

Cálculo Variacional ou das Variações



<https://doi.org/10.56238/sevened2023.006-155>

Henrique Mariano do Amaral

Universidade Estadual do Maranhão (UEMA), Mestrado em Engenharia Mecânica - UNICAMP, Mestrado em Engenharia da Computação UFRJ. Atua no CCT, São Luís, Maranhão, Brasil. Campus São Luís: Cidade Universitária Paulo VI, Av. Avenida Lourenço Vieira da Silva; N.º 1000, CEP: 65.055-310, Jardim São Cristóvão – São Luís/ MA, Brasil.

E-mail: hmca13@gmail.com

Ivanildo Silva Abreu

Pós Doutorado em Automação e Controle (EESC - USP), Professor Adjunto IV, DEMATI - CECEN; Universidade Estadual do Maranhão (UEMA), São Luís, Maranhão, Brasil. Campus São Luís: Cidade Universitária Paulo VI, Av. Avenida Lourenço Vieira da Silva; N.º 1000, CEP: 65.055-310, Jardim São Cristóvão – São Luís/ MA, Brasil.

E-mail: ivanildoabreu@yahoo.com.br

Cristovam Filho Dervalmar Rodrigues Teixeira

Universidade Estadual do Maranhão (UEMA), Mestrado em Engenharia da Computação e Sistemas - UEMA, Professor Assistente II, São Luís, Maranhão, Brasil. Campus São Luís: Cidade Universitária Paulo VI, Av. Avenida Lourenço Vieira da Silva; N.º 1000, CEP: 65.055-310, Jardim São Cristóvão – São Luís/ MA, Brasil.

E-mail: cristovamfilho17013@gmail.com

Elder Abreu Junior

Mestrado em Matemática (PROFMAT), Professor Substituto da Universidade Federal do Maranhão (UFMA) Cidade Universitária do Bacanga, s/n – Bacanga, São Luís, Ma, Av. dos Portugueses, N.º 1966, CEP: 65080-805, Vila Bacanga; São Luís Maranhão Brasil.

E-mail: elderabreu@yahoo.com.br

Rildenir Ribeiro Silva

Mestrado em Engenharia da Computação e Sistemas-UEMA, Instituto Federal de Educação Ciência e Tecnologia do Maranhão-IFMA Campus Rosário, Avenida Principal, S/N, Residencial VAÍparaiso, Sapucaia, Rosário-MA/CEP: 65150-000 –Campus Rosário;

E-mail: ril.ifmatematico@gmail.com

Alan Jefferson Lima Aragão

Licenciado em Matemática pela Universidade Estadual do Piauí - UESPI e Mestrando em Matemática pelo programa PROFMAT - UEMA, Professor efetivo da Seduc - MA e Professor da rede particular de ensino de Teresina - PI, Rua Antônio Guimarães, n.º2514, CEP: 65636-460; Parque Piauí - Timon/MA, Brasil.

E-mail: alanjeffersonlima@gmail.com

Vankys Ferreira Reis

Licenciado em Ciências - Habilitação em Matemática pela Universidade Estadual do Maranhão, Mestrando em Matemática pelo PROFMAT - UEMA, Professor efetivo em SEDUC - MA e SEMED - Alto Alegre do Pindaré – MA; Rua Rio Branco, S/N, Bairro: Trizidela, CEP: 65.398-000.

E-mail: vankysferreira@gmail.com

Karillos Alexandre Sousa Pereira

Especialista em Metodologia do Ensino da Matemática (UFPI) e Mestrando em Matemática pelo PROFMAT - UEMA, Professor Efetivo em Magalhães de Almeida - MA, SEMECT Av Getulio Vargas, 155 – centro, CEP 65 560 000, Magalhães de Almeida MA.

E-mail: karillos-alexandre@hotmail.com

Marcus Vinicius Viegas Rodrigues

Licenciado em Matemática pela Universidade Estadual do Maranhão, Especialista em Ensino da Matemática (IFMA) e Mestrando em Matemática pelo PROFMAT - UEMA, Professor Efetivo em Pedro do Rosário-MA, U.I. Nelson Marques, MA-006, s/n, CEP 65206-000, Povoado Anta.

E-mail: marcus.vinicius.viegas.2015@gmail.com

Danilo Furtado Veras

Discente do Programa de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT) na Universidade Estadual Brasil. Campus São Luís: Cidade Universitária Paulo VI, Av. Avenida Lourenço Vieira da Silva; N.º 1000, CEP: 65.055-310, Jardim São Cristóvão – São Luís/ MA, Brasil.

E-mail: daniloveras10@gmail.com

ABSTRACT

Este material de pesquisa sugere a exploração de abordagens para lidar com problemas variacionais por meio de técnicas de aproximação. Em contextos matemáticos, problemas variacionais envolvem otimização de funções, e os métodos de aproximação buscam encontrar soluções aproximadas para esses



problemas. Essas abordagens podem ser essenciais em situações em que encontrar uma solução exata é desafiador ou impraticável, permitindo a análise e resolução eficaz de questões complexas por meio de técnicas aproximativas.

Keywords: Métodos de aproximação, séries infinitas, MATLAB, condições de contorno.

1 INTRODUÇÃO

Anteriormente (vide volume 1 desta obra) definiu-se funcional como toda função numérica estabelecida sobre um espaço linear \mathcal{L} . Poder-se-ia dizer também, que funcionais são as magnitudes variáveis cujos valores se determinam mediante a eleição ou escolha de uma ou de várias funções.

Este capítulo foi acrescentado a este volume para dar uma fundamentação mais forte quando da apresentação dos métodos variacionais e energéticos.

Antes de se começar a descrever o *Cálculo das Variações*, serão apresentadas alguns conceitos e definições que se acredita sejam úteis para desenvolver melhor o entendimento do objeto deste capítulo.

Definição 1: Seja $x \rightarrow x_0$ denotar x se aproxima de x_0 pela esquerda e seja $x \rightarrow x_0^+$ denotar x se aproxima de x_0 pela direita. Se

$$\lim_{x \rightarrow x_0} f(x) \neq \lim_{x \rightarrow x_0^+} f(x) \quad (1.1.1)$$

diz-se que $f(x)$ é descontínua em x_0 , caso contrário é contínua em x_0 .

Definição 2: Uma função é dita contínua por partes em um intervalo qualquer, se ela possui um número finito de descontinuidades no intervalo considerado.

Definição 3: Uma função é dita ser diferenciável em x_0 se o limite existe. Ela é dita ser diferenciável por parte num intervalo qualquer, se ela possui uma derivada à direita e à esquerda para todo ponto interior do intervalo em consideração e, além disso, estas derivadas são iguais, com exceção em um número finito de pontos.

$$\lim_{x \rightarrow x_0} \left\{ \frac{f(x) - f(x_0)}{x - x_0} \right\} \quad (1.1.2)$$



Definição 4: Seja a função $v = f(x_1, x_2, \dots, x_n)$ onde cada variável é função de outras variáveis, assim $x_i = x_i(u_1, u_2, \dots, u_n)$. Então a derivada parcial de v em relação a uma determinada variável u_i é dada por:

$$\frac{\partial v}{\partial u_i} = \sum_{j=1}^n \frac{\partial f}{\partial x_j} \cdot \frac{\partial x_j}{\partial u_i} \quad (1.1.3)$$

Definição 5: A quantidade

$$p(x, y) + q(x, y) \frac{dy}{dx} \quad (1.1.4)$$

é a derivada dg/dx de alguma função $g(x, y)$ de alguma função $y = y(x)$. Neste evento, $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$.

isto é

$$p = \frac{\partial g}{\partial x} \text{ e } q = \frac{\partial g}{\partial y}$$

$$\frac{dg}{dx} = \frac{\partial g(x, y)}{\partial x} + \frac{\partial g(x, y)}{\partial y} \cdot \frac{dy}{dx} \quad (1.1.5)$$

Definição 6: Se

$$I = I(\epsilon) = \int_{x_1(\epsilon)}^{x_2(\epsilon)} f(x, \epsilon) dx \quad (1.1.6)$$

Então

$$\frac{dI}{d\epsilon} = I'(\epsilon) = f(x_2, \epsilon) \frac{dx_2}{d\epsilon} - f(x_1, \epsilon) \frac{dx_1}{d\epsilon} + \int_{x_1(\epsilon)}^{x_2(\epsilon)} \frac{\partial f}{\partial \epsilon} dx \quad (1.1.7)$$

se, e somente se,

garante que $\partial f / \partial \epsilon$ é uma função contínua de ϵ e de x em $[x_1, x_2]$. No caso de x_1 e x_2 serem estritamente constantes, isto é, independentes de ϵ , o lado direito da expressão acima se reduz ao seu termo final uma vez que $\partial x_1 / \partial \epsilon \equiv 0$; $\partial x_2 / \partial \epsilon \equiv 0$.

Definição 7: Para se empregar repetidamente a regra de integração por partes é necessário e suficiente que as funções f e g sejam apenas diferenciáveis por partes no intervalo considerado

$$\int_{x_1}^{x_2} g \frac{df}{dx} dx = g f \Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} f \frac{dg}{dx} dx \quad (1.1.8)$$



Definição 8: Uma função $F(x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_{m+n})$ dita ser homogênea de grau n , nas variáveis x_{m+1}, \dots, x_{m+n} se, para uma constante arbitrária h , tem-se:

$$F(x_1, x_2, \dots, x_m, hx_{m+1}, \dots, hx_{m+n}) = h^n F(x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_{m+n}) \quad (1.1.9)$$

Qualquer função para a qual a expressão acima é válida satisfaz o teorema de Euler:

$$x_1 \frac{\partial F}{\partial x_1} + \dots + x_{m+n} \frac{\partial F}{\partial x_{m+n}} = nF(x_1, x_2, \dots, x_m, \dots, x_{m+n}) \quad (1.1.10)$$

Definição 9: A condição necessária para um mínimo (ou máximo) de uma função $F(x_1, x_2, \dots, x_n) = 0$ em relação às variáveis $x_i = x_1, x_2, \dots, x_n$ é que satisfaça às relações:

$$G(x_1, x_2, \dots, x_n) = C_k, \forall k = 1, 2, \dots, N \quad (1.1.11)$$

e

$$\frac{\partial F}{\partial x_i} = 0, \forall i = 1, 2, \dots, n \quad (1.1.12)$$

onde C_k são constantes e $F^+ = F + \sum_{k=1}^n \lambda_k G_k$. As constantes $\lambda_1, \dots, \lambda_n$ introduzidas como incógnitas e denominadas de multiplicadores de Lagrange, são calculadas juntamente com os valores de minimização (ou maximização) de $x_i, \forall i = 1, 2, \dots, n$, por meio de um conjunto de equações formadas por (1.1.11) e (1.1.12).

Definição 10: A integral de linha de uma função $f(x, y, z)$ desde o ponto P1 até o ponto P2 ao longo de uma curva C é definida por:

$$I = \lim_{n \rightarrow \infty} S_n = \int_C f(x, y, z) ds \quad (1.1.13)$$

Onde $S_n = \sum F(x_k, y_k, z_k) \Delta S_k$ sendo ΔS_k comprimento dos arcos da curva C entre os pontos

$(x_{k-1}, y_{k-1}, z_{k-1})$ e (x_k, y_k, z_k) . A integral (1.1.13) pode também ser representada na forma:

$$\begin{aligned} I_x &= \int_C f(x, y, z) dx \\ I_y &= \int_C f(x, y, z) dy \\ I_z &= \int_C f(x, y, z) dz \end{aligned} \quad (1.1.14)$$



Introduzindo-se equações paramétricas $x = x(t)$, $y = y(t)$, $z = z(t)$ onde t cresce na direção do crescimento de s , pode-se calcular (3) pela integral definida

$$I = \int_{t_1}^{t_2} f[x(t), y(t), z(t)] \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt \quad (1.1.15)$$

Com $t_1 < t_2$.

Um importante exemplo de integral de linha é feita no sentido anti-horário sobre uma curva fechada no plano xy . Neste caso o parâmetro t é escolhido de tal forma que no ponto $[x(t), y(t)]$ percorra a curva C no sentido anti-horário quando t cresce de t_1 para t_2 , A integral acima é igual à área contida em C .

$$I = \frac{1}{2} \int_C (x dy - y dx) = \frac{1}{2} \int_{t_1}^{t_2} \left(x \frac{dy}{dt} - y \frac{dx}{dt} \right) dt \quad (1.1.16)$$

Definição 11: A mudança de variáveis $x = x(u, v, w)$, $y = y(u, v, w)$ e $z = z(u, v, w)$ no cálculo de integral tripla é feita por:

$$\iiint_{\Omega} F(x, y, z) dx dy dz = \iiint_{\Omega^*} f(u, v, w) \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| \quad (1.1.17)$$

onde f é a função F expressa em termo de u, v, w sendo Ω^* é a região Ω descrita pelas variáveis u, v, w e onde $\left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right|$ é o Jacobiano.

Definição 12: Se $z = z(x, y)$: uma função continuamente diferenciável de x e y , a área de uma porção da superfície representada por esta função é dada por

$$I = \iint_{\Omega} \left[1 + \left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2 \right]^{1/2} dx dy \quad (1.1.18)$$

onde a integração é realizada sobre um domínio Ω do plano xy sobre o qual é dada uma porção da superfície projetada.



Seja Ω uma região fechada e limitada no plano xy cujo contorno $\partial\Omega$ consiste de finitas curvas bem suaves. Sejam $f(x,y)$ e $g(x,y)$ funções contínuas tendo derivadas parciais em relação a x e y contínuas em algum subdomínio contido em Ω . Então a integral existe ao longo de todo o contorno $\partial\Omega$ de Ω , tal que esteja à esquerda quando se avança na direção da integração, isto é, a integral de linha deve ser avaliada no sentido anti-horário.

$$\iint_{\Omega} \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx dy = \oint_{\partial\Omega} f dx + g dy \quad (1.1.19)$$

Exemplo

Seja $w(x,y)$ uma função contínua com derivadas parciais de até segunda ordem contínuas em um domínio Ω do plano xy , do tipo indicado pelo teorema 1.

$$f = -\frac{\partial w}{\partial y}; g = \frac{\partial w}{\partial x} \quad (1.1.20)$$

Sejam então $\partial f/\partial y$ e $\partial g/\partial x$ existem e estão contidas em Ω . Seja o Laplaciano de w :

$$\nabla^2 w = \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \quad (1.1.21)$$

O segundo membro da equação integral que define o teorema de Green no plano, pode ser desenvolvido pela definição de integral de linha, assim

$$\oint_{\partial\Omega} (f dx + g dy) = \int_{\partial\Omega} \left(f \frac{dx}{ds} + g \frac{dy}{ds} \right) ds = \int_{\partial\Omega} \left(-\frac{\partial w}{\partial y} \frac{dx}{ds} + \frac{\partial w}{\partial x} \frac{dy}{ds} \right) ds \quad (1.1.22)$$

onde s é o comprimento de arco de $\partial\Omega$. O integrando da ultima integral acima, pode ser escrita como o produto escalar de dois vetores:

$$\nabla w = \text{grad } w = \frac{\partial w}{\partial x} \mathbf{i} + \frac{\partial w}{\partial y} \mathbf{j} \quad (1.1.23)$$



e

$$\mathbf{n} = \frac{dy}{ds} \mathbf{i} + \frac{dx}{ds} \mathbf{j} \quad (1.1.24)$$

isto é,

$$\nabla w \cdot \mathbf{n} = -\frac{\partial w}{\partial y} \frac{dx}{ds} + \frac{\partial w}{\partial x} \frac{dy}{ds} \quad (1.1.25)$$

onde \mathbf{n} é um vetor normal a $\partial\Omega$, uma vez que o vetor tangente a $\partial\Omega$

$$\mathbf{u} = \left(\frac{d\mathbf{r}}{ds} \right) = \frac{dx}{ds} \mathbf{i} + \frac{dy}{ds} \mathbf{j} \quad (1.1.26)$$

é ortogonal a \mathbf{n} , isto é $\mathbf{n} \cdot \mathbf{u} = 0$. Por outro lado, o produto escalar $\nabla w \cdot \mathbf{n} = \text{grad } w \cdot \mathbf{n}$ é a derivada direcional de w na direção de \mathbf{n} ; denotando-se este produto escalar por $\partial w / \partial \mathbf{n}$ e levando as expressões encontradas na fórmula do teorema de Gauss, se tem

$$\iint_{\Omega} \nabla^2 w d\Omega = \int_{\partial\Omega} \frac{\partial w}{\partial \mathbf{n}} ds = \int_{\partial\Omega} \nabla w \cdot \mathbf{n} ds \quad (1.1.27)$$

Ainda, sobre a expressão do Teorema de Green, se $g = \eta P$ e $f = -\eta Q$, se obtém:

$$\iint_{\Omega} \left(P \frac{\partial \eta}{\partial x} + Q \frac{\partial \eta}{\partial y} \right) d\Omega = - \iint_{\Omega} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \right) \eta d\Omega + \int_{\partial\Omega} P dy - Q dx \quad (1.1.28)$$

que é a fórmula de integração por parte para integrais duplas. Se na expressão acima si fizer $\eta = \psi$; $P = \partial\phi/\partial x$; $Q = \partial\phi/\partial y$ e de módulo análogo como se fez acima, encontramos um importante resultado do teorema de Green:

$$\iint_{\Omega} \psi \nabla^2 \phi d\Omega = - \iint_{\Omega} \left(\frac{\partial \phi}{\partial x} \frac{\partial \psi}{\partial x} + \frac{\partial \phi}{\partial y} \frac{\partial \psi}{\partial y} \right) d\Omega + \int_{\partial\Omega} \psi \frac{\partial \phi}{\partial \mathbf{n}} ds \quad (1.1.29)$$

1.1 TEOREMA FUNDAMENTAL DO CÁLCULO

Usar-se-á duas formas do teorema fundamental do cálculo, a primeira para o par função-gradiente e a outra para o par gradiente-Hessiano.



Lema 1^a: Seja f duas vezes continuamente diferenciável $f \in \mathcal{C}^2$ x na vizinhança de um segmento de linha entre os pontos ξ e $\mathbf{x} = \xi + \varepsilon, \varepsilon \in \mathbb{R}^n$; então se tem

$$f(\mathbf{x}) = f(\xi) + \int_0^1 \nabla f(\xi + t\varepsilon)^T \varepsilon dt; \quad (1.1.30)$$

$$\nabla f(\mathbf{x}) = \mathbf{g}(\mathbf{x}) = \nabla f(\xi) + \int_0^1 \nabla^2 f(\xi + t\varepsilon) \varepsilon dt$$

Lema 1^b: Seja f duas vezes continuamente diferenciável $f \in \mathcal{C}^2$ x na vizinhança de um ponto $\xi \in \mathbb{R}^n$ então para $\mathbf{h} \in \mathbb{R}^n$ com $\|\mathbf{h}\| \ll 1$ suficientemente pequeno, se tem

$$f(\xi + \mathbf{h}) = f(\xi) + \nabla f(\xi)^T \mathbf{h} + \frac{1}{2} \mathbf{h}^T \nabla^2 f(\xi) \mathbf{h} + o(\|\mathbf{h}\|^2)$$

$$= f(\xi) + \mathbf{g}(\xi)^T \mathbf{h} + \frac{1}{2} \mathbf{h}^T H(\xi) \mathbf{h} + o(\|\mathbf{h}\|^2) \quad (1.1.31)$$

Neste ponto seria bom o leitor fazer uma leitura do capítulo 2 do volume 1 desta obra (referencia [4]), especificamente nas seções 2.6 a 2.8. A seguir apresenta-se, a título de recordação, algumas definições já estabelecidas na referência supracitada.

1.2 FUNCIONAL

Definição 13: Seja um conjunto $S \subseteq \mathbb{R}^n$, um subconjunto $X \subseteq S$, e \mathbb{R} , e o conjunto dos números reais. Define-se funcional f sobre X a o mapeamento $f: X \rightarrow \mathbb{R}$. **Logo funcional é uma função numérica** cujo conjunto de valores de f é dado por R_f :

$$R_f = \{k \in \mathbb{R} \mid f(\mathbf{x}) = k \text{ } \therefore \mathbf{x} \in X\} \quad (1.1.32)$$

Neste capítulo ter-se-á sempre $S = \mathbb{R}^n$ o espaço vetorial de n dimensões. Assim, um típico elemento de $S \subseteq \mathbb{R}^n$ é escrito da seguinte forma:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \text{ } \therefore x_i \in \mathbb{R}$$



Definição 14: Para vizinhança de \mathbf{x} com raio ε como um subconjunto $V \subseteq S$, dado por:

$$V_{\mathbf{x}, \varepsilon} = \{ \mathbf{y} \in \mathbb{R}^n \mid \mathbf{0} \leq \|\mathbf{x} - \mathbf{y}\| < \varepsilon \} \quad (1.1.33)$$

Onde $\|\cdot\|$ denota a norma do espaço Euclidiano definido para todos os $\mathbf{x} \in \mathbb{R}^n$ por:

$$\|\mathbf{x}\| = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}$$

Definição 15: Seja um funcional f definido sobre $X \subset \mathbb{R}^n$. Então $\xi \in X$ é um minimizador local de f se existir um $\varepsilon > 0$ tal que

$$f(\xi) \leq f(\mathbf{x}), \forall \mathbf{x} \in V_{\xi, \varepsilon} \cap X \quad (1.1.34)$$

Se $f(\xi) < f(\mathbf{x}), \forall \mathbf{x} \in V_{\xi, \varepsilon} \cap X, \mathbf{x} \neq \xi$ então ξ é um minimizador forte local de f . De forma similar pode-se definir maximizador local e maximizador forte local.

Definição 16: (minimizador global). Seja um funcional f definido sobre $X \subset \mathbb{R}^n$. Então $\xi \in X$ é um inimizador global de f se existir um $\varepsilon > 0$ tal que

$$f(\xi) \leq f(\mathbf{x}), \forall \mathbf{x} \in X \quad (1.1.35)$$

1.3 GRADIENTE

Definição 17: (Gradiente). Seja $f \in C^1 X$. Então o gradiente de f em $\mathbf{x} \in X$ é o vetor dado por

$$g_{\mathbf{x}}^T = \nabla f^T = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \dots \quad \frac{\partial f}{\partial x_n} \right] \quad (1.1.36)$$

Se $f(\xi) < f(\mathbf{x}), \forall \mathbf{x} \in X, \mathbf{x} \neq \xi$ então ξ é um minimizador forte global de f .

De forma similar pode-se definir maximizador global e maximizador forte global.

Analisando o comportamento do funcional quadrático na vizinhança aberta (*bola aberta*) de um *minimizador local* ξ utilizando-se como ferramenta básica a expansão de Taylor sobre ξ , ver-se-á aparecerem duas entidades dependendo de que o funcional dado pertença a $C^1 X$ ou $C^2 X$ –



onde $C^m X$ representa o conjunto dos funcionais para os quais as derivadas parciais de ordem $\leq m$ existem e são contínuas em $X \subset \mathbb{R}^n$.

A primeira entidade é chamada de gradiente (definido por $\nabla f \mathbf{x} \in C^1 X$) e a segunda é denominada Hessiano ($H \mathbf{x} = \nabla^2 f \mathbf{x} \in C^2 X$).

1.4 HESSIANO

Definição 18: (Hessiano) Seja $f \in C^2 X$. Então o Hessiano de f em $\mathbf{x} \in X$ é a matriz simétrica dada por

$$H \mathbf{x} = \nabla^2 f = [h_{ij}]_{i,j=1}^n, \quad \therefore h_{ij} = \partial^2 f / \partial x_i \partial x_j \quad (1.1.37)$$

Ou ainda

$$H \mathbf{x} = \nabla^2 f = \begin{bmatrix} \partial^2 f / \partial x_1 \partial x_1 & \partial^2 f / \partial x_1 \partial x_2 & \cdots & \partial^2 f / \partial x_1 \partial x_n \\ \partial^2 f / \partial x_2 \partial x_1 & \partial^2 f / \partial x_2 \partial x_2 & \cdots & \partial^2 f / \partial x_2 \partial x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2 f / \partial x_n \partial x_1 & \partial^2 f / \partial x_n \partial x_2 & \cdots & \partial^2 f / \partial x_n \partial x_n \end{bmatrix}$$

A expansão de Taylor do funcional f sobre \mathbf{x} , é convenientemente expressa:

Se $f \in C^1$ ter-se-á

$$f \mathbf{x} + \mathbf{h} = f \mathbf{x} + \mathbf{g}^T \mathbf{x} \mathbf{h} + o \|\mathbf{h}\| \quad (1.1.38)$$

Se $f \in C^2 X$, por

$$f \mathbf{x} + \mathbf{h} = f \mathbf{x} + \mathbf{g}^T \mathbf{x} \mathbf{h} + \frac{1}{2} \mathbf{h}^T H \mathbf{x} \mathbf{h} + o \|\mathbf{h}\|^2 \quad (1.1.39)$$

Observar que de uma forma geral, se $f \in C^m X$, ter-se-á

$$f \mathbf{x} + \tau \mathbf{y} = \sum_{k=0}^{m-1} \frac{\tau^k}{k!} f^{(k)} \mathbf{x}; \mathbf{y} + \frac{\tau^m}{m!} f^{(m)} \mathbf{x} + \theta \tau \mathbf{y}; \mathbf{y} \quad (1.1.40)$$

$$\therefore f^{(m)} \mathbf{x}; \mathbf{y} = \left(y_1 \frac{\partial}{\partial x_1} + y_2 \frac{\partial}{\partial x_2} + \cdots + y_n \frac{\partial}{\partial x_n} \right)^m f \mathbf{x}$$



Definição 19: Seja $f \in C^1 X$. Diz-se que f é estacionário em $\xi \in X$ se tivermos o gradiente neste ponto nulo, isto é, $\mathbf{g}(\xi) = 0$. Isto é equivalente a dizer que se em um ponto qualquer de um $\xi \in X$ funcional $f \in C^1 X$ o gradiente (primeira derivada) é nulo, então em $\xi \in X$ se tem um ponto estacionário.

Teorema 2: Seja $f \in C^1 X$. Se $\xi \in X$ é um minimizador local de f , então f é estacionária em $\xi \in X$

Prova:

Em $f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \mathbf{g}^T(\mathbf{x}) \mathbf{h} + o(\|\mathbf{h}\|)$, colocando $\mathbf{x} = \xi$ e $\mathbf{h} = -\eta \mathbf{g}(\xi) \therefore \eta \in \mathbb{R}$, então $f(\mathbf{x} + \mathbf{h}) = f(\xi) - \eta \|\mathbf{g}(\xi)\|^2 + o(\|\mathbf{h}\|)$. Assumindo que f não é estacionário em $\xi \rightarrow \mathbf{g}(\xi) \neq 0$, então para valores de η suficientemente pequenos se tem que $-\eta \|\mathbf{g}(\xi)\|^2 + o(\|\mathbf{h}\|) < 0$, ou seja $f(\mathbf{x} + \mathbf{h}) < f(\xi)$, caso em que não pode ser um minimizador local.

O teorema 2 acima é uma condição necessária, mas não suficiente para que ξ seja um minimizador local. A condição suficiente é dada pelo teorema a seguir:

Teorema 3: Seja $f \in C^1 X$ e seja f estacionária em $\xi \in X$.

Então $\xi \in X$ é um minimizador forte local de f se a matriz do Hessiano $H(\xi)$ for positiva-definida (SPD)

Lembrem-se que se aprendeu na Álgebra Linear que as condições necessárias e suficientes para que uma matriz A simétrica seja positiva definida são:

- $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0, \forall \mathbf{x} \neq \mathbf{0}$

- $\forall \lambda_i = \text{eig } A \rightarrow \lambda_i > 0$

- Todas as submatrizes superiores A_k de A devem ter $\det A_k > 0$

- Todos os pivot's d_i (sem troca de linhas) tem que ser $d_i > 0$

Prova:

De $f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \mathbf{g}^T(\mathbf{x}) \mathbf{h} + \frac{1}{2} \mathbf{h}^T H(\mathbf{x}) \mathbf{h} + o(\|\mathbf{h}\|^2)$, assumindo que o gradiente é nulo

$\mathbf{g}(\xi) = 0$ si tem



$$\begin{aligned}
 f(\xi + \mathbf{h}) &= f(\xi) + \cancel{\mathbf{g}^T(\xi) \mathbf{h}} + \frac{1}{2} \mathbf{h}^T H(\xi) \mathbf{h} + o(\|\mathbf{h}\|^2) \\
 &= f(\xi) + \frac{1}{2} \mathbf{h}^T H(\xi) \mathbf{h} + o(\|\mathbf{h}\|^2)
 \end{aligned}
 \tag{1.1.41}$$

Como por hipótese $H(\xi)$ é positiva definida, então existe um número positivo λ_ξ (que pode ser dado pelo menor autovalor de $H(\xi)$) tal que

$$\mathbf{h}^T H(\xi) \mathbf{h} \geq \lambda_\xi \|\mathbf{h}\|^2, \quad \forall \mathbf{h} \in \mathbb{R}^n.$$

Assim, $f(\xi + \mathbf{h}) = f(\xi) + \frac{1}{2} \lambda_\xi \|\mathbf{h}\|^2 + o(\|\mathbf{h}\|^2)$ e a sua parte direita precisa ser positiva para valores suficientemente pequenos de $\|\mathbf{h}\|$, donde se tira $f(\xi + \mathbf{h}) > f(\xi)$ em alguma vizinhança de ξ .

Outra forma do Teorema 3 anterior é:

Teorema 3^a: Seja $f \in C^1 X$ e seja f estacionária em $\xi \in X$.

Então $\xi \in X$ é um minimizador forte local de f se o gradiente $\mathbf{g}(\xi) = 0$ e a matriz do Hessiano $H(\xi)$ for positiva-definida. Seja então $\mathbf{h} \in \mathbb{R}^n$ pelo teorema de Taylor, para pequenos valores de $t \in \mathbb{R}$, de forma que $f(\xi + t\mathbf{h}) - f(\xi) \geq 0$, logo:

$$\begin{aligned}
 f(\xi + t\mathbf{h}) &= f(\xi) + t\mathbf{g}^T(\xi) \mathbf{h} + \frac{1}{2} t^2 \mathbf{h}^T H(\xi) \mathbf{h} + o(t^2) \\
 \therefore \begin{cases} \mathbf{g}^T(\xi) = \nabla f(\xi)^T \\ H(\xi) = \nabla^2 f(\xi) \end{cases} \\
 \rightarrow \mathbf{g}^T(\xi) \mathbf{h} + \frac{1}{2} t^2 \mathbf{h}^T H(\xi) \mathbf{h} + o(t) \geq 0
 \end{aligned}
 \tag{1.1.42}$$

Se $t = 0$ e $\mathbf{h} = -\mathbf{g}(\xi)$ se obtém: $\|\mathbf{g}(\xi)\|^2 = 0 \Rightarrow \mathbf{g}(\xi) \equiv 0$, logo

$$\frac{1}{2} t^2 \mathbf{h}^T H(\xi) \mathbf{h} \geq 0 \xrightarrow[t \text{ é escalar}]{} \frac{1}{2} \mathbf{h}^T H(\xi) \mathbf{h} \geq 0, \quad \forall \mathbf{h} \in \mathbb{R}^n.$$

Pode-se definir analogamente o caso de maximizador local de f , assim, se $\xi \in X$ é um maximizador local, então $\xi \in X$ precisa ser um ponto estacionário, e se $H(\xi)$ for negativo definido ele será um maximizador forte local.

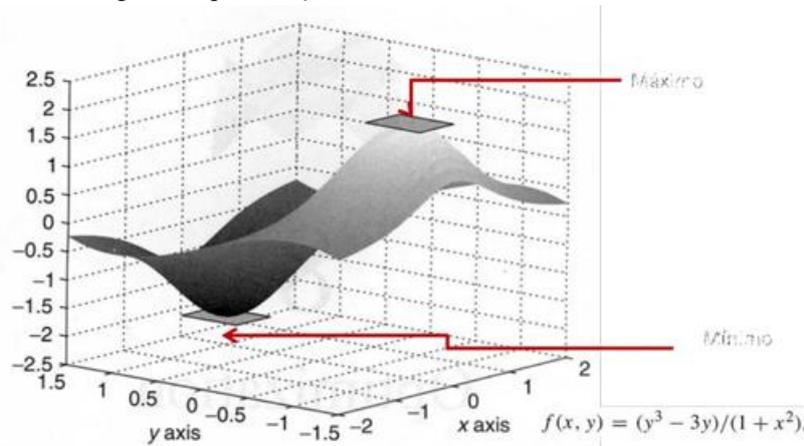
É fácil mostrar que se é estacionário, então

1. se $H(\xi)$ tiver autovalores positivos e negativos, ξ não é nem um minimizador local nem um maximizador local;



2. se H_ξ é positivo (negativo) semi-definido ele poder ser ou não ser um minimizador (maximizador) local.

Figura 1 representação e máximo e mínimo



1.4.1 Derivada Direcional

Definição 20: (derivada direcional). Seja $f \in C^m \ X$, $\mathbf{x} \in X \subseteq \mathbb{R}^n$ e seja $\mathbf{y} \in \mathbb{R}^n$ onde $\|\mathbf{y}\| = 1$. A derivada direcional de m -ésima ordem de f em \mathbf{x} na direção \mathbf{y} é dada por:

$$f^{(m)}(\mathbf{x}; \mathbf{y}) \equiv \left. \frac{d^m f(\mathbf{x} + \tau \mathbf{y})}{d\tau^m} \right|_{\tau=0} \quad (1.1.43)$$

A derivada direcional de f pode ser calculada pela regra da cadeia de diferenciação: assim, sendo se $f \in C^1 \ X$ si tem

$$\begin{aligned} f^{(1)}(\mathbf{x}; \mathbf{y}) &= \left. \frac{df}{d\tau} (x_1 + \tau y_1, x_2 + \tau y_2 + \dots + x_n + \tau y_n) \right|_{\tau=0} \\ &= \sum_{i=1}^n \frac{\partial}{\partial x_i} f(x_1, x_2, \dots, x_n) y_i = \mathbf{g}^T(\mathbf{x}) \mathbf{y} \end{aligned} \quad (1.1.44)$$

Se $f \in C^2 \ X$, si tem

$$f^{(2)}(\mathbf{x}; \mathbf{y}) = \mathbf{y}^T H(\mathbf{x}) \mathbf{y} \quad (1.1.45)$$

Reescrevendo a expressão de $f^{(1)}(\mathbf{x}; \mathbf{y})$ usando-se a desigualdade de Cauchy- Schwarz

($|\mathbf{x}^T \mathbf{y}| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$) se tem



$$\begin{aligned} \max_{y: \|y\|=1} |f^1 \mathbf{x}; \mathbf{y}| &= |f^1 \mathbf{x}; \psi| = |\mathbf{g}^T \mathbf{x} \psi| \leq \|\mathbf{g}^T \mathbf{x}\| \cdot \|\psi\|, \\ \text{mas } \|\psi\| &= 1 \text{ e } \|\mathbf{g}^T \mathbf{x}\| = \|\mathbf{g} \mathbf{x}\| \rightarrow \\ |f^1 \mathbf{x}; \psi| &= \|\mathbf{g} \mathbf{x}\| \Rightarrow \psi = \mathbf{g} \mathbf{x} / \|\mathbf{g} \mathbf{x}\| \end{aligned} \quad (1.1.46)$$

Onde ψ é a direção y em que $f^1 \mathbf{x}; \mathbf{y}$ é máxima.

Em outras palavras: $f^1 \mathbf{x}; \mathbf{y} = 0$ em todas as direções y se e somente se $\mathbf{g} \mathbf{x} = \mathbf{0}$, isto é, se o ponto \mathbf{x} for um ponto estacionário de f .

De modo análogo: $f^2 \mathbf{x}; \mathbf{y}$ para todas as direções y se e somente se $H \mathbf{x} > \mathbf{0}$ (a matriz Hessiana é positivo-definido).

Estas conclusões permitem que se enuncie dois teoremas fundamentais (que não serão demonstrados aqui).

Teorema 4: Seja $f \in C^1 X$. Se, $\xi \in X$ for um minimizador local de f , então $f^1 \mathbf{x}; \mathbf{y} = \mathbf{0}$ em todas as direções y .

Teorema 5: Seja $f \in C^2 X$. Suponha que para algum $\xi \in X$ se tenha $f \in C^1 X = \mathbf{0}$ em todas as direções y . Então $\xi \in X$ é um minimizador forte local de f se $f^2 \mathbf{x}, \mathbf{y} > \mathbf{0}$ em todas as direções y .

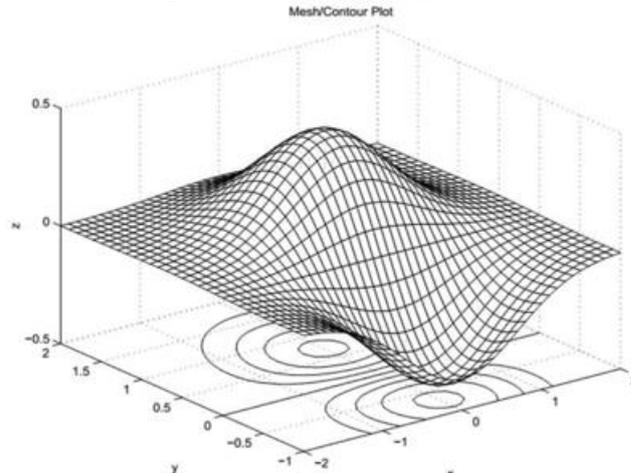
Definição 21: Seja f definido sobre $X \subseteq \mathbb{R}^n$ e seja $k \in \mathbb{R}$, onde R_f é um intervalo de f ; Então o conjunto

Usando o Matlab para gerar curvas de nível, pode-se escrever (por exemplo):

```
% plots de uma função de 2 variáveis
x = -2:0.1:2; y = -1:0.1:2;
[X Y] = meshgrid(x,y);
Z = Y.*exp(-(X.^2 + Y.^2));
meshc(X, Y, Z), ...
title('Mesh/Contour Plot'), xlabel('x'), ...
ylabel('y'), zlabel('z')
```



Figura 2: curvas de nível gerada pelo matlab



Um fato conhecido é o de que se ξ pertence a uma superfície (ou curva) de nível L_k , então o vetor gradiente $\nabla f(\mathbf{x})$ ou $g(\mathbf{x})$ é perpendicular a L_k em ξ e em todos os pontos em que a direção do funcional aumente mais rapidamente.

Portanto, gradientes próximos a um minimizador forte local aponta para fora enquanto que aqueles próximos a um maximizador forte local apontam para dentro.

Em geral, os métodos numéricos para achar um minimizador local forte de um funcional são mais ágeis quando as superfícies de nível na proximidade do minimizado são esféricas e são mais difíceis quando elas mostram uma acentuada distorção da forma de uma esfera. Se define a medida de distorção de L_k da forma esférica (forma esférica perfeita: $D_k = 1$) como a quantidade

$$D_k = \inf_{\xi \in S_k} \left\{ \frac{\sup_{x \in L_k} \|x - \xi\|}{\inf_{x \in L_k} \|x - \xi\|} \right\} \geq 1 \quad (1.1.47)$$

Onde S_k é o conjunto de todos os pontos interiores a L_k .

2 MÉTODO VARIACIONAL EM PROBLEMAS COM FRONTEIRAS FIXAS

2.1 PRELIMINARES

Equation Section (Next)O cálculo variacional visa fundamentalmente investigar os máximos e mínimos dos funcionais e se assemelha bastante à investigação de máximos e mínimos de funções. Portanto, neste item, procura-se introduzir os conceitos fundamentais e as principais propriedades do método variacional.

Um funcional é representado, em geral, por uma integral do tipo:



$$J f = \int_a^b I f, \dots, x dx \quad (2.1.1)$$

Definição 1: Denomina-se variação ou incremento δf do argumento do funcional $J f$, à diferença entre duas funções $f_1(x)$ e $f_0(x)$, pertencentes a uma mesma classe de funções M , considerada para o funcional $J f$:

$$\delta f = f_1(x) - f_0(x) \quad (2.1.2)$$

Para uma classe de funções de $C_k[a, b]$, isto é, k vezes diferenciáveis, obtém-se

$$\delta f^{(k)} = \delta f^{(k)} \quad (2.1.3)$$

Definição 2: Diz-se que as funções $f_1(x)$ e $f_0(x)$ definidas em um intervalo $[a, b]$ são próximas de ordem zero ou nula se no intervalo de definição tem se

$$|\delta f| = |f_1(x) - f_0(x)| \ll 1 \quad (2.1.3)$$

De uma forma geral, diz-se que estas funções são próximas de ordem k , se:

$$\begin{aligned} |\delta f| &= |f_1(x) - f_0(x)| \ll 1 \\ |\delta f| &= |f_1'(x) - f_0'(x)| \ll 1 \\ |\delta f| &= |f_1''(x) - f_0''(x)| \ll 1 \\ &\dots \\ |\delta f| &= |f_1^{(k)}(x) - f_0^{(k)}(x)| \ll 1 \end{aligned} \quad (2.1.4)$$

e assim se $f_1(x)$ e $f_0(x)$ forem próxima de ordem k , elas são próximas de qualquer ordem $j \leq k$.

2.1.1 Distância entre funções de um funcional

Definição 3: Denomina-se distância entre $f_0(x)$ e $f_1(x) \in C[a, b]$, com $x \in [a, b]$ à métrica definida nesse espaço por:

$$\rho_{f_0, f_1} = \max_{a < x < b} |f_1(x) - f_0(x)| \quad (2.1.5)$$



Generalizando, denomina-se distancia de n-ésima ordem entre $f_0(x)$ e $f_1(x) \in C_k[a, b]$ com $x \in [a, b]$ ao maior dos máximos das expressões do tipo

$$\max_{a < x < b} \left| f_1^k(x) - f_0^k(x) \right|, \forall k = 1, 2, \dots, n \quad (2.1.6)$$

ou ainda

$$\rho_n(f_0, f_1) = \max_{0 < k < n} \left[\max_{a < x < b} \left| f_1^k(x) - f_0^k(x) \right| \right] \quad (2.1.7)$$

2.1.2 Vizinhaça

Definição 4: Denomina-se ϵ -vizinhaça de n-ésima ordem (ver item 1.3.6) da função $f_0(x)$ com $x \in [a, b]$ ao conjunto f_i de funções cuja distancia de n-ésima ordem delas a $f_0(x)$ sejam menores que ϵ :

$$\rho_n(f_0(x), f_1(x)) < \epsilon \quad (2.1.8)$$

Se a ϵ -vizinhaça é de ordem zero, diz que ela é uma ϵ -vizinhaça forte de $f_0(x)$ enquanto a ϵ -vizinhaças de primeira ordem é chamada de ϵ -vizinhaça fraca.

Definição 5: Um funcional $J(f)$ definida em uma classe M de funções, se chama contínua em $f^* = f^*(x)$ no sentido da proximidade de n-ésima ordem, se qualquer que seja o número $\epsilon > 0$, existe um número $\delta > 0$, tal que a desigualdade

$$\left| J(f) - J(f^*) \right| < \epsilon \quad (2.1.9)$$

se cumpre para todas as funções admissíveis $f(x)$, ou seja, para todas as funções de satisfazem as condições

$$\begin{aligned} & \left| f(x) - f^*(x) \right| < \delta \\ & \dots \\ & \left| f^n(x) - f^{*n}(x) \right| < \delta \end{aligned} \quad (2.1.10)$$



Em outras palavras, tem-se $|J f - J f^*| < \varepsilon$ sempre que

$$\rho_n(f, f^*) < \delta \tag{2.1.11}$$

Exemplo:

Demonstrar que o funcional $J(f) = \int_0^1 (y + 2y') dx$ considerado no espaço $C_1[0,1]$ é contínuo na função $f^*(x) = x$ no sentido da proximidade de primeira ordem.

Solução: Seja $\varepsilon > 0$ e demonstraremos que existe um número $\delta > 0$ tal que

$$|J f - J f^*| < \varepsilon \tag{a)}$$

sempre que

$$\begin{aligned} |f(x) - f^*(x)| &= |f(x) - x| = |y - x| < \delta \\ |f'(x) - f^{*'}(x)| &= |f'(x) - 1| = |y' - 1| < \delta \end{aligned} \tag{b)}$$

Assim

$$|J(f) - J(x)| = \left| \int_0^1 (y + 2y' - x - 2) dx \right| \leq \int_0^1 |y - x| dx + 2 \int_0^1 |y' - 1| dx$$

Seja então $\delta = \varepsilon/3$. Logo para todas as funções $f(x) \in C_1[0,1]$ tais que as condições b) forem satisfeitas com $\delta = \varepsilon/3$, a condição (a) também será satisfeita, isto é

$$J(f) - J(x) \leq \int_0^1 |y - x| dx + 2 \int_0^1 |y' - 1| dx = \frac{\varepsilon}{3} + 2 \cdot \frac{\varepsilon}{3} = \varepsilon.$$

2.1.3 Incremento funcional

Definição 6: Dado um funcional $J f$ definido sobre uma classe M de funções $f(x)$, a magnitude:

$$\Delta J = \Delta J f = J f + \delta f - J f \tag{2.1.12}$$



se denomina incremento funcional correspondente ao incremento δf do argumento.

Definição 7: Se o incremento funcional ΔJ de $J f$ pode ser representado na forma

$$\Delta J = L f x, \delta f + \beta f x, \delta f \|\delta f\| \quad (2.1.13)$$

onde L é um funcional linear em relação a δf e $\beta \rightarrow 0$ quando $\|\delta f\| \rightarrow 0$, então o incremento linear em relação δf . isto é, $L f x, \delta f$ se chama variação do funcional e se representa por δJ . Neste caso si diz ser o funcional $J f$ diferenciável no ponto $f x$.

Definição 8: Seja um funcional linear $J f, x$ dependente dos elementos $f x$ e x

$$J = J f, x = \int_a^b I f, x dx \quad (2.1.14)$$

2.1.4 Forma Bilinear

Definição 9: Diz-se que J é uma forma bilinear se ele é um funcional linear em relação a f para x fixo e um funcional para x com $f x$ fixo, isto é

$$\begin{aligned} J f, \alpha x_1 + \beta x_2 &= \alpha J f, x_1 + \beta J f, x_2 \\ J \alpha f_1 + \beta f_2, x &= \alpha J f_1, x + \beta J f_2, x \end{aligned} \quad (2.1.15)$$

2.1.5 Funcional Quadrático

Definição 10: Um funcional linear $J x, x$ dependente dos elementos x e x , isto é, uma forma bilinear com $f x = x$, é chamada de funcional quadrático.

Definição 11: Um funcional quadrático é dito funcional positivo definido se $J x, x > 0$ qualquer que seja o elemento x diferente de zero.

Definição 12: Um funcional linear $J f$ definido sobre um espaço linear normado tem segunda variação se o seu incremento

$$\Delta J = J f + \delta f - J f \quad (2.1.16)$$



pode ser representado na forma

$$\Delta J = L_1 \delta f + \frac{1}{2} L_2 \delta f + \beta \|\delta f\|^2 \quad (2.1.17)$$

onde L_1 é um funcional linear, L_2 um funcional quadrático e $\beta \rightarrow 0$ quando $\|\delta f\| \rightarrow 0$. Ao funcional $L_2 \delta f$ se denomina segunda variação ou segunda diferencial do funcional $J f$ e se designa por $\delta^2 J$. Estudar-se-á os funcionais quadráticos com mais detalhes numa seção mais a frente.

2.1.6 Extremos de funcionais

Definição 13: Diz-se que um funcional $J f$ alcança seu máximo na função $f_0 x$ se os valores que toma o funcional $J f$ em qualquer função próxima de $f_0 x$ não são maiores que $J f_0$, ou seja,

$$\Delta J = J f x - J f_0 x \leq 0 \quad (2.1.18)$$

Como $\Delta J \leq 0 \rightarrow \Delta J = 0 \Leftrightarrow f = f_0$, diz-se que $J f$ alcança o máximo estrito

para a função $f_0 x$. Analogamente se define o mínimo do funcional em uma função $f_0 x$, quando $\Delta J \geq 0$ para todas as curvas próximas a $f_0 x$.

Exemplo

Demonstrar que o funcional $J f = \int_0^1 x^2 + y^2 dx$ alcança mínimo estrito na função $y x = 0$.

Solução: Qualquer que seja $y x \in [0,1]$ tem-se

$$\Delta J = J(f) - J(0) = \int_0^1 (x^2 + y^2) dx - \int_0^1 (x^2 + 0^2) dx = \int_0^1 y^2 dx \geq 0$$

E além do mais, $\Delta J = 0 \Leftrightarrow y x = 0$.

Definição 14: Diz-se que um funcional $J f$ alcança seu máximo relativo forte na função f_0 se $J f \leq J f_0$ em todas as funções admissíveis f pertencentes a uma ε -vizinhança de ordem nula da função f_0 . Analogamente se define mínimo relativo forte de um funcional quando $J f \geq J f_0$.

Definição 15: Diz-se que um funcional $J f$ alcança seu máximo relativo débil ou fraco na função f_0 se $J f \leq J f_0$ em todas as funções admissíveis f pertencentes a uma ε -vizinhança de primeira ordem da função f_0 . Analogamente se define mínimo relativo débil ou fraco de um funcional quando $J f \geq J f_0$.



Os máximos e mínimos (fortes e débeis ou fracos) de um funcional $J f$ se denominam extremos relativos. O extremo referente à totalidade das funções em que está definido o funcional é chamado extremo absoluto.

Todo extremo absoluto é ao mesmo tempo extremo relativo forte e débil, porém nem todo extremo relativo será extremo absoluto.

Exemplo

Seja o funcional $J f = \int_0^{\pi} y^2 (1 - y^2) dx$, no espaço de funções $y = f(x) \in C_1[0, \pi]$, que satisfazem as condições $y(0) = y(\pi) = 0$.

No segmento $[0, \pi]$ do eixo dos x tem mínimo débil de $J f$, isto é, $J f \geq J f_0$ em uma ε -vizinhança de primeira ordem. De fato, tem-se $J = 0$ se $y = 0$, por outro lado, as funções pertencentes a ε -vizinhança de primeira ordem, se tem $|y'| < 1$, de modo que o integrando é positivo para $y \neq 0$ e por conseguinte, o funcional se anula somente se $y = 0$. Isto é, o funcional alcança mínimo débil na curva $y = 0$.

Teorema 1: Condição Necessária de Extremo de Funcional. Se o funcional diferenciável $J f$ alcança seu valor extremo em uma curva $f = f_0$, sendo f_0 um ponto interior do campo de definição do funcional, então em $f = f_0$ se tem

$$\delta J f_0 = 0 \tag{11.2.19}$$

As funções para as quais $\delta J = 0$, denominam-se funções estacionárias.

2.1.6.1 Exercícios

1. Determinar a ordem de proximidade das curvas abaixo:

$$y(x) = \cos\left(\frac{nx}{n^2 + 1}\right) \quad \text{e} \quad y_1(x) = 0 \in [0, 2\pi]$$

$$y(x) = \operatorname{sen}\left(\frac{x}{n}\right) \quad \text{e} \quad y_1(x) = 0 \in [0, \pi]$$

2. Determinar a distância entre as curvas abaixo:

$$y(x) = x^{e^x} \quad \text{e} \quad y_1(x) = 0 \in [0, 2]$$

$$y(x) = \operatorname{sen} 2x \quad \text{e} \quad y_1(x) = 0 \in \left[0, \frac{\pi}{2}\right]$$

3. Achar a distância de primeira ordem entre as curvas $y(x) = \ln x$ e $y(x) = x$ no segmento $[e^{-1}, e]$.



4. Analisar a continuidade dos seguintes funcionais:

$J f x = f x_0 \rightarrow f x \in C[a,b] \wedge x \in [a,b]$ no sentido de proximidade de ordem nula.

$J f x = \int_0^1 |y'| dx, y \in C[0,1]$, no sentido da proximidade de ordem nula e de primeira ordem.

5. Analisar se são diferenciáveis os seguintes funcionais:

$$J f x = f a \in C[a,b]$$

$$J f x = f a \in C_1[a,b]$$

$$J f x = |f a| \in C[a,b]$$

6. Para os funcionais abaixo determinar, nos espaços correspondentes, suas variações.

$$J f x = \int_a^b x + y dx$$

$$J f x = \int_0^x y' \sin x dx$$

2.1.7 Problema Elementar do Cálculo Variacional

Seja o funcional

$$J(f) = \int_a^b I(f, f_x, x) dx \tag{11.2.20}$$

Onde $f \in [a,b] \wedge f_x = \frac{df}{dx} \in [a,b]$. O problema elementar do cálculo variacional consiste em achar a função que oferece extremo débil ao funcional (11.2.20) e que satisfaça as condições de contorno:

$$\begin{aligned} f a &= f_a \\ f b &= f_b \end{aligned} \tag{11.2.21}$$

Seja $f x$, por hipótese, a função solução do problema, e seja $h x$ outra função a qual difere de $f x$ certa quantidade:

$$\delta f x = \varepsilon \eta x \tag{11.2.22}$$

Onde $h x \in C_1[a,b]$, ε é um parâmetro que varia continuamente, e ηx é uma função arbitrária que satisfaz as condições de contorno abaixo:

$$\eta a = \eta b = 0 \tag{11.2.23}$$



De (11.2.22) e da figura abaixo, tem-se

$$h(x) = f(x) + \varepsilon \eta(x) \quad (11.2.24)$$

Dessa forma, pode-se reescrever o funcional (11.2.20) para a função $h(x)$, como:

$$J(h) = \int_a^b I(h, h_x, x) dx \quad (11.2.25)$$

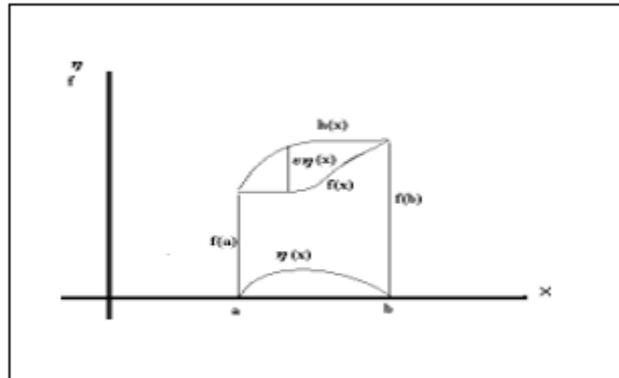


FIGURA 2.1

Onde de (11.2.24) vê-se

$$h_x = \frac{\partial h}{\partial x} = \frac{\partial f}{\partial x} + \varepsilon \frac{\partial \eta}{\partial x} \quad (11.2.26)$$

Como o valor do funcional J varia continuamente com ε , pela fórmula de Taylor, pode-se desenvolver J , assim

$$J(h) = J(f) + \varepsilon \left. \frac{\partial J}{\partial \varepsilon} \right|_{\varepsilon=0} + \frac{1}{2!} \varepsilon^2 \left. \frac{\partial^2 J}{\partial \varepsilon^2} \right|_{\varepsilon=0} + \dots \quad (11.2.27)$$

ou na notação variacional

$$J(h) = J(f) + \delta J(f) + \frac{1}{2!} \delta^2 J(f) + \dots \quad (11.2.28)$$

Onde $\delta J(f) = \varepsilon \left. \frac{\partial J}{\partial \varepsilon} \right|_{\varepsilon=0}$ e $\delta^2 J(f) = \varepsilon^2 \left. \frac{\partial^2 J}{\partial \varepsilon^2} \right|_{\varepsilon=0}$

Sabe-se que a condição de estacionariedade de um funcional é $\delta J(f) = 0$ ou $\left. \frac{\partial J}{\partial \varepsilon} \right|_{\varepsilon=0} = 0$. Logo, diferenciando (11.2.25), tem-se:

$$\delta J = \left(\frac{dJ}{d\varepsilon} \right)_{\varepsilon=0} = \left\{ \int_a^b \frac{dI}{d\varepsilon} dx \right\}_{\varepsilon=0} = \left\{ \int_a^b \left(\frac{\partial I}{\partial h} \frac{\partial h}{\partial \varepsilon} + \frac{\partial I}{\partial h_x} \frac{\partial h_x}{\partial \varepsilon} \right) dx \right\}_{\varepsilon=0} \quad (11.2.29)$$



De (11.2.23) e (11.2.26) tem-se que

$$\begin{aligned} \frac{\partial h}{\partial \varepsilon} &= \eta \quad x \\ \frac{\partial h_x}{\partial x} &= \eta_x \quad x \end{aligned} \quad (11.2.30)$$

Levando-se (11.2.30) em (11.2.29) encontra-se

$$\left(\frac{dJ}{d\varepsilon} \right)_{\varepsilon=0} = \left\{ \int_a^b \left(\frac{\partial I}{\partial h} \eta + \frac{\partial I}{\partial h_x} \eta_x \right) dx \right\}_{\varepsilon=0} = \int_a^b \left(\frac{\partial I}{\partial f} \eta + \frac{\partial I}{\partial f_x} \eta_x \right) dx = 0 \quad (11.2.31)$$

Pois $\left(\frac{\partial I}{\partial h} \right)_{\varepsilon=0} = \frac{\partial I}{\partial f}$ e $\left(\frac{\partial I}{\partial h_x} \right)_{\varepsilon=0} = \frac{\partial I}{\partial f_x}$

Integrando por parte o segundo termo da equação (11.2.31), se obtém

$$\left(\frac{dJ}{d\varepsilon} \right)_{\varepsilon=0} = \int_a^b \left(\frac{\partial I}{\partial f} \eta - \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) \eta \right) dx + \left. \frac{\partial I}{\partial f_x} \eta \right|_a^b = 0 \quad (11.2.32)$$

O último termo do segundo membro de (11.2.32) é nulo visto que $\eta|_a = \eta|_b = 0$ o que é verdadeiro e válido para qualquer e todas as funções admissíveis. Assim a equação (11.2.32) torna-se:

$$\left(\frac{dJ}{d\varepsilon} \right)_{\varepsilon=0} = \int_a^b \left(\frac{\partial I}{\partial f} - \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) \right) \eta dx = 0 \quad (11.2.33)$$

Enuncia-se agora o **Lema Básico** do Cálculo Variacional:

2.1.7.1 Lema Básico do Cálculo Variacional

Lema: Se a e b $b > a$ são constantes fixas e $G(x)$ é uma função contínua pertencente a $C(a,b)$ e se

$$\int_a^b \eta(x) G(x) dx = 0 \quad (11.2.34)$$

para toda função continuamente diferenciável $\eta(x)$ que satisfaça as condições $\eta|_a = \eta|_b = 0$ pode-se concluir que $G(x) = 0$ para todo o intervalo $a \leq x \leq b$.

Do lema básico e de (11.2.33) vê-se que

$$\frac{\partial I}{\partial f} - \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) = 0 \quad (11.2.35)$$



que é a *equação de Euler* ou ainda de *Euler-Lagrange*. As curvas integrais da equação de Euler se denominam extremos ou curvas de Lagrange.

2.1.7.2 Equação de Euler-Lagrange

Teorema 1: A condição necessária para que o funcional

$$J(f) = \int_a^b I(f, f_x, x) dx \quad (11.2.36)$$

definido no conjunto $C_1[a, b]$ e que satisfazem as condições de contorno $f_1 = f(a)$ e $f_2 = f(b)$, alcance seu valor extremo na função $f(x)$ é que esta função verifique a equação de Euler-Lagrange (11.2.35):

$$\frac{\partial I}{\partial f} - \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) = 0.$$

Esta condição é necessária para o extremo débil; mas como todo extremo forte é também ao mesmo tempo débil, qualquer condição para o extremo débil também será condição para o extremo forte.

Exemplo:

Seja o funcional $J(f) = \int_1^2 (y'^2 - 2xy) dx$ $\therefore y(1) = 0 \wedge y(2) = -1$.

A equação de Euler-Lagrange será: $y'' + x = 0$. Resolvendo a equação diferencial (eq. de *Euler-Lagrange*), tem-se

$$y = -\left(\frac{x^3}{6}\right) + \alpha x + \beta$$

Usando-se as condições de contorno, encontra-se, $\alpha = \frac{1}{6}$ e $\beta = 0$.

Por conseguinte, o extremo pode ser alcançado na função: $y = \left(\frac{x}{6}\right) - x^2$.

2.1.7.3 Exercícios

Achar a equação de Euler-Lagrange e a função em que o funcional dado pode alcançar extremo, para:



$$1. J f = \int_0^{\frac{\pi}{2}} f'^2 + f^2 dx \quad \therefore f(0) = 0 \wedge f\left(\frac{\pi}{2}\right) = 1.$$

$$2. J f = \int_1^3 (3x - f) f dx \quad \therefore f(1) = 1 \wedge f(3) = 4 \frac{1}{2}.$$

$$3. J f = \int_0^{2\pi} f'^2 + f^2 dx \quad \therefore f(0) = f(2\pi) = 1.$$

Teorema 2: Seja $f(x)$ solução da equação de Euler-Lagrange. Se a função $I(x, f, f_x)$ tem derivadas parciais contínuas até segunda ordem inclusive, então a função $f = f(x)$ tem segunda derivada contínua em todos os pontos x, y para os quais

$$\frac{\partial^2 I}{\partial f_x^2} \neq 0$$

Teorema 3: (Bernstein). Se na equação $f_{xx} = I(x, f, f_x)$ as funções I e $\frac{\partial I}{\partial f}$ forem contínuas em todo ponto finito x, f , para qualquer valor finito de f_x e ainda exista uma constante $k > 0$ e funções

$$\alpha = \alpha(x, f) \geq 0 \quad (11.2.38)$$

$$\beta = \beta(x, f) \geq 0$$

limitadas em qualquer porção finita do plano, tais que

$$\frac{\partial I}{\partial f} = I_f > k \quad (11.2.39)$$

$$|I| \leq \alpha f_x^2 + \beta$$

então, por dois pontos quaisquer do plano (x_1, f_1) e (x_2, f_2) de abscissas distintas $x_1 \neq x_2$ passa uma e somente uma curva integral $f = \varphi(x)$ de (11.2.39).

Exemplos

1. Demonstrar que por dois pontos quaisquer do plano de abscissas distintas passa uma única função extrema do funcional

$$J(y) = \int e^{2y^2} (y'^2 - 1) dx$$

A equação de Euler para o funcional acima é: $y'' = 2y - 1 + y'^2$ que está no formato (11.2.37) do Teorema de Bernstein, logo se pode aplicá-lo. De fato, tem-se:



$$f_{xx} = I_{x,y,y'} = 2y \sqrt{1+y'^2}$$

$$\frac{\partial I}{\partial f} = \frac{\partial I}{\partial y} = I_y = 2 \sqrt{1+y'^2}$$

Na expressão de I_y , vê-se que qualquer que seja o valor de y', y'^2 é positivo, logo $I_y \geq 2 = k$. Além disso, $|I| = |2y \sqrt{1+y'^2}| \leq 2|y| + 2|y|y'^2$. Comparando com a segunda expressão de (11.2.39) do Teorema de Bernstein, vê-se que:

$$\alpha = 2|y| > 0$$

$$\beta = 2|y| > 0$$

e que também satisfaz a condição (11.2.38) [11.2.38], passa uma única função extrema xy [11.2.38] Jy .

2. Demonstrar que não existe extremo do funcional $I(y) = \int y^2 + \sqrt{1+y'^2} dx$ que passe por dois pontos quaisquer do plano de abscissas distintas.

Solução

A equação de Euler para o funcional dado tem a forma:

$$y'' = 2y \sqrt{1+y'^2}^{3/2}$$

Analisando a expressão acima, vê-se que se pode aplicar o teorema de Bernstein, já que a expressão cumpre a condição (11.2.39) extrema do funcional dado.

Assim, sejam dois pontos quaisquer, por exemplo, $(0,0)$ e $(1/2, 2)$. Reescrevendo a equação de Euler do funcional dado, fazendo $y' = p; y'' = p \frac{dp}{dy}$.

$$p \frac{dp}{dy} = 2y \sqrt{1+p^2}^{3/2}$$

Separando as variáveis tem-se: $p \sqrt{1+p^2}^{3/2} dp = 2y dy$ que integrando se tem

$$- \sqrt{1+p^2}^{-1/2} = y^2 - C \Rightarrow C - \sqrt{1+p^2}^{-1/2} = y^2$$

de modo que

$$\frac{dx}{dy} = \sqrt{C - y^2}^{-1} \left[1 - \sqrt{C - y^2} \right]^{1/2}$$

onde C é uma constante real. Separando as variáveis e integrando de $(0,0)$ a $(1/2, 2)$, obtém-se:



$$\frac{1}{2} = \int_0^2 \left(\frac{C - y^2}{\sqrt{1 - (C - y^2)^2}} \right) dy$$

Analisando a expressão acima, vê-se que qualquer que seja C real, o denominador será complexo em certo intervalo $a, b \subset [0, 2]$ da variação de y . Logo, a igualdade é impossível. Isto quer dizer que não se pode traçar nenhum extremo pelos pontos considerados, logo, falha aqui o teorema de Bernstein.

2.1.7.4 Exercício

A equação de uma elipse é conhecida por ser derivada da condição de que a soma das distâncias de qualquer ponto A sobre ela a dois outros pontos fixos $F_1(x_1 = -c, y = 0)$ e $F_2(x_2 = +c, y = 0)$ é constante. Ache a tangente e a normal a uma elipse num ponto A qualquer. Ache os ângulos entre as linhas F_1A , F_2A e a normal no ponto A .

2.1.8 Generalizações do Problema Elementar do Cálculo Variacional

Seja agora determinar a equação de Euler-Lagrange para o funcional do tipo

$$J = \int_a^b I(f, f_x, f_{xx}, x) dx \quad (11.2.40)$$

onde f, f_x, f_{xx} são as variáveis dependentes. Seja então

$$\begin{aligned} f_a &= f \quad a \\ f_b &= f \quad b \\ f_{xa} &= f_x \quad a \\ f_{xb} &= f_x \quad b \end{aligned} \quad (11.2.41)$$

valores conhecidos. Aqui, neste caso, não apenas η (ou $\delta f = \varepsilon \eta$) é zero em a e b , mas também η_x ($\delta f_x = \varepsilon \eta_x$) é nulo nos pontos de contorno, então

$$\delta J = \int_a^b \left(\frac{\partial I}{\partial f} \delta f + \frac{\partial I}{\partial f_x} \delta f_x + \frac{\partial I}{\partial f_{xx}} \delta f_{xx} \right) dx = 0 \quad (11.2.42)$$

sendo $\delta f_{xx} = \varepsilon \eta_{xx}$. Integrando por parte os dois últimos termos a fim de colocar sob o sinal da integral apenas os termos multiplicados por δf , tem-se:

$$\int_a^b \left(\frac{\partial I}{\partial f} \delta f \right) dx = - \int_a^b \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) \delta f_x dx + \frac{\partial I}{\partial f_x} \delta f \Big|_a^b \quad (11.2.43)$$



e

$$\int_a^b \left(\frac{\partial I}{\partial f_{xx}} \delta f_{xx} \right) dx = \int_a^b \frac{d^2}{dx^2} \left(\frac{\partial I}{\partial f_{xx}} \right) \delta f dx - \frac{d}{dx} \frac{\partial I}{\partial f_{xx}} \delta f \Big|_a^b + \frac{\partial I}{\partial f_{xx}} \delta f_x \Big|_a^b \quad (11.2.44)$$

2.1.8.1 Equações de Euler-Lagrange de Quarta Ordem

Então, pode-se reescrever a expressão de δJ da seguinte maneira:

$$\begin{aligned} \delta J = & \int_a^b \left[\frac{\partial I}{\partial f} - \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) + \frac{d^2}{dx^2} \left(\frac{\partial I}{\partial f_{xx}} \right) \right] \delta f dx \\ & + \left[\frac{\partial I}{\partial f_x} - \frac{d}{dx} \left(\frac{\partial I}{\partial f_{xx}} \right) \right] \delta f \Big|_a^b + \left[\frac{\partial I}{\partial f_{xx}} \right] \delta f_x \Big|_a^b \end{aligned} \quad (11.2.45)$$

como as funções admissíveis são tais que $\delta f = \delta f_x = 0$ em $x = a$ e $x = b$, então os termos do contorno, na expressão acima, são nulos e a equação de Euler-Lagrange tem a seguinte forma:

$$\frac{\partial I}{\partial f} - \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) + \frac{d^2}{dx^2} \left(\frac{\partial I}{\partial f_{xx}} \right) = 0 \quad (11.2.46)$$

que é uma equação diferencial de quarta ordem.

2.1.8.2 Equações de Euler-Lagrange de Ordem Superior

Generalizando, para funcionais do tipo:

$$J f = \int_a^b I(f, f_x, f_{xx}, \dots, f_{n-1}, x) dx \quad (11.2.47)$$

para o qual assumir-se-á conhecidos em $x = a$ e $x = b$ a função f e as $n - 1$ derivadas de f . Neste caso tem-se:

$$\delta J = \int_a^b \left\{ \frac{\partial I}{\partial f} \delta f + \frac{\partial I}{\partial f_x} \delta f_x + \dots + \frac{\partial I}{\partial f_{x(n)}} \delta f_{x(n)} \right\} dx = 0 \quad (11.2.48)$$

Ou

$$\delta J = \int_a^b \left\{ \sum_{k=0}^{k=n} \frac{\partial I}{\partial f_{x(k)}} \delta f_{x(k)} \right\} dx = 0 \quad (11.2.49)$$

onde se deve integrar os diferentes termos por partes até que o integrando esteja multiplicado apenas por δf . Um termo típico de integração por partes é:



$$\int_a^b \frac{\partial I}{\partial f_{x^{(j)}}} \delta f_{x^{(j)}} dx = \left[\frac{\partial I}{\partial f_{x^{(j)}}} \delta f_{x^{(j-1)}} \right]_a^b - \left[\frac{d}{dx} \frac{\partial I}{\partial f_{x^{(j-1)}}} \delta f_{x^{(j-2)}} \right]_a^b + \left[\frac{d^2}{dx^2} \left(\frac{\partial I}{\partial f_{x^{(j-2)}}} \right) \delta f_{x^{(j-3)}} \right]_a^b + \dots + (-1)^{j-1} \left[\frac{d^{j-1}}{dx^{j-1}} \left(\frac{\partial I}{\partial f_x} \right) \delta f \right]_a^b + (-1)^j \int_a^b \left[\frac{d^j}{dx^j} \left(\frac{\partial I}{\partial f_{x^{(j)}}} \right) \delta f \right] dx \quad (11.2.50)$$

ou numa forma mais compacta:

$$\int_a^b \frac{\partial I}{\partial f_{x^{(j)}}} \delta f_{x^{(j)}} dx = \sum_{k=0}^{j-1} \left\{ (-1)^k \frac{d^k}{dx^k} \left(\frac{\partial I}{\partial f_{x^{(j-k)}}} \right) \delta f_{x^{(j-k-1)}} \right\} \Big|_a^b + (-1)^j \int_a^b \left[\frac{d^j}{dx^j} \left(\frac{\partial I}{\partial f_{x^{(j)}}} \right) \delta f \right] dx \quad (11.2.51)$$

A equação de Euler para funcionais deste tipo, tem então, a seguinte forma:

$$\frac{\partial I}{\partial f} - \frac{d}{dx} \frac{\partial I}{\partial f_x} + \dots + (-1)^j \frac{d^j}{dx^j} \left(\frac{\partial I}{\partial f_{x^{(j)}}} \right) + \dots + (-1)^n \frac{d^n}{dx^n} \left(\frac{\partial I}{\partial f_{x^{(n)}}} \right) = 0 \quad (11.2.52)$$

ou ainda

$$\sum_{k=0}^{k=n} \left\{ (-1)^k \frac{d^k}{dx^k} \left(\frac{\partial I}{\partial f_{x^{(k)}}} \right) \right\} = 0 \quad (11.2.53)$$

Exemplo

Achar a função extrema do funcional:

$$J y = \int_0^1 360x^2y - y''^2 dx \quad \therefore y(0) = y(1) = 0 \wedge y'(0) = 1 \wedge y'(1) = 2.5.$$

Solução:

A equação de Euler para este problema tem a forma:

$$\delta J = \frac{\partial I}{\partial y} - \frac{d}{dx} \left(\frac{\partial I}{\partial y_x} \right) + \frac{d^2}{dx^2} \left(\frac{\partial I}{\partial y_{xx}} \right) = 0$$

Ou seja



$$\delta J = 360x^2 + \frac{d^2}{dx^2} - 2y'' = 360x^2 - 2y'' = 0$$

Donde $y'' = 180x^2$, que integrando se tem:

$$y = \frac{1}{2}x^6 + \alpha x^3 + \beta x^2 + \chi x + \gamma$$

Levando em consideração as condições de contorno, achamos

$$\alpha = \frac{3}{2}$$

$$\beta = -3$$

$$\chi = 1$$

$$\gamma = 0$$

logo a função extrema procurada é: $y = \frac{1}{2}x^6 + \frac{3}{2}x^3 - 3x^2 + x$.

Sejam, agora, funcionais tendo diversas funções como variáveis dependentes, o que acontece com muitos problemas de engenharia. Considere o funcional

$$J(f, g) = \int_a^b I(f, g, f_x, g_x, x) dx \quad (11.2.54)$$

com as seguintes condições de contorno:

$$f_a = f \quad a \quad (11.2.55)$$

$$f_b = f \quad b$$

$$g_a = g \quad a$$

$$g_b = g \quad b$$

Sejam $\eta(x)$ e $\zeta(x)$ duas funções admissíveis que satisfaçam as condições de contorno acima e são tais que definem as seguintes funções:

$$h(x) = f(x) + \varepsilon \eta(x) = f + \delta f \quad (11.2.56)$$

$$z(x) = g(x) + \varepsilon \zeta(x) = g + \delta g$$

Com $\eta(a) = \eta(b) = \zeta(a) = \zeta(b) = 0$ e onde $f(x)$ e $g(x)$ são funções extremais de $J(f, g)$ e δf e δg são suas variações. Dessa forma:

$$\delta J = \varepsilon \left(\frac{dJ(f, g)}{d\varepsilon} \right)_{\varepsilon=0} = 0 \quad (11.2.57)$$

isto é:



$$\delta J = \varepsilon \left\{ \int_a^b \left[\frac{\partial I}{\partial h} \frac{\partial h}{\partial \varepsilon} + \frac{\partial I}{\partial h_x} \frac{\partial h_x}{\partial \varepsilon} + \frac{\partial I}{\partial z} \frac{\partial z}{\partial \varepsilon} + \frac{\partial I}{\partial z_x} \frac{\partial z_x}{\partial \varepsilon} \right] dx \right\}_{\varepsilon=0} \quad (11.2.58)$$

Para $\varepsilon = 0$ lembrando que $\varepsilon\eta = \delta f$ e $\varepsilon\zeta = \delta g$, tem-se:

$$\begin{aligned} \delta J &= \varepsilon \left\{ \int_a^b \left[\frac{\partial I}{\partial h} \eta + \frac{\partial I}{\partial h_x} \eta_x + \frac{\partial I}{\partial z} \zeta + \frac{\partial I}{\partial z_x} \zeta_x \right] dx \right\} \quad (11.2.59) \\ \delta J &= \left\{ \int_a^b \left[\frac{\partial I}{\partial h} \varepsilon\eta + \frac{\partial I}{\partial h_x} \varepsilon\eta_x + \frac{\partial I}{\partial z} \varepsilon\zeta + \frac{\partial I}{\partial z_x} \varepsilon\zeta_x \right] dx \right\} \\ \delta J &= \left\{ \int_a^b \left[\frac{\partial I}{\partial h} \delta f + \frac{\partial I}{\partial h_x} \delta f_x + \frac{\partial I}{\partial z} \delta g + \frac{\partial I}{\partial z_x} \delta g_x \right] dx \right\} \end{aligned}$$

Integrando os termos em δf_x , δg_x da expressão acima, obtém-se:

$$\begin{aligned} \delta J &= \int_a^b \left\{ \left[\frac{\partial I}{\partial f} - \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) \right] \delta f + \left[\frac{\partial I}{\partial g} - \frac{d}{dx} \left(\frac{\partial I}{\partial g_x} \right) \right] \delta g \right\} dx + \\ &+ \left. \frac{\partial I}{\partial f_x} \delta f \right|_a^b + \left. \frac{\partial I}{\partial g_x} \delta g \right|_a^b \quad (11.2.60) \end{aligned}$$

Os últimos dois termos da expressão acima são zero devido às condições de contorno e os termos multiplicados por δf e δg são também nulos, uma vez que δf e δg são arbitrários. Desta forma as equações de Euler são:

$$\begin{aligned} \frac{\partial I}{\partial f} - \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) &= 0 \\ \frac{\partial I}{\partial g} - \frac{d}{dx} \left(\frac{\partial I}{\partial g_x} \right) &= 0 \end{aligned} \quad (11.2.61)$$

Generalizando para funcionais da forma:

$$J = \int_a^b I(f_1, f_2, \dots, f_n, f_{1x}, f_{2x}, \dots, f_{nx}, x) dx \quad (11.2.62)$$

com as funções f_i distintas e com as seguintes condições de contorno:

$$\begin{aligned} f_{1x} &= f_1 \quad a ; f_{1x} = f_1 \quad b ; \\ &\dots \\ f_{nx} &= f_n \quad a ; f_{nx} = f_n \quad b ; \end{aligned} \quad (11.2.63)$$

então, obtém-se o seguinte sistema de equações de Euler:



$$\frac{\partial I}{\partial f^k} - \frac{d}{dx} \left(\frac{\partial I}{\partial f^k_x} \right) = 0 \quad \therefore k = 1, 2, \dots, n$$

Exemplos

1. Achar os extremos do funcional $J(y, z) = \int_0^2 (y'^2 + z^2 + z'^2) dx$ com as seguintes condições de contorno: $y(0) = 1; y(2) = 2; z(0) = 0; z(2) = 1$.

Solução:

Aplicando a teoria acima, encontra-se o seguinte sistema de equações diferenciais ordinárias:

$$\begin{aligned} y'' &= 0 \\ z - z'' &= 0 \end{aligned}$$

cuja solução é do tipo:

$$\begin{aligned} y &= \alpha x + \beta \\ z &= \chi e^x + \kappa e^{-x} \end{aligned}$$

Aplicando as condições de contorno, se tem: $\alpha = 1; \beta = 0; \chi = \frac{1}{e^2 - 1}; \kappa = -\frac{e^2}{e^2 - 1}$

de modo que o extremos pedido é:

$$\begin{aligned} y &= x \\ z &= \frac{\sinh x - 1}{\sinh 1} \end{aligned}$$

2. Achar os extremos do funcional $J(y, z) = \int_0^\pi (2yx - 2y^2 + y'^2 - z'^2) dx$

com as seguintes condições de contorno:

$$y(0) = 0; y(\pi) = 1; z(0) = 0; z(\pi) = -1.$$

Solução:

De forma semelhante ao que se fez no exemplo anterior, as equações de *Euler* são:

$$\begin{aligned} y'' + 2y - z &= 0 \\ z'' + y &= 0 \end{aligned}$$

donde, eliminando a função z , se obtém:



$$y^{iv} + 2y'' + y = 0$$

que é uma equação diferencial ordinária de quarta ordem, cuja solução genérica tem a forma:

$$y = A \cos x + B \sin x + x C \cos x + D \sin x$$

Aplicando as condições de contorno $y(0) = 0; y(\pi) = 1 \rightarrow A = 0 \wedge C = -\frac{1}{\pi}$, de modo que:

$$y = B \sin x + Dx \sin x - \frac{x}{\pi} \cos x$$

A função z se determina da condição $z = y'' + 2y$. Assim,

$$z = B \sin x + D(2 \cos x + x \sin x) + \left(\frac{1}{\pi}\right) (2 \sin x - x \cos x)$$

Aplicando agora as condições de contorno $z(0) = 0; z(\pi) = 1$, obtém-se: $D = 0$ e $B =$ um número arbitrário qualquer. Assim:

$$z = B \sin x + \left(\frac{1}{\pi}\right) (2 \sin x - x \cos x)$$

Donde se conclui que os extremos do funcional dado são a família de curvas:

$$y = B \sin x - \left(\frac{x}{\pi}\right) \cos x$$

$$z = B \sin x + \left(\frac{1}{\pi}\right) (2 \sin x - x \cos x)$$

com B arbitrário.

2.1.8.3 Exercícios:

Achar os extremos dos funcionais:



$$1. J y = \int_0^1 y^2 + 2y'^2 + y''^2 dx$$

$$\text{com } y(0) = y(1) = 0; y'(0) = 1; y'(1) = -\sinh 1.$$

$$2. J x = \left(\frac{1}{2}\right) \int_0^1 y''^2 dx \therefore y(0) = y'(0) = 0; y'(1) = 1.$$

$$3. J y(x), z(x) = \int_0^{\pi/2} y'^2 + z'^2 - 2yz dx \therefore y(0) = z(0) = 0; y\left(\frac{\pi}{2}\right) = z\left(\frac{\pi}{2}\right) = 1$$

$$4. J y(x), z(x) = \int_0^1 y'^2 + z'^2 - 2yz dx \therefore y(0) = z(0) = 0; y(1) = \frac{3}{2}; z(1) = 1$$

2.1.9 Princípio da Ação Mínima

Como si sabe da física, o sucesso do princípio universal da energia potencial mínima usada para determinar a posição de equilíbrio de um sistema estimula a busca de um princípio análogo universal com a ajuda do qual pode ser possível determinar os movimentos possíveis de um sistema. Isto levou à descoberta do princípio da ação mínima.

Seja primeiro um caso especial: supondo-se uma partícula de massa m em um movimento ao longo do eixo x sob a ação de uma força com potencial $U(x)$. Como se sabe da física a equação do movimento da partícula é

$$m \frac{d^2x}{dt^2} = -U'(x) \rightarrow m \frac{d^2x}{dt^2} + U'(x) = 0 \quad (11.2.65)$$

É fácil escolher um funcional para o qual a última equação é justamente uma equação de Euler.

Denotando $\frac{dx}{dt} = \dot{x}$ pode-se reescrever

$$\frac{dU}{dx} + \frac{d}{dt} m\dot{x} = 0 \rightarrow \frac{d}{dx} \left[-U - \frac{d}{dt} \left(\frac{m\dot{x}^2}{2} \right) \right] = 0 \quad (11.2.66)$$

Esta última forma é uma remontagem da equação de Euler, a qual no caso da função desejada $x(t)$ precisa ter a forma:

$$\frac{\partial}{\partial x} F(t, x, \dot{x}) - \frac{d}{dt} \left[\frac{\partial}{\partial \dot{x}} F(t, x, \dot{x}) \right] = 0 \quad (11.2.67)$$



Como tanto na expressão acima, (11.2.67), como na anterior (11.2.66) a função F é derivável em ambos os membros e considerando o sinal da equação acima, vê-se que as derivadas são nulas, logo sendo

$$F_{t,x,\dot{x}} = \left[\frac{m\dot{x}^2}{2} - U(x) \right]$$

aplicando (11.2.67), se tem

$$\frac{\partial}{\partial x} \left[\frac{m\dot{x}^2}{2} - U(x) \right] - \frac{d}{dt} \left[\frac{\partial}{\partial \dot{x}} \left[\frac{m\dot{x}^2}{2} - U(x) \right] \right] = 0 \quad (11.2.68)$$

Então o funcional desejado é

$$J = \int_{t_1}^{t_2} \left[\frac{m\dot{x}^2}{2} - U(x) \right] dt \quad (11.2.69)$$

Observe que o termo $\frac{m\dot{x}^2}{2}$ é justamente a energia cinética E do movimento da partícula. Assim, denotando o integrando como $L = E - U$, si tem a chamada função Lagrangeana. Desta forma o problema variacional consiste em buscar o valor estacionário da integral:

$$J = \int_{t_1}^{t_2} L dt$$

Que é conhecido como o nome de “ação”; aqui t_1 e t_2 são o tempo inicial e final do movimento. Pode-se verificar que num número grande de casos o intervalo entre t_1 e t_2 é muito pequeno e nesta situação a integral acima tem um valor mínimo e não meramente um valor estacionário para o movimento real. Por esta razão, a possibilidade de achar que o movimento pelo procedimento do problema variacional para a integral acima é chamado de princípio da ação mínima.

É importante se verificar que o princípio variacional da ação mínima é de natureza universal e permanece válido para qualquer sistema fechado não envolvendo dissipação de energia, por exemplo, via atrito; incidentalmente, um sistema com dissipação pode, sob certo sentido, ser considerado aberto. De acordo com este princípio, de todos os modos concebidos (sob dadas restrições) de passar de um estado num tempo t_1 para outro estado no tempo t_2 , o sistema escolhe o modo para o qual a ação assume um valor estacionário (mínimo, como regra).

Aqui, a função Lagrangeana L é a diferença entre a energia cinética e a energia potencial do sistema, cada qual dessas energias expressa em coordenadas generalizadas do sistema e de suas derivadas temporais. Assim, o princípio das ações mínimas é aplicável tanto a sistemas com um



número finito de graus de liberdade como também para meios contínuos, e não apenas mecânicos, mas também eletromagnéticos e outros fenômenos.

Exemplos

Seja a aplicação do princípio da ação mínima para determinar a equação das oscilações transversais de uma membrana que satisfaz a equação de Laplace, com a condição de contorno $z|_{\Omega} = \varphi$ (dado). A energia cinética de um elemento $d\sigma$ da membrana

é igual a $\frac{1}{2}\rho d\sigma \left(\frac{dz}{dt}\right)^2$; onde ρ é a densidade da superfície da membrana, logo a energia cinética total da membrana é

$$E = \frac{1}{2} \rho \iint_{\sigma} \left(\frac{dz}{dt}\right)^2 dx dy$$

Considerando que a tensão na membrana permanece inalterada no processo de vibração (tensão T), então a energia potencial acumulada é

$$P = \frac{T}{2} \iint_{\sigma} \left[\left(\frac{dz}{dx}\right)^2 + \left(\frac{dz}{dy}\right)^2 \right] dx dy$$

Então, a função Lagrangeana e a ação são:

$$L = \frac{1}{2} \iint_{\sigma} \left\{ \rho \left(\frac{dz}{dt}\right)^2 - T \left[\left(\frac{dz}{dx}\right)^2 + \left(\frac{dz}{dy}\right)^2 \right] \right\} dx dy$$

$$S = \frac{1}{2} \int_{t_1}^{t_2} \iint_{\sigma} \left\{ \rho \left(\frac{dz}{dt}\right)^2 - T \left[\left(\frac{dz}{dx}\right)^2 + \left(\frac{dz}{dy}\right)^2 \right] \right\} dx dy dt$$

Aplicando a equação de Euler com relação às variáveis independentes (vide equações (11.2.65) e (11.2.72)), encontra-se a equação da oscilação da membrana:

$$-\frac{\partial}{\partial t} \left(\rho \frac{\partial z}{\partial t} \right) + \frac{\partial}{\partial x} \left(T \frac{\partial z}{\partial x} \right) + \frac{\partial}{\partial y} \left(T \frac{\partial z}{\partial y} \right) = 0$$

$$\frac{\partial^2 z}{\partial t^2} = \frac{T}{\rho} \left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} \right) = a^2 \left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} \right) \therefore a = \sqrt{\frac{T}{\rho}}$$

Seja um sistema dinâmico como o pêndulo que se caracteriza por um corpo de massa m sujeito a um elemento de atuação representado por uma vareta de comprimento l em que o torque aplicado no



atuador é u . Como se sabe da física, o movimento do pêndulo é executado de forma circular onde uma variável θ mede a posição angular em relação a vertical e sua derivada $\dot{\theta}$ representa a velocidade angular; tem-se ainda agindo sobre o sistema, uma força de gravidade que age diretamente sobre a massa e um arrasto que é proporcional à velocidade angular.

Da teoria do cálculo variacional e da física, vale lembrar que, no caso de um sistema dinâmico com n graus de liberdade e com coordenadas generalizadas $q \in \mathbb{R}^n$ e com forças externas generalizadas $Q \in \mathbb{R}^n$, este é descrito pela seguinte equação de Euler- Lagrange:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = Q$$

$\therefore L = T - V$

Dadas estas informações, pode-se modelar o pêndulo, definindo que a energia cinética dele é, considerando que a velocidade linear em relação a sua velocidade angular é $s = r \dot{\theta} = l \dot{\theta}$ (r o raio da circunferência de seu movimento):

$$T = \frac{1}{2} m s^2 = \frac{1}{2} m l^2 \dot{\theta}^2$$

Por outro lado, a energia potencial do sistema pêndulo é dada pela diferença de altura do pêndulo em relação a sua posição vertical, dado por:

$$V = mgl (1 - \cos \theta)$$

O arrasto do pêndulo devido seu movimento no ar é proporcional à função de perda de Rayleigh vê à velocidade angular, e é dada por:

$$F_{\dot{\theta}} = \frac{1}{2} k l^2 \dot{\theta}^2$$

Assim, dado que o funcional lagrangeano é dado por $L = T - V$, então a equação de *Euler-Lagrange* é dada por:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} + \frac{\partial F}{\partial \dot{\theta}} = u$$

$$\therefore \frac{\partial L}{\partial \dot{\theta}} = ml^2 \dot{\theta}; \quad \frac{\partial L}{\partial \theta} = -mgl \sin \theta; \quad \frac{\partial F}{\partial \dot{\theta}} = kl^2 \dot{\theta};$$

$$\therefore \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) = ml^2 \ddot{\theta}$$



Donde se tem a equação do modelo não linear do pendulo atuado

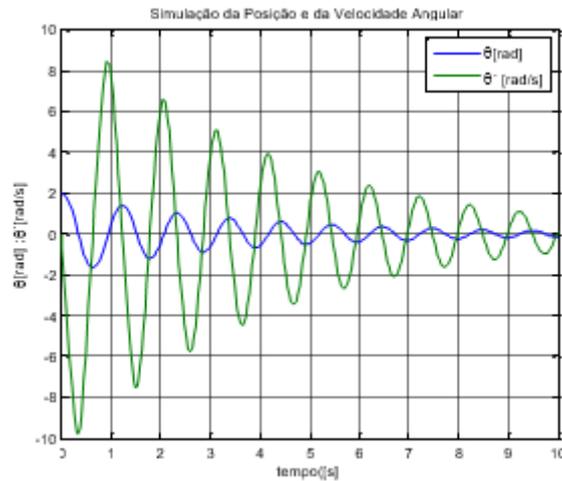
$$ml^2\ddot{\theta} + mgl \sin \theta + kl^2\dot{\theta} = u$$

Abaixo apresenta-se um programa para solucionar a equação do modelo não linear:

```
function pendulo(L,M,U)
% =====
% Problema: pêndulo de comprimento L com massa M
% entrada:   L = comprimento da haste do pendulo [m]
%           M = massa do pendulo em [kg]
%           U = torque aplicado no atuador [Nm]
% =====
if nargin == 0
    L=0.25; M=0.2; U=0.0;
end
clc; global k l m u g;
%
k = 0.1; % coeficiente de Rayleigh
if L==0
    l = 0.25; % comprimento da haste do pendulo em [m]
else
    l=L;
end
if M==0
    m = 0.20; % massa do pendulo em [kg]
else
    m = M;
end
if U == 0
    u = 0 ;    % torque aplicado no atuador em [Nm]
else
    u = U;
end
g = 9.8; % aceleração da gravidade [m/s2]
x0 = [2 0]; % condição inicial = \theta=2 radiano;\theta'=0
tsim = [0 10]; % tempo de simulação entre 0 e 10 segundos
%
%options = odeset('RelTol',1e-4,'AbsTol',[1e-4 1e-4 1e-5]);
[t,y] = ode45(@naolinear,tsim,x0);
figure
plot(t,y,'-'),grid on
title('Simulação da Posição \theta e da Velocidade Angular \theta\'')
xlabel('tempo ([s]')
ylabel('\theta[rad] ;\theta\'[rad/s]')

legend('\theta[rad]','\theta\' [rad/s]');
end
%
%%
function xdot = naolinear(t,x)
global k l m u g;
xdot = [x(2); -(g/l)*sin(x(1))-(k/m)*x(2)+u/(m*l*l) ];
end
```

Executando este programa se tem a seguinte saída gráfica para a posição θ e a velocidade angular $\dot{\theta}$:



2.1.9.1 Funcionais que Dependem de Funções de Várias Variáveis Independentes

Seja o funcional

$$J = \iint_{\Omega} I(f_x, f_y, x, y) dx dy \quad (11.2.70)$$

para o qual queremos analisar o extremo, e cujas condições de contorno na fronteira $\partial\Omega$ da região Ω são conhecidas.

Seja $\eta(x, y)$ uma função admissível que satisfaça as condições de contorno e tenha derivadas contínuas até o grau desejado e que

$$\begin{aligned} h(x, y) &= f(x, y) + \varepsilon \eta(x, y) = f + \delta f \\ h_x(x, y) &= f_x(x, y) + \varepsilon \eta_x(x, y) = f_x + \delta f_x \\ h_y(x, y) &= f_y(x, y) + \varepsilon \eta_y(x, y) = f_y + \delta f_y \end{aligned}$$

Assim

$$\delta J = \varepsilon \left(\frac{dI}{d\varepsilon} \right)_{\varepsilon=0} = 0 \quad (11.2.71)$$

$$\delta J = \varepsilon \left(\iint_{\Omega} \left\{ \frac{\partial I}{\partial h} \frac{\partial h}{\partial \varepsilon} + \frac{\partial I}{\partial h_x} \frac{\partial h_x}{\partial \varepsilon} + \frac{\partial I}{\partial h_y} \frac{\partial h_y}{\partial \varepsilon} \right\} dx dy \right)_{\varepsilon=0} = 0 \quad (11.2.72)$$

$$\delta J = \varepsilon \left(\iint_{\Omega} \left\{ \frac{\partial I}{\partial h} \delta f + \frac{\partial I}{\partial h_x} \delta f_x + \frac{\partial I}{\partial h_y} \delta f_y \right\} dx dy \right) = 0 \quad (11.2.73)$$

Integrando o segundo e terceiro termos da expressão acima pelo Teorema de Green, obtém-se:



$$\begin{aligned} \iint \frac{\partial I}{\partial f_x} \delta f_x dx dy &= - \iint \frac{\partial}{\partial x} \left(\frac{\partial I}{\partial f_x} \right) \delta f dx dy + \oint \frac{\partial I}{\partial f_x} \delta f \frac{dy}{ds} ds \\ \iint \frac{\partial I}{\partial f_y} \delta f_y dx dy &= - \iint \frac{\partial}{\partial y} \left(\frac{\partial I}{\partial f_y} \right) \delta f dx dy + \oint \frac{\partial I}{\partial f_y} \delta f \frac{dx}{ds} ds \end{aligned} \quad (11.2.74)$$

Levando-se estes resultados em δJ , tem-se:

$$\begin{aligned} \delta J &= \iint_{\Omega} \left\{ \frac{\partial I}{\partial f} - \frac{\partial}{\partial x} \left(\frac{\partial I}{\partial f_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial I}{\partial f_y} \right) \right\} \delta f . dx . dy + \\ &+ \oint_{\partial \Omega} \left\{ \frac{\partial I}{\partial f_x} \frac{dy}{ds} - \frac{\partial I}{\partial f_y} \frac{dx}{ds} \right\} \delta f . ds = 0 \end{aligned} \quad (11.2.75)$$

ou ainda, como $\delta f = \varepsilon \eta$ satisfaz as condições de contorno em $\partial \Omega$, a segunda integral da expressão acima é nula, logo a condição de estacionariedade é:

$$\delta J = \iint_{\Omega} \left\{ \frac{\partial I}{\partial f} - \frac{\partial}{\partial x} \left(\frac{\partial I}{\partial f_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial I}{\partial f_y} \right) \right\} \delta f . dx . dy = 0 \quad (11.2.76)$$

Como a variação de δf é arbitrária (a δf se impõe apenas restrições de carácter geral sobre a continuidade e derivabilidade, anulação no contorno, etc.) e o primeiro fator da integral acima é contínuo, então do Lema Básico a função $f_{x,y}$ que realiza extremo no funcional dado é:

$$\frac{\partial I}{\partial f} - \frac{\partial}{\partial x} \left(\frac{\partial I}{\partial f_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial I}{\partial f_y} \right) = 0 \quad (11.2.77)$$

Esta equação diferencial de segunda ordem em derivadas parciais (11.2.77) que é a equação de Euler do funcional (11.2.70) recebe o nome particular de Equação de *Euler-Ostrogradski*.

Exemplos deste tipo de funcional são:

$J = \iint f_x^2 + f_y^2 dx dy$ cuja equação correspondente de *Euler-Ostrogradski* é da forma $\nabla^2 f = 0$, que é a equação de Laplace. Além disso, determinar uma solução contínua desta equação em Ω , com valores de contorno em $\partial \Omega$ conhecidos, é um problema chamado de Problema de *Dirichlet*, que é fundamental para a física- matemática e para a engenharia; é um problema típico de condução de calor estacionária.

$J = \iint f_x^2 + f_y^2 + 2fp_{x,y} dx dy$ cuja equação correspondente de *Euler-Ostrogradski* é da forma $\nabla^2 f = p_{x,y}$, que é a equação de Poisson.



$J = \iint \sqrt{1 + f_x^2 + f_y^2} dx dy$ que é o funcional cujo mínimo resolve o problema da determinação de uma superfície de área mínima limitada por um contorno dado, tem sua equação correspondente de Euler-Ostrogradski da forma

$$\frac{\partial^2 f}{\partial x^2} \left[1 + \left(\frac{\partial f}{\partial y} \right)^2 \right] - 2 \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2} \left[1 + \left(\frac{\partial f}{\partial x} \right)^2 \right] = 0.$$

Generalizando para um funcional do tipo

$$J = \iiint \dots \int I f_{x_1} f_{x_2} \dots f_{x_n} x_1 x_2 \dots x_n dx_1 dx_2 \dots dx_n \quad (11.2.78)$$

cuja equação de Euler é:

$$\frac{\partial I}{\partial f} - \sum_{i=1}^n \left\{ \frac{\partial^2 I}{\partial x_i \partial f_{x_i}} + \frac{\partial^2 I}{\partial f \partial f_{x_i}} + \sum_{j=1}^n \left[\frac{\partial I}{\partial f_{x_i} \partial f_{x_j}} \cdot \frac{\partial f_{x_i}}{\partial x_j} \right] \right\} = 0 \quad (11.2.79)$$

Outro tipo de funcional que aparece nos problemas de engenharia é do tipo:

$$J = \iint I f_{x_x} f_{y_y} f_{x_x x_x} f_{y_y y_y} f_{x_y} x y dx dy \quad (11.2.80)$$

cuja equação de Euler é do tipo:

$$\begin{aligned} \frac{\partial I}{\partial f} - \frac{\partial}{\partial x} \left(\frac{\partial I}{\partial f_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial I}{\partial f_y} \right) + \frac{\partial^2}{\partial x^2} \left(\frac{\partial I}{\partial f_{xx}} \right) + \\ + \frac{\partial^2}{\partial y^2} \left(\frac{\partial I}{\partial f_{yy}} \right) - \frac{\partial^2}{\partial x \partial y} \left(\frac{\partial I}{\partial f_{xy}} \right) = 0 \end{aligned} \quad (11.2.81)$$

Um exemplo para este tipo de funcional é $J = \iint f_{xx}^2 + f_{yy}^2 + 2f_{xy}^2 - 2fc dx dy$

cuja equação de Euler é $f_{xxxx} + 2f_{xxyy} + f_{yyyy} = c$, ou ainda $\nabla^4 f = c$. Se $f = w(x,y)$ representar a deflexão vertical de uma placa isotrópica sob flexão, e $c = p/d$ (p = carregamento uniformemente distribuído sobre a placa e $d = Eh^3/12(1-\nu^2)$ a rigidez da placa, então a equação $\nabla^4 w = p/d$ representa a equação da flexão de uma placa isotrópica sob carga vertical p uniformemente distribuída sobre ela, Se por outro lado, o funcional anterior dependesse de derivadas superiores à de segunda ordem, ter-se-ia:

$$J = \iint I f_{x_x} f_{y_y} f_{x_x x_x} f_{y_y y_y} f_{x_y} \dots f_{x_n} f_{y_n} f_{x_{m-n} y_m} x dx dy \quad (11.2.82)$$

a equação de Euler é:



$$\sum_{k=0}^{k=n} \sum_{m=0}^k -1^k \frac{\partial^k}{\partial x^{k-m} \partial y^m} \left(\frac{\partial I}{\partial f_{x(k-m)y(m)}} \right) = 0 \quad (11.2.83)$$

Num funcional do tipo:

$$J = \iint I f, g, f_x, g_x, f_y, g_y, x, y \, dx dy \quad (11.2.84)$$

a equação de Euler representado pelo seguinte sistema:

$$\begin{aligned} \frac{\partial I}{\partial f} - \frac{\partial}{\partial x} \left(\frac{\partial I}{\partial f_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial I}{\partial f_y} \right) &= 0 \\ \frac{\partial I}{\partial g} - \frac{\partial}{\partial x} \left(\frac{\partial I}{\partial g_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial I}{\partial g_y} \right) &= 0 \end{aligned} \quad (11.2.85)$$

Um exemplo deste tipo de funcional bidimensional é aquele que representa a energia interna de uma placa no estado plano de tensão:

$$F(u, v) = \left(\frac{Eh}{2(1-\nu^2)} \right) \iint_{\Omega} \left\{ u_x^2 + u_y^2 + 2\nu u_x v_y + \right. \\ \left. + \frac{1}{2}(1-\nu) [u_y^2 + v_x^2 + 2\nu u_y v_x] \right\} dx dy \quad (11.2.86)$$

onde E = módulo de elasticidade; h = espessura da placa; ν = coeficiente de Poisson; u e v os deslocamentos nas direções x e y . As equações de Euler para o funcional acima, representa as equações de equilíbrio para a placa em termos de x, y , e são as seguintes:

$$\begin{aligned} \frac{Eh}{1-\nu^2} \left[u_{xx} + \nu v_{xy} + \left(\frac{1-\nu}{2} \right) (u_{yy} + v_{xy}) \right] &= 0 \\ \frac{Eh}{1-\nu^2} \left[v_{xx} + \nu u_{xy} + \left(\frac{1-\nu}{2} \right) (v_{yy} + u_{xy}) \right] &= 0 \end{aligned}$$

2.1.9.2 Exercícios

1. Mostre que as condições necessárias para o carácter estacionário das integrais:

a - $\int_{x_1}^{x_2} F(x, y, z, \dots, y', \dots, z', \dots) dx$

b - $\int_{\Omega} F(x, y, u, u_x, u_y) dx dy$

c - $\int_{\Omega} F(x, y, u, u_x, u_y, v_x, v_y, \dots) dx dy$



são respectivamente

$$a - F_{y'} = F_{z'} = 0 \therefore x = x_1 \wedge x = x_2$$

$$b - F_{u_x} \left(\frac{dy}{ds} \right) - F_{u_y} \left(\frac{dx}{ds} \right) = 0$$

$$c - \begin{cases} F_{u_x} \left(\frac{dy}{ds} \right) - F_{u_y} \left(\frac{dx}{ds} \right) = 0 \\ F_{v_x} \left(\frac{dy}{ds} \right) - F_{v_y} \left(\frac{dx}{ds} \right) = 0 \end{cases}$$

no contorno de Ω onde s é o comprimento deste contorno. Achar os extremos dos funcionais

$$a - J(y, x) = \int_0^1 y^2 + 2y'^2 + y''^2 dx \quad \text{com} \quad \begin{cases} y(0) = y(1) = 0 \\ y'(0) = 1; y'(1) = -\sinh 1 \end{cases}$$

$$b - J(y, x) = \int_0^1 2y'^2 + y''^2 dx \quad \text{com} \quad \begin{cases} y(0) = 0; \quad y(1) = 1 \\ y'(0) = 1; \quad y'(1) = \cosh 1 \end{cases}$$

□

2.1.10 Funcionais Quadráticos

Definição 1: Pode-se definir também como sendo um funcional quadrático, ao funcional $J : V \rightarrow \mathbb{R}$, definido sobre uma classe V de funções $f(x)$, se ele satisfaz a identidade (veja definição 11 acima):

$$J(u + \tau\eta) = J(u) + \tau\delta J(u; \eta) + \frac{1}{2}\tau^2\delta^2 J(u; \eta) \quad (11.2.87)$$

$$\text{Com } \forall u \in V, \quad \forall \eta \in \vec{V}, \quad \|\eta\| = 1, \quad \forall \tau \in \mathbb{R}.$$

Definição 2: Um funcional quadrático é um funcional na forma

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T H \mathbf{x} - \mathbf{b}^T \mathbf{x} + c, \quad (11.2.88)$$

$$\therefore \mathbf{x} \in \mathbb{R}^n, H \in \mathbb{R}^{n \times n} (\text{com } H^T = H), b \in \mathbb{R}^n, c \in \mathbb{R}$$

Como já visto anteriormente, o gradiente e o Hessiano do funcional acima são facilmente encontrados:

$$\begin{cases} \mathbf{g}(\mathbf{x}) = H\mathbf{x} - \mathbf{b} \\ H(\mathbf{x}) = H \end{cases} \quad (11.2.89)$$



Assim, se vê que, os Funcionais Quadráticos têm o Hessiano constante.

Tomando ξ como um ponto estacionário de f , então o gradiente se anula nele e se tem

$$\mathbf{g}(\xi) = \nabla f(\xi) = H\xi - \mathbf{b} = 0 \Rightarrow \xi = H^{-1}\mathbf{b}$$

Fazendo $\mathbf{x} = \mathbf{x} - \xi$ no funcional quadrático se tem

$$\begin{aligned} f(\mathbf{x} - \xi) &= \frac{1}{2} (\mathbf{x} - \xi)^T H (\mathbf{x} - \xi) - \mathbf{b}^T (\mathbf{x} - \xi) + c, \quad \therefore \mathbf{x} \in \mathbb{R}^n, \\ &= \frac{1}{2} (\mathbf{x} - \xi)^T H (\mathbf{x} - \xi) + \chi, \quad \therefore \chi = -\mathbf{b}^T (\mathbf{x} - \xi) + c \end{aligned} \quad (11.2.90)$$

] Sendo H simétrico e λ_i, \mathbf{v}_i $_{i=1}^n$ o conjunto de auto soluções de H então

$$\lambda_i, \mathbf{v}_i \quad _{i=1}^n \rightarrow \begin{cases} H\mathbf{v}_i = \lambda_i \mathbf{v}_i, i = 1, 2, \dots, n \\ \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n \\ \mathbf{v}_i \mathbf{v}_j = \delta_{ij}, i, j = 1, 2, \dots, n \end{cases} \quad (11.2.91)$$

Definindo-se a matriz diagonal $\Lambda_{n \times n}$ cujas diagonais são os autovalores de H e a matriz $V_{n \times n}$ n cujas colunas são os autovetores de H tem-se:

$$\Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_n]; \quad V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$$

Como definido $\Lambda_{n \times n}$ e $V_{n \times n}$ vê-se que a matriz V é ortogonal, isto é $V^{-1} = V^T$ e $HV = V\Lambda$.

Assim, chega-se a seguinte expressão:

$$f(\mathbf{x} - \xi) = \frac{1}{2} (\mathbf{x} - \xi)^T H (\mathbf{x} - \xi) + \chi, \quad \therefore \chi = -\mathbf{b}^T (\mathbf{x} - \xi) + c$$

Fazendo $\mathbf{z} = V^T (\mathbf{x} - \xi)$ na expressão anterior se tem

$$\begin{aligned} \hat{f}(\mathbf{z}) &\equiv f(V\mathbf{z} - \xi) = \frac{1}{2} \mathbf{z}^T V^T H V \mathbf{z} + \chi, \quad \therefore \chi = -\mathbf{b}^T (\mathbf{x} - \xi) + c \\ &= \frac{1}{2} \mathbf{z}^T \Lambda \mathbf{z} + \chi \\ &= \frac{1}{2} \sum_{i=1}^n \lambda_i z_i^2 + \chi \end{aligned} \quad (11.2.92)$$

Olhando para a expressão(11.2.92), verifica-se que

- Se H for SPD (simétrica e positiva definida) então seus autovalores são todos positivos e o intervalo de f é $[\chi, \infty)$ e \mathbf{z} é um *minimizador global* de f .
- Se H for negativa definida o intervalo de f onde se localizam os autovalores é



$-\infty, \chi\}$, e $\mathbf{z} = 0$ é um maximizador forte global.

- Se H tiver autovalores positivos e negativos o intervalo de f é $-\infty, \infty$, em $\mathbf{z} = 0$ f estacionário, pois não é nem máximo nem mínimo.

Observar que quando H for SPD (simétrica e positiva definida) e quando $k > \chi$ a superfície de nível L_k (equação (11.1.47).) é a elipsoide:

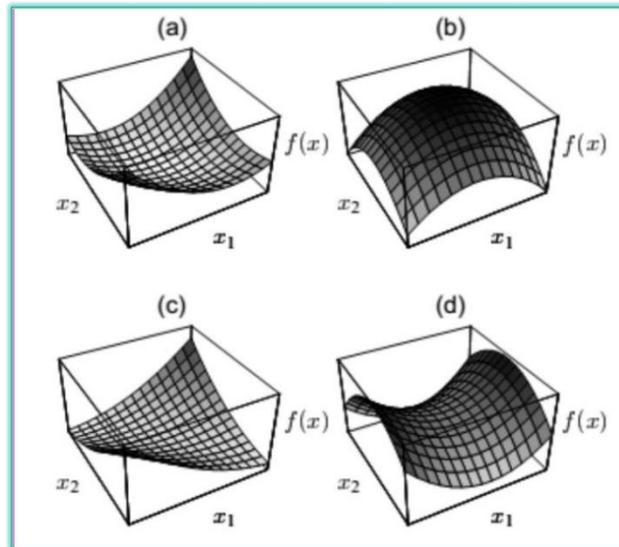
$$\hat{k} = \sum_{i=1}^n \lambda_i z_i^2 \therefore \hat{k} = 2 k - \chi \quad (11.2.93)$$

A medida de distorção D_k pode ser avaliada pela chamada número de condição espectral de H^l , dado por:

$$\kappa H = \frac{\lambda_n}{\lambda_1} \equiv \frac{\lambda_{\max} H}{\lambda_{\min} H} \quad (11.2.94)$$

Olhando-se a figura a seguir que contém gráficos para diversas formas quadráticas, identifica-se o seguinte:

FIGURA 3 GRÁFICOS DE FORMAS QUADRÁTICAS²



A figura (a) representa o gráfico de uma Forma Quadrática para uma matriz positiva-definida;

A figura (b) é o gráfico representativo de uma forma quadrática que tem uma matriz negativa-definida.

¹ Não confundir número condição espectral definido por (11.2.94) com o número condição de uma matriz \mathbf{A} com relação a uma norma matricial, que é dado por $\kappa A = \|A\| \|A^{-1}\|$.

² Fonte: [211] Jonathan Richard Shewchuk, An Introdycion to Conjugate Gradient Method without the Agonising Pain, publicado pela School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, em 1994



A figura (c) mostra o gráfico para uma matriz singular (e positiva-indefinida). A linha que corre na base do vale é o conjunto de soluções.

O gráfico (d) é representativo de uma forma com matriz indefinida. A solução é um ponto de sela.

As propriedades dos Funcionais Quadráticos são relevantes para funcionais não-quadráticos em $C^2 X$. Para se ver isto deve-se comparar a expansão de Taylor para os Funcionais Quadráticos (onde H é constante) e a expansão de Taylor para funcionais Não-Quadráticos:

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \mathbf{g}^T(\mathbf{x}) \mathbf{h} + \frac{1}{2} \mathbf{h}^T H \mathbf{h} + o(\|\mathbf{h}\|^2) \quad (11.2.95)$$

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \mathbf{g}^T(\mathbf{x}) \mathbf{h} + \frac{1}{2} \mathbf{h}^T H(\mathbf{x}) \mathbf{h} + o(\|\mathbf{h}\|^2)$$

Vê-se assim que numa vizinhança de \mathbf{x} o funcional arbitrário se comporta como um funcional quadrático. Como já foi visto, se si fizer $\mathbf{h} = \tau \mathbf{y}$ com $\|\mathbf{y}\| = 1$ e recordando as derivadas direcionais, sabe-se que

$$f(\mathbf{x} + \tau \mathbf{y}) = \sum_{k=0}^{m-1} \frac{\tau^k}{k!} f^{(k)}(\mathbf{x}; \mathbf{y}) + \frac{\tau^m}{m!} f^{(m)}(\mathbf{x}) + o(\tau^m)$$

$$\therefore f^{(m)}(\mathbf{x}; \mathbf{y}) = \left(y_1 \frac{\partial}{\partial x_1} + y_2 \frac{\partial}{\partial x_2} + \dots + y_n \frac{\partial}{\partial x_n} \right)^m f(\mathbf{x})$$

para $k=0,1,2$ se tem a identidade

$$f(\mathbf{x} + \tau \mathbf{y}) = f(\mathbf{x}) + \tau f^{(1)}(\mathbf{x}; \mathbf{y}) + \frac{1}{2} \tau^2 f^{(2)}(\mathbf{x}; \mathbf{y}) + o(\tau^3)$$

$$\text{com } f^{(1)}(\mathbf{x}; \mathbf{y}) = \mathbf{g}^T(\mathbf{x}) \mathbf{y}, \quad \text{e } f^{(2)}(\mathbf{x}; \mathbf{y}) = \mathbf{y}^T H(\mathbf{x}) \mathbf{y}$$

3 MÉTODO DO PASSO DESCENDENTE

Um dos métodos mais importantes para avaliação assintótica de certos tipos de problemas com funcionais, inclusive de problemas integrais, é conhecido como o Método do Passo Descendente. Este método tem sua origem na observação realizada por Laplace em conexão com a estimação de uma integral vinda da teoria da probabilidade, na forma de:

$$I_n = \int_a^b f(x) [g(x)]^n dx = \int_a^b f(x) e^{n \psi(x)} dx, \quad n \rightarrow +\infty \quad (11.3.1)$$

Onde $f(x)$ e $g(x)$ são funções reais contínuas e definidas num intervalo $[a, b]$, O qual pode também ser infinito, com $g(x) > 0$ e $\psi(x) = \log g(x)$ Laplace arguiu que a contribuição dominante desta



integral quando $n \rightarrow \infty$ precisaria vir da vizinhança do ponto onde $g(x)$ (ou $\psi(x)$) teria seu máximo valor. No caso mais simples é a situação onde $\psi(x)$ possuir um máximo do ponto $x = \xi \in [a, b]$ de modo que $\psi'(\xi) = 0$, $\psi''(\xi) < 0$, e $f(\xi) \neq 0$.

Maiores detalhes pode ser visto na obra de Paris, R. B. (Richard Bruce), intitulada “Hadamard Expansions and Hyperasymptotic Evaluation : An Extension of the Method, of Steepest Descents”, veja referencia [241]

Aqui, como o foco é um tratamento matricial de equações provenientes de funcionais, apresentar-se-á um método iterativo para si achar um minimizador forte local de um funcional quadrático e com Hessiano SPD, do tipo mostrado em (11.2.88).

Mostrar-se-á que o número de iterações requeridas para garantir que o erro, sob a norma energética (norma da energia), seja menor que um número ε vezes o erro inicial é limitado por:

$$\frac{1}{2} \kappa H \ln 1/\varepsilon \quad (11.3.2)$$

Onde κH é o número condição espectral de H .

O conceito de condicionamento foi introduzido no capítulo 7 do volume 1 desta obra, e ele será aqui utilizado para reduzir o limite acima.

Em geral os métodos para determinar um minimizador de um funcional f no \mathbb{R}^n tem a forma

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \mathbf{d}_k \quad (11.3.3)$$

Onde \mathbf{d}_k é a direção de busca e τ_k é escolhido de maneira a minimizar, ou no mínimo reduzir, f sobre algum intervalo da linha que passa através de \mathbf{x}_k na direção \mathbf{d}_k .

Dessa forma existem dois problemas associados:

1. a escolha de \mathbf{d}_k , e
2. a inspeção de f sobre a linha $\mathbf{x} = \mathbf{x}_k + \tau \mathbf{d}_k$, com $\tau \in (-\infty, \infty)$.

Definição 1: (direção descendente). Seja um funcional f e sejam os vetores \mathbf{x} e \mathbf{d} , então existe um número $\tau_0 > 0$ tal que

$$f(\mathbf{x} + \tau \mathbf{d}) < f(\mathbf{x}), \quad 0 < \tau \leq \tau_0 \quad (11.3.4)$$

onde \mathbf{d} é uma direção descendente para f em \mathbf{x} .

Teorema 1: Seja $f \in C^1 \mathbb{R}^n$ e seja $\mathbf{g}(\mathbf{x})$ (ou $\nabla f(\mathbf{x})$) o gradiente de f em \mathbf{x} . Se um vetor \mathbf{d} satisfizer a $\mathbf{g}^T(\mathbf{x}) \mathbf{d} < 0$, então \mathbf{d} é uma direção descendente para f em \mathbf{x} .



$$f(\mathbf{x} + \mathbf{d}) = f(\mathbf{x}) + \mathbf{g}^T(\mathbf{x}) \mathbf{d} + \frac{1}{2} \mathbf{d}^T H(\mathbf{x}) \mathbf{d} + o(\|\mathbf{d}\|^2)$$

$\therefore \mathbf{g}^T(\mathbf{x}) \mathbf{d} = 0 \Rightarrow$ ponto estacionário

$$\mathbf{d}^T H(\mathbf{x}) \mathbf{d} < 0 \Rightarrow f(\mathbf{x} + \mathbf{d}) = f(\mathbf{x}) - \frac{1}{2} \mathbf{d}^T H(\mathbf{x}) \mathbf{d}$$

Vê-se assim que um acréscimo na direção \mathbf{d} há um decréscimo no valor da função, portanto \mathbf{d} é uma direção descendente.

Teorema 3: Seja $f \in C^1 \mathbb{R}^n$. Então dentre todas as direções de pesquisa \mathbf{d} em um mesmo ponto \mathbf{x} , a direção para a qual f descende mais rapidamente é numa vizinhança de \mathbf{x} na qual $\mathbf{d} = -\mathbf{g}^T(\mathbf{x})$.

Prova:

Como se quer minimizar a derivada direcional de f em \mathbf{x} em todas as direções se tem

$f'(\mathbf{x}; \mathbf{y}) = \mathbf{g}^T(\mathbf{x}) \mathbf{y}$ $\therefore \|\mathbf{y}\| = 1$, isto é o mesmo que minimizar $\mathbf{g}^T(\mathbf{x}) \mathbf{y}$ para todo \mathbf{y} de maneira que se tenha $\|\mathbf{y}\| = 1$. Dessa forma se tem

$$|\mathbf{g}^T(\mathbf{x}) \mathbf{y}| \leq \|\mathbf{g}^T(\mathbf{x})\| \cdot \|\mathbf{y}\| = \|\mathbf{g}^T(\mathbf{x})\| \Rightarrow \mathbf{y}_{\min} = -\mathbf{g}(\mathbf{x}) / \|\mathbf{g}(\mathbf{x})\| \quad (11.3.5)$$

Analisando o problema de determinar τ_k dados \mathbf{x}_k e \mathbf{d}_k (revise teorema 5 acima) em que $f(\mathbf{x})$ é minimizado sobre a linha $\mathbf{x} = \mathbf{x}_k + \tau_k \mathbf{d}_k$, $\tau_k \in (-\infty, \infty)$.

Qualquer procedimento para se determinar τ_k é denominado de “linha de busca” ou “linha de pesquisa”.

Para funcionais genéricos, as linhas de pesquisa são, em geral, bastante complicadas e envolve algum processo iterativo. No caso de funcionais quadráticos com o Hessiano SPD, o processo é simplificado e pode ser deduzido uma fórmula simples para τ_k . Se

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{2} \mathbf{x}^T H \mathbf{x} - \mathbf{b}^T \mathbf{x} + c \rightarrow \\ f(\mathbf{x} + \tau \mathbf{d}) &= \frac{1}{2} \tau^2 \mathbf{d}^T H \mathbf{d} + \tau \mathbf{d}^T \mathbf{g}(\mathbf{x}) + \chi \therefore \chi \text{ independente de } \tau \end{aligned} \quad (11.3.6)$$

Se H é SPD e $\mathbf{d} \neq 0$ então $\mathbf{d}^T H \mathbf{d} > 0$ e $f(\mathbf{x} + \tau \mathbf{d})$ é uma parábola na variável τ e é aberta para cima de modo que $f(\mathbf{x} + \tau \mathbf{d})$ é unicamente minimizado por (independentemente da escolha de \mathbf{d}_k):

$$\tau = -\mathbf{d}^T \mathbf{g}(\mathbf{x}) / \mathbf{d}^T H \mathbf{d} \Rightarrow \tau_k = -\mathbf{d}_k^T \mathbf{g}(\mathbf{x}_k) / \mathbf{d}_k^T H \mathbf{d}_k \quad (11.3.7)$$

3.1 ALGORITMO DO MÉTODO DO PASSO DESCENDENTE

Baseado no Teorema 3 acima, a nossa escolha natural para a direção descendente \mathbf{d} é $\mathbf{d} = -\mathbf{g}(\mathbf{x}_k)$.

Assim, o processo iterativo dar-se-á por:



$$\mathbf{x}_{k+1} = \mathbf{x}_k - \tau_k \mathbf{g}_k \quad (11.3.8)$$

Esta relação, com alguma estratégia de escolha da “linha de pesquisa”, define o método do passo descendente, para minimizar um funcional genérico. Para um funcional quadrático, viu-se que a maneira de determinar τ_k é:

$$\tau_k = -\mathbf{d}_k^T \mathbf{g}_k / \mathbf{d}_k^T H \mathbf{d}_k \quad (11.3.9)$$

Baseado no que se apresentou aqui, o método do passo descendente para um funcional quadrático segue o seguinte algoritmo:

$$\left. \begin{aligned} \mathbf{g}_k &= H\mathbf{x}_k - \mathbf{b} \\ \tau_k &= \mathbf{g}_k^T \mathbf{g}_k / \mathbf{g}_k^T H \mathbf{g}_k \\ \mathbf{x}_{k+1} &= \mathbf{x}_k - \tau_k \mathbf{g}_k \end{aligned} \right\} \therefore k = 0, 1, \dots; \wedge \mathbf{g}_k = \mathbf{g}_k \quad (11.3.10)$$

Na (k+1)-ésima iteração, a linha que liga \mathbf{x}_k a \mathbf{x}_{k+1} é tangente ao ponto \mathbf{x}_{k+1} pertencente a superfície de nível elipsoidal dada por $\mathbf{x} \in \mathbb{R}^n; f(\mathbf{x}) = f(\mathbf{x}_{k+1})$

Tomando o processo iterativo $\mathbf{x}_{k+1} = \mathbf{x}_k - \tau_k \mathbf{g}_k$ pré-multiplicando a expressão por H e subtraindo \mathbf{b} em ambos os lados se tem:

$$\underbrace{H\mathbf{x}_{k+1} - \mathbf{b}}_{\mathbf{g}_{k+1}} = \underbrace{H\mathbf{x}_k - \mathbf{b}}_{\mathbf{g}_k} - \tau_k H \mathbf{g}_k \Rightarrow \mathbf{g}_{k+1} = \mathbf{g}_k - \tau_k H \mathbf{g}_k \quad (11.3.11)$$

Redefinindo-se o algoritmo anterior se tem:

$$\left. \begin{aligned} \tau_k &= \mathbf{g}_k^T \mathbf{g}_k / \mathbf{g}_k^T H \mathbf{g}_k \\ \mathbf{x}_{k+1} &= \mathbf{x}_k - \tau_k \mathbf{g}_k \\ \mathbf{g}_{k+1} &= \mathbf{g}_k - \tau_k H \mathbf{g}_k \end{aligned} \right\} \therefore k = 0, 1, \dots; \wedge \mathbf{g}_0 = H\mathbf{x}_0 - \mathbf{b} \quad (11.3.12)$$

3.2 ANÁLISE DE CONVERGÊNCIA

3.2.1 Escolha do Tamanho do Passo para a Linha de Pesquisa

Para aproximações de primeira ordem, cada passo diminui o valor de f por aproximadamente

$\tau_k \mathbf{g}_k = \tau_k \|\nabla f(\mathbf{x}_k)\|^2$. Se τ_k for muito pequeno, o algoritmo convergirá muito lentamente. Por outro lado, se o tamanho do passo não for convenientemente pequeno o algoritmo pode falhar na redução de f . Uma maneira adequada é adotar um tamanho de passo de modo a ser suficiente para a redução de f devendo o algoritmo progredir o mais rápido possível. Este procedimento é conhecido como linha de pesquisa e é empregado em muitos outros algoritmos de otimização multivariáveis.



3.2.2 Análise

É necessário medir a convergência quantitativa de todo processo iterativo. Pode-se usar a noção de quanto \mathbf{x} se aproxima de ξ , utilizando-se a norma euclidiana $\|\mathbf{x} - \xi\|_2$ que é o erro definido pela clássica distância entre \mathbf{x} e ξ .

Como trata-se de funcionais e estes representam, em geral quando provenientes de um problema físico, um sistema de energia, é mais apropriado se medir este erro pela quantidade $f(\mathbf{x}) - f(\xi)$. Assim, é mais importante tornar $f(\mathbf{x}) - f(\xi)$ pequeno que fazer $\|\mathbf{x} - \xi\|_2$ pequeno. Precisa-se então, definir uma norma para medir a energia: a norma energética.

Definição 2: O produto interno energético e a norma energética correspondendo a uma matriz H positiva definida são respectivamente

$$\langle \mathbf{x}, \mathbf{y} \rangle_H = \mathbf{x}^T H \mathbf{y} \quad (11.3.13)$$

$$\|\mathbf{x}\|_H = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_H} = \sqrt{\mathbf{x}^T H \mathbf{x}} \quad (11.3.14)$$

Observar que estas definições obedecem respectivamente aos axiomas de produto interno e de norma.

Note que quando H é a matriz identidade, o produto interno energético e a norma energética se reduzem ao produto interno euclidiano e à norma euclidiana respectivamente. Designando o quadrado da norma energética por $E(\mathbf{x})$, se tem

$$E(\mathbf{x}) = \|\mathbf{x}\|_H^2 = \left(\mathbf{x}^T H \mathbf{x} \right) = \mathbf{x}^T H \mathbf{x}, \quad \forall \mathbf{x} \in \mathbb{R}^n \quad (11.3.15)$$

Tomando a primeira expressão de (11.3.6) e transformando-a, encontra-se

$$\left. \begin{aligned} f(\mathbf{x}) &= \frac{1}{2} \mathbf{x}^T H \mathbf{x} - b^T \mathbf{x} + c \\ f(\xi) &= \frac{1}{2} \xi^T H \xi - b^T \xi + c \end{aligned} \right\} \rightarrow f(\mathbf{x}) - f(\xi) = \frac{1}{2} \mathbf{x}^T H \mathbf{x} - b^T \mathbf{x} + c - \left[\frac{1}{2} \xi^T H \xi - b^T \xi + c \right]$$

$$f(\mathbf{x}) - f(\xi) = \frac{1}{2} \mathbf{x}^T H \mathbf{x} - \frac{1}{2} \xi^T H \xi + b^T \xi - b^T \mathbf{x}$$

$$f(\mathbf{x}) - f(\xi) = \frac{1}{2} (\mathbf{x} - \xi)^T H (\mathbf{x} - \xi) + b^T (\xi - \mathbf{x})$$

mas $E(\mathbf{x}) = \mathbf{x}^T H \mathbf{x}$ e $b^T (\xi - \mathbf{x}) \rightarrow 0$ logo

$$f(\mathbf{x}) - f(\xi) = \frac{1}{2} E(\mathbf{x} - \xi) \Rightarrow \left\{ \begin{aligned} E(\mathbf{x} - \xi) &= 2(f(\mathbf{x}) - f(\xi)) \\ \|\mathbf{x} - \xi\|_H &= \sqrt{2(f(\mathbf{x}) - f(\xi))} \end{aligned} \right.$$

Logo

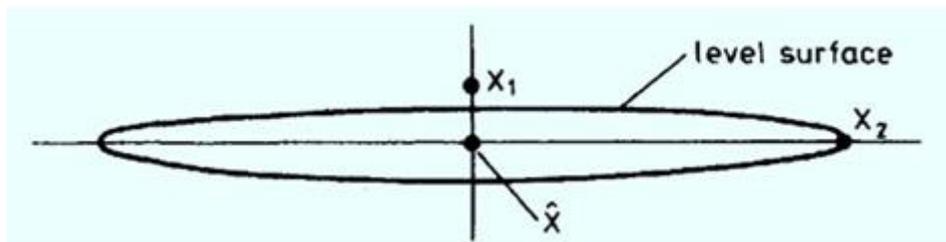


$$E \mathbf{x} - \xi = 2 f \mathbf{x} - f \xi \quad (11.3.16)$$

$$\rightarrow \|\mathbf{x} - \xi\|_H = \sqrt{2 f \mathbf{x} - f \xi} \quad (11.3.17)$$

são as relações entre a norma energética e $E \mathbf{x}$ e a energia ou funcional $f \mathbf{x}$. Observar que $E \mathbf{x} - \xi$ e $\|\mathbf{x} - \xi\|_H$ são constantes em toda superfície de nível de f .

Por que usar normas energéticas em vez de usar a norma euclidiana? Como pode-se verificar, as superfícies de nível são elipsoidais, e quando a excentricidade aumenta (tornando a excentricidade maior) os valores destas normas citadas ficam cada vez mais pronunciadas: $\|\mathbf{x}_2 - \xi\|_H < \|\mathbf{x}_1 - \xi\|_H$ enquanto $\|\mathbf{x}_1 - \xi\| < \|\mathbf{x}_2 - \xi\|$ como pode ser visto na figura a seguir:



Para se evitar isto, pode-se utilizar a técnica de condicionamento, onde se troca H por uma matriz cujas superfícies de nível são significativamente menos excêntricas do que aquelas de H . Assim, se está interessado em saber qual a taxa de convergência do MPD – Método do Passo Descendente ou Método do Gradiente.

Teorema 4: Sendo H uma matriz simétrica e positiva definida, então o método do passo descendente (ou método do gradiente) é convergente para qualquer escolha do datum inicial \mathbf{x}_0 , e mais, tendo

$$\|\mathbf{x}_k - \xi\|_H \leq \left(\frac{\kappa H - 1}{\kappa H + 1} \right)^k \|\mathbf{x}_0 - \xi\|_H \quad (11.3.18)$$

Onde $\kappa H = \lambda_n / \lambda_1 \geq 1$ é o número condição espectral da matriz H e $\|\cdot\|_H$ é a norma de energia definida mais acima. Se além disso, se $p \varepsilon$ é definido para qualquer $\varepsilon > 0$, como sendo o menor inteiro k tal que

$$\|\mathbf{x}_k - \xi\|_H \leq \varepsilon \|\mathbf{x}_0 - \xi\|_H \quad \forall \mathbf{x}_0 \in \mathbb{R}^n \quad (11.3.19)$$

Então o menor inteiro k que atende a (11.3.18) é dado por

$$\text{menor } k = p \varepsilon \leq \frac{1}{2} \kappa H \ln \left(\frac{1}{\varepsilon} \right) + 1. \quad (11.3.20)$$



3.3 PRECONDICIONAMENTO

O que é condicionamento?

Para se conhecer este mecanismo sugerimos estar o tópico 7.7 do livro do prof. Henrique Mariano “Análise e Métodos Numéricos em Engenharia”, que será lembrado aqui: “Para se reduzir o número condicional espectral (relação entre o maior autovalor e menor autovalor) de uma matriz \mathbf{A} , $\kappa \mathbf{A}$, e desta forma melhorar o desempenho dos algoritmos iterativos, normalmente se troca o sistema original por outro sistema que tenha a mesma solução. Sabe-se que quando o número condicional de \mathbf{A} é grande a matriz tende a ser mal condicionada, necessitando-se condicioná-la.”

Por hipótese seja uma matriz \mathbf{A} simétrica e positivo-definida (SPD) e que um pré- condicionador \mathbf{M} seja disponível.

Definição 3: Um pré-condicionado \mathbf{M} é uma matriz que se aproxima de uma matriz \mathbf{A} em algum sentido, por exemplo $\mathbf{M}^{-1}\mathbf{A} = \mathbf{I}$.

3.3.1 Produto Interno e Norma de Energia

Definição 4: O produto interno de energia (ou \mathbf{H} -produto interno) e a norma de energia (ou H -norma), correspondente a uma matriz positiva definida qualquer, são respectivamente dadas por, à semelhança das expressões (11.3.13) e (11.3.14) para a matriz Hessiana e que coincidem quando

$\mathbf{H} = \mathbf{M}$:

$$\mathbf{x}, \mathbf{y}_H = \mathbf{x}^T \mathbf{M} \mathbf{y} \quad (11.3.21)$$

$$\|\mathbf{x}\|_H = \mathbf{x}, \mathbf{x}^{\frac{1}{2}} = \mathbf{x}^T \mathbf{M} \mathbf{x}^{\frac{1}{2}}$$

e satisfazem a todas as propriedades válidas para produto interno euclidiano e norma euclidiana. Note que quando $\mathbf{M} = \mathbf{I}$, o H -produto interno coincide com o produto interno euclidiano e a H -norma com a norma euclidiana (como já explicado anteriormente).

Assumindo que a matriz \mathbf{M} seja simétrica e positivo-definida então, do ponto de vista prático, o único requerimento para \mathbf{M} ser um pré-condicionador é que ele induza uma solução fácil para um sistema linear $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Um sistema pré-condicionado tem a seguinte forma (vide eq. (7.7.1) de[4]:

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{x} = \mathbf{M}^{-1}\mathbf{b} \quad (11.3.22)$$

A seguir aplicar-se-á o conceito de condicionamento a um funcional quadrático.

Seja \mathbf{M} uma matriz simétrica positiva definida fatorada na forma $\mathbf{M} = \mathbf{L}\mathbf{L}^T$

(decomposição de Cholesky inferior) e seja o funcional quadrático, como definido em (11.2.88):



$$f_{\mathbf{x}} = \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} - \mathbf{b}^T \mathbf{x} + \chi, \quad \mathbf{x} \in \mathbb{R}^n$$

onde \mathbf{H} é positiva definida. Definindo-se um segundo funcional $\tilde{f}_{\mathbf{y}}$ pela transformação $\mathbf{y} = \mathbf{L}^T \mathbf{x}$, se tem

$$\tilde{f}_{\mathbf{y}} = f_{\mathbf{L}^T \mathbf{y}} = \frac{1}{2} \mathbf{y}^T \mathbf{H} \mathbf{y} - \mathbf{b}^T \mathbf{y} + \tilde{\chi} \quad (11.3.23)$$

onde

$$\tilde{\mathbf{H}} = \mathbf{L}^{-1} \mathbf{H} \mathbf{L}^T, \quad \tilde{\mathbf{b}} = \mathbf{E}^{-1} \mathbf{b}, \quad \tilde{\chi} = \chi \quad (11.3.24)$$

Como, por definição, \mathbf{H} é positiva definida, $\tilde{\mathbf{H}} = \mathbf{L}^{-1} \mathbf{H} \mathbf{L}^T$ é uma transformação de similaridade, logo $\tilde{\mathbf{H}}$ é também simétrica e positiva definida

A transformação de similaridade

$$\mathbf{L}^{-T} \tilde{\mathbf{H}} \mathbf{L}^T = \mathbf{L}^{-T} \mathbf{L}^{-1} \mathbf{H} \mathbf{L}^T \mathbf{L}^T = \underbrace{\mathbf{L}^{-T} \mathbf{L}^{-1}}_{=\mathbf{I}} \mathbf{H} = \mathbf{M}^{-1} \mathbf{H} \quad (11.3.25)$$

Revela que tanto $\tilde{\mathbf{H}}$ e $\mathbf{M}^{-1} \mathbf{H}$ têm os mesmos autovalores, $0 < \lambda_1 < \lambda_2 < \dots < \lambda_n$ de modo que número condição espectral $\kappa \mathbf{H} = \lambda_n / \lambda_1$ é completamente determinado por \mathbf{M} e \mathbf{H} visto que $\tilde{\mathbf{H}}$ depende da fatoração de \mathbf{M} .

Aplicando o método do passo descendente para o problema (11.3.23), vê-se que é conveniente o uso do cálculo direto dos gradientes e então o processo iterativo fica

$$\tilde{\mathbf{g}}_k = \tilde{\mathbf{H}} \mathbf{y}_k - \tilde{\mathbf{b}} \quad (11.3.26)$$

$$\tilde{\tau}_k = \tilde{\mathbf{g}}_k^T \tilde{\mathbf{g}}_k / \tilde{\mathbf{g}}_k^T \tilde{\mathbf{H}} \tilde{\mathbf{g}}_k \quad (11.3.27)$$

$$\mathbf{y}_{k+1} = \mathbf{y}_k - \tilde{\tau}_k \tilde{\mathbf{g}}_k \quad (11.3.28)$$

Para $k = 0, 1, \dots$, e sendo o ponto inicial \mathbf{y}^0 escolhido arbitrariamente se tem

$$\lim_{k \rightarrow \infty} \mathbf{y}^k = \psi \equiv \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{b}} \quad (11.3.29)$$

A taxa de convergência de (11.3.29) depende de $\kappa \tilde{\mathbf{H}}$.

Fazendo $\mathbf{x}^k = \mathbf{L}^T \mathbf{y}^k$ e $\mathbf{g}^k = \mathbf{H} \mathbf{x}^k - \mathbf{b}$ para $k = 0, 1, 2, \dots$

4 PROBLEMAS DE VALORES DE CONTORNO: FORMULAÇÃO VARIACIONAL

Em engenharia a grande maioria dos problemas é formulada como problemas de valores de contorno (PVC) os quais podem ser expressos na determinação de uma função que satisfaz alguma



equação diferencial em uma região Ω de definição e que deve satisfazer a condições específicas no contorno Γ da região Ω . Numa visão geral, a grande maioria destes problemas é relacionada com uma solução que minimiza um funcional $J f$ definido para funções f pertencentes a um conjunto de funções V .

Esta minimização (ou maximização) requer que o funcional $J f$ seja estacionário. Deste modo a tarefa de resolver um PVC é equivalente a achar uma função uma função $f \in V$ que faz $J f$ estacionário. Esta é a formulação variacional de um PVC.