fréchet diferenciável neste ponto. Isto pode ser verificado no exemplo (2.1.5) pois, como vimos, considerando $C[0,1]$ com a norma L^2, J é descontínuo em todo o domínio. Logo não é Fréchet diferenciável em nenhum ponto, apesar de possuir variação de Gâteaux em todo seu domínio. No entanto, sob certas condições adicionais, a recíproca do teorema anterior torna-se válida. Isto é o que diz o teorema abaixo:

Teoreoma 2.2.4. *Seja J um funcional definido em um aberto D do espaço vetorial normado H que possui variação de Gâteaux numa vizinhança $B(x_0, r)$ de x_0. Suponha que as condições abaixo são satisfeitas:*

a) A variação de Gâteaux $\delta J(x_0)$ é linear e contínua.

b) Quando $x \to x_0$, $\delta J(x, v) \longrightarrow \delta J(x_0, v)$ uniformemente para $v \in S = \{v \in H : \|v\| = 1\}$.

Então J é Fréchet diferenciável em x_0.

Demonstração. Por hipótese existe uma aplicação linear e contínua L tal que $\delta J(x_0, v) = L(v)$. Inicialmente, observemos que $B(x_0, r) = \{x_0 + tv / \ 0 \leq t < r, v \in S\}$. Fixado $v \in S$, definamos a função $f(t) = J(x_0 + tv)$, $t \in (-r, r)$. Provemos que f é diferenciável em $(-r, r)$. De fato, dado $t \in (-r, r)$, temos

$$\frac{f(t+s) - f(t)}{s} = \frac{J(x_0 + (t+s)v) - J(x_0 + tv)}{h} \longrightarrow \delta J(x_0 + tv, v)$$

quando $s \to 0$. Assim, $f'(t) = \delta J(x_t, v)$, onde $x_t = x_0 + tv$.

Seja $x \in B(x_0, r)$. Então existem $t \in [0, r)$ e $v \in S$ tais que $x = x_0 + tv$. Daí, $J(x) - J(x_0) = J(x_0 + tv) - J(x_0) = f(t) - f(0)$. Como f é diferenciável em $[0, t]$ segue, pelo teoreoma do valor médio, que $f(t) - f(0) = f'(t_0)t$ para algum $t_0 \in (0, t)$. Da linearidade de L obtemos $J(x) - J(x_0) - L(x - x_0) = f'(t_0)t - tL(v) = [\delta J(x_{t_0}, v) - \delta J(x_0, v)]t$. Notemos também que

$$\left| \frac{J(x) - J(x_0) - L(x - x_0)}{\|x - x_0\|} \right| = |\delta J(x_{t_0}, v) - \delta J(x_0, v)|, \quad x \neq x_0.$$

Então, dado $\epsilon > 0$, a condição b) nos informa que existe $\rho > 0$ tal que $0 < \|y - x_0\| < \rho$ implica

$$|\delta J(y, w) - \delta J(x_0, w)| < \epsilon, \quad \forall\ w \in S.$$

Assim,

$$\begin{aligned} 0 < \|x - x_0\| < \rho \quad &\Rightarrow \quad 0 < \|x_{t_0} - x_0\| < \rho \\ &\Rightarrow \quad |\delta J(x_{t_0}, v) - \delta J(x_0, v)| < \epsilon \\ &\Rightarrow \quad \left| \frac{J(x) - J(x_0) - L(x - x_0)}{\|x - x_0\|} \right| < \epsilon, \end{aligned}$$

e isto prova que J é Fréchet diferenciável em x_0 □

Exemplo 2.2.1. A função do exemplo (2.1.1)

$$J(y) = \int_a^b \{\text{sen}^3(x) + y(x)^2\} dx$$

está definida em $C[a, b]$. Consideremos a norma

$$\|y\|_{C[a,b]} = \max\{|y(x)| : x \in [a, b]\},$$

calculando a variação de J obtemos

$$\begin{aligned} \delta J(y, v) &= \lim_{t \to 0} \frac{\int_a^b \text{sen}^3 x + (y(x) + tv(x))^2\, dx - \int_a^b \text{sen}^3 x + y(x)^2 dx}{t} \\ &= \lim_{t \to 0} \frac{\int_a^b 2ty(x)v(x) + t^2 v(x)^2 dx}{t} \\ &= 2\int_a^b y(x)v(x)dx + \lim_{t \to 0} t \int_a^b v(x)^2 dx \\ &= 2\int_a^b y(x)v(x)dx. \end{aligned}$$

Logo $\delta J(y)$ é linear. Assim, para estabelecer a continuidade de $\delta J(y_0)$ é suficiente mostrar a continuidade em 0. Mas como $\delta J(y_0, 0) = 0$,

$$\begin{aligned}
|\delta J(y_0, v) - \delta J(y_0, 0)| &= |J(y_0, v)| \\
&= 2\left|\int_a^b y_0(x)v(x)dx\right| \\
&\leq 2\int_a^b |y_0(x)|\,|v(x)|\,dx \\
&\leq 2\|y_0\|_{C[a,b]}\|v\|_{C[a,b]}(b-a).
\end{aligned}$$

Fazendo $v \to 0$ segue que o último termo $\longrightarrow 0$. Assim J satisfaz a condição $a)$ do teorema (2.2.4).

Para a condição $b)$ suponhamos que $u \in C[a,b]$, $\|u\|_{C[a,b]} = 1$. Estimando similarmente obtemos

$$\begin{aligned}
|\delta J(y, u) - \delta J(y_0, u)| &= 2\left|\int_a^b [y(x) - y_0(x)]u(x)dx\right| \\
&\leq 2\|y - y_0\|_{C[a,b]}\|u\|_{C[a,b]}(b-a) \\
&= 2\|y - y_0\|_{C[a,b]}(b-a). \qquad (2.2.3)
\end{aligned}$$

Observamos que o último termo $\longrightarrow 0$ quando $y \to y_0$ e é independente de u. Consequentemente o lado esquerdo de (2.2.3) $\longrightarrow 0$ uniformemente em u quando $\|u\|_{C[a,b]} = 1$. Portanto J é Fréchet diferenciável em cada $y_0 \in C[a,b]$.

2.3 Extremos com Restrições

Como vimos anteriomente, todo extremo local necessariamente anula a variação de Gâteaux. Para provar este fato usamos fortemente o fato do funcional estar definido em um conjunto aberto. De modo análogo se restringirmos este aberto ainda assim podemos caracterizar os extremos.

Nesta seção, trataremos desta questão ao apresentar o teorema dos multiplicadores de Lagrange. Mais precisamente, através deste teorema, faremos uma caracterização dos extremos locais de um funcional J quando restrito a um ou mais conjuntos de nível. Entendemos por conjunto de nível para um funcional K definido em D, um conjunto da forma $\{x \in D : K(x) = k\}$. No nosso contexto, esses conjuntos são chamados de **restrições**.

Seja K um funcional definido em um aberto $\mathcal{D}$ do espaço vetorial normado H. Dado $k \in \mathbb{R}$, denotamos $\{x \in \mathcal{D} : K(x) = k\}$ por $\mathcal{D}\,[K = k]$. Mais geralmente, $\mathcal{D}\,[K_i = k_i : i = 1, \ldots, m]$ representa a interseção $\mathcal{D}\,[K_1 = k_1] \cap \ldots \cap \mathcal{D}\,[K_m = k_m]$. Além disso, ao longo do texto suporemos que as restrições são não-vazias, isto é, $\mathcal{D}\,[K = k] \neq \emptyset$.

Exemplo 2.3.1. Se x^* é um extremo local em $\mathcal{D}\,[K = k]$ para J não significa que a variação $\delta J(x^*)$ se anula. De fato, considere

$$J(x) = x^2, \ K(x) = x^2 - 1, \ x \in H = \mathbb{R}.$$

Como $\mathcal{D}\,[K = 0] = \{-1, 1\}$ segue que $J(x) = 1$ para todo $x \in \mathcal{D}\,[K = 0]$. No entanto, a variação $\delta J(x, h) = 2xh$ não se anula em $\mathcal{D}\,[K = 0] = \{-1, 1\}$. Isto aconteceu, evidentemente, pelo fato de $\mathcal{D}\,[K = 0]$ não ser aberto.

O teorema dos multiplicadores de Lagrange necessita não somente da existência das variações dos funcionais envolvidos como também da continuidade das variações em cada direção. Este conceito é introduzido pela definição abaixo:

Definição 2.3.1. Se J é um funcional que tem variação de Gâteaux em um subconjunto aberto $\mathcal{D}$ do espaço vetorial normado H e para algum vetor $x_0 \in \mathcal{D}$ tem-se

$$\lim_{x \to x_0} \delta J(x, v) = \delta J(x_0, v)$$

para todo $v \in H$, então nós diremos que a variação de J é **fracamente contínua** em x_0.

Exemplo 2.3.2. Seja $K : C^1[a, b] \longrightarrow \mathbb{R}$ dado por $K(y) = y(x_0)$, onde x_0 é um ponto fixado em $[a, b]$. A variação de Gâteaux de K é fracamente contínua em todo ponto. Com efeito, note que $K(y + \epsilon h) = y(x_0) + \epsilon h(x_0)$. Logo $\delta K(y, h) = h(x_0)$, e obviamente, a variação é fracamente contínua em y.

Teoreoma 2.3.1. *(**Teorema dos multiplicadores de Lagrange**) Sejam $J, K_1, \ldots, K_m$ funcionais definidos em um aberto $\mathcal{D}$ do espaço vetorial normado H, tais que admitem variação fracamente contínua em $\mathcal{D}$. Suponhamos que x^* é um extremo local em $\mathcal{D}[K_i = k_i : i = 1, \ldots, m]$ para J. Então ao menos uma das possibilidades acontecem:*

i) Para todo $v_1, \ldots, v_m \in H$ tem-se

$$det(\delta K_i(x^*, v_j)) = 0. \tag{2.3.1}$$

ii) Existem $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ tais que

$$\delta J(x^*, v) = \sum_{i=1}^{m} \lambda_i \delta K_i(x^*, v) \tag{2.3.2}$$

para todo $v \in H$.

Demonstração. Suponhamos que a condição $i)$ é falsa, isto é, existem $\overline{v}_1, \ldots, \overline{v}_m \in H$ tais que

$$\det(\delta K_i(x^*, \overline{v}_j)) \neq 0.$$

Devemos mostrar que a condição *ii*) é válida. Para isto provaremos a afirmação a seguir:

$$\det \begin{bmatrix} \delta J(x^*, v) & \delta J(x^*, v_1) & \cdots & \delta J(x^*, v_m) \\ \delta K_1(x^*, v) & \delta K_1(x^*, v_1) & \cdots & \delta K_1(x^*, v_m) \\ \vdots & \vdots & \ddots & \vdots \\ \delta K_m(x^*, v) & \delta K_m(x^*, v_1) & \cdots & \delta K_m(x^*, v_m) \end{bmatrix} = 0 \tag{2.3.3}$$

para todo $v, v_1, \ldots, v_m \in H$.

Demonstremos por absurdo. Suponha que existem $v, v_1, \ldots, v_m \in H$ tais que o determinante acima é não-nulo. Como x^* é extremo local em $\mathcal{D}[K_i = k_i : i = 1, \ldots, m]$, existe $\rho > 0$ tal que x^* é um extremo em $B(x^*, \rho) \cap \mathcal{D}[K_i = k_i : i = 1, \ldots, m]$.

Considere a função

$$\begin{array}{rrcl} f: & \mathbb{R}^{m+1} & \longrightarrow & H \\ & (r, r_1, \ldots, r_m) & \longmapsto & x^* + rv + r_1 v_1 + \ldots + r_m v_m \end{array}$$

Note que f é uma transformação afim definida em um espaço de dimensão finita por isso f é contínua. Então existe uma vizinhança U de 0 tal que $f(U) \subseteq B(x^*, \rho)$.

Agora definimos a função $F : U \subseteq \mathbb{R}^{m+1} \longrightarrow \mathbb{R}^{m+1}$ cujas funções coordenadas são dadas por:

$$\begin{aligned} F_1(r, r_1 \ldots, r_m) &= J(x^* + rv + r_1 v_1 + \ldots + r_m v_m) \\ F_2(r, r_1 \ldots, r_m) &= K_1(x^* + rv + r_1 v_1 + \ldots + r_m v_m) \\ &\vdots \\ F_{m+1}(r, r_1 \ldots, r_m) &= K_m(x^* + rv + r_1 v_1 + \ldots + r_m v_m) \end{aligned}$$

E em seguida, vamos obter suas derivadas parcias:

$$\begin{aligned} D_1 F_1(r, r_1 \ldots, r_m) &= \lim_{\epsilon \to 0} \frac{F_1(r + \epsilon, r_1 \ldots, r_m) - F_1(r, r_1 \ldots, r_m)}{\epsilon} \\ &= \lim_{\epsilon \to 0} \{[J(x^* + rv + r_1 v_1 + \ldots + r_m v_m + \epsilon v) \\ &\quad - J(x^* + rv + r_1 v_1 + \ldots + r_m v_m)]\epsilon^{-1}\} \\ &= \delta J(x^* + rv + r_1 v_1 + \ldots + r_m v_m, v). \end{aligned}$$

De modo análogo, obtemos

$$\begin{aligned} D_2 F_1(r, r_1 \ldots, r_m) &= \delta J(x^* + rv + r_1 v_1 + \ldots + r_m v_m, v_1) \\ &\vdots \\ D_m F_1(r, r_1 \ldots, r_m) &= \delta J(x^* + rv + r_1 v_1 + \ldots + r_m v_m, v_{m-1}) \\ D_{m+1} F_1(r, r_1 \ldots, r_m) &= \delta J(x^* + rv + r_1 v_1 + \ldots + r_m v_m, v_m), \end{aligned}$$

e assim sucessivamente, até adquirirmos

$$\begin{aligned}
D_1F_{m+1}(r,r_1\ldots,r_m) &= \delta K_m(x^*+rv+r_1v_1+\ldots+r_mv_m,v)\\
&\vdots\\
D_mF_{m+1}(r,r_1\ldots,r_m) &= \delta K_m(x^*+rv+r_1v_1+\ldots+r_mv_m,v_{m-1})\\
D_{m+1}F_{m+1}(r,r_1\ldots,r_m) &= \delta K_m(x^*+rv+r_1v_1+\ldots+r_mv_m,v_m).
\end{aligned}$$

$J, K_1, \ldots, K_m$ possuem variações fracamente contínuas e a função f é contínua. Logo as derivadas parcias de F são composições de funções contínuas e portanto F é de classe C^1. A matriz jacobiana de F no ponto 0 tem determinante não-nulo, isto é,

$$\det JF(0) = \begin{vmatrix}
\delta J(x^*,v) & \delta J(x^*,v_1) & \cdots & \delta J(x^*,v_m)\\
\delta K_1(x^*,v) & \delta K_1(x^*,v_1) & \cdots & \delta K_1(x^*,v_m)\\
\vdots & \vdots & \ddots & \vdots\\
\delta K_m(x^*,v) & \delta K_m(x^*,v_1) & \cdots & \delta K_m(x^*,v_m)
\end{vmatrix} \neq 0.$$

Assim, aplicando o teorema da função inversa, existem abertos $V \subseteq U$ (vizinhança de 0) e W (vizinhança de $F(0)$) tais que $F : V \longrightarrow W$ é um difeomorfismo. Como $F(0,\ldots,0) = (J(x^*), k_1, \ldots, k_m) \in W$ e W é aberto, podemos encontrar $j_1 > J(x^*)$ e $j_2 < J(x^*)$ tais que $(j_1, k_1, \ldots, k_m)$ e $(j_2, k_1, \ldots, k_m) \in W$. Pela bijetividade de $F\Big|_V$ concluimos que existem

$$(j_1^*, r_1^*, \ldots, r_m^*), (j_2^{**}, r_1^{**}, \ldots, r_m^{**}) \in V$$

tais que

$$\begin{aligned}
F(j_1^*, r_1^*, \ldots, r_m^*) &= (j_1, k_1, \ldots, k_m)\\
F(j_2^{**}, r_1^{**}, \ldots, r_m^{**}) &= (j_2, k_1, \ldots, k_m).
\end{aligned}$$

Fazendo

$$\begin{aligned}
v^* &= (x^* + j_1^* v + r_1^* v_1 + \ldots + r_m^* v_m)\\
v^{**} &= (x^* + j_1^{**} v + r_1^{**} v_1 + \ldots + r_m^{**} v_m),
\end{aligned}$$

temos que $J(v^*) = j_1$, $J(v^{**}) = j_2$ com $v^*, v^{**} \in B(x^*,\rho) \cap D[K_i = k_i : i = 1,\ldots,m]$. Mas isto é uma contradição, pois x^* é um extremo em $B(x^*,\rho) \cap D[K_i = k_i : i = 1,\ldots,m]$.

Uma vez que provamos a afirmação (2.3.3), podemos voltar para nosso objetivo, que é mostrar a condição *ii*). Pois bem, calculando o determinante através da primeira coluna segue que (2.3.3) implica

$$\delta J(x^*,v)\det\left(\delta K_i(x^*,v_j)\right) + \sum_{i=1}^{m} \delta K_i(x^*,v)\mu_i(v_1,\ldots,v_m) = 0.$$

Observemos que se $\det(\delta K_i(x^*, v_j)) = 0$ nada podemos afirmar sobre a identidade acima. Agora, tomando $v_i = \overline{v}_i$, para cada $i \in \{1, \ldots, m\}$, temos para todo $v \in H$ que

$$\delta J(x^*, v) = \sum_{i=1}^{m} \lambda_i \delta K_i(x^*, v)$$

onde $\lambda_i = \dfrac{-\mu_i(\overline{v}_1, \ldots, \overline{v}_m)}{\det(\delta K_i(x^*, \overline{v}_j))}$. E com isto o teorema está demonstrado. □

Exemplo 2.3.3. Considere o problema de minimização do funcional

$$J(y) = \int_a^b \{\mathrm{sen}^3(x) + y(x)^2\} dx$$

sobre

$$D = \{y \in C[0,1] : \int_0^1 x[y(x)]^{4/3} dx = 1\},$$

em outras palavras queremos minimizar J sujeito a restrição

$$G(y) = \int_0^1 x[y(x)]^{4/3} dx = 1.$$

Nós sabemos pelo exemplo (2.2.1) que $\delta J(y, v)$ existe para todo $y, v \in C[a, b]$ e é dada por

$$\delta J(y, v) = 2 \int_0^1 y(x)v(x) dx.$$

Similarmente, $\delta G(y, v)$ existe para todo $y, v \in C[0, 1]$ e é dada por

$$\delta G(y, v) = \frac{4}{3} \int_0^1 x[y(x)]^{1/3} v(x) dx.$$

$\delta J(y, v)$ e $\delta G(y, v)$ são fracamente contínuas em $C[0, 1]$ com a norma $\| \ \|_{C[0,1]}$. Assim, pelo teorema dos multiplicadores de Lagrange, uma função $y_0 \in C[0, 1]$ que minimiza J em D deve satisfazer ou

a) $\delta G(y_0, w) = \dfrac{4}{3} \int_0^1 x y_0(x)^{1/3} w(x) dx = 0, \ \forall\ w \in C[0, 1]$,(e esta condição não é válida pois ela implicaria $x y_0(x)^{1/3} = 0, \forall\ x \in [0.1]$, e, por continuidade, $y_0(x) = 0$ para todo $x \in [0, 1]$, no entanto a função nula não pertence a D) ou

b) $\exists\ \lambda \in \mathbb{R}$ tal que

$$\delta J(y_0, v) + \lambda \delta G(y_0, v) = \int_0^1 [2y_0(x) + \frac{4}{3} \lambda x y_0(x)^{1/3}] v(x) dx = 0, \ \forall\ v \in C[0, 1].$$

A condição acima implica $2y_0(x) + \dfrac{4}{3}\lambda x y_0(x)^{1/3} = 0, \forall\ x \in [0, 1]$, e por continuidade $y_0 = 0$ (que não pertence a D) ou $y_0(x) = \pm \left(\frac{-2}{3}\lambda x\right)^{3/2}$. A constante $-\lambda > 0$ deve ser escolhida de modo a satisfazer $\int_0^1 x y_0(x)^{4/3} = 1$, que requer o valor $\lambda = -3$. Assim as únicas possibilidades para funções de mínimo em D para J são $y_0(x) = \pm(2x)^{3/2}$.

Consideremos o mesmo contexto do teorema dos multiplicadores de Lagrange. Mas, façamos agora $k_1, \ldots, k_m$ variar nos reais, sendo assim, $x^*(k_1, \ldots, k_n)$ é extremo local no conjunto de nível $\mathcal{D}[K_i = k_i : i = 1, \ldots, m]$. Então, nesse caso, com algumas condições adicionais, pode-se determinar os coeficientes $\lambda_1, \ldots, \lambda_n$. Para isso basta derivar parcialmente a composição

$$J\left(x^*(k_1, \ldots, k_m)\right).$$

Vejamos como isso acontece no teorema abaixo:

Teoreoma 2.3.2. *Sejam $J, K_1, \ldots, K_m$ funcionais definidos em um aberto D do espaço vetorial normado H. Suponha que as condições abaixo são satisfeitas:*

i) *$J, K_1, \ldots, K_m$ são Fréchet diferenciáveis e possuem variações fracamente contínuas em $\mathcal{D}$.*

ii) *$\mathcal{D}[K_i = k_i : i = 1, \ldots, m]$ é não-vazio para quaisquer $k_1, \ldots, k_m$ escolhidos.*

iii) *Para cada $k_1, \ldots, k_m$, tem-se um vetor extremo local $x^* = x^*(k_1, \ldots, k_m)$ em $\mathcal{D}[K_i = k_i : i = 1, \ldots, m]$ para J e o determinante (2.3.1) não é identicamente nulo. Além disso, assuma que a dependência de x^* em relação a $k_1, \ldots, k_m$ ocorre diferenciavelmente, isto é, a função $x^* = x^*(k_1, \ldots, k_m)$ é Fréchet diferenciável.*

Então

$$DJ(x^*)(v) = \sum_{i=1}^{m} \lambda_i DK_i(x^*)(v) \tag{2.3.4}$$

onde

$$\lambda_i = \frac{\partial}{\partial k_i} J\left(x^*(k_1, \ldots, k_m)\right) \tag{2.3.5}$$

e o símbolo $\dfrac{\partial}{\partial k_i}$ denota a derivada parcial em relação a variável k_i.

Demonstração. A partir do teorema dos multiplicadores de Lagrange concluimos que existem $\lambda_1, \ldots, \lambda m \in \mathbb{R}$ tais que

$$\delta J(x^*, v) = \sum_{i=1}^{m} \lambda_i \delta K_i(x^*, v).$$

Como os funcionais são Fréchet diferenciáveis, aplicando o teorema (2.2.3) podemos reescrever a equação (2.3.2) acima na forma

$$DJ(x^*)(v) = \sum_{i=1}^{m} \lambda_i DK_i(x^*)(v)$$

Agora provemos a igualdade (2.3.5). Novamente utilizando o teorema (2.2.3), a regra da cadeia e a expressão (2.3.4) obtemos, para cada $i \in \{1, \ldots, m\}$,

$$\begin{aligned}
\frac{\partial J}{\partial k_i} J\left(x^*(k_1, \ldots, k_m)\right) &= D(J \circ x^*)(k)(e_i) \\
&= DJ\left(x^*(k)\right)\left(Dx^*(k)(e_i)\right) \\
&= \sum_{j=1}^{m} \lambda_j DK_j(x^*(k))\left(Dx^*(k)(e_i)\right) \\
&= \sum_{j=1}^{m} \lambda_j D(K_j \circ x^*)(k)(e_i) \\
&= \sum_{j=1}^{m} \lambda_j \frac{\partial}{\partial k_i}(K_j \circ x^*)(k). \qquad (2.3.6)
\end{aligned}$$

onde $k = (k_1, \ldots, k_m)$ e $\{e_1, \ldots, e_m\}$ é a base canônica do $\mathbb{R}^m$. Afirmamos que

$$\frac{\partial}{\partial k_i}(K_j \circ x^*)(k) = \begin{cases} 1, & se \quad j = i, \\ 0, & se \quad j \neq i. \end{cases}$$

De fato, note que $x^*(k) \in \mathcal{D}[K_i = k_i : i = 1, \ldots, m]$. Consequentemente,

$$K_j\left(x^*(k + \epsilon e_i)\right) = \begin{cases} k_j + \epsilon, & se \quad j = i, \\ k_j, & se \quad j \neq i. \end{cases}$$

Disto seque que

$$\begin{aligned}
\frac{\partial}{\partial k_i}\left(K_j \circ x^*\right)(k) &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \left(K_j \circ x^*\right)(k + \epsilon e_i). \\
&= \begin{cases} 1, & se \quad j = i, \\ 0, & se \quad j \neq i. \end{cases}
\end{aligned}$$

Agora, substituindo os valores das derivadas parciais acima na equação (2.3.6) prova-se (2.3.5). □

2.4 O Problema de Extremidades Fixas e a Equação de Euler-Lagrange

Um problema clássico do cálculo variacional pode ser formulado como se segue: Seja $f : [a, b] \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ uma função com derivadas parciais f_y, f_z contínuas. Dada uma função $y \in C^1[a, b]$, definimos o funcional integral:

$$F(y) = \int_a^b f[y(x)]dx, \qquad (2.4.1)$$

onde $f[y(x)]$ denota a composição $f(x, y(x), y'(x))$. Considere as funções de $C^1[a, b]$ que satisfazem as condições de fronteira

$$y(a) = a_1, \; y(b) = b_1. \qquad (2.4.2)$$

Entre todas estas funções, quais atingem um valor extremo local para F?

Antes de iniciarmos o estudo deste problema, precisamos mostrar que o funcional (2.4.1) possui variação de Gâteaux em $C^1[a,b]$. Para alcançar este objetivo necessitamos do seguinte resultado básico:

Teoreoma 2.4.1 (Leibniz). *Se $f = f(x,y)$ e a derivada parcial f_y é contínua em $[a,b]\times[u,v]$ então*

$$g(y) = \int_a^b f(x,y)dx$$

é C^1 no intervalo (u,v) com derivada

$$g'(y) = \frac{d}{dy}\int_a^b f(x,y)dx = \int_a^b f_y(x,y)dx.$$

Demonstração. Veja Troutman (2012, pág 426). □

O cálculo da variação é obtido derivando $F(y+\epsilon v) = \int_a^b f\left[(y+\epsilon v)(x)\right]dx$ com relação a ϵ sob o sinal da integral. Então pela regra da cadeia, para todo x, y, z, w:

$$\frac{\partial}{\partial\epsilon}f(x,y+\epsilon v,z+\epsilon w) = f_y(x,y+\epsilon v,z+\epsilon w)v + f_z(x,y+\epsilon v,z+\epsilon w)w.$$

Logo

$$\begin{aligned}\frac{\partial}{\partial\epsilon}f\left[(y+\epsilon v)(x)\right] &= \frac{\partial}{\partial\epsilon}f\left(x,y(x)+\epsilon v(x),y'(x)+\epsilon v'(x)\right)\\ &= f_y[(y+\epsilon v)(x)]v(x) + f_z[(y+\epsilon v)(x)]v'(x)\end{aligned}$$

é contínua em $[a,b]\times\mathbb{R}$. Assim, pelo teorema de Leibniz, temos que

$$\frac{\partial}{\partial\epsilon}F(y+\epsilon v) = \int_a^b f_y[(y+\epsilon v)(x)]v(x) + f_z[(y+\epsilon v)(x)]v'(x)dx,$$

de modo que, fazendo $\epsilon = 0$, obtemos

$$\delta F(y,v) = \int_a^b \{f_y[y(x)]v(x) + f_z[y(x)]v'(x)\}dx.$$

Portanto, F admite variação de Gâteaux. Ademais, F é Fréchet diferenciável e possui variação fracamente contínua.

Propocição 2.4.1. *Se $f = f(x,y,z)$, f_y e $f_z \in C([a,b]\times\mathbb{R}^2)$, então*

$$F(y) = \int_a^b f(x,y(x),y'(x))dx$$

é Fréchet diferenciável e tem variação fracamente contínua em cada $y_0 \in C^1[a,b]$ na norma $\|y\|_M = máx\{|y(x)| + |y'(x)| : x \in [a,b]\}$.

Demonstração. Como calculamos anteriormente, a variação de Gâteuax de F é dada por

$$\delta F(y,v) = \int_a^b f_y[y(x)]v(x) + f_z[y(x)]v'(x)dx.$$

Nessa demonstração vamos usar o teorema 2.2.4. Fixado $y_0 \in C[a,b]$, comecemos por verificar a condição a). É fácil ver que a diferencial de Gâteaux é linear, então só precisamos mostrar a continuidade. Na verdade, mais que contínua, a diferencial é Lipschtziana. De fato,

$$\begin{aligned}
|\delta F(y,w) - \delta F(y,v)| &\leq \int_a^b |f_y[y(x)](w(x)-v(x))|\,dx + \int_a^b |f_z[y(x)](w'(x)-v'(x))|\,dx \\
&\leq \int_a^b |f_y[y(x)]|\,\|w-v\|_M + |f_z[y(x)]|\,\|w-v\|_M dx \\
&= \int_a^b (|f_y[y(x)]| + |f_z[y(x)]|)\,\|w-v\|_M dx \\
&= K\|w-v\|_M,
\end{aligned}$$

onde $K = \int_a^b |f_y[y(x)]| + |f_z[y(x)]|dx$.

Para provarmos a condição b) em y_0, note que

$$\begin{aligned}
|\delta F(y,v) - \delta F(y_0,v)| &= \left|\int_a^b (f_y[y(x)] - f_y[y_0(x)])\,v(x) + (f_z[y(x)] - f_z[y_0(x)])\,v'(x)dx\right| \\
&\leq \int_a^b [|f_y[y(x)] - f_y[y_0(x)]| + |f_z[y(x)] - f_z[y_0(x)]|]\,\|v\|_M dx.
\end{aligned}$$

Considere as funções abaixo:

$$\begin{aligned}
F_1 :\ C^1[a,b] &\longrightarrow C[a,b] \\
y &\longmapsto F_1(x) = f_y[y(x)]
\end{aligned}$$

$$\begin{aligned}
F_2 :\ C^1[a,b] &\longrightarrow C[a,b] \\
y &\longmapsto F_2(x) = f_z[y(x)].
\end{aligned}$$

Afirmamos que F_1 e F_2 são contínuas em y_0. Com efeito, faça $c = 1 + \|y_0\|_M$. Então

$$\begin{aligned}
\|y-y_0\|_M \leq 1 &\Rightarrow \|y\|_M \leq \|y-y_0\| + \|y\|_M \leq 1 + \|y_0\|_M = c \\
&\Rightarrow |y(x)|, |y'(x)|, |y_0(x)|, |y_0'(x)| \leq c,\ \forall\ x \in [a,b].
\end{aligned}$$

f_y é contínua no compacto $[a,b]\times[-c,c]^2$ logo f_y é uniformemente contínua em $[a,b]\times[-c,c]^2$. Consequentemente, dado $\epsilon > 0$ existe $r > 0$ tal que

$$|f_y(x,y,z) - f_y(x,y_1,z_1)| < \epsilon$$

desde que

$$|(x,y,z) - (x,y_1,z_1)| < r;\ |y|, |y_1|, |z|, |z_1| \leq c,\ x \in [a,b].$$

Agora tome $\delta =$mín$\{1, r\}$. Daí,

$$\begin{aligned} \|y - y_0\| < \delta \;\; &\Rightarrow \;\; |y(x)|, |y'(x)|, |y_0(x)|, |y_0'(x)| \leq c, \;\; |(x, y(x), y'(x)) - (x, y_0(x), y_0'(x)| < r \\ &\quad\;\; \forall \; x \in [a, b] \\ &\Rightarrow \;\; |f_y[y(x)] - f_y[y_0(x)]| < \epsilon, \;\; \forall \; x \in [a, b] \\ &\Rightarrow \;\; \|F_1(y) - F_1(y_0)\| < \epsilon. \end{aligned}$$

E isto prova a continuidade de F_1 em y_0. A prova para F_2 é análoga.

Assim dado $\epsilon > 0$, existe $r > 0$ tal que

$$\begin{aligned} \|y - y_0\| < r \;\; &\Rightarrow \;\; \|F_1(y) - F_1(y_0)\|, \|F_2(y) - F_2(y_0)\| < \epsilon \\ &\Rightarrow \;\; |\delta F(y, v) - \delta F(y_0, v)| \leq (\epsilon + \epsilon)(b - a)\|v\|_M \\ &\Rightarrow \;\; |\delta F(y, v) - \delta F(y_0, v)| \leq 2\epsilon(b - a), \;\; \forall \; \|v\|_M = 1. \end{aligned}$$

E isto prova a condição b), o que implica em particular que a variação de F é fracamente contínua em y_0. E por sua vez, pelo teorema (2.2.4), F é Fréchet diferenciável em y_0. □

A fim de encontrarmos uma condição necessária para as funções que estamos à procura, faremos uso de alguns resultados básicos do cálculo elementar, os quais impulsionaram fortemente o desenvolvimento do cálculo variacional.

Lema 2.4.1. ***(du-Bois-Reymond)*** *Se $h \in C[a, b]$ e $\int_a^b h(x)v'(x)dx = 0$, $\forall \; v \in C_0^1[a, b]$, onde*

$$C_0^1[a, b] = \{v \in C^1[a, b] : v(a) = v(b) = 0\},$$

então h é constante em $[a, b]$.

Demonstração. Dada uma constante c, definimos a função

$$v(x) = \int_a^x (h(t) - c)dt.$$

Pelo teoreoma fundamental do cálculo v é derivável com $v'(x) = h(x) - c$, e satisfaz $v(a) = 0$. Disto segue que $v \in C_0^1[a, b]$ se $v(b) = 0$. Fazendo

$$c = \frac{1}{b - a} \int_a^b h(t)dt,$$

temos $v(b) = 0$. Então utilizando a hipótese temos que

$$\begin{aligned} 0 \;\; &\leq \;\; \int_a^b (h(t) - c)^2 dt = \int_a^b (h(t) - c)v'(t)dt \\ &= \;\; \int_a^b h(t)v'(t)dt - cv\Big|_a^b \\ &= \;\; 0 \end{aligned}$$

Portanto $(h(t) - c)^2 = 0$, isto é, $h \equiv c$. □

Propocição 2.4.2. *Se $g,h \in C[a,b]$ e $\int_a^b (g(x)v(x) + h(x)v'(x))dx = 0$, $\forall\ v \in C_0^1[a,b]$, então $h \in C^1[a,b]$ e $h' = g$.*

Demonstração. Seja $G(x) = \int_a^b g(t)dt$. Então $G \in C^1[a,b]$ e $G' = g$. Agora, integrando por partes o primeiro termo da integral, obtemos

$$\begin{aligned}\int_a^b \{g(x)v(x) + h(x)v'(x)\}dx &= G(x)v(x)\Big|_a^b - \int_a^b G(x)v'(x)dx + \int_a^b h(x)v'(x)dx \\ &= \int_a^b (h(x) - G(x))\, v'(x)dx + G(x)v(x)\Big|_a^b.\end{aligned}$$

Daí,

$$\int_a^b (h(x) - G(x))\, v'(x)dx = 0,$$

para todo $v \in C_0^1[a,b]$, e pelo lema (2.4.1) tem-se $h(x) - G(x) = c$ para alguma constante c. Então $h = G + c \in C^1[a,b]$ e $h' = G' = g$. □

Corolario 2.4.1. *Se $g \in C[a,b]$ e $\int_a^b g(x)v(x)dx = 0$, $\forall\ v \in C_0^1[a,b]$, então $g \equiv 0$ em $[a,b]$.*

Demonstração. Basta fazer $h \equiv 0$ na proposição (2.4.2). □

Propocição 2.4.3. ***(Lagrange)*** *Seja $g \in C[a,b]$ tal que para todo $v \in C_0^k[a,b] = \{v \in C^k[a,b] : v^{(j)}(a) = v^{(j)}(b) = 0, j = 0, 1, \ldots k-1\}$ tenhamos*

$$\int_a^b g(t)v(t)dt = 0.$$

Então $g \equiv 0$ em $[a,b]$.

Demonstração. Suponha que $g(t_0) > 0$ para algum $t_0 \in (a,b)$. Como g é contínua existe intervalo $[c,d] \subset (a,b)$ contendo t_0 tal que

$$g(t) \geq \frac{1}{2}g(t_0) > 0.$$

Definindo a função

$$v(t) = \begin{cases} [(t-c)(d-t)]^{k+1}, & \text{se } \ t \in [c,d] \\ 0, & \text{se } \ t \in [a,b]/[c,d] \end{cases}$$

observamos que v é não negativa e $v \in C_0^k[a,b]$. Como gv é contínua, não-negativa e não-identicamente nula, segue que

$$\int_a^b g(t)v(t)dt > 0,$$

mas isto contradiz a hipótese. Logo $g(t) \leq 0$ em (a,b) e pela continuidade de g concluímos que $g \leq 0$ em $[a,b]$.

Analogamente, supondo a existência de t_0 em (a,b) com $f(t_0) < 0$, provamos que $f \geq 0$ em $[a,b]$. Portanto $f \equiv 0$ em $[a,b]$. □

A proposição a seguir é necessária para o resultado principal desta seção.

Propocição 2.4.4. *Se $y \in C^1[a,b]$ é tal que $\delta F(y,v) = 0$ para todo $v \in C_0^1[a,b]$ então*

$$f_z[y(x)] \in C^1[a,b] \quad e \quad \frac{d}{dx} f_z[y(x)] - f_y[y(x)] = 0. \tag{2.4.3}$$

Demonstração. Por hipótese,

$$\int_a^b f_y[y(x)]v(x) + f_z[y(x)]v'(x)dx = 0$$

para todo $v \in C_0^1[a,b]$. Fazendo $g(x) = f_y[y(x)]$ e $h(x) = f_z[y(x)]$ na proposição (2.4.2) temos que $f_z[y(x)] \in C^1[a,b]$ e $\dfrac{d}{dx} f_z[y(x)] = f_y[y(x)]$. □

A equação (2.4.3) é conhecida como equação de **Euler-Lagrange**.

Teoreoma 2.4.2. *Se y^* é um extremo local em $\{y \in C^1[a,b]/y(a) = a_1,\ y(b) = b_1\}$ para F então y^* satisfaz a equação de Euler-Lagrange, isto é,*

$$\frac{d}{dx} f_z[y(x)] - f_y[y(x)] = 0.$$

Além disso, existem λ_0, $\lambda_1 \in \mathbb{R}$ tais que

$$\delta F(y^*, \Delta y) = \lambda_0 \Delta y(a) + \lambda_1 \Delta y(b),$$

onde $\lambda_0 = -f_z[y^*(a)]$ e $\lambda_1 = f_z[y^*(b)]$.

Demonstração. Definimos $K_0(y) = y(a)$, $K_1(y) = y(b)$, para todo $y \in C^1[a,b]$. Assim y^* trata-se de um extremo local no conjunto $C^1[a,b][K_0 = a_1, K_1 = b_1]$ para F. Como vimos anteriormente, as variações de Gâteaux dos funcionais K_0, K_1 e F são dadas por

$$\begin{aligned} \delta K_0(y,\Delta y) &= \Delta y(a), \\ \delta K_1(y,\Delta y) &= \Delta y(b), \\ \delta F(y,\Delta y) &= \int_{x_0}^{x_1} f_y[y(x)]\Delta y(x) + f_z[y(x)]\Delta y'(x)dx, \end{aligned} \tag{2.4.4}$$

e ambas são fracamente contínuas em $C^1[a,b]$. Agora note que o determinante

$$\det(y, \Delta y_0, \Delta y_1) = \begin{vmatrix} \delta K_0(y,\Delta y_0) & \delta K_0(y,\Delta y_1) \\ \delta K_1(y,\Delta y_0) & \delta K_1(y,\Delta y_1) \end{vmatrix}$$

não é identicamente nulo, para qualquer y em $C^1[a,b]$. De fato,

$$\det(y, \Delta y_0, \Delta y_1) = \Delta y_0(a)\Delta y_1(b) - \Delta y_0(b)\Delta y_1(a),$$

tomando $\Delta y_0(x) = 1$ e $\Delta y_1(x) = \dfrac{x-a}{b-a}$ segue que $\det(y, \Delta y_0, \Delta y_1) = 1$. Então, pelo teorema dos multiplicadores de Lagrange, existem λ_0, $\lambda_1 \in \mathbb{R}$ tais que

$$\delta F(y^*, \Delta y) = \lambda_0 \delta K_0(y^*, \Delta y) + \lambda_1 \delta K_1(y^*, \Delta y).$$

Consequentemente,

$$\begin{aligned} \delta F(y^*, \Delta y) &= \lambda_0 \Delta y(a) + \lambda_1 \Delta y(b) \\ &= 0 \end{aligned} \tag{2.4.5}$$

para todo $\Delta y \in C_0^1[a,b]$. E, de acordo com a proposição (2.4.4), isso implica

$$f_z[y^*(x)] \in C^1[a,b] \text{ e } \frac{d}{dx} f_z[y^*(x)] - f_y[y^*(x)] = 0.$$

Usando a equação de Euler-Lagrange em (2.4.4), obtemos:

$$\begin{aligned} \delta F(y^*, \Delta y) &= \int_{x_0}^{x_1} f_y[y^*(x)]\Delta y(x) + f_z[y^*(x)]\Delta y'(x) dx \\ &= \int_{x_0}^{x_1} \frac{d}{dx} f_z[y^*(x)]\Delta y(x) + f_z[y^*(x)]\Delta y'(x) dx \\ &= f_z[y^*(x)]\Delta y(x)\Big|_a^b \end{aligned}$$

Agora, segue de (2.4.5) que

$$(\lambda_0 + f_z[y^*(a)])\Delta y(a) + (\lambda_1 - f_z[y^*(b)])\Delta y(b) = 0. \tag{2.4.6}$$

Finalmente, tomando $\Delta y(x) = \dfrac{x-b}{a-b}$ em (2.4.6), concluimos que $\lambda_0 = -f_z[y^*(a)]$. Similarmente, $\lambda_1 = f_z[y^*(b)]$. □

Exemplo 2.4.1. Sejam (t_0, y_0) e (t_1, y_1) dois pontos do plano. Vejamos que a função $y^*(t) = (t_1 - t)\frac{y_0}{t_1 - t_0} + (t - t_0)\frac{y_1}{t_1 - t_0}$ é a função de $C^1[t_0, t_1]$ de menor comprimento entre todas as que satisfazem $y(t_0) = y_0$ e $y(t_1) = y_1$.

Como vimos na introdução se trata de minimizar o funcional

$$J(y) = \int_a^b \sqrt{1 + y'(t)^2} dt$$

sobre o conjunto $\{y \in C^1[t_0, t_1] : \ y(t_0) = y_0, \ y(t_1) = y_1\}$.

Neste caso, a função $f(x, y, z) = \sqrt{1 + z^2}$ possui derivadas parciais f_y e f_z contínuas e a equação de Euler-Lagrange é dada por

$$\frac{d}{dt}\left[\frac{y'(t)}{(1 + y'(t)^2)^{1/2}}\right] = 0 \tag{2.4.7}$$

Portanto, se y satisfaz (2.4.7) temos que

$$\frac{y'(t)}{(1 + y'(t)^2)^{1/2}} = const$$

O que implica y' ser constante, ou seja, y ser linear. Entre as funções lineares a única que satisfaz a condição $y(t_0) = y_0, \ y(t_1) = y_1$ é $y^*(t) = (t_1 - t)\frac{y_0}{t_1 - t_0} + (t - t_0)\frac{y_1}{t_1 - t_0}$.

Exemplo 2.4.2. (Superfíce mínima de revolução) Sejam $P_0 = (x_0, y_0)$ e $P_1 = (x_1, y_1)$ dois pontos no plano com $x_0 < x_1$ e $y_0, y_1 > 0$. Considere o problema de encontrar uma curva suave $\gamma : y = y(x),\ x \in [x_0, x_1]$ ligando os pontos P_0 e P_1, cuja a rotação em torno do eixo x gera uma superfície de aréa mínima. A aréa da superfície obtida pela rotação de γ sobre o eixo x é dada por

$$A(y) = 2\pi \int_{x_0}^{x_1} y(x)\sqrt{1 + y'(x)^2}dx$$

Assim o problema consiste em encontrar um vetor mínimo para o funcional A no conjunto

$$\{y \in C^1[x_0, x_1] :\ y(x_0) = y_0,\ y(x_1) = y_1\}.$$

Note que a função $f(x, y, z) = 2\pi y\sqrt{1 + z^2}$ possui derivadas parciais contínuas e a equação de Euler-Lagrange é dada por

$$\frac{d}{dx}\left(2\pi y(x)\frac{y'(x)}{\sqrt{1 + y'(x)^2}}\right) - 2\pi\sqrt{1 + y'(x)^2} = 0. \tag{2.4.8}$$

Após algumas manipulações concluimos que a equação (2.4.8) implica

$$2\pi\frac{d}{dx}\left(\frac{y(x)}{\sqrt{1 + y'(x)^2}}\right) = 0$$

e, consequentemente,

$$2\pi\left(\frac{y(x)}{\sqrt{1 + y'(x)^2}}\right) = C.$$

A solução geral da equação diferencial acima tem a forma

$$y(x) = a\ \cosh\left(\frac{x - b}{a}\right),\ \ a, b \in \mathbb{R}.$$

As curvas acima são chamadas de família de catenárias e são naturalmente candidatas a solução do problema. As superfícies de revolução resultantes são denominadas catenóides.

2.5 O Problema de Extremidades Fixas com Vínculos

Muitos problemas variacionais são modelados somente impondo condições na fronteira. No entanto, é frequente o acréscimo de outras restrições dadas geralmente por uma função integral

$$G(y) = \int_a^b g\left(x, y(x), y'(x)\right) dx = \int_a^b g\left[y(x)\right] dx,$$

assumindo um valor preestabelecido. Sendo assim, podemos aplicar o teorema dos multiplicadores de Lagrange e a proposição (2.4.4) para provar o seguinte teorema:

Teoreoma 2.5.1. *Suponha que as funções* $f = f(x,y,z)$*,* $g_i = g_i(x,y,z)$*,* $i = 1,\ldots,n$*, e suas derivadas parciais em relação a* y *e* z *são contínuas em* $[a,b]\times\mathbb{R}^2$*. Seja* y_0 *um extremo local para o funcional*

$$F(y) = \int_a^b f(x,y(x),y'(x))dx$$

em

$$D = \{y \in C^1[a,b]/y(a) = a_1, y(b) = b_1\},$$

quando restrito ao conjunto

$$G_{y_0} = \{y \in C^1[a,b]/G_i(y) = \int_a^b g_i[y(x)]dx = G_i(y_0), i = 1,\ldots,n\}.$$

Então uma das possibilidades abaixo acontece:

a) o determinante $n\times n$

$$|\delta G_i(y_0,v_j)| = 0,$$

quando $v_j \in C_0^1[a,b]$*,* $j = 1,\ldots,n$*.*

b) existem $\lambda_1,\ldots,\lambda_n \in \mathbb{R}$ *tais que* y_0 *é solução da equação de Euler-Lagrange para*

$$\widetilde{f} = f + \sum_{i=1}^n \lambda_i g_i,$$

isto é, y_0 *é solução da equação*

$$\frac{d}{dx}\widetilde{f}_z[y_0(x)] - \widetilde{f}_y[y_0(x)] = 0.$$

Demonstração. Definimos $K_1(y) = y(a)$, $K_2(y) = y(b)$, para todo $y \in C^1[a,b]$. Assim y_0 é extremo local em $C^1[a,b][K_1 = a_1, K_2 = b_1, G_i = G_i(y_0), i = 1,\ldots,n]$ para F. Então, pelo teorema de Euler-Lagrange, temos duas possibilidades a considerar: ou o determinante $(n+2)\times(n+2)$

$$\begin{vmatrix} \delta G_1(y_0,v_1) & \cdots & \delta G_1(y_0,v_n) & \delta G_1(y_0,v_{n+1}) & \delta G_1(y_0,v_{n+2}) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \delta G_n(y_0,v_1) & \cdots & \delta G_n(y_0,v_n) & \delta G_n(y_0,v_{n+1}) & \delta G_n(y_0,v_{n+2}) \\ \delta K_1(y_0,v_1) & \cdots & \delta K_1(y_0,v_n) & \delta K_1(y_0,v_{n+1}) & \delta K_1(y_0,v_{n+2}) \\ \delta K_2(y_0,v_1) & \cdots & \delta K_2(y_0,v_n) & \delta K_2(y_0,v_{n+1}) & \delta K_2(y_0,v_{n+2}) \end{vmatrix} = 0, \tag{2.5.1}$$

para quaisquer $v_1,\ldots,v_n,v_{n+1},v_{n+2} \in C^1[a,b]$, ou existem $\mu_1,\ldots,\mu_n,\mu_{n+1},\mu_{n+2} \in \mathbb{R}$ tais que

$$\delta F(y_0,v) = \mu_1\delta G_1(y_0,v)+\mu_2\delta G_2(y_0,v)+\ldots+\mu_n\delta G_n(y_0,v)+\mu_{n+1}K_1(y_0,v)+\mu_{n+2}\delta K_2(y_0,v) \tag{2.5.2}$$

para todo $v \in C^1[a,b]$.

Suponha que a primeira alternativa seja verdadeira. Sejam $v_1, \ldots, v_n \in C_0^1[a,b]$, tomando $v_{n+1}(x) = \dfrac{x-b}{a-b}$ e $v_{n+2}(x) = \dfrac{x-a}{b-a}$, temos que a equação (2.5.1) implica em

$$\begin{aligned} 0 &= \begin{vmatrix} \delta G_1(y_0, v_1) & \cdots & \delta G_1(y_0, v_n) & \delta G_1(y_0, v_{n+1}) & \delta G_1(y_0, v_{n+2}) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \delta G_n(y_0, v_1) & \cdots & \delta G_n(y_0, v_n) & \delta G_n(y_0, v_{n+1}) & \delta G_n(y_0, v_{n+2}) \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 \end{vmatrix} \\ &= \left|\delta G_i(y_0, v_j)\right|_{i,j=1,\ldots,n} \end{aligned}$$

Agora vejamos o que acontece se a segunda alternativa for válida. Sabemos que as variações de K_1 e K_2 em y_0 desaparecem nos vetores que se anulam nas extremidades por isso, para todo $v \in C_0^1[a,b]$, tem-se

$$\delta\left(F + \sum_{i=1}^{n} \lambda_i G_i\right)(y_0, v) = 0$$

onde $\lambda_i = -\mu_i$, $i = 1, \ldots, n$.
Na verdade, isto quer dizer que

$$\delta\widetilde{F}(y_0, v) = 0 \text{ para todo } v \in C_0^1[a,b],$$

onde

$$\widetilde{F}(y) = \left(F + \sum_{i=1}^{n} \lambda_i G_i\right)(y) = \int_a^b \widetilde{f}[y(x)]dx.$$

Portanto, pela proposição (2.4.4), concluimos que $\dfrac{d}{dx}\widetilde{f}_z[y_0(x)] - \widetilde{f}_y[y_0(x)] = 0.$ □

3 Aplicação ao problema de autovalores de Sturm-Liouville

No capítulo anterior vimos, no problema de extremidades fixas, que equações diferenciais aparecem naturalmente por meio da equação de Euler-Lagrange. Desse modo, reduzimos o problema variacional ao estudo de um problema de equações diferenciais. Nesse capítulo faremos o caminho inverso, isto é, estudaremos um tópico de equações diferenciais, a saber, o problema de Sturm-Liouville, através de uma abordagem variacional.

3.1 Problema de Sturm-Liouville

Chamamos equação de Sturm-Liouvile a uma equação com coeficientes reais da forma

$$(p(x)u')' + (\lambda\rho(x) - q(x))u = 0 \tag{3.1.1}$$

onde λ é um parâmetro real e as funções $\rho(x)$ e $q(x)$ são positivas.

Um problema de Sturm-Liouville regular em $[a, b]$ consiste de uma equação do tipo (3.1.1) com condições de contorno, ditas auto-adjuntas, dadas por

$$\alpha u(a) + \alpha' u'(a) = 0 \tag{3.1.2}$$

$$\beta u(b) + \beta' u'(b) = 0 \tag{3.1.3}$$

onde $(\alpha, \alpha') \neq 0$ e $(\beta, \beta') \neq 0$, ou pelas condições periódicas

$$u(a) = u(b) = 0 \tag{3.1.4}$$

$$u'(a) = u'(b) = 0 \tag{3.1.5}$$

quando $p(a) = p(b)$. Os valores de λ para os quais o problema admite solução não trivial, isto é, para os quais (3.1.1) admite solução não trivial satisfazendo as condições de contorno fixadas são ditos autovalores do problema. As soluções não triviais correspondentes a um autovalor λ são ditas autofunções do problema associadas a λ. Se u é uma autofunção associada ao autovalor λ chamamos o par (λ, u) de autopar.

Propocição 3.1.1. *Consideremos um problema de Sturm-Liouville regular em $[a, b]$ para a equação (3.1.1). Se λ_1, λ_2 são autovalores distintos do problema e u_1, u_2 são autofunções a eles associadas então u_1 e u_2 são ortogonais em $[a, b]$ em relação ao peso $\rho(x)$, ou seja,*

$$\int_a^b u_1(x)u_2(x)\rho(x)dx = 0.$$

Teoreoma 3.1.1. *Todo autovalor do problema de Sturm-liouville regular com condições auto-adjuntas dadas por (3.1.2),(3.1.3) tem multiplicidade 1, isto é, o espaço vetorial das autofunções correspondentes tem dimensão 1.*

Exemplo 3.1.1. Consideremos o problema de Sturm-Liouville regular com condições de contorno periódicas

$$\begin{aligned}&y''(t) + \lambda y(t) = 0, \text{ para } t \in [-\pi, \pi]\\&y(-\pi) = y(\pi)\\&y'(-\pi) = y'(\pi)\end{aligned}$$

Note que $p(t) = 1$ e então $p(-\pi) = p(\pi)$. Para $\lambda > 0$ a solução geral da equação é da forma $y(t) = C_1\cos(\sqrt{\lambda}t)+C_2\text{sen}(\sqrt{\lambda}t)$ usando as condições $y(-\pi) = y(\pi)$ e $y'(-\pi) = y'(\pi)$ obtemos

$$\begin{aligned}&C_2\text{sen}(\sqrt{\lambda}\pi) = 0\\&2C_1\sqrt{\pi}\text{sen}(\sqrt{\pi}) = 0.\end{aligned}$$

Assim, para que exista soluções não triviais, devemos ter $\text{sen}(\sqrt{\lambda}\pi) = 0$, ou seja, $\sqrt{\lambda}\pi = n\pi$ $(n = 1, 2, \ldots)$. Logo $\lambda_n = n^2$ $(n = 1, 2, \ldots)$ são autovalores. Para cada $\lambda_n = n^2$ existem duas soluções linearmente independentes $\varphi_n(t) = \cos(nt)$ e $\psi_n(t) = \text{sen}(nt)$ (diferentemente do problema de Sturm-Liouville regular com as condições (3.1.2),(3.1.3)).

Além disso, $\lambda = 0$ é autovalor e a correspondente autofunção é a função constante $\varphi_0(t) = 1$. Para $\lambda < 0$ o problema só possui solução trivial.

Portanto os autovalores são

$$0, 1, 4, 9, \ldots, n^2, \ldots$$

e as autofunções são

$$1, \cos(t), \text{sen}(t), cos(2t), sen(2t), \ldots, cos(nt), sen(nt), \ldots$$

Vejamos como as equações de Sturm-Liouville aparecem na Física Matemática através do problema a seguir.

Exemplo 3.1.2. O problema da corda vibrante consiste em descrever as oscilações de uma corda com extremidades fixas em dois pontos. O modelo físico é o seguinte:

a) As vibrações ocorrem em um plano de coordenadas (x, y) e a função

$$y(x, t),\ t \geq 0,\ a \leq x \leq b$$

denota a posição do ponto x da corda de extremidades a e b no instante t.

b) As vibrações são transversais, isto é, o movimento de cada ponto da corda é vertical.

c) A oscilação da corda é devida unicamente a tensão e esta é constante T. Além disso, a corda é flexível de forma que em cada ponto T age segundo a tangente à corda.

Consideremos agora um elemento Δl da corda correspondente aos pontos entre x e $x + \Delta x$. Como as vibrações são verticais, a força resultante F de todas a forças que age sobre Δl é vertical. Pela segunda lei de Newton F é dada por

$$F = \rho(x)\Delta x \frac{\partial^2 y}{\partial t^2} \tag{3.1.6}$$

onde $\rho(x) > 0$, $a \leq x \leq b$ é a densidade linear de massa da corda. Por outro lado, T age em cada ponto segundo a tangente à corda, então

$$F = T\text{sen}\theta_2 - T\text{sen}\theta_1$$

onde θ_1 e θ_2 são os ângulos da tensão com relação à horizontal no ponto x e $x + \Delta x$, respectivamente. Como as vibrações são pequenas podemos fazer as aproximações

$$\text{sen}\theta_1 \simeq \text{tg}\theta_1 = \left(\frac{\Delta y}{\Delta x}\right)_x \text{ e } \text{sen}\theta_2 \simeq \text{tg}\theta_2 = \left(\frac{\Delta y}{\Delta x}\right)_{x+\Delta x},$$

logo

$$\begin{aligned} F &= T\left(\frac{\Delta y}{\Delta x}\right)_x - T\left(\frac{\Delta y}{\Delta x}\right)_{x+\Delta x} \\ F &= T\Delta\left(\frac{\Delta y}{\Delta x}\right) \end{aligned}$$

Usando (3.1.6) obtemos então

$$T\frac{\Delta}{\Delta x}\left(\frac{\Delta y}{\Delta x}\right) = \rho(x)\frac{\partial^2 y}{\partial t^2}$$

que fazendo $\Delta x \to 0$ nos dá a equação da corda vibrante

$$\frac{\partial^2 y}{\partial x^2} = \frac{\rho(x)}{T}\frac{\partial^2 y}{\partial t^2} \tag{3.1.7}$$

Suponhamos que no instante $t = 0$ a corda está em repouso e sua posição é descrita por uma função $f(x)$ de classe C^2 então o seu movimento é obtido por (3.1.7) com as condições

$$\begin{aligned} y(a,t) = y(b,t) = 0 &\quad 0 \leq t \\ \frac{\partial y}{\partial t}(x,0) = 0 &\quad a \leq x \leq b \\ y(x,0) = f(x) &\quad a \leq x \leq b \end{aligned} \tag{3.1.8}$$

Para resolver o problema apliquemos o método da separação de variáveis, isto é, procuremos soluções de (3.1.7) da forma

$$y(x,t) = u(x)v(t).$$

Substituindo em (3.1.7) obtemos

$$\frac{1}{\rho(x)}\frac{u''(x)}{u(x)} = \frac{1}{T}\frac{v''(t)}{v(t)}$$

o que implica a existência de uma constante λ tal que

$$u'' + \lambda\rho(x)u = 0 \quad (3.1.9)$$

$$u(a) = u(b) = 0 \quad (3.1.10)$$

e

$$v'' + \lambda T v = 0 \quad (3.1.11)$$

$$v'(0) = 0 \quad (3.1.12)$$

Veremos mais adiante que o problema de Sturm-Liouvile regular (3.1.9),(3.1.10) admite uma seguência infinita de autovalores $\lambda_0 < \lambda_1 < \ldots < \lambda_n < \ldots$. Além disso, as autofunções correspondentes $u_1, \ldots, u_n, \ldots$ satisfazem

$$\int_a^b u_m(x)u_n(x)\rho(x)dx = 0$$

e podemos supor que elas estão normalizadas, ou seja,

$$\int_a^b u_n(x)^2\rho(x)dx = 1$$

.

Multiplicando a equação (3.1.9) por u e integrando em $[a, b]$ resulta que

$$\lambda\int_a^b \rho(x)u(x)^2dx = -\left[u'(x)u(x)\Big|_a^b - \int_a^b u'(x)^2dx\right].$$

Assim, se u é um autovetor correspondente ao autovalor λ do problema (3.1.9),(3.1.10) então

$$\lambda = \frac{\int_a^b u'(x)^2dx}{\int_a^b \rho(x)u(x)^2dx},$$

consequentemente os autovalores do problema (3.1.9)(3.1.10) são positivos. Portanto devemos considerar apenas o caso $\lambda > 0$, dessa forma para cada n a solução de (3.1.11),(3.1.12) com $\lambda = \lambda_n$ é múltiplo de $v_n(t) = \cos(\sqrt{\lambda_n T}t)$. Daí segue-se que as funções

$$y_n(x,t) = u_n(x)\cos(\sqrt{\lambda_n T}t)$$

são soluções da equação (3.1.7) e verificam as duas primeiras condições em (3.1.8). Como isso também é válido para combinações lineares dessas funções, procuramos uma solução de (3.1.7) que satisfaça todas as condições em (3.1.8) na forma de uma série

$$y(x,t) = \sum_{n=0}^{\infty} a_n u_n(x)\cos(\sqrt{\lambda_n T}t) \quad (3.1.13)$$

onde $a_n \in \mathbb{R}$. Se isso é possível, a condição $f(x) = y(x, 0)$ equivale a

$$f(x) = \sum_{n=0}^{\infty} a_n u_n(x) \tag{3.1.14}$$

que por sua vez implica

$$a_n = \int_a^b f(x) u_n(x) \rho(x) dx \tag{3.1.15}$$

pois a seguência $u_1, \ldots, u_n \ldots$ é ortonormal em $[a, b]$ em relação ao peso $\rho(x)$. Futuramente, veremos que a expansão (3.1.14) é válida para toda função f de classe C^2 satisfazendo a condição $u(a) = u(b) = 0$, sendo a convergência uniforme em $[a, b]$.

Consideremos a equação diferencial

$$p(x)u'' + r(x)u' + q(x)u = \lambda u \tag{3.1.16}$$

onde os coeficientes são contínuos num intervalo $[a, b]$ e λ é um parâmetro real. Se $C^0[a, b]$ é o espaço das funções reais contínuas em $[a, b]$, $C^2[a, b]$ é o subespaço das funções de classe C^2 e $L : C^2[a, b] \longrightarrow C^0[a, b]$ é o operador linear dado por

$$L(u) = p(x)u'' + r(x)u' + q(x)u \tag{3.1.17}$$

então (3.1.16) assume a forma $L(u) = \lambda u$.

Isto nos sugere utilizar a teoria dos operadores lineares em espaços de dimensão infinita para estudar a equação (3.1.16). Vejamos alguns resultados e definições desta teoria que utilizaremos.

Definição 3.1.1. Um operador linear $T : E \longrightarrow F$ entre espaços normados é dito compacto se $\overline{T(B_E)}$ é compacto em F, onde $B_E = \{x \in E : \|x\| \leq 1\}$.

Propocição 3.1.2. *Todo operador compacto é contínuo*

Propocição 3.1.3. *Sejam E e F espaços normados. As seguintes afirmações são equivalentes para um operador linear $T : E \longrightarrow F$.*

a) T é compacto

b) $\overline{T(A)}$ é compacto em F para todo limitado A em E.

c) Para toda sequência limitada (x_n) em E, a sequência $(T(x_n)$ tem subsequência convergente em F.

Definição 3.1.2. Um espaço vetorial real H no qual está definido um produto interno $\langle\ \rangle$, ou seja uma forma bilinear $\langle\ \rangle : H \times H \longrightarrow \mathbb{R}$ tal que

$$\begin{aligned} \langle x, y \rangle &= \langle y, x \rangle && x, y \in H \\ \langle x, x \rangle &\geq 0 && x \in H \\ \langle x, x \rangle &= 0 \Longleftrightarrow x = 0 \end{aligned}$$

é dito um espaço pré-Hilbertiano.

Todo espaço pré-Hilbertiano H é um espaço normado com a norma canônica

$$\|x\| = \sqrt{\langle x, x \rangle}$$

que satisfaz a desigualdade de Cauchy-Schwarz

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad x, y \in H$$

Dois vetores $x, y \in H$ com $\langle x, y \rangle = 0$ são ditos ortogonais. Uma sequência $e_1, \ldots, e_n, \ldots$ de vetores de H dois a dois orgonais é dita ortonormal se $\|e_n\| = 1$ para todo n. Uma sequência ortonormal de vetores de um subespaço $H_1 \subset H$ é dita uma base ortonormal para H_1 se as combinações lineares dos vetores da sequência formam um subespaço denso de H_1.

Propocição 3.1.4. *Se $e_1, \ldots, e_n, \ldots$ é uma base ortonormal para um espaço pré-Hilbertiano H então*

$$x = \sum_{n=1}^{\infty} \langle x, e_n \rangle \; e_n$$

para todo $x \in H$.

Um operador $T : H \longrightarrow H$ é dito auto-adjunto se

$$\langle Tx, y \rangle = \langle x, Ty \rangle \text{ para todo } x, y \in H$$

Se E é um espaço vetorial dizemos que $\lambda \in \mathbb{R}$ é um autovalor de um operador $T : E \longrightarrow E$ se existe $x \in E, \; x \neq 0$ tal que $Tx = \lambda x$. Então x é dito um autovetor de T associado a λ.

Propocição 3.1.5. *Se x, y são autovetores de um operador auto-adjunto associados a autovalores distintos então x e y são ortogonais.*

Enunciemos agora um resultado que é uma versão, adaptada aos nossos interesses, do teorema espectral para operadores compactos auto-adjuntos.

Teoreoma 3.1.2. *Seja H um espaço pré-Hilbertiano de dimensão infinita e $T : H \longrightarrow H$ um operador compacto auto-adjunto com $T(H)$ denso em H. Então H admite uma base ortonormal $e_1 \ldots, e_n, \ldots$ tal que cada e_n é um autovalor de T. Além disso a sequência correspondente $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ de autovalores tende para 0 e $\lambda_n \neq 0$ para todo n.*

Lema 3.1.1. *Seja $(H, \langle \;\; \rangle)$ um espaço pré-Hilbertiano e $T : (H, \langle \;\; \rangle) \longrightarrow (H, \langle \;\; \rangle)$ um operador autoadjunto. Seja $(H, \| \;\; \|_1)$ uma outra estrutura de espaço normado em H tal que a identidade $Id : (H, \| \;\; \|_1) \longrightarrow (H, \langle \;\; \rangle)$ seja contínua. Se $T : (H, \langle \;\; \rangle) \longrightarrow (H, \| \;\; \|_1)$ é* *compacto então* $T : (H, \langle \ \ \rangle) \longrightarrow (H, \langle \ \ \rangle)$ *também o é. Além disso, para todo* $y \in T(H)$ *a expansão*

$$y = \sum_{n=1}^{\infty} \langle y, e_n \rangle \ e_n$$

na base ortonormal dada pelo teorema (3.1.2)

Consideremos agora $C^0[a, b]$ com a estrutura de espaço pré-hilbertiano dado pelo produto interno

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

e denotemos a norma canônica associada por $\| \ \|_2$.

Voltando ao operador $L : C^2[a, b] \longrightarrow C^0[a, b]$ definido em (3.1.17) com a hipótese adicional que $p(x) > 0$ em $[a, b]$. Pelo teorema da existência de soluções de uma equação diferencial linear segue-se que L é sobrejetivo. Fixemos um complemento M do núcleo $N = \{f \in C^2[a, b] : Lf = 0\}$. A restrição de L a M é um operador $L/M : M \longrightarrow C^0[a, b]$ injetivo e sobrejetivo, daí sua inversa $(L/M)^{-1}$ é um operador $(L/M)^{-1} : C^0[a, b] \longrightarrow C^0[a, b]$. Tentamos agora aplicar o teorema (3.1.2) a $(L/M)^{-1}$.

Para encontrar um M conveniente procuramos condições em L para que tenhamos $\langle Lf, g \rangle = \langle f, Lg \rangle$ para todo $f, g \in C^2[a, b]$.

Integrando por partes temos

$$\begin{aligned}
\langle Lf, g \rangle &= \int_a^b pf''gdx + \int_a^b rf'gdx + \int_a^b qfgdx \\
&= pgf'\Big|_a^b - \int_a^b f'(pg)'dx + rgf\Big|_a^b - \int_a^b f(rg)'dx + \int_a^b qfgdx \\
&= pgf'\Big|_a^b - f(pg)'\Big|_a^b + \int_a^b f(pg)''dx + rgf\Big|_a^b - \int_a^b f(rg)'dx + \int_a^b qgfdx \\
&= \int_a^b f[(pg)'' - (rg)' + qg]dx + [pgf' - f(pg)' + rgf]_a^b \\
&= \int_a^b f[pg'' + (2p' - r)g' + (p'' - r' + q)g]dx + [pgf' - f(pg)' + rgf]_a^b
\end{aligned}$$

.

Definindo por L^* o operador

$$L^*g = pg'' + (2p' - r)g' + (p'' - r' + q)g$$

e por B a aplicação bilinear

$$B(f, g) = pgf' - f(pg)' + rgf$$

temos que $\langle Lf, g \rangle = \langle f, L^*g \rangle + B(f, g)\Big|_a^b$.

O operador L^* é dito o adjunto formal de L. Se $L = L^*$ dizemos que L é formalmente auto-adjunto. É fácil verificar que L é formalmente auto-adjunto se e só se é da forma

$$Lf = (pf')' + qf$$

e neste caso

$$B(f,g) = p(f'g - fg').$$

Portanto, L é formalmente auto-adjunto se e só se L é um operador de Sturm-Liouville.

Para que L seja um operador auto-adjunto em $C^2[a,b]$ além de L ser um operador de Sturm-Liouville devemos restringir L a um subespaço de $C^2[a,b]$ cujo elementos sastifaçam condições de contorno que anulem $B(f,g)\Big|_a^b$.

Vejamos alguns exemplos de espaços que possuem essa característica:

$$\begin{aligned} M_1 &= \{f \in C^2[a,b] : f(a) = f(b) = 0\} \\ M_2 &= \{f \in C^2[a,b] : f'(a) = f'(b) = 0\} \end{aligned}$$

Se $p(a) = p(b)$ temos também o espaço

$$M_3 = \{f \in C^2[a,b] : f(a) = f(b) \text{ e } f'(a) = f'(b)\}$$

Notemos que os espaços M_1, M_2 e M_3 tem codimensão 2 pois cada um deles é o núcleo de uma aplicação linear sobrejetiva de $C^0[a,b]$ em $\mathbb{R}^2$.

O lema seguinte mostra que substituindo L por um operador da forma $(L-\mu)f = Lf - \mu f$ os subespaços M_1, M_2 e M_3 são complementos de N

Lema 3.1.2. *Seja M um dos espaços M_1, M_2 ou M_3 acima e sejam $\mu \geq 1 + sup_{x\in[a,b]}\ q(x)$ e $m = inf_{x\in[a,b]}\ p(x)$.*

i) $\langle(\mu - L)f, f\rangle \geq m\|f'\|_2^2 + \|f\|_2^2$ *para todo* $f \in M$

ii) $(L-\mu) : M \longrightarrow C^0[a,b]$ *é injetiva e sobrejetiva.*

No lema seguinte denotaremos por $\|\ \|_\infty$ a norma uniforme em $C^0[a,b]$, isto é, $\|f\|_\infty = \sup\{|f(x)| : x \in [a,b]\}$.

Lema 3.1.3. *Nas condições do lema (3.1.2) temos que a inversa S de*

$$(L-\mu) : M \longrightarrow C^0[a,b]$$

é um operador compacto

$$S : (C^0[a,b], \langle\ \ \rangle) \longrightarrow (C^0[a,b], \|\ \|_\infty).$$

Notemos que se $v, w \in C^0[a,b]$ e $f, g \in M$ com $(L-\mu)f = v$ e $(L-\mu)g = w$ então

$$\begin{aligned}\langle Sv, w\rangle &= \langle f, (L-\mu)g\rangle \\ &= \langle (L-\mu)f, g\rangle = \langle v, Sw\rangle\end{aligned}$$

provando que o operador S é auto-adjunto no espaço pré-Hilbertiano $C^0[a,b]$.

Propocição 3.1.6. *Seja a equação de Sturm-Liouville*

$$(p(x)u')' + q(x)u + \lambda u = 0, \tag{3.1.18}$$

onde p, p' e q são contínuas e $p > 0$ em $[a,b]$, com uma das condições de contorno seguintes:

a) u(a)=u(b)=0

b) u'(a)=u'(b)=0

e, caso tenhamos p(a)=p(b),

c) u(a)=u(b), u'(a)=u'(b).

Então podemos afimar:

i) Em cada um dos três casos, os autovalores formam uma seguência $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ que tende para ∞.

ii) Em cada um dos três casos, existe uma base ortonormal $f_1, f_2, \ldots, f_n, \ldots$ de $C^0[a,b]$ formada por autofunções.

iii) Se $f \in C^2[a,b]$ satisfaz uma das condições de contorno indicadas então a expressão de f na base ortogonal correspondente dada em ii)

$$f = \sum_{n=1}^{\infty} \langle f, f_n\rangle\; f_n$$

converge uniformemente em $[a,b]$.

Observação 3.1.1. A proposição é válida se substituímos (3.1.6) por $(p(x)u')' + q(x)u + \lambda\rho(x)u = 0$ onde $\rho(x) > 0$ em $[a,b]$. Prova-se de modo semelhante se em $C^0[a,b]$ considerarmos o produto interno $\langle f, g\rangle = \int_a^b f(x)g(x)\rho(x)dx$ e estudarmos o operador $f \mapsto \rho^{-1}Lf$.

3.2 Princípio de Rayleigh e o Autovalor Mínimo

Consideremos o operador de Sturm-liouville L no intervalo fechado $[a, b]$

$$L(u) = -(\tau u')' - qu,$$

onde $\tau \in C^1[a, b]$, $q \in C[a, b]$, $\tau > 0$.

Nós estamos interessados no problema de autovalores de Sturm-Liouville regular em $[a, b]$ com condições de contorno de Dirichlet:

$$Lu = \lambda u, \ \ \lambda \in \mathbb{R} \tag{3.2.1}$$

$$u(a) = u(b) = 0 \tag{3.2.2}$$

No nosso contexto vamos lidar com uma situação mais abrangente que é o problema (3.2.1), (3.2.2) generalizado da seguinte forma: encontrar λ e u, $u \neq 0$, tais que

$$-(\tau(x)u')' - q(x)u = \lambda\rho(x)u \tag{3.2.3}$$

$$u(a) = u(b) = 0. \tag{3.2.4}$$

ρ é uma função contínua e estritamente positiva em $[a, b]$, enquanto τ e q satisfazem as mesmas condições acima.

Se multiplicarmos por $u(x)$ em ambos os lados de (3.2.3), obtemos

$$\lambda\rho(x)u(x)^2 = -u(x)\left[\tau(x)u'(x)\right]' - q(x)u(x)^2$$

e então

$$\begin{aligned} \lambda\int_a^b \rho(x)u(x)^2dx &= \int_a^b -u(x)\left[\tau(x)u'(x)\right]' - q(x)u(x)^2dx \\ &= -\left[u(x)\tau(x)u'(x)\Big|_a^b - \int_a^b \tau(x)u'(x)^2dx\right] - \int_a^b q(x)u(x)^2dx \\ &= \int_a^b \tau(x)u'(x)^2 - q(x)u(x)^2dx. \end{aligned}$$

Logo, para toda autofunção u correspondente a λ tem-se

$$\lambda = \frac{\int_a^b \tau(x)u'(x)^2 - q(x)u(x)^2dx}{\int_a^b \rho(x)u(x)^2dx}.$$

Por simplicidade, definimos os funcionais

$$\begin{aligned} D(u) &= \int_a^b \tau(x)u'(x)^2 - q(x)u(x)^2dx \\ H(u) &= \int_a^b \rho(x)u(x)^2dx \end{aligned}$$

para todo $u \in C_0^1[a,b]$, o espaço das funções continuamente diferenciáveis no intervalo $[a,b]$ que se anulam nas extremidades. O **quociente de Rayleigh** é definido pela razão abaixo

$$R(u) = \frac{D(u)}{H(u)},$$

para todo $u \in C_0^1[a,b] - \{0\}$. Portanto, todo autopar (λ, u) satisfaz a relação $\lambda = R(u)$.

A partir de agora, nosso objetivo é mostrar que todo autovalor do problema de Sturm-Liouville (3.2.3),(3.2.4) pode ser obtido minimizando o quociente de Rayleigh sob restrições adequadas.

Sejam $\lambda_1 < \ldots < \lambda_n < \ldots$ os autovalores do problema de Sturm-Liouville regular (3.2.3),(3.2.4). A caracterização variacional dos autovalores será feita através de uma iteração, por essa razão iniciamos naturalmente pelo autovalor mínimo λ_1:

Teoreoma 3.2.1. *Seja λ_1 o autovalor mínimo. Então*

$$\lambda_1 = mín\ \{R(u) : \ u \in C_0^1[a,b], \ H = 1\}.$$

Demonstração. Como τ, ρ são positivas e ρ, q são contínuas, segue que

$$D(u) \geq \int_a^b -q(x)u(x)^2 dx \geq -\int_a^b \|q\|_{C[a,b]} u(x)^2 dx = -\|q\|_{C[a,b]} \int_a^b u(x)^2 dx,$$
$$H(u) \leq \int_a^b \|\rho\|_{C[a,b]} u(x)^2 = \|\rho\|_{C[a,b]} \int_a^b u(x)^2 dx,$$

e daí o quociente de Rayleigh é limitado inferiormente por $\dfrac{-\|q\|_{C[a,b]}}{\|\rho\|_{C[a,b]}}$. Assim, pode-se provar que R possui um mínimo, veja (CORMANI; RYHAM, 2002)

Suponha que u_0 minimiza R em

$$C_0^1[a,b][H=1] = \{u \in C_0^1[a,b] : \ H(u) = 1\}.$$

Convém observar que $R \equiv D$ em $C_0^1[a,b][H=1]$, por este motivo podemos assumir que

$$R(u) = \int_a^b f[u(x)]dx,$$

onde $f[u(x)] = f(x, u(x), u'(x))$, e $f(x,y,z) = \tau(x)z^2 - q(x)y^2$ é uma função com derivadas parciais contínuas em relação às variáveis y e z. Notemos também que

$$H(u) = \int_a^b h[u(x)]dx$$

onde $h(x,y,z) = \rho(x)y^2$ é tal que $h_y, h_z \in C([a,b] \times \mathbb{R}^2)$.
Portanto recaímos em um problema de extremidades fixas com vínculos. A variação de Gâteaux de H em u_0, dada por

$$\delta H(u,v) = 2\int_a^b \rho(x)u(x)v(x)dx,$$

é não-identicamente nula pois $\delta H(u_0, u_0) > 0$. Sendo assim temos, de acordo com o teorema (2.5), que existe $\lambda^* \in \mathbb{R}$ tal que u_0 é solução da equação de Euler-Lagrange para $\widetilde{f} = f - \lambda^* g$, isto é,

$$\frac{d}{dx}\widetilde{f}_z[u_0(x)] - \widetilde{f}_y[u_0(x)] = 0.$$

Calculando as derivadas parciais temos

$$\begin{aligned}\widetilde{f}_z[u_0(x)] &= 2\tau(x)u_0'(x),\\ \widetilde{f}_y[u_0(x)] &= -2q(x)u_0(x) - 2\lambda^*\rho(x)u_0(x)\end{aligned}$$

e então $(\tau(x)u_0'(x))' + q(x)u_0(x) + \lambda^*\rho(x)u_0(x) = 0$. Logo (λ^*, u_0) é um autopar para o problema (3.2.3), (3.2.4).

Se λ é um autovalor com autofunção normalizada u ($\int_a^b \rho(x)u(x)^2dx = 1$) então $\lambda^* = R(u_0) \leq R(u) = \lambda$. Consequentemente $\lambda^* = \lambda_1$. □

Corolario 3.2.1. ***(Princípio de Rayleigh)*** *O menor autovalor λ_1 é igual ao valor mínimo do quociente de Rayleigh:*

$$\lambda_1 = \textit{mín}\ \{R(u) :\ u \in C_0^1[a,b] - \{0\}\}.$$

Demonstração. Suponha que $u \in C_0^1[a,b] - \{0\}$, tome $c = H(u)$. Então

$$R(u_0) \leq R(c^{-1/2}u) = R(u)$$

. □

Exemplo 3.2.1. Considere o problema de Sturm-Liouville

$$u'' - xu = -\lambda u$$
$$u(0) = u(1) = 0.$$

O quociente de Rayleigh associado é dado por

$$R(u) = \frac{\int_0^1 u'(x)^2 + xu(x)^2dx}{\int_0^1 u(x)^2dx} = \frac{\int_0^1 u'(x)^2dx}{\int_0^1 u(x)^2dx} + \frac{\int_0^1 xu(x)^2dx}{\int_0^1 u(x)^2dx},$$

então o autovalor mínimo λ_1 satisfaz

$$\min \frac{\int_0^1 u'(x)^2dx}{\int_0^1 u(x)^2dx} \leq \lambda_1 = \min\ R(u) \leq 1 + \min \frac{\int_0^1 u'(x)^2dx}{\int_0^1 u(x)^2dx}.$$

Agora podemos estimar λ_1 pois $\dfrac{\int_0^1 u'(x)^2dx}{\int_0^1 u(x)^2dx}$ é o quociente de Rayleigh do problema

$$u''(x) = -\mu u(x)$$

com condições de contorno

$$u(0) = u(1) = 0.$$

Seja μ_1 o autovalor mínimo do problema de Sturm-Liouville acima, então $\mu_1 \leq \lambda_1 \leq 1 + \mu_1$. Fazendo um cálculo simples obtemos os autovalores $\mu_n = n^2\pi^2, n \in \mathbb{N}$. Logo $\pi^2 \leq \lambda_1 \leq 1 + \pi^2$.

3.3 O Quociente de Rayleigh e os demais Autovalores

Sabemos, pela proposição (3.1.5), que quaisquer duas autofunções com autovalores distintos são ρ-ortogonais. Assim, em busca de uma caracterização para o segundo autovalor λ_2, parece uma boa tentativa minimizar o quociente de Rayleigh R em $\{u \in C_0^1[a,b] : H(u) = 1\}$ com a restrição

$$H_1(u) = \int_a^b \rho(x)u(x)u_1(x)dx = 0,$$

onde u_1 é autofunção associada a λ_1

A partir de uma generalização desta idéia, obtemos o seguinte resultado:

Teoreoma 3.3.1. *Se $\lambda_1 < \ldots < \lambda_{n-1}$ são os primeiros autovalores com autofunções normalizadas $u_1, \ldots, u_{n-1}$ então*

$$\lambda_n = m\acute{\imath}n\{R(u) : u \in C_0^1[a,b] : H(u) = 1, H_i(u) = \int_a^b \rho(x)u(x)u_i(x)dx = 0,\ i = 1, \ldots, n-1\}.$$

Demonstração. Suponha que u^* é um mínimo em $\{u \in C_0^1[a,b] :\ H(u) = 1,\ H_i(x) = 0,\ i = 1, \ldots, n-1\}$. para R. Por definição,

$$H_i(u) = \int_a^b h_i[u(x)]dx,\ i = 1, \ldots, n$$

onde $h_i(x,y,z) = \rho(x)u_i(x)y$, $(h_i)_y$, $(h_i)_z \in C\left([a,b] \times \mathbb{R}^2\right)$ e H_i possui variação de Gâteaux dada por

$$\delta H_i(u,v) = \int_a^b \rho(x)u_i(x)v(x)dx.$$

Agora note que

$$\delta H_i(u^*, u_j) = \begin{cases} H(u_j), & \text{se} \quad i = j \\ 0, & se \quad i \neq j. \end{cases}$$

e portanto o determinante $(n-1) \times (n-1)$

$$\det\left(\delta H_i(u^*, u_j)\right) = H(u_1) \ldots H(u_{n-1}) > 0. \tag{3.3.1}$$

Como, em particular, u^* é um mínimo local em $C_0^1[a,b][H = 1, H_i = 0,\ i = 1, \ldots, n-1]$ para R e o determinante (3.3.1) é não identicamente nulo em $C_0^1[a,b]$, segue do teorema (2.5) que existem $-\lambda^*, \lambda_1^*, \ldots, \lambda_{n-1}^* \in \mathbb{R}$ tais que u^* é solução da equação de Euler-Lagrange para $\widetilde{f} = f - \lambda^* h + \lambda_1^* h_1 + \ldots + \lambda_{n-1}^* h_{n-1}$, isto é,

$$\frac{d}{dx}\widetilde{f}_z[u^*(x)] - \widetilde{f}_y[u^*(x)] = 0. \tag{3.3.2}$$

Substituindo as derivadas parciais $\widetilde{f}_z[u^*(x)]$, $\widetilde{f}_y[u^*(x)]$ em (3.3.2), ficamos com a seguinte equação:

$$\frac{d}{dx}\left(2\tau(x)\frac{d}{dx}u^*(x)\right) + 2q(x)u^*(x) + 2\lambda^*\rho(x)u^*(x) - \lambda_1^*\rho(x)u_1(x) - \ldots - \lambda_{n-1}^*\rho(x)u_{n-1}(x) = 0.$$

Com o uso de duas integrações por partes e das condições $u(a) = u(b) = 0,\ H_i(u) = 0,\ i = 1, \ldots, n-1\}$, obtemos

$$\begin{aligned}
\lambda_j^* \int_a^b \rho(x) u_j(x)^2 dx &= 2\int_a^b \left[\frac{d}{dx}\left(\tau(x)\frac{d}{dx}u^*(x)\right)u_j(x) + q(x)u^*(x)u_j(x)\right]dx \\
&= -\int_a^b \tau(x)\frac{d}{dx}u^*(x)\frac{d}{dx}u_j(x)dx + \int_a^b q(x)u^*(x)u_j(x)dx \\
&= \int_a^b \frac{d}{dx}\left(\tau(x)\frac{d}{dx}u_j(x)\right)u^*(x)dx + \int_a^b q(x)u^*(x)u_j(x)dx \\
&= \int_a^b \left[\frac{d}{dx}\left(\tau(x)\frac{d}{dx}u_j(x)\right) + q(x)u_j(x)\right]u^*(x)dx \\
&= -\lambda_j \int_a^b \rho(x)u_j(x)u^*(x)dx \\
&= 0.
\end{aligned}$$

Logo $\lambda_j^* = 0,\ j = 1, \ldots, n-1$ e consequentemente

$$\frac{d}{dx}\left(\tau(x)\frac{d}{dx}u^*(x)\right) + q(x)u^*(x) = -\lambda^*\rho(x)u^*(x).$$

Portanto $\lambda^* = R(u^*)$ é autovalor com autofunção u^*.

Assim, fica pendente a prova de que o autovalor λ^* é de fato o n-ésimo autovalor λ_n. Pois bem, suponha que λ^* é um dos primeiros $n-1$ autovalores $\lambda_1, \ldots, \lambda_{n-1}$. Então, como todo autovalor tem multiplicidade 1, existe $c \in \mathbb{R}$ tal que $u^* = cu_j$ para algum $j \in \{1, \ldots, n-1\}$. e isto implica

$$\int_a^b \rho(x)u^*(x)u_j(x)dx > 0,$$

o que é uma contradição. Se porém, λ_m é um autovalor com $m \geq n$, segue da ortogonalidade das autofunções que existe $u_m \in \{u \in C_0^1[a,b] :\ H(u) = 1,\ H_i(x) = 0,\ i = 1, \ldots, n-1\}$ tal que $R(u_m) = \lambda_m$. Pela minimalidade de u^* concluimos que $\lambda^* \leq \lambda_m$ e portanto $\lambda^* = \lambda_n$.

□

Observação 3.3.1. Note que a condição $H = 1$ é irrelevante assim como no princípio de Rayleigh, pois o quociente de Rayleigh é invariante por multiplicação por escalar em u.

3.4 O Princípio Minimax de Courant

Dados $\phi_1, \ldots, \phi_{n-1} \in C[a,b]\}$, definimos

$$C(\phi_1, \ldots, \phi_{n-1}) = \text{mín}\{R(u) : u \in C_0^1[a,b] - \{0\},\ \Phi_i(u) = \int_a^b \phi_i(x)u(x)dx = 0\ \forall\ i\}.$$

Lema 3.4.1. *$C(\phi_1, \ldots, \phi_{n-1}) \leq \lambda_n$ para todo $\phi_1, \ldots, \phi_{n-1} \in C[a,b]$.*

Demonstração. Dado $\phi_1, \ldots, \phi_n \in C[a,b]$, a prova consiste em construir uma função u^* em

$$\{u \in C_0^1[a,b] - \{0\},\ \Phi_i(x) = \int_a^b \phi_i(x)u(x)dx = 0,\ i = 1, \ldots, n-1\}$$

tal que $R(u^*) \leq \lambda_n$. Nós podemos considerar

$$u^* = \sum_{i=1}^{n} c_i u_i$$

e escolher $c_1, \ldots, c_n$ não todas nulas tais que u^* satisfaça as condições de ortogonalidade

$$\Phi_j(u^*) = \sum_{i=1}^{n} c_i \Phi_j(u_i) = 0, \quad j = 1 \ldots, n-1, \tag{3.4.1}$$

pois (3.4.1) é um sistema linear homogêneo com n incógnitas e $n-1$ equações e portanto admite solução não trivial. O quociente de Rayleigh de u^* é dado por

$$R(u^*) = \frac{D\left(\sum_{i=1}^{n} c_i u_i\right)}{H\left(\sum_{i=1}^{n} c_i u_i\right)} = \frac{\sum_{1 \leq i,j \leq n} c_i c_j \alpha_{ij}}{\sum_{1 \leq i,j \leq n} c_i c_j \beta_{ij}}, \tag{3.4.2}$$

onde

$$\begin{aligned}
\alpha_{ij} &= \int_a^b [\tau(x)u_i'(x)u_j'(x) - q(x)u_i(x)u_j(x)]dx \\
&= -\int_a^b \left\{\frac{d}{dx}[\tau(x)\frac{d}{dx}u_i(x)] + q(x)u_i(x)\right\} u_j(x)dx \\
&= \lambda_i \int_a^b \rho(x)u_i(x)u_j(x)dx, \\
\beta_{ij} &= \int_a^b \rho(x)u_i(x)u_j(x)dx.
\end{aligned} \tag{3.4.3}$$

Da ortogonalidade das autofunções e de (3.4.3) segue que

$$R(u^*) = \frac{\sum_{i=1}^{n} c_i^2 \lambda_i \beta_{ii}}{\sum_{i=1}^{n} c_i^2 \beta_{ii}}. \tag{3.4.4}$$

Como $\lambda_1 < \ldots < \lambda_n$ e $\beta_{ii} = \int_a^b \rho(x)u_i(x)^2 dx > 0$ temos finalmente que

$$R(u^*) \leq \frac{\sum_{i=1}^{n} c_i^2 \lambda_n \beta_{ii}}{\sum_{i=1}^{n} c_i^2 \beta_{ii}} = \lambda_n.$$

Portanto, pela minimalidade concluimos que

$$C(\phi_1, \ldots, \phi_{n-1}) \leq \lambda_n$$

□

Em particular, se $\phi_i(x) = \rho(x)u_i(x),\ i = 1, \ldots, n-1$, onde $u_1, \ldots, u_{n-1}$ são as primeiras $(n-1)$ autofunções, então $C(\rho u_1, \ldots, \rho u_{n-1}) = \lambda_n$.

Assim nós provamos que λ_n é o máximo sobre o conjunto da funções contínuas $\phi_1, \ldots, \phi_{n-1}$ do mínimo do quociente de Rayleigh das funções ortogonais a ϕ_j, $j = 1, \ldots, n$:

Teoreoma 3.4.1. ***(Princípio minimax de Courant)***

$$\lambda_n = máx\,\{C(\phi_1, \ldots, \phi_{n-1}) :\ \phi_1, \ldots, \phi_{n-1} \in C[a,b]\}.$$

Agora vamos utilizar o princípio minimax para comparar os autovalores de dois problemas de Sturm-Liouville distintos.

Sejam λ_n, λ_n^* os n-ésimos autovalores dos problemas de Sturm-Liouville

$$\frac{d}{dx}\left[\tau(x)\frac{d}{dx}u(x)\right] + q(x)u(x) = -\lambda\rho(x)u(x), \tag{3.4.5}$$

$$\frac{d}{dx}\left[\tau^*(x)\frac{d}{dx}u(x)\right] + q^*(x)u(x) = -\lambda^*\rho^*(x)u(x), \tag{3.4.6}$$

respectivamente com condições de contorno $u(a) = u(b) = 0$. Relembramos que os quocientes de Rayleigh associados a (3.4.5) e (3.4.6) são dados por

$$\begin{aligned} R(u) &= \frac{\int_a^b \tau(x)u'(x)^2 - q(x)u(x)^2 dx}{\int_a^b \rho(x)u(x)^2 dx} \\ R^*(u) &= \frac{\int_a^b \tau^*(x)u'(x)^2 - q^*(x)u(x)^2 dx}{\int_a^b \rho^*(x)u(x)^2 dx}. \end{aligned}$$

Essencialmente, para comparar os autovalores basta investigar o que acontece com o quociente de Rayleigh:

Lema 3.4.2. *Se $R(u) \leq R^*(u)$ para todo $u \in C_0^1[a,b] - \{0\}$, então*

$$\lambda_n \leq \lambda_n^*,\ n \in \mathbb{N}. \tag{3.4.7}$$

Demonstração. Fixados $\phi_1, \ldots, \phi_{n-1} \in C[a,b]$, temos, em particular, que

$$R(u) \leq R^*(u)\ \forall\ u \in C_0^1[a,b] - \{0\}[\Phi_i = 0,\ i = 1, \ldots, n-1]$$

Logo $C(\phi_1, \ldots, \phi_{n-1}) \leq C^*(\phi_1, \ldots, \phi_{n-1})$. Portanto, pelo princípio minimax, segue que

$$\lambda_n \leq \lambda_n^*.$$

□

No entanto, na prática não é conveniente comparar em cada ponto o quociente de Rayleigh. Por esse motivo é bem mais útil considerar as funções τ, ρ e q que caracterizam o problema de Sturm- Liouville e assim obter um critério a partir delas, como no corolário seguinte:

Corolario 3.4.1. *Se* $\tau(x) \leq \tau^*(x),\ q(x) \geq q^*(x),\ \rho(x) \geq \rho^*(x)$, *para todo* $x \in [a,b]$, *então*

$$\lambda_n \leq \lambda_n^*,\ n \in \mathbb{N}.$$

Demonstração. Basta notar que $R(u) \leq R^*(u)$ em $C_0^1[a,b] - \{0\}$. □

Em particular, tomando

$$\begin{aligned} \tau^* &= \tau_M = \text{máx}\ \tau(x), \\ q^* &= q_m = \text{mín}\ q(x), \\ \rho^* &= \rho_m = \min\ \rho(x), \end{aligned}$$

tem-se $\lambda_n \leq \lambda_n^*$. Note que o problema de Sturm-Liouville associado a λ^* é dado por

$$u''(x) = -\frac{q_m + \lambda^* \rho_m}{\tau_M} u(x) \tag{3.4.8}$$

$$u(a) = u(b) = 0, \tag{3.4.9}$$

e seu n-ésimo autovalor é $\lambda_n^* = \dfrac{-q_m}{\rho_m} + \dfrac{n^2\pi^2\tau_M}{(b-a)^2}$ com autofunção

$$u_n^*(x) = \text{sen}\left[\sqrt{\frac{q_m + \lambda_n^* \rho_m}{\tau_M}}(x-a)\right].$$

Com isso conseguimos a cota superior

$$\lambda_n \leq \frac{n^2\pi^2}{(b-a)^2}\frac{\tau_M}{\rho_m} - \frac{q_m}{\rho_m}.$$

Analogamente, tomando

$$\tau_m = \text{mín}\tau(x),\ q_M = \text{máx}\ q(x),\ \rho_M = \text{máx}\rho(x)$$

obtemos a cota inferior

$$\lambda_n \geq -\frac{q_M}{\rho_M} + \frac{n^2\pi^2}{(b-a)^2}\frac{\tau_m}{\rho_M}.$$

Portanto, o n-ésimo autovalor satisfaz

$$-\frac{q_M}{\rho_M} + \frac{n^2\pi^2}{(b-a)^2}\frac{\tau_m}{\rho_M} \leq \lambda_n \leq \frac{n^2\pi^2}{(b-a)^2}\frac{\tau_M}{\rho_m} - \frac{q_m}{\rho_m}.$$

Referências

BAGGETT, L. W. *Functional Analysis: A Primer*. [S.l.]: Dekker, 1992. Citado na página 7.

CAVALHEIRO, A. C. *Minicurso: O Problema de Sturm-Liouville*. [S.l.]: Departamento de Matemática Universidade Estadual de Londrina, 2012. Citado na página 7.

CORMANI, V.; RYHAM, R. Variational Techniques for Sturm-Liouville Eigenvalue Problems. 2002. Citado na página 46.

FALSET, J. G. *Análisis Funcional No Lineal*. [s.n.], 2007. Disponível em: <http://www.uv.es/garciaf/apuntes/analv.PDF>. Acesso em: 25 de jun. 2016. Citado na página 7.

LI, Y. *Lectures on Variational Methods*. [s.n.], 2012. Disponível em: <http://www.math.jhu.edu/~yli/Variational%20methods.pdf>. Acesso em: 24 de jun. 2016. Citado na página 7.

NACHBIN, L. *Introdução à Análise Funcional, Espaços de Banach e Cálculo Diferencial*. [S.l.]: Oea, 1976. Citado na página 7.

SAGAN, H. *Introduction to the Calculus of Variations*. [S.l.]: Courier Corporation, 2012. Citado na página 7.

SOTOMAYOR, J. *Lições de Equações Diferenciais Ordinárias*. [S.l.]: Instituto de Matemática Pura e Aplicada, CNPq, 1979. v. 11. Citado na página 7.

TROUTMAN, J. L. *Variational Calculus and Optimal Control: Optimization with Elementary Convexity*. [S.l.]: Springer Science & Business Media, 2012. Citado 2 vezes nas páginas 7 e 27.