Capítulo 2

Bases matemáticas

Na maioria dos projetos não é direta a determinação de sua melhor solução. Alguns poucos privilegiados tem a capacidade de vislumbrar a melhor solução por intuição, heuristicamente. A grande maioria dos projetistas precisa de alguma ferramenta que os oriente entre as muitas possibilidades em geral existentes. Essa ferramenta é a matemática. Em particular, o cálculo diferencial, a técnica que trata da medição do efeito da variação de parâmetros sobre o valor de uma função. Uma de suas virtudes é possibilitar a determinação de valores dos parâmetros que maximizam ou minimizam uma função. O cálculo variacional também é bastante utilizado, principalmente em problemas dinâmicos e também nos métodos energéticos.

O cálculo diferencial foi o grande legado de dois dos maiores gênios que a humanidade produziu em todos os tempos, os contemporâneos Newton e Leibntiz. Quanto ao cálculo variacional pode-se distinguir as grandes contribuições de Lagrange e Hamilton.

Este capítulo trata especialmente dos conceitos matemáticos envolvidos nos processos de otimização. Aqueles que não tiveram sua educação descuidada na matemática podem, se quiserem, pular este Capítulo e ir em frente, voltando a ele apenas quando tiverem dúvida em alguma ferramenta matemática

2.1 Vetores e matrizes

Um vetor **x**, ou um **ponto** no espaço afim, é, de modo simplista, um conjunto ordenado de n valores x_i (i = 1,...,n), geralmente representados em uma única coluna, que exprimem o estado de um sistema. Assim, a posição de um ponto no espaço tridimensional em que vivemos pode ser expressa sem nenhuma ambiguidade por três coordenadas,

$$\mathbf{x} = \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} \tag{2.1}$$

Se essas coordenadas variarem no tempo, isto é, se a partícula está em movimento, essas coordenadas são funções do tempo, não constantes e exprimem uma **trajetória** nesse espaço. Nesse caso, é comum se referir ao vetor como um **campo**.

Neste texto, vetores serão representados por letras latinas minúsculas em negrito.
O produto escalar de dois vetores é definido como

$$(\mathbf{x} \bullet \mathbf{y}) = \mathbf{x}^{\boldsymbol{T}}\mathbf{y} = \sum_{i=1}^{n} x_i\, y_i \tag{2.2}$$

Na chamada convecção de Einstein, a repetição de índices em um monômio implica em somatória para todos os valores desse índice. Assim, Eq. (2.2) é, simplesmente,

$$(\mathbf{x} \bullet \mathbf{y}) = x_i y_i \tag{2.3}$$

A norma ou comprimento de um vetor é dado por

$$\|\mathbf{x}\| = \sqrt{x_i x_i} = \sqrt{\mathbf{x} \bullet \mathbf{x}} \qquad (2.4)$$

Uma matriz é um conjunto ordenado de valores com um certo número *n* de linhas e *m* de colunas, e é representada por letras latinas maiúsculas em negrito.

2.2 Funções e suas derivadas

2.2.1 Funções de uma variável

Uma função de uma variável *f* = *f(x)*, é um procedimento que transforma o valor da variável *x* em um outro número *f*.

Em uma função de uma única variável, sem restrições, a determinação de máximos e mínimos é um problema clássico do cálculo diferencial, e os leitores são remetidos aos textos dessa disciplina. Note-se que a função deve ser necessariamente não linear, uma vez que a busca de máximos e mínimos de uma função linear, sem restrições, não tem sentido.

2.2.2 Funções de várias varáveis e funções vetoriais

Uma função *f* = *f(***x***)*, ou

$$f(\mathbf{x}) = f(x_1, x_2, \dots, x_n), \qquad (2.5)$$

é um procedimento que transforma os valores de um vetor ou ponto **x** em um número *f*. Da mesma forma, pode-se ter um vetor de *m* funções de *n* variáveis, na forma

$$\mathbf{g}(\mathbf{x}) = \begin{Bmatrix} g_{\mathbf{1}}(\mathbf{x}) \\ g_2(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{Bmatrix} \text{ em que } g_j(\mathbf{x}) = g_i(x_1, x_2, \dots, x_n), \quad j = 1, \dots, m \qquad (2.6)$$

Na pesquisa de máximos e mínimos de funções desses tipos surge a necessidade do cálculo de derivadas parciais. As derivadas parciais de primeira ordem de (2.5) são

$$\frac{\partial f(\mathbf{x})}{\partial x_i}; \; i = 1, \dots, n \qquad (2.7)$$

que podem ser arranjadas em um vetor coluna chamado **vetor gradiente** da função:

$$\mathbf{c}(\mathbf{x}) = \vec{\nabla} f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial x_i} = \begin{Bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{Bmatrix} \tag{2.8}$$

em que foi utilizado o operador$\vec{\nabla}$, **nabla**.

Cada componente de (2.8) pode ser novamente diferenciado com respeito a uma variável para obter as derivadas parciais de segunda ordem

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j};\ i,j = 1,\dots,n \tag{2.9}$$

A Eq. (2.9) pode ser arranjada em uma matriz *n*², denominada matriz Hessiana:

$$\mathbf{H}(\mathbf{x}) = \vec{\nabla}\vec{\nabla}^T f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_n} \end{bmatrix}. \tag{2.10}$$

Há casos em que se necessita diferenciar um vetor de *m* funções de *n* variáveis como (2.6), com relação às próprias variáveis. Tem-se, nesse caso, a chamada matriz Jacobiana:

$$\mathbf{J}(\mathbf{x}) = \vec{\nabla}\mathbf{g}(\mathbf{x})^T = \begin{bmatrix} \frac{\partial g_1(\mathbf{x})}{\partial x_1} & \frac{\partial g_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial g_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial g_2(\mathbf{x})}{\partial x_1} & \frac{\partial g_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial g_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_m(\mathbf{x})}{\partial x_1} & \frac{\partial g_m(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial g_m(\mathbf{x})}{\partial x_n} \end{bmatrix} \tag{2.11}$$

Note-se que a matriz Jacobiana tem dimensões *m* x *n*, ou seja, ela não é necessariamente quadrada!

2.3 Expansão em série de Taylor

Uma ferramenta básica deste capítulo são as séries de Taylor, devidas ao matemático britânico Brook Taylor (1685-1731). No caso de uma única função (escalar) de uma única variável, $f(x)$, conhecido seu valor e de suas derivadas num certo ponto a, pode-se aproximar o valor da função em um outro ponto próximo *x*, definida a "distância" *d* = *x* – *a*, pela expansão em série:

$$f(x) = \frac{1}{0!}d^0 f(a) + \frac{1}{1!}d^1 f'(a) + \frac{1}{2!}d^2 f''(a) + \frac{1}{3!}d^3 f'''(a) + \cdots \qquad (2.12)$$

onde as linhas a direita e acima denotam derivas sucessivas em *x*, calculadas no ponto *a*.

Já no caso de uma única função (escalar) de várias variáveis (um vetor de *n* variáveis) $f(\mathbf{x})$, conhecido seu valor num ponto **a** (*n* x 1), pode-se aproximar o valor da função em um outro ponto próximo **x** (*n* x 1), tendo-se o "vetor distância" **d** = **x** – **a** (*n* x 1)**,** pela expansão em série:

$$f(\mathbf{x}) = f(\mathbf{a}) + \mathbf{c}(\mathbf{a})^T\mathbf{d} + \frac{1}{2!}\mathbf{d}^T\mathbf{H}(\mathbf{a})\mathbf{d} + \cdots \qquad (2.13)$$

onde se utiliza a transposta do *vetor gradiente* da Eq. (2.8) e a matriz Hessiana (*n* x *n*), da Eq. (2.10).

Finalmente, no caso de um vetor de *n* funções de várias variáveis (um vetor de *n* variáveis) $\boldsymbol{f}(\mathbf{x})$, (*n* x 1), conhecido seu valor num ponto ***a*** (*n* x 1), pode-se aproximar o valor dessas funções em um outro ponto próximo **x** (*n* x 1), dado o "vetor distância" **d** = **x** – **a** (*n* x 1)**,** pela expansão em série:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \mathbf{J}(\mathbf{a})\mathbf{d} + \cdots \qquad (2.14)$$

onde se utiliza a matriz Jacobiana da Eq. (2.11), (*n* x *n* neste caso).

A título de informação, quando se faz uma expansão em torno da origem, isto é, ***a*** = 0, a série de Taylor passa a ser denominada série de MacLaurin, devida a ao escocês Colin MacLaurin (1698-1746).

2.4 Formas quadráticas e matrizes definidas

Uma forma quadrática é uma forma especial de função não linear com somente termos de segunda ordem, na forma geral

$$F(\mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} x_i x_j = \mathbf{x}^{\boldsymbol{T}}\boldsymbol{P}\mathbf{x} \qquad (2.15)$$

em que **P** é a matriz da forma quadrática. Há infinitas tais matrizes, todas assimétricas, menos uma simétrica dada por

$$\mathbf{A} = \frac{1}{2}(\mathbf{P} + \mathbf{P}^{\boldsymbol{T}}) \qquad (2.16)$$

Como **A** é simétrica, sabe-se, da álgebra linear, que todos seus autovaloresλ_i são reais.

Uma forma quadrática $F(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ pode ser positiva, negativa ou zero para qualquer **x**, e sua matriz é dita positivo definida, positivo semidefinida, negativa definida, negativa semidefinida ou indefinida, como se segue.

1. Positivo definida: $F(\mathbf{x}) > 0$para todo $\mathbf{x} \neq 0$; $\lambda_i > 0$.
2. Positivo semidefinida: $F(\mathbf{x}) \geq 0$para todo $\mathbf{x} \neq 0$; $\lambda_i \geq 0$ (pelo menos um valor nulo).
3. Negativo definida: $F(\mathbf{x}) < 0$para todo $\mathbf{x} \neq 0$; $\lambda_i < 0$.
4. Negativo semidefinida: $F(\mathbf{x}) \leq 0$para todo $\mathbf{x} \neq 0$; $\lambda_i \leq 0$ (pelo menos um valor nulo).
5. Indefinida: se a forma for positiva para alguns valores de **x** e negativa para outros; $\lambda_i > 0$, para alguns valores de **x** e $\lambda_i < 0$para outros.

2.5 Mínimos e máximos de funções

2.5.1 Otimização Sem Restrições

Considere-se o problema de minimizar uma função $f(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^n$. O problema pode ser colocado na forma

$$\text{minimizar} \quad f(\mathbf{x}) \qquad \mathbf{x} \in \mathbf{R}^n. \tag{2.17}$$

Admite-se que a função $f(\mathbf{x}) \in C_2$. A função $f(\mathbf{x})$ é denominada função objetivo.

Nesta seção, serão estabelecidas condições que devem ser satisfeitas por um ponto, para que seja um mínimo local do problema (2.17). Também serão descritas as propriedades de convexidade da função objetivo que asseguram que o ponto encontrado seja um ponto de mínimo global.

Para que um ponto **x*** seja um ponto de mínimo local do problema (2.17) é suficiente que o gradiente da função objetivo em **x*** seja nulo, ou seja,$\nabla f(\mathbf{x}*) = 0$ e que a matriz Hessiana $\boldsymbol{\nabla^2 f}(\mathbf{x}*)$ seja definida positiva, isto é

$$\mathbf{d}^T \nabla^2 f(\mathbf{x}^*)\mathbf{d} > 0 \quad \forall \mathbf{d} \neq 0.$$

Essa notação indica produto escalar dos vetores.

Cabe aqui uma explicação desses critérios de mínimo de uma função sem restrições. Um ponto de mínimo é aquele em que o valor de uma função só pode crescer nos pontos em uma vizinhança **d** pequena. Para determinar esses valores pode-se utilizar a expansão de Taylor da Eq. (2.13). Transferindo o primeiro termo do lado direito para o esquerdo, obtém-se a diferença no valor da função entre o ponto de mínimo e um ponto vizinho

$\Delta f = \mathbf{c}(\mathbf{x}^*)^T\mathbf{d} + \frac{1}{2!}\mathbf{d}^T\mathbf{H}(\mathbf{x}^*)\mathbf{d} + \cdots$

Como a distância **d** é pequena, o primeiro termo dessa expressão, relacionado às derivadas primeiras da função, é dominante sobre o segundo. Essa distância **d** pode ser positiva ou negativa, mas esse termo tem que ser nulo se o ponto é de mínimo. Em consequência, o gradiente nesse ponto tem que ser nulo. Para uma só variável, simplesmente a primeira derivada deve ser nula num ponto de mínimo.

Resta examinar o segundo termo. Esse deve ser positivo para que a diferença obtida sempre cresça em torno do ponto que se supõem de mínimo, para qualquer distância **d** pequena. Isso implica em que a matriz Hessiana, das derivadas segundas, deva ser positivo definida ou, ao menos, positivo semidefinida. Para uma só variável, simplesmente a segunda derivada deve ser positiva num ponto de mínimo.

Seja $f(\mathbf{x})$ uma função convexa definida em $\mathbf{R}^n$ e seja Ω o conjunto dos pontos $\mathbf{x} \in \mathbf{R}^n$ onde $f(\mathbf{x})$ atinge seu mínimo. Então Ω é convexo e todo mínimo local é um ponto de mínimo global.

Seja $f(\mathbf{x}) \in C_1$ uma função convexa. Se existir um $\mathbf{x}^* \in \mathbf{R}^n$ tal que para todo $\mathbf{y} \in R^n$, o produto escalar

$$\nabla f(\mathbf{x}^*)^T(\mathbf{y} - \mathbf{x}^*) \geq 0$$

então $\mathbf{x}^*$ é um ponto de mínimo global de $f(\mathbf{x})$.

Exemplo com uma variável de projeto: um recipiente cilíndrico de volume fixo V para um fluido, feito de chapa metálica de espessura fixa, tem raio R e altura H, as variáveis de projeto. O consumo de chapa metálica, em área, a função objetivo, ou custo, é

$$A = \pi(R^2 + RH)$$

com a restrição do volume fixo

$V = \pi R^2 H.$

Essas duas equações podem ser unidas numa função objetivo equivalente em uma única variável, R:

$$f = R^2 + \frac{V}{\pi R}$$

A condição necessária de ótimo, de primeira ordem, é:

$$f' = 2R - \frac{V}{\pi R^2} = 0$$

correspondendo aos valores das variáveis

$$R* = \sqrt[3]{\frac{V}{2\pi}} \quad \text{e} \quad H* = \sqrt[3]{\frac{4V}{\pi}}$$

os candidatos a ponto de mínimo. A verificação pela condição suficiente é dada pela curvatura

$$f'' = 2 + \frac{2V}{\pi R^3} = 6$$

que é positiva, indicando ponto de mínimo.

2.5.2 Otimização com Restrições

Considere-se o problema geral de otimização da forma

$$\text{minimizar} \quad f(\mathbf{x}),\ \mathbf{x} \in \mathbf{R}^n,$$

$$\text{sujeito a} \quad h_j(\mathbf{x}) = 0, \quad j \in E$$

$$g_i(\mathbf{x}) \leq 0, \quad i \in I \tag{2.18}$$

onde a função $f(\mathbf{x})$ é denominada função objetivo. E é o conjunto dos índices das restrições de igualdade $h_j(\mathbf{x}) = 0, j = 1, \cdots, p$, e I é o conjunto dos índices das restrições de desigualdade $g_i(\mathbf{x}) \leq 0, i = 1, \cdots, m$. A solução de (2.18) é denominada de solução ou ponto ótimo e será denotada por $\mathbf{x}$*.

Quando um ponto $\mathbf{x} \in \mathbf{R}^n$ satisfaz todas as restrições, diz-se que ele é **viável,** e o conjunto de todos os pontos viáveis é denominado de **região viável**.

Admite-se que as funções $f(\mathbf{x})$ e de restrições $\in C_2\,[\mathbf{R}^n]$.

As restrições $\{g_i(\mathbf{x}), i \in I \mid g_i(\mathbf{x}) = 0\}$ são denominadas **restrições ativas** em $\mathbf{x}$.

Iniciando pelo problema com apenas restrições de igualdade, diz-se que um ponto $\mathbf{x}$, que satisfaz as restrições $h_j(\mathbf{x})$ é um ponto regular do conjunto viável se os vetores gradientes das restrições forem linearmente independentes, isto é, nenhum par de vetores gradientes são paralelos entre si, e nenhum gradiente pode ser obtido por combinação linear de outros.

Denomina-se a função Lagrangiana de um problema com restrições de igualdade a função definida por:

$$L(\mathbf{x}, \mathbf{v}) = f(\mathbf{x}) + \sum_{j=1}^{p} v_j h_j(\mathbf{x}), \tag{2.19}$$

onde $v_i, j = 1, \cdots, p$ são os multiplicadores de Lagrange.

A função Lagrangiana é um artifício matemático para transformar um problema de otimização com restrições de igualdade em um problema sem restrições em que já se conhecem as condições de mínimo.

Indica-se por $\nabla L(\mathbf{x}, \mathbf{v}) = \begin{bmatrix} \nabla_x L(\mathbf{x}, \mathbf{v}) \\ \nabla_v L(\mathbf{x}, \mathbf{v}) \end{bmatrix}$ o vetor gradiente de $L(\mathbf{x},\mathbf{v})$, onde ∇_x indica as derivadas parciais em relação à x_i, $i = 1,...,n$, e ∇_v indica as derivadas em relação aos multiplicadores de Lagrange $v_j, j = 1, \cdots, p$.

O vetor gradiente da função Lagrangiana em relação à $\mathbf{x}$ em $(\mathbf{x}^*,\mathbf{v}^*) \in \mathbf{R}^{n+p}$ é dado por:

$$\nabla_x L(\mathbf{x}^*, \mathbf{v}^*) = \nabla f(\mathbf{x}^*) + \sum_{j=1}^{p} v_j^* \nabla h_j(\mathbf{x}\,*),$$

onde **v*** é o vetor dos multiplicadores de Lagrange no ponto ótimo.

Analogamente ao vetor gradiente, a matriz Hessiana da função Lagrangiana em relação à **x** no ponto (**x***,**v***) ∈ $\mathbf{R}^{n+p}$ é dada por:

$$\nabla_x^2 L(\mathbf{x}*, \mathbf{v}*) = \nabla^2 f(\mathbf{x}*) + \sum_{j=1}^{p} v_j \nabla^2 h_j(\mathbf{x}*).$$

Exemplo com duas variáveis de projeto: um recipiente cilíndrico de volume fixo V para um fluido, feito de chapa metálica de espessura constante, tem raio R e altura H, as variáveis de projeto. O consumo de chapa metálica, em área, a função objetivo, ou custo, é

$$f = A = \pi(R^2 + RH)$$

com a restrição do volume fixo

$h = \pi R^2 H - V = 0.$

O Lagrangiano é

$L = \pi(R^2 + RH) + v(\pi R^2 H - V)$, sendo v o multiplicador de Lagrange.

Impondo que seja estacionário:

$$\frac{\partial L}{\partial R} = 2R + H + 2\pi v R H = 0$$

$$\frac{\partial L}{\partial H} = R + \pi v R^2 = 0$$

$$\frac{\partial L}{\partial v} = \pi R^2 H - V = 0$$

chega-se às soluções do problema

$$R *= \sqrt[3]{\frac{V}{2\pi}} \;\; ; \quad H *= \sqrt[3]{\frac{4V}{\pi}}; \quad v *= -\frac{1}{\pi R} = \sqrt[3]{\frac{2}{\pi^2 V}}$$

2.5.3 Condições de Karush-Kuhn-Tucker (KKT): restrições de igualdade

Seja **x*** um ponto de mínimo local do problema (2.18), **apenas com restrições de igualdade**. Se **x*** é um ponto regular, então existem multiplicadores de Lagrange, de tal forma que **x*** e **v*** satisfazem o seguinte sistema de equações:

$$\nabla f(\mathbf{x}^*) + \sum_{j=1}^{p} v_j^* \nabla h_j(\mathbf{x}^*) = 0$$

$$h_j(\mathbf{x}^*) = 0, \qquad i \in E$$

$$v_j^* > 0, \qquad i \in E \tag{2.20}$$

As equações acima são denominadas Condições de Karush-Kuhn-Tucker, para o caso de apenas restrições de igualdade.

Para que o ponto **x*** seja um ponto de mínimo local do problema (2.18) é necessário que seja um ponto regular, que satisfaça as condições de Karush-Kuhn-Tucker e ainda que o produto escalar

$$\mathbf{d}^T \nabla_x^2 L(\mathbf{x}^*, \mathbf{v}^*)\mathbf{d} \geq 0 \tag{2.21}$$

Considere-se o problema:

minimizar $f(\mathbf{x})$

sujeito às restrições

$$h_j(\mathbf{x}) = 0, \; j = 1, \ldots, p$$

$$\mathbf{x} \in \mathbf{R}^n.$$

Se a função objetivo e as restrições forem convexas o problema é dito de programação convexa. Para problema de programação convexa valem os seguintes resultados:

- Toda solução **x*** de um problema de programação convexa é uma solução global e o conjunto das soluções globais *S* é um conjunto convexo;
- Se no problema de programação convexa a função objetivo for estritamente convexa em $\mathbf{R}^n$ então toda solução global é única;
- Se num problema de programação convexa, as funções $f(\mathbf{x})$ e $h_j(\mathbf{x})$ são contínuas com derivadas parciais contínuas até primeira ordem, e se as condições de Karush-Kuhn-Tucker estão satisfeitas em **x***, então o ponto **x*** é uma solução global do problema de programação convexa.

2.5.4 Condições K-K-T: Problemas com restrições gerais utilizando variáveis de folga

Considere-se o problema geral de otimização de várias variáveis **x**, com função objetivo *f*(**x**), sujeito a p restrições de igualdade $h(\mathbf{x})_j = 0, j = 1$ a p, e m restrições de desigualdade $g_i(\mathbf{x}) \leq 0, i = 1$ a m, que podem ser escritas na forma de um vetor

$$\mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x}) \quad g_2(\mathbf{x}) \quad \dots \quad g_m(\mathbf{x})]^T \tag{2.23}$$

Essas últimas podem ser transformadas em restrições de igualdade pela adição de *slack variables* (variáveis de folga) s_i, tais que

$$g_i(\mathbf{x}) + {s_i}^2 = 0 \tag{2.24}$$

que constituem o vetor

$$\mathbf{s} = [s_1 \quad s_2 \quad \dots \quad s_m]^T \tag{2.25}$$

Para essas novas relações de igualdade são adotados multiplicadores de Lagrange

$$u_i \geq 0 \tag{2.26}$$

que constituem o vetor

$$\mathbf{u} = [u_1 \quad u_2 \quad \dots \quad u_m]^T \tag{2.27}$$

Pode-se, agora, escrever uma **função Lagrangiana**, ou Lagrangiano, escalar, na forma

$$L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = f(\mathbf{x}) + \mathbf{v}^{\boldsymbol{T}}\mathbf{h}(x) + \mathbf{u}^{\boldsymbol{T}}(\mathbf{g}(\mathbf{x}) + \mathbf{s}^{\mathbf{2}}) \tag{2.28}$$

Demonstra-se um Teorema, as Condições de Karush-Kuhn-Tacker (KKT) do problema geral de otimização para um ponto **x*** de mínimo local dentro do conjunto viável. O procedimento é, em resumo:

1. Escrever o Lagrangiano

$$L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = f(\boldsymbol{x}) + \sum_{j=1}^{p} v_j h_j(\mathbf{x}) + \sum_{i=1}^{m} u_i(g_i(\mathbf{x}) + s_i^2) \tag{2.29}$$

2. Calcular as Condições Gradientes

$$\frac{\partial L}{\partial x_k} = \frac{\partial f}{\partial x_k} + \sum_{j=1}^{p} v_j * \frac{\partial h_j}{\partial x_k} + \sum_{i=1}^{m} u_i * \frac{\partial g_i}{\partial x_k} = 0; \quad k = 1 \text{ a } n \tag{2.30}$$

$$\frac{\partial L}{\partial v_j} = 0 \quad \text{ou} \quad h_j(\boldsymbol{x}*) = 0; \quad j = 1 \text{ a } p \tag{2.31}$$

$$\frac{\partial L}{\partial u_i} = 0 \quad \text{ou} \quad g_i(\boldsymbol{x}*) - s_i^2 = 0; \quad i = 1 \text{ a } m \tag{2.32}$$

3. Calcular as condições de “chaveamento”

$\frac{\partial L}{\partial s_i} = 0$ ou $2u_i * s_i = 0; \quad i = 1$ a m (2.33)

4. Verificar a viabilidade para as inequações

$s_i^2 \geq 0$ equivalente a $g_i(\mathbf{x}) \leq 0; \quad i = 1$ a m (2.34)

5. Verificar a não negatividade dos multiplicadores de Lagrange das inequações

$u_i \geq 0; \quad i = 1$ a m (2.35)

2.5.5 Convexidade

Um **conjunto convexo** *S* é uma coleção de pontos **x** em que se dois pontos $\mathbf{x}_1$ e $\mathbf{x}_2$estão nele contidos, então todo o segmento $\mathbf{x}_1 - \mathbf{x}_2$ também está contido em *S*. Um segmento de reta é sempre um conjunto convexo pois todos os pontos no interior do segmento obedecem à definição acima. Se numa função de uma só variável a linha que une dois pontos está sempre acima da curva da função, essa é uma **função convexa**.

Uma função de várias variáveis $f(\mathbf{x})$, definida em um conjunto convexo, é uma função convexa se somente se sua Hessiana é positivo semi-definida ou definida em todos os pontos do conjunto. Nesse último caso é estritamente convexa.

Pode-se provar um teorema em que se define um **problema convexo de programação**, como se segue. Seja um problema geral de otimização cujo conjunto viável é dado pela seguinte expressão:

$$S = \left\{\mathbf{x} | h_j(\mathbf{x}) = 0, j = 1 \text{ a } p; g_i(\mathbf{x}) = 0, i = 1 \text{ a } m\right\}$$

Então, o conjunto *S* é um conjunto convexo se as funções g_i forem convexas e as funções h_j forem lineares. Restrições de igualdade não lineares sempre levam a conjuntos não convexos. Restrições de igualdade ou de desigualdade lineares sempre levam a conjuntos convexos.

Outro teorema diz que se$f(\mathbf{x}^*)$ é um **mínimo local** de uma função convexa $f(\mathbf{x})$ definida em um conjunto viável *S* convexo, então é também um **mínimo global**.

Em virtude dos teoremas acima, pode-se afirmar que se uma função convexa $f(\boldsymbol{x})$ uma função convexa definida em um conjunto viável *S* convexo, então as condições **necessárias** KKT são também **suficientes** para um mínimo global.

2.6 Exemplos

2.6.1 Exemplo 1: Condições KKT

Minimizar $f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$

Sujeito a $g_1 = -2x_1 - x_2 + 4 \leq 0 \qquad g_2 = -x_1 - 2x_2 + 4 \leq 0$

1. Função Lagrangina

$$L = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2 + u_1(-2x_1 - x_2 + 4 + s_1^2) + u_2(-x_1 - 2x_2 + 4 + s_2^2)$$

2. Condições KKT

$$\frac{\partial L}{\partial x_1} = 2x_1 - 2 - 2u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 2 - u_1 - 2u_2 = 0$$

As derivadas parciais do Lagrangiano com respeito aos multiplicadores de Lagrange recuperam as próprias equações de restrições:

$$g_1 = -2x_1 - x_2 + 4 + s_1^2 = 0; \quad s_1^2 \geq 0, \quad u_1 \geq 0$$

$$g_2 = -x_1 - 2x_2 + 4 + s_2^2 = 0; \quad s_2^2 \geq 0, \quad u_2 \geq 0$$

As derivadas parciais do Lagrangiano com respeito às variáveis de folga resultam sempre as mesmas condições de chaveamento:

$$u_i s_i = 0; \quad i = 1{,}2$$

Análise de possíveis soluções

Caso 1. $u_1 = 0, \quad u_1 = 0$, implica em que $s_1^2 = -1, \quad s_2^2 = -1$, impossível.

Caso 2. $u_1 = 0, \quad s_2 = 0$, implica em que $x_1 = 1{,}2; \quad x_2 = 1{,}2; \quad u_1 = 0; \quad u_2 = 0{,}4$; mas $s_2^2 = -0{,}2$, impossível.

Caso 3. $s_1 = 0, \quad u_2 = 0$, implica em que $s_2^2 = -0{,}2$, impossível.

Caso 4. $s_1 = 0, \quad s_2 = 0$, implica em que $x_1 = 4/3; \quad x_2 = 4/3; \quad u_1 = 2/9; \quad u_2 = 2/9$. É um candidato a mínimo válido! E a função custo nesse caso vale $f = 2/9$.

2.6.2 Exemplo 2

Dois geradores elétricos são interconectados para alimentar uma carga de pelo menos 60 unidades de um certo consumidor. O custo de operação de cada gerador é função de sua produção de energia e é dado pelas expressões abaixo, com base em custo por unidade.

Formule o problema de custo mínimo para determinar as potências P_1 e P_2 que cada gerador deve fornecer. Formular as condições KKT.

Custo por unidade de potência do gerador 1 – $C_1 = 1 - P_1 + P_1^2$

Custo por unidade de potência do gerador 2 – $C_2 = 1 + 0{,}6\, P_2 + P_2^2$

Variáveis de projeto: $x_1 = P_1$ e $x_2 = P_2$

Função objetivo: $f(\mathbf{x}) = C_1 + C_2 = 2 - x_1 + x_1^2 + 0{,}6x_2 + x_2^2$

Sujeita a:

$$g_1 = -x_1 - x_2 + 60 \leq 0$$

$$g_2 = -x_1 \leq 0$$

$$g_3 = -x_2 \leq 0$$

1. Lagrangiano

$$L = 2 - x_1 + x_1^2 + 0{,}6x_2 + x_2^2 + u_1(-x_1 - x_2 + 60 + s_1^2) + u_2(-x_1 + s_2^2) + u_3(-x_2 + s_3^2)$$

2. Condições KKT

$$\frac{\partial L}{\partial x_1} = -1 + 2x_1 - u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial x_2} = 0{,}6 + 2x_2 - u_1 - u_3 = 0$$

As derivadas parciais do Lagrangiano com respeito aos multiplicadores de Lagrange recuperam as próprias equações de restrições:

$$\frac{\partial L}{\partial u_1} = -x_1 - x_2 + 60 + s_1^2 = 0$$

$$\frac{\partial L}{\partial u_2} = -x_1 + s_2^2 = 0$$

$$\frac{\partial L}{\partial u_5} = -x_2 + s_3^2 = 0$$

3. As derivadas parciais do Lagrangiano com respeito às variáveis de folga resultam sempre as mesmas condições de chaveamento:

$$2u_1 s_1 = 0$$

$$2u_2 s_2 = 0$$

$$2u_3 s_3 = 0$$

4. Viabilidade

$$s_i^2 \geq 0; \quad i = 1 \text{ a } 3$$

5. Não negatividade dos multiplicadores de Lagrange

$$u_1 \geq 0; \quad i = 1 \text{ a } 3$$

2.6.3 Exemplo

A treliça de duas barras da Fig. 2.1 (triângulo retângulo 30:40:50 cm) deve ser projetada para suportar um peso W = 1200 KN no nó A, sem que as barras excedam a tensão normal admissível do material $\bar{\sigma}$=16 KN/cm². O volume total de material deve ser minimizado.

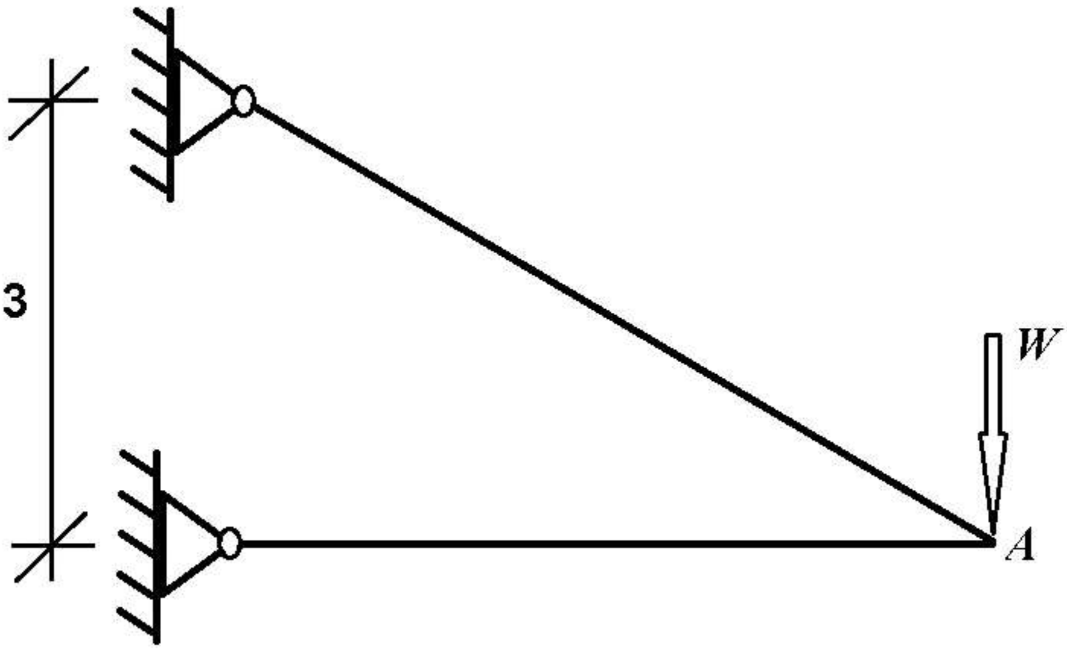

Figura 2.1 – Treliça exemplo 2.6.3

Equilíbrio do nó A:

$$\sum V = 0{,}6F_1 - 1200 = 0 \quad \therefore \quad F_1 = 2000KN$$

$$\sum H = -0{,}8F_1 - F_2 = 0 \quad \therefore \quad F_2 = 1600KN$$

Verificação da tensão admissível: $\sigma_i = \frac{F_i}{A_i} \leq \bar{\sigma}$

Variáveis de projeto: $x_1 = A_1$; $x_2 = A_2$, em cm²

Função objetivo: $f(\mathbf{x}) = 50x_1 + 40x_2$, em cm³

Sujeita a $g_1(\mathbf{x}) = \frac{2000}{x_1} - 16 \leq 0$; $g_2(\mathbf{x}) = \frac{1600}{x_2} - 16 \leq 0$;
$g_3(\mathbf{x}) = -x_1 \leq 0$; $g_4(\mathbf{x}) = -x_2 \leq 0$

Lagrangiano:

$$L = 50x_1 + 40x_2 + u_1\left(\frac{2000}{x_1} - 16 + s_1^2\right) + u_2\left(\frac{1600}{x_1} - 16 + s_2^2\right) + u_3(-x_1 + s_3^2) + u_4(-x_2 + s_4^2)$$

Condições necessárias KKT:

$$\frac{\partial L}{\partial x_1} = 0 = 50 - u_1 \frac{2000}{x_1^2} - u_3$$

$$\frac{\partial L}{\partial x_2} = 0 = 40 - u_2 \frac{1600}{x_2^2} - u_4$$

As derivadas parciais do Lagrangiano com respeito aos multiplicadores de Lagrange recuperam as próprias equações de restrições.

As derivadas parciais do Lagrangiano com respeito às variáveis de folga resultam sempre as mesmas condições de chaveamento:

$$u_i s_i = 0; \quad u_i \geq 0; \quad g_i + s_i^2 = 0; \quad s_i^2 \geq 0; \quad i = 1 \text{ a } 4$$

Casos em que $s_3 = s_4 = 0$ levam a áreas das seções transversais nulas, o que não é fisicamente aceitável.

Outros casos:

Caso 1: $u_1 = 0$; $u_2 = 0$; $u_3 = 0$; $u_4 = 0$; resultam $50 = 0$ e $40 = 0$, o que é absurdo!

Caso 2: $s_1 = 0$; $u_2 = 0$; $u_3 = 0$; $u_4 = 0$; resulta $50 = 0$, o que é absurdo!

Caso 3: $u_1 = 0$; $s_2 = 0$; $u_3 = 0$; $u_4 = 0$; resulta $40 = 0$, o que é absurdo!

Caso 4: $s_1 = 0$; $s_2 = 0$; $u_3 = 0$; $u_4 = 0$; resultam $x_1^* = 125$ cm², $x_2^* = 100$ cm², $u_1 = 0{,}391$ e $u_2 = 0{,}25$, atendendo às condições necessárias KKT de candidato a mínimo local.

Como $f(\mathbf{x})$, $g_3(\mathbf{x})$ e $g_4(\mathbf{x})$ são lineares, e as matrizes Hessianas de $g_1(\mathbf{x})$ e $g_2(\mathbf{x})$ são positivo semi-definidas, o problema é convexo e essa solução atende à condição suficiente para mínimo global.

2.6.4 Exemplo

A viga em balanço da Fig. 2.2, de seção transversal retangular *b* (largura) x *d* (tal que a altura não exceda 2 vezes a largura), tem vão *L* = 400/15 cm e suporta um peso *V* = 150 KN em sua extremidade livre, sem que exceda a tensão normal admissível do material $\bar{\sigma}$=1 KN/cm² e a tensão de cisalhamento admissível do material $\bar{\tau}$ =0,2 KN/cm². O volume total de material deve ser minimizado.

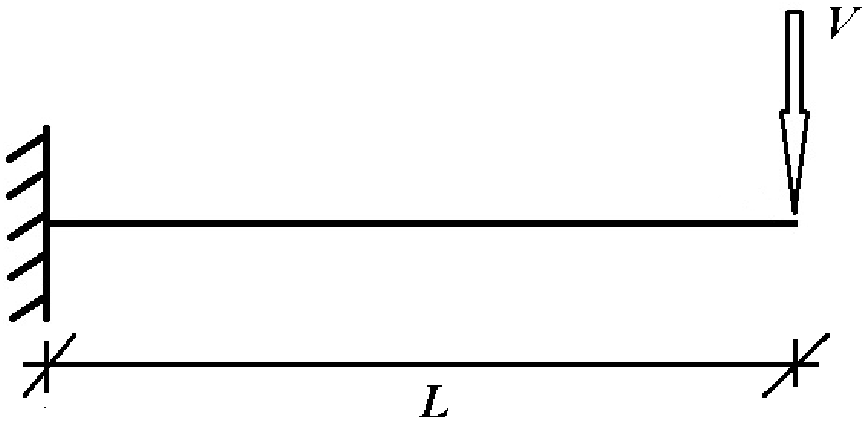

Figura 2.2 – Treliça exemplo 2.6.4

Momento fletor máximo: 150 x 400/15 = 4000 KNcm

Força cortante máxima: 150 KN

Verificação da tensão normal: $\sigma = \frac{6M}{bd^2} \leq \bar{\sigma}$

Verificação da tensão de cisalhamento: $\tau = \frac{3}{2}\frac{V}{bd} \leq \bar{\tau}$

Variáveis de projeto: $x_1 = b; \quad x_2 = d,$ em cm

Função objetivo: $f(\mathbf{x}) = x_1 x_2,$ em cm²

Sujeita a $\quad g_1(\mathbf{x}) = \frac{6x4000}{x_1 x_2^2} - 1 \leq 0; \qquad g_2(\mathbf{x}) = \frac{3x150}{2x_1 x_2} - 0,2 \leq 0;$

$$g_3(\mathbf{x}) = x_2 - 2x_1 \leq 0; \qquad g_4(\mathbf{x}) = -x_1 \leq 0; \qquad g_5(\mathbf{x}) = -x_2 \leq 0$$

Lagrangiano:

$$L = x_1 x_2 + u_1\left(\frac{24000}{x_1 x_2^2} - 1 + s_1^2\right) + u_2\left(\frac{225}{x_1 x_2} - 0,2 + s_2^2\right) + u_3(x_2 - 2x_1 + s_3^2) + u_4(-x_1 + s_4^2) + u_5(-x_2 + s_5^2)$$

Condições necessárias KKT:

$$\frac{\partial L}{\partial x_1} = 0 = x_2 - u_1\frac{24000}{x_1^2 x_2^2} - u_2\frac{225}{x_1^2 x_2} - 2u_3 - u_4$$

$$\frac{\partial L}{\partial x_2} = 0 = x_1 - u_1\frac{48000}{x_1^2 x_2^2} - u_2\frac{225}{x_1 x_2^2} + u_3 - u_5$$

As derivadas parciais do Lagrangiano com respeito aos multiplicadores de Lagrange recuperam as próprias equações de restrições.

As derivadas parciais do Lagrangiano com respeito às variáveis de folga resultam sempre as mesmas condições de chaveamento:

$$u_i s_i = 0; \quad u_i \geq 0; \quad g_i + s_i^2 = 0; \quad s_i^2 \geq 0; \quad i = 1 \text{ a } 5$$

Casos em que $s_4 = 0$ ou $s_5 = 0$ ou $s_5 = s_4 = 0$ não servem. Assim, é necessário que $u_5 = u_4 = 0$ em todos os casos.

Caso 1: $u_1 = 0; \quad u_2 = 0; \quad u_3 = 0; \quad u_4 = 0; \quad u_5 = 0;$ resultam $b = 0$ e $d = 0$, o que não é aceitável.

Caso 2: $u_1 = 0; \quad u_2 = 0; \quad s_3 = 0; \quad u_4 = 0; \quad u_5 = 0;$ resultam $b = 0$ e $d = 0$, o que não é aceitável.

Caso 3: $u_1 = 0; \quad s_2 = 0; \quad u_3 = 0; \quad u_4 = 0; \quad u_5 = 0;$ resultando infinitas soluções válidas em faixas de valores para as dimensões da seção transversal:

$23{,}717 \leq b \leq 52{,}734$ cm; $\quad 21{,}333 \leq d \leq 47{,}433$ cm, sujeito a $bd = 1125$ cm²

Caso 4: $s_1 = 0; \quad u_2 = 0; \quad u_3 = 0; \quad u_4 = 0; \quad u_5 = 0;$ não tem soluções consistentes.

Caso 5: $u_1 = 0; \quad s_2 = 0; \quad s_3 = 0; \quad u_4 = 0; \quad u_5 = 0;$ corresponde a um dos limites do caso 3, especificamente $b = 23{,}717$ cm e $d = 47{,}434$ cm.

Caso 6: $s_1 = 0; \quad s_2 = 0; \quad u_3 = 0; \quad u_4 = 0; \quad u_5 = 0;$ corresponde a um dos limites do caso 3, especificamente $b = 52{,}374$ cm e $d = 21{,}333$ cm.

Caso 7: $s_1 = 0; \quad u_2 = 0; \quad s_3 = 0; \quad u_4 = 0; \quad u_5 = 0;$ resulta em que $u_3 < 0$, o que não é válido.

Caso 8: $s_1 = 0; \quad s_2 = 0; \quad s_3 = 0; \quad u_4 = 0; \quad u_5 = 0;$ resulta um sistema de 3 equações em 2 incógnitas, que não tem solução.

2.6.5 Exemplo 5, proposto.

Determinar a massa mínima de um tripé de alumínio de altura *H* para suportar uma carga vertical *V* = 60 kN. A base é um triângulo equilátero com lados *B* = 120 cm. As barras têm seção circular maciça de diâmetro *D* e não devem ultrapassar a tensão admissível do material, nem a carga crítica de flambagem de Euler, com coeficiente de segurança 2. As faixas de valores das variáveis de projeto são $50 \text{ cm} \leq H \leq 500$ cme $0{,}5 \text{ cm} \leq D \leq 50$ cmDados: tensão admissível à compressão 150 MPa; Módulo de Elasticidade 75 GPa; densidade 2800 kg/m³.

Respostas: *H* = 50 cm; *D* = 3,42 cm; massa mínima 6,6 kg.

2.7 Funcionais e seus máximos e mínimos

Procurar-se-á, nesta apresentação, manter um certo paralelismo entre o cálculo variacional e o cálculo diferencial clássico.

Seja B um espaço vetorial de funções. Chama-se funcional a aplicação Π que associa a cada elemento f de B um único elemento y de **R**. A notação utilizada é $\Pi:\ B\ \alpha\ \mathbf{R}$, tal que se $f \in B$ então $y = \Pi(f)$.

Um funcional $\Pi:\ B\ \alpha\ \mathbf{R}$ é dito **convexo** se

$$\Pi((1-\theta) f_a + \theta f_b) \leq (1-\theta)\Pi(f_a) + \theta\Pi(f_b), \quad \forall f_a, f_b \in B, \ \forall\theta \in [0,1]. \qquad (2.36)$$

Um funcional $\Pi:\ B\ \alpha\ \mathbf{R}$ é dito **estritamente convexo** se

$$\Pi((1-\theta) f_a + \theta f_b) < (1-\theta)\Pi(f_a) + \theta\Pi(f_b), \quad \forall f_a, f_b \in B, \ \forall\theta \in (0,1). \qquad (2.37)$$

Considere $V_h(f_0) = \{\Omega \subset B \mid \forall f \in \Omega,\ d(f, f_0) < \varepsilon\}$ uma vizinhança de f_0.

Diz-se que o funcional $\Pi:\ B\ \alpha\ \mathbf{R}$ passa por um **mínimo local** em f_0 se existir uma vizinhança de f_0 na qual

$$\Pi(f) \geq \Pi(f_0), \quad \forall f \in V_h(f_0). \qquad (2.38)$$

Diz-se que este mínimo é **global** se

$$\Pi(f) \geq \Pi(f_0), \quad \forall f \in B. \qquad (2.39)$$

Diz-se que este mínimo é **estrito** se

$$\Pi(f) > \Pi(f_0), \quad \forall f \in V_h(f_0) \mid d(f, f_0) \neq 0. \qquad (2.40)$$

Diz-se que o funcional $\Pi:\ B\ \alpha\ \mathbf{R}$ passa por um **máximo local** em f_0 se existir uma vizinhança de f_0 na qual

$$\Pi(f) \leq \Pi(f_0), \quad \forall f \in V_h(f_0). \qquad (2.41)$$

Diz-se que este máximo é **global** se

$$\Pi(f) \leq \Pi(f_0), \quad \forall f \in B. \qquad (2.42)$$

Diz-se que este máximo é **estrito** se

$$\Pi(f) < \Pi(f_0), \quad \forall f \in V_h(f_0) \mid d(f, f_0) \neq 0. \qquad (2.43)$$

Observações:
Funcionais convexos possuem pelo menos um mínimo global. Quando eles são estritamente convexos este mínimo não só existe mas é único.

Diz-se também que $\Pi(f_0)$ é um **extremo** de Π e que f_0 é um **extremante**.