

Otimização ELE037

Condições de Otimalidade

Jaime A. Ramírez
Felipe Campelo
Frederico G. Guimarães
Lucas S. Batista
Ricardo H.C. Takahashi

Universidade Federal de Minas Gerais
Departamento de Engenharia Elétrica

- 1 Introdução
- 2 Caracterização de Funções
- 3 Problema Exemplo
- 4 Condições Analíticas

Tema abordado:

- Caracterização das funções objetivo e de restrições;
- Condições de otimalidade que deverão ser atendidas em $\mathbf{x}^*$:

$$\begin{aligned} \mathbf{x}^* &= \arg \min_{\mathbf{x}} f(\mathbf{x}) \\ \text{sujeito a: } &\begin{cases} g_i(\mathbf{x}) \leq 0; \quad i = 1, \dots, p \\ h_j(\mathbf{x}) = 0; \quad j = 1, \dots, q \end{cases} \end{aligned} \quad (1)$$

sendo que $\mathbf{x} \in \mathbb{R}^n$, $f(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^1$, $g(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^p$, e $h(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^q$.

Introdução

- A escolha da técnica de otimização depende da natureza das funções $f(\mathbf{x})$, $g(\mathbf{x})$, $h(\mathbf{x})$.
- Não há uma técnica de otimização que seja universal.
- Mas existem informações que guiam essa escolha.
- Questões sobre otimalidade a serem respondidas:
 - 1 Dado o funcional $f(\cdot)$, o que são os pontos de mínimo desse funcional?
 - 2 O que são os pontos de mínimo local desse funcional, se são dadas também as restrições $g_i(\mathbf{x}) \leq 0$ e $h_j(\mathbf{x}) = 0$?
 - 3 Dado um ponto $\mathbf{x} \in \mathbb{R}^n$, que tipo de testes podem ser realizados para determinar se esse ponto é ou não um ponto de mínimo de $f(\cdot)$, nos dois casos anteriores?

Caracterização de Funções

Introdução

- Quais condições de otimalidade empregar?
- Algumas caracterizações úteis:
 - 1 Função, funcional, continuidade e diferenciabilidade;
 - 2 Curvas de nível, superfície de nível, região subnível;
 - 3 Convexidade, quasi-convexidade, e não convexidade;
 - 4 Unimodalidade e multimodalidade.

Caracterização de Funções

Função e Funcional

Uma **função** é uma relação que associa de maneira única membros de um conjunto A com membros de um conjunto B .

Matematicamente:

Sejam A e B dois conjuntos com membros $a_1, \dots, a_m$ e $b_1, \dots, b_n$, respectivamente. Uma função f que associa de maneira única membros de A em B é definida como:

$$f : A \mapsto B \quad (f : \mathbb{R}^m \mapsto \mathbb{R}^n) \quad (2)$$

Caracterização de Funções

Função e Funcional

Um **funcional** é uma função que retorna um único valor, i.e., um número escalar.

Formalmente:

Se $f(\cdot)$ é um funcional então:

$$f : \mathbb{R}^n \mapsto \mathbb{R}^1 \quad (3)$$

Caracterização de Funções

Superfície de Nível e Região Subnível

Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$. A **superfície de nível** $S(f, \alpha)$, associada ao nível α , é definida como:

$$S(f, \alpha) = \{\mathbf{x} \in C \mid f(\mathbf{x}) = \alpha\} \quad (4)$$

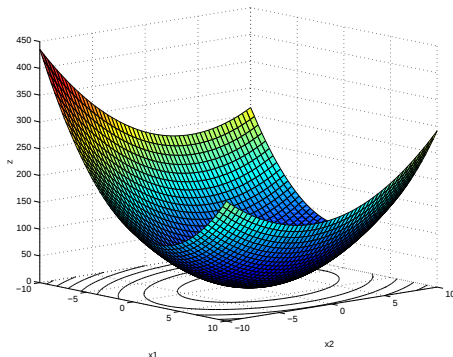
Caracterização de Funções

Superfície de Nível e Região Subnível

Ilustração do conceito de **superfície de nível**.

As curvas de nível estão representadas no plano $x_1 \times x_2$.

Cada curva contém os pontos que possuem o mesmo valor de função.

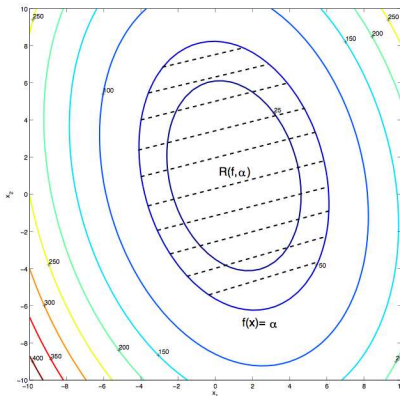


Caracterização de Funções

Superfície de Nível e Região Subnível

Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$. A **região de sub-nível** $R(f, \alpha)$, associada ao nível α , é definida como:

$$R(f, \alpha) = \{\mathbf{x} \in C \mid f(\mathbf{x}) \leq \alpha\} \quad (5)$$



Caracterização de Funções

Superfície de Nível e Região Subnível

Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$. As regiões de sub-nível dessa função obedecem a:

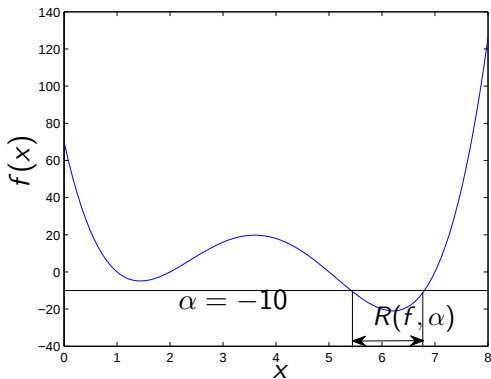
$$R(f, \alpha_1) \supset R(f, \alpha_2) \Leftrightarrow \alpha_1 > \alpha_2 \quad (6)$$

Pode-se pensar os problemas de otimização como sendo equivalentes a um problema de determinar pontos que estejam sucessivamente no interior de regiões de sub-nível cada vez menores.

Caracterização de Funções

Unimodalidade e Multimodalidade

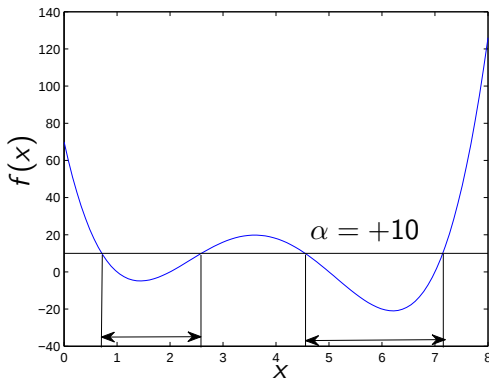
Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$. Diz-se que $f(\cdot)$ é **unimodal** se $R(f, \alpha)$ é conexo para todo $\alpha \in \mathbb{R}$.



Caracterização de Funções

Unimodalidade e Multimodalidade

Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$. Diz-se que $f(\cdot)$ é **multimodal** se existe $\alpha \in \mathbb{R}$ tal que $R(f, \alpha)$ não é conexo.



Caracterização de Funções

Unimodalidade e Multimodalidade

Note-se que **uma função unimodal pode possuir múltiplos mínimos**, desde que o conjunto deste seja conexo. Por exemplo:

$$f(x) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (7)$$

Caracterização de Funções

Bacias de Atração

Região Conexa de Sub-Nível: Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$, seja a região de sub-nível $R(f, \alpha)$, associada ao nível α , e seja um ponto $\mathbf{x}_0 \in R(f, \alpha)$. A região conexa de sub-nível $R(f, \alpha, \mathbf{x}_0)$ é definida como o maior subconjunto conexo de $R(f, \alpha)$ que contém $\mathbf{x}_0$.

Bacia de Atração: Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$, e seja $\mathbf{x}^* \in C$ um mínimo local de $f(\cdot)$. A bacia de atração de $\mathbf{x}^*$ é definida como a maior região conexa de sub-nível associada a $\mathbf{x}^*$, sendo α^* o nível correspondente, tal que a função restrita a essa região

$$f(\cdot) : R_c(f, \alpha^*, \mathbf{x}^*) \mapsto \mathbb{R} \quad (8)$$

é unimodal.

Caracterização de Funções

Continuidade e Diferenciabilidade

Uma **função contínua** é aquela para a qual uma pequena variação na entrada gera uma pequena variação no resultado da função.

Formalmente, uma função $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$ é contínua se $\forall \mathbf{x}_0 \in C$:

- 1 $f(\mathbf{x}_0)$ é definido;
- 2 $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$.

Caracterização de Funções

Continuidade e Diferenciabilidade

Função diferenciável: Uma função $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$ é diferenciável se $\forall \mathbf{x}_0 \in C$ existe o vetor gradiente:

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix} \quad (9)$$

Caracterização de Funções

Continuidade e Diferenciabilidade

Outras propriedades importantes:

1) Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$. Se $f(\cdot)$ é contínua no domínio C , então

$$\text{dist}(S(f, \alpha_1), S(f, \alpha_2)) > 0 \quad \forall (\alpha_1, \alpha_2) \mid |\alpha_1 - \alpha_2| > 0 \quad (10)$$

sendo $\text{dist}(\cdot, \cdot)$ a função distância.

2) Superfícies de nível de funções contínuas não se tocam nem se cruzam.

3) Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$. Se $f(\cdot)$ é diferenciável no domínio C , então toda superfície de nível $S(f, \alpha)$ é suave, sendo o hiperplano tangente à superfície em cada ponto perpendicular ao gradiente da função no ponto.

Caracterização de Funções

Continuidade e Diferenciabilidade

4) Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$ uma função diferenciável no domínio C , seja $\mathbf{x}_0$ um ponto pertencente à superfície de nível $S(f, \alpha)$, e seja $\nabla f(\mathbf{x}_0)$ o gradiente de $f(\cdot)$ no ponto $\mathbf{x}_0$. Seja ainda um vetor $\mathbf{d} \in \mathbb{R}^n$. Então, se

$$\mathbf{d} \cdot \nabla f(\mathbf{x}_0) < 0 \quad (11)$$

então existe $\epsilon > 0$ tal que:

$$f(\mathbf{x}_0 + \epsilon \mathbf{d}) < f(\mathbf{x}_0) \quad (12)$$

Diz-se que $\mathbf{d}$ é uma direção minimizante de $f(\cdot)$ no ponto $\mathbf{x}_0$.

Caracterização de Funções

Continuidade e Diferenciabilidade

No caso de funções não diferenciáveis, define-se,

Subgradiente: Seja $f(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}$. Um funcional linear f^{sb} é um subgradiente de $f(\cdot)$ no ponto $\mathbf{x}_0$ se:

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + f^{sb}(\mathbf{x} - \mathbf{x}_0), \quad \forall \mathbf{x} \quad (13)$$

Por exemplo, a derivada da função $f(x) = |x|$ é:

$$f'(x) = \begin{cases} 1, & x > 0 \\ -1, & x < 0 \end{cases} \quad (14)$$

No ponto $x = 0$ a derivada não é definida, entretanto pode-se definir o subgradiente como qualquer número real no intervalo $[-1, 1]$.

Exemplo em duas dimensões...

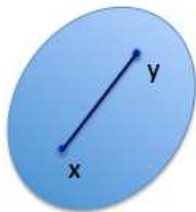
Caracterização de Funções

Convexidade e Quasi-Convexidade

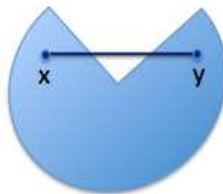
Conjunto Convexo: Diz-se que um conjunto $C \in \mathbb{R}^n$ é convexo se para quaisquer vetores $\mathbf{x}, \mathbf{y} \in C$,

$$\alpha \mathbf{x} + (1 - \alpha) \mathbf{y} \in C \quad (15)$$

para todo $\alpha \in [0, 1]$.



(a)



(b)

Figura: Representação: (a) Conjunto convexo, (b) Conjunto não convexo

Caracterização de Funções

Convexidade e Quasi-Convexidade

Função Convexa: Diz-se que uma função $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$ definida sobre um conjunto convexo C é convexa se para quaisquer $\mathbf{x}, \mathbf{y} \in C$,

$$f(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha) f(\mathbf{y}) \quad (16)$$

para todo $\alpha \in [0, 1]$. Se para quaisquer $\mathbf{x}, \mathbf{y} \in C$, sendo $\mathbf{x} \neq \mathbf{y}$ e $0 < \alpha < 1$, a desigualdade é estrita, então $f(\cdot)$ é estritamente convexa.

Caracterização de Funções

Convexidade e Quasi-Convexidade

Caracterizações de Funções Convexas: Seja $f(\cdot)$ uma função duas vezes diferenciável, sobre um conjunto convexo $C \subset \mathbb{R}^n$.

Então são equivalentes as afirmativas a seguir:

❶ $f(\alpha \mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y}) \quad \forall \alpha \in [0, 1]$

❷ $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})'(\mathbf{y} - \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in C$

❸ $H(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in C$

Caracterização de Funções

Convexidade e Quasi-Convexidade

Combinações Convexas: Sejam $f_i(\cdot) : C_i \subset \mathbb{R}^n \mapsto \mathbb{R}$ funções convexas definidas sobre conjuntos convexos C_i , $i = 1, \dots, m$. Então:

① $\alpha f_i(\cdot)$ é convexa sobre C_i , $\forall \alpha \geq 0$

② $\sum_{i=1}^m \alpha_i f_i(\cdot)$ é convexa sobre $\bigcap_{i=1}^m C_i$ para $\alpha_i \geq 0$, $i = 1, \dots, m$

Caracterização de Funções

Convexidade e Quasi-Convexidade

Função Quasi-Convexa: Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$ uma função tal que suas regiões de sub-nível $R(f, \alpha)$ são convexas para todo $\alpha \in \mathbb{R}$. Neste caso, diz-se que $f(\cdot)$ é quasi-convexa no domínio C .

Se $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$ é uma função quasi-convexa, então:

$$f(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}) \leq \max \{f(\mathbf{x}), f(\mathbf{y})\} \quad \forall \mathbf{x}, \mathbf{y} \in C, \quad \forall \alpha \in [0, 1] \quad (17)$$

Caracterização de Funções

Convexidade e Quasi-Convexidade

A convexidade de funções pode ser relacionada com as regiões de sub-nível, superfícies de nível e bacias de atração.

- 1) Todas as regiões de sub-nível de uma função convexa num domínio convexo são conjuntos convexos.
- 2) Uma função convexa em um domínio convexo possui uma única bacia de atração, a qual é um conjunto convexo.

Caracterização de Funções

Mínimo Local e Mínimo Global

Introduz-se o conceito de mínimo local como o ponto $\mathbf{x}^*$, para o qual qualquer vetor $\mathbf{x}$ na vizinhança ϵ de $\mathbf{x}^*$ implica em $f(\mathbf{x}^*) \leq f(\mathbf{x})$.

Matematicamente:

Mínimo Local: Seja $f(\cdot) : C \subset \mathbb{R}^n \mapsto \mathbb{R}$. Um ponto $\mathbf{x}^*$ é um mínimo local de $f(\cdot)$ sobre C se existe $\epsilon > 0$ tal que

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \quad , \quad \forall \mathbf{x} \in V(\mathbf{x}^*, \epsilon) \cap C \quad (18)$$

onde $V(\mathbf{x}^*, \epsilon) \triangleq \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon\}$. O ponto $\mathbf{x}^* \in C$ é um mínimo local estrito se vale a desigualdade estrita.

Caracterização de Funções

Mínimo Local e Mínimo Global

O conjunto C é o subconjunto do espaço $\mathbb{R}^n$ definido pelas restrições:

$$C \triangleq \{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \leq 0 ; i = 1, \dots, p ; h_j(\mathbf{x}) = 0 ; j = 1, \dots, q\} \quad (19)$$

Mínimo global do funcional: Se for possível escolher $\epsilon > 0$ tal que $V(\mathbf{x}^*, \epsilon) \cap C = C$, então $\mathbf{x}^*$ é um mínimo global de $f(\cdot)$ sobre C . O mínimo global é ainda estrito se a desigualdade for satisfeita de modo estrito.

Problema Exemplo

Introdução

Consideremos o problema:

$$\begin{aligned} \mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) &= (x_1 - 3)^2 + 2(x_2 - 3)^2 \\ \text{Sujeito a: } \begin{cases} g_1(x_1, x_2) : 3x_1 + 2x_2 \leq 12 \\ h_1(x_1, x_2) : x_1 + x_2 = 5 \\ 0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6 \end{cases} \end{aligned} \quad (20)$$

Problema Exemplo

Introdução

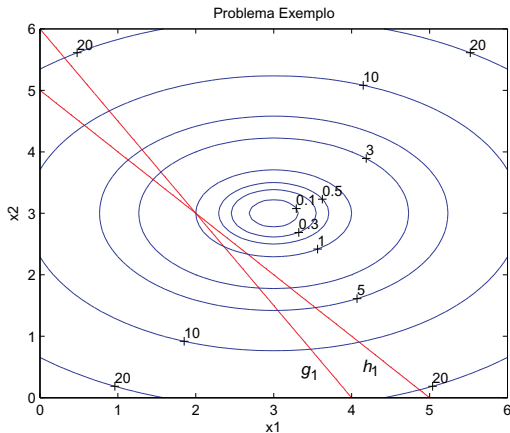


Figura: Ilustração gráfica do problema exemplo.

Problema Exemplo

Problema irrestrito

A partir de (20) pode-se definir o problema irrestrito:

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) = (x_1 - 3)^2 + 2(x_2 - 3)^2 \quad (21)$$

Sujeito a: $\{ 0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6$

- Não existe nenhuma função de restrição imposta a $f(\cdot)$.
- Os limites inferiores e superiores de $\mathbf{x}$ definem a região factível.

Problema Exemplo

Problema irrestrito

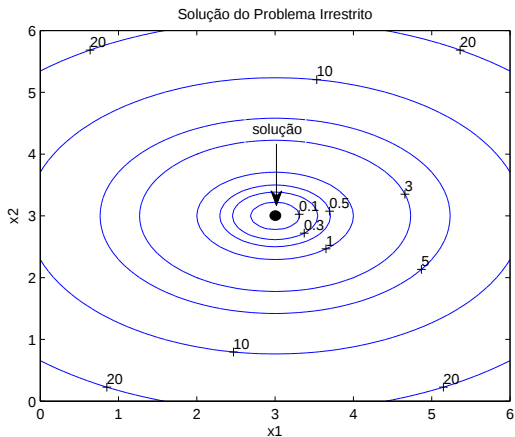


Figura: Solução gráfica do problema exemplo – irrestrito.

Problema Exemplo

Restrição de desigualdade

A partir de (20) pode-se definir o problema com restrição de desigualdade:

$$\begin{aligned} \mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) &= (x_1 - 3)^2 + 2(x_2 - 3)^2 \\ \text{sujeito a: } \begin{cases} g_1(x_1, x_2) : 3x_1 + 2x_2 \leq 12 \\ 0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6 \end{cases} \end{aligned} \quad (22)$$

- Ao incluir $g_1(\cdot) \leq 0$, a solução deverá satisfazer tal restrição.
- Por inspeção, pode-se identificar que o mínimo é o ponto (x_1^*, x_2^*) definido na curva de nível de $f(\cdot)$ que tangencia $g_1(\cdot)$.
- No ponto solução (x_1^*, x_2^*) , $\nabla f(\cdot)$ está, exatamente, no sentido oposto de $\nabla g_1(\cdot)$. Essa relação entre os gradientes é a base para estabelecer as condições de otimalidade de primeira ordem.

Problema Exemplo

Restrição de desigualdade

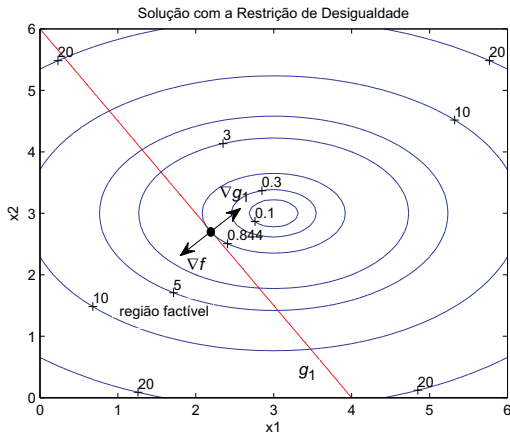


Figura: Solução gráfica do problema exemplo – restrição de desigualdade.

Problema Exemplo

Restrição de igualdade

A partir de (20), pode-se definir o problema:

$$\begin{aligned} \mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) &= (x_1 - 3)^2 + 2(x_2 - 3)^2 \\ \text{sujeito a: } \begin{cases} h_1(x_1, x_2) : x_1 + x_2 = 5 \\ 0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6 \end{cases} \end{aligned} \quad (23)$$

- Ao incluir $h_1(\cdot)$, força-se que a solução esteja sobre a reta $h_1(\cdot)$.
- Por inspeção, pode-se identificar que o mínimo é o ponto (x_1^*, x_2^*) definido na curva de nível de $f(\cdot)$ que tangencia $h_1(\cdot)$.
- Nesse ponto $\nabla f(\cdot)$ está, exatamente, no sentido oposto de $\nabla h_1(\cdot)$.

Problema Exemplo

Restrição de igualdade

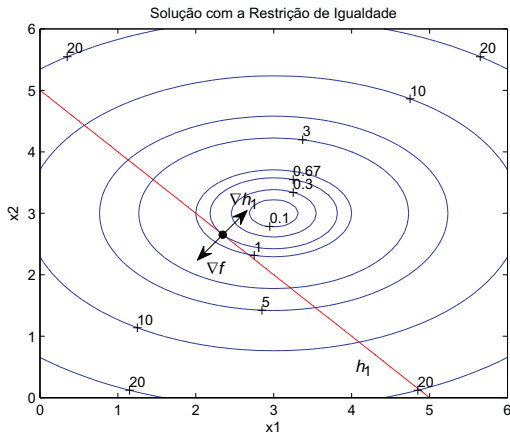


Figura: Solução gráfica do problema exemplo – restrição de igualdade.

Problema Exemplo

Restrições de desigualdade e igualdade

A partir de (20), pode-se definir o problema:

$$\begin{aligned} \mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) &= (x_1 - 3)^2 + 2(x_2 - 3)^2 \\ \text{sujeito a: } \begin{cases} g_1(x_1, x_2) : 3x_1 + 2x_2 \leq 12 \\ h_1(x_1, x_2) : x_1 + x_2 = 5 \\ 0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6 \end{cases} \end{aligned} \quad (24)$$

- A região factível deve satisfazer simultaneamente $g_1(\cdot) \leq 0$ e $h_1(\cdot) = 0$, respeitando-se os limites de x_1 e x_2 .
- Por inspeção, tem-se que o mínimo (x_1^*, x_2^*) é o ponto de interseção entre $g_1(\cdot)$ e $h_1(\cdot)$.
- Observe que, no ponto solução, o somatório dos gradientes de $f(\cdot)$, $g_1(\cdot)$ e $h_1(\cdot)$ não se anula automaticamente.

Problema Exemplo

Restrições de desigualdade e igualdade

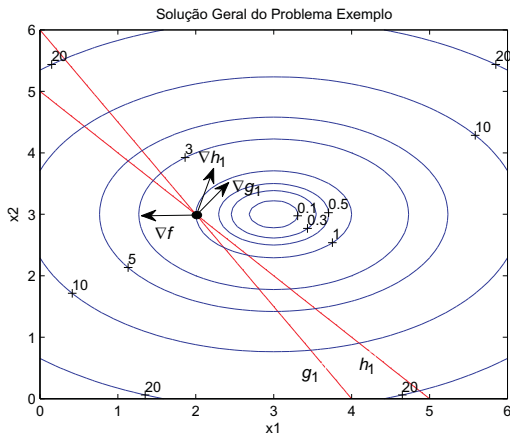


Figura: Solução gráfica do problema – restrição de desigualdade e igualdade.

Condições Analíticas

Problemas irrestritos

- Condições analíticas necessárias e suficientes;
- Permitem afirmar se a solução encontrada $\mathbf{x}$ é de fato a solução ótima;
- Essas condições serão usadas como critérios de parada e convergência.

Seja o problema irrestrito definido por:

$$\begin{aligned} \mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) &= (x_1 - 3)^2 + 2(x_2 - 3)^2 \\ \text{sujeito a: } \{ \quad &0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6 \end{aligned} \tag{25}$$

Condições Analíticas

Problemas irrestritos

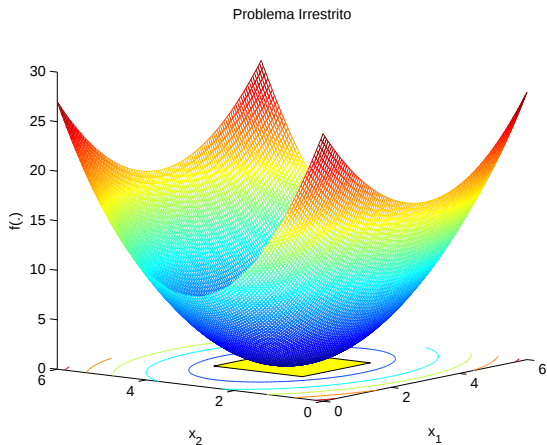


Figura: Solução gráfica 3D do problema irrestrito.

Analisando a Figura 7, observa-se:

- O mínimo ocorre em $(x_1^* = 3, x_2^* = 3)$, e nesse ponto $f(x_1, x_2) = 0$.
- Qualquer variação em $\mathbf{x}^*$ leva a um aumento de $f(\cdot)$.

Supondo a variação na vizinhança do ponto ótimo como $\Delta \mathbf{x}$, e a variação do valor ótimo da função como $\Delta f(\cdot)$:

- Fica evidente que o *mínimo* deve ser um ponto que satisfaça:

$$\Delta f > 0, \quad \forall \quad \Delta \mathbf{x} \quad (26)$$

Condições Analíticas

Problemas irrestritos: condições de 1ª ordem

- O conceito desenvolvido em (26) pode ser aplicado no limite, isto é, para incrementos infinitesimais dx_1 e dx_2 sobre $\mathbf{x}^*$.
- A função $f(\cdot)$ pode ser aproximada por um plano tangente em $\mathbf{x}^*$, por exemplo usando os primeiros termos da série de Taylor.
- Supondo esta aproximação para $f(\cdot)$, o seu valor será o mesmo em qualquer ponto deste plano (i.e., $df = 0$).
- Por outro lado, qualquer variação no plano a partir do ponto mínimo implica que dx_1 e dx_2 *não são zero* (i.e. $dx_1 \neq 0$ e $dx_2 \neq 0$).

Condições Analíticas

Problemas irrestritos: condições de 1ª ordem

Matematicamente:

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 = 0 \quad (27)$$

ou

$$df = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \right] \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} = 0 \quad (28)$$

A equação (28) deve ser satisfeita para todos os pontos do plano.

Sabendo-se que $dx_1 \neq 0$ e $dx_2 \neq 0$, obtém-se consequentemente:

$$\frac{\partial f}{\partial x_1} = 0 \quad ; \quad \frac{\partial f}{\partial x_2} = 0 \quad (29)$$

ou

$$\nabla f(x_1^*, x_2^*) = 0 \quad (30)$$

Condições Analíticas

Problemas irrestritos: condições de 1ª ordem

Condições Necessárias de 1ª Ordem: Seja $\Omega \subset \mathbb{R}^n$ e $f(\cdot)$ uma função diferenciável sobre Ω . Se $\mathbf{x}^*$ é um mínimo local de $f(\cdot)$ sobre Ω , então tem-se que:

$$\nabla f(\mathbf{x}^*) = 0 \quad (31)$$

Considerações adicionais devem ser impostas para assegurar que a solução encontrada pela condição de primeira ordem seja de fato ótima.

Condições Analíticas

Problemas irrestritos: condições de 2ª ordem

- As condições de segunda ordem são normalmente conhecidas como *condições suficientes*.
- As condições de segunda ordem são obtidas através da expansão de Taylor da função.
- Se $\mathbf{x}^*$ é ótima e $\Delta \mathbf{x}$ representa a variação no ponto solução, a qual resulta em uma variação em Δf , então:

$$\Delta f = f(\mathbf{x}^* + \Delta \mathbf{x}) - f(\mathbf{x}^*) = \nabla f(\mathbf{x}^*)^T \Delta \mathbf{x} + \frac{1}{2} \Delta \mathbf{x}^T H(\mathbf{x}^*) \Delta \mathbf{x} \quad (32)$$

Condições Analíticas

Problemas irrestritos: condições de 2ª ordem

- No ponto de ótimo $(\mathbf{x}^*)$, $\Delta f \geq 0$, $\forall \Delta \mathbf{x}$.
- Pelas condições necessárias de primeira ordem:

$$\nabla f(\mathbf{x}^*)^T \Delta \mathbf{x} = 0 \quad (33)$$

$$\Delta f = f(\mathbf{x}^* + \Delta \mathbf{x}) - f(\mathbf{x}^*) = 0.5 \Delta \mathbf{x}^T H(\mathbf{x}^*) \Delta \mathbf{x} \geq 0 \quad (34)$$

- Para que (34) seja verdadeira, a matrix $H(\mathbf{x}^*)$ deve ser positiva semidefinida.

Condições Analíticas

Problemas irrestritos: condições de 2ª ordem

Há três maneiras para determinar se H é positiva definida:

- 1 Para todos os valores possíveis de $\Delta \mathbf{x}$, $\Delta \mathbf{x}^T H(\mathbf{x}^*) \Delta \mathbf{x} > 0$.
- 2 Todos os autovalores de $H(\mathbf{x}^*)$ devem ser positivos.
- 3 Os determinantes de todas as submatrizes que envolvem a diagonal principal de $H(\mathbf{x}^*)$ devem ser positivos.

Condições Analíticas

Problemas irrestritos: condições de 2ª ordem

De forma geral, uma matriz H é:

positiva definida sss	todos os autovalores são positivos
positiva semi-definida sss	todos os autovalores são não-negativos
negativa definida sss	todos os autovalores são negativos
negativa semi-definida sss	todos os autovalores são não-positivos
indefinida sss	possui autovalores positivos e negativos

Condições Analíticas

Problemas irrestritos: condições de 2ª ordem

Classificação dos pontos estacionários de $f(x)$:

H positiva definida	ponto de mínimo local
H negativa definida	ponto de máximo local
H indefinida	ponto de sela
H semi-definida	inconclusivo (ponto de sela ou extremo local)

Condições Analíticas

Problemas irrestritos: condições de 2ª ordem

Seja o problema de minimização definido por:

$$\begin{aligned} \mathbf{x}^* &= \arg \min_{\mathbf{x}} f(\mathbf{x}) = (x_1 - 3)^2 + 2(x_2 - 3)^2 \\ \text{sujeito a: } &\{ 0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6 \end{aligned} \quad (35)$$

As condições necessárias de primeira ordem requerem:

$$\frac{\partial f(\cdot)}{\partial x_1} = 2(x_1 - 3) = 0 \quad \text{e} \quad \frac{\partial f(\cdot)}{\partial x_2} = 4(x_2 - 3) = 0 \quad (36)$$

- Resultando na solução $x_1^* = 3$ e $x_2^* = 3$.
- Este ponto é de mínimo, máximo ou inflexão?

Condições Analíticas

Problemas irrestritos: condições de 2ª ordem

As condições de segunda ordem requerem que a matriz Hessiana seja positiva definida:

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \quad (37)$$

- Se verdadeiro, a solução encontrada é de fato o ponto de mínimo da função.

Condições Analíticas

Problemas irrestritos: condições de 2ª ordem

Condições Necessárias de 2a Ordem: Seja $\Omega \subset \mathbb{R}^n$ e $f(\cdot)$ uma função duas vezes diferenciável sobre Ω . Se $\mathbf{x}^*$ é um mínimo local de $f(\cdot)$ sobre Ω , então tem-se que:

1 $\nabla f(\mathbf{x}^*) = 0$

2 $H(\mathbf{x}^*) \geq 0$

Condições Analíticas

Problemas irrestritos: condições de 1ª e 2ª ordem

- Assim, conclui-se a dedução analítica das **condições necessárias e suficientes** que um ponto deve satisfazer para ser considerado um mínimo de uma função $f(\cdot)$ sem restrições.

Otimização Irrestrita: Seja $f(\cdot) \in C^2$ e $\mathbf{x}^* \in \mathbb{R}^n$. Se forem simultaneamente satisfeitas:

① $\nabla f(\mathbf{x}^*) = 0$

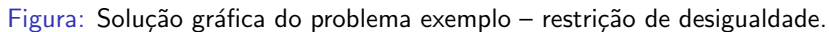
② $H(\mathbf{x}^*) > 0$

então $\mathbf{x}^*$ é um mínimo local estrito de $f(\cdot)$ sobre $\mathbb{R}^n$.

O problema sujeito à restrição de desigualdade é definido como:

$$\begin{aligned} \mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) &= (x_1 - 3)^2 + 2(x_2 - 3)^2 \\ \text{sujeito a: } \begin{cases} g_1(\mathbf{x}) : 3x_1 + 2x_2 \leq 12 \\ 0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6 \end{cases} \end{aligned} \quad (38)$$

Problemas com restrição de desigualdade



Condições Analíticas

Problemas com restrição de desigualdade

- No ponto solução, o vetor $\nabla f(\cdot)$ está na mesma direção e no sentido oposto do vetor $\nabla g_1(\cdot)$.
- Essa relação entre $\nabla f(\cdot)$ e $\nabla g_1(\cdot)$ só é possível no ponto solução.
- Existe no ponto solução uma **relação proporcional** entre $\nabla f(\cdot)$ e $\nabla g_1(\cdot)$.
- Representando a constante de proporcionalidade por β_1 , pode-se expressar a relação entre os gradientes por:

$$\nabla f(\cdot) = -\beta_1 \nabla g_1(\cdot) \quad \text{ou} \quad \nabla f(\cdot) + \beta_1 \nabla g_1(\cdot) = 0 \quad (39)$$

Condições Analíticas

Problemas com restrição de desigualdade: condições de 1a ordem

Método de Lagrange (transforma o problema original restrito em irrestrito):

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) \quad (40)$$

$$\text{sujeito a: } \begin{cases} g_1(\mathbf{x}) \leq 0 \\ x_1^{\min} \leq x_1 \leq x_1^{\max} \text{ e } x_2^{\min} \leq x_2 \leq x_2^{\max} \end{cases}$$

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \bar{f}(\mathbf{x}, \beta_1, z_1^2) = f(\mathbf{x}) + \beta_1[g_1(\mathbf{x}) + z_1^2] \quad (41)$$

$$\text{sujeito a: } \begin{cases} g_1(\mathbf{x}) + z_1^2 = 0 \\ x_1^{\min} \leq x_1 \leq x_1^{\max} \text{ e } x_2^{\min} \leq x_2 \leq x_2^{\max} \end{cases}$$

Condições Analíticas

Problemas com restrição de desigualdade: condições de 1a ordem

Considerando o Lagrangeano, $\bar{f}(\mathbf{x}, \beta_1, z_1^2)$, como uma função irrestrita, as condições de primeira ordem tem que ser satisfeitas, ou seja:

$$\frac{\partial \bar{f}(\cdot)}{\partial x_1} = \frac{\partial f(\cdot)}{\partial x_1} + \beta_1 \frac{\partial g_1(\cdot)}{\partial x_1} = 0 \quad (42)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial x_2} = \frac{\partial f(\cdot)}{\partial x_2} + \beta_1 \frac{\partial g_1(\cdot)}{\partial x_2} = 0 \quad (43)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial z_1} = 2\beta_1 z_1 = 0 \quad (44)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial \beta_1} = g_1(\cdot) + z_1^2 = 0 \quad (45)$$

Condições Analíticas

Problemas com restrição de desigualdade: condições de 1a ordem

As equações (44) e (45) podem ser combinadas para eliminar a variável de folga z_1 . Assim:

$$\frac{\partial \bar{f}(\cdot)}{\partial x_1} = \frac{\partial f(\cdot)}{\partial x_1} + \beta_1 \frac{\partial g_1(\cdot)}{\partial x_1} = 0 \quad (46)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial x_2} = \frac{\partial f(\cdot)}{\partial x_2} + \beta_1 \frac{\partial g_1(\cdot)}{\partial x_2} = 0 \quad (47)$$

$$\beta_1 g_1(\cdot) = 0 \quad (48)$$

Condições Analíticas

Problemas com restrição de desigualdade: condições de 1a ordem

Em linhas gerais, **as condições de primeira ordem para um problema com restrição de desigualdade** são:

$$\nabla f(\cdot) + \beta_1 \nabla g_1(\cdot) = 0 \quad (49)$$

$$\beta_1 g_1(\cdot) = 0 \quad (50)$$

A generalização para p restrições de desigualdade será apresentada no final da unidade.

Condições Analíticas

Problemas com restrição de desigualdade: condições de 1a ordem

Substituindo os valores do problema original no sistema de equações (46)–(48), obtém-se:

$$\frac{\partial \bar{f}(\cdot)}{\partial x_1} = 2x_1 - 6 + 3\beta_1 = 0 \quad (51)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial x_2} = 4x_2 - 12 + 2\beta_1 = 0 \quad (52)$$

$$\beta_1 g_1(\cdot) = \beta_1(3x_1 + 2x_2 - 12) = 0 \quad (53)$$

Condições Analíticas

Problemas com restrição de desigualdade: condições de 1a ordem

O sistema de equações simultâneas (51)–(53) requer que as condições no multiplicador β_1 e na restrição $g_1(\cdot)$ sejam satisfeitas simultaneamente.

❶ Caso a: $\beta_1 = 0$ [$g_1 < 0$]

❷ Caso b: $\beta_1 \neq 0$ [$g_1 = 0$]

Condições Analíticas

Problemas com restrição de desigualdade: condições de 1a ordem

- A solução do sistema de equações (51)–(53) é inviável no caso a.
- Considerando o caso b, obtém-se $x_1^* = 2.18$, $x_2^* = 2.73$ e $\beta_1^* = 0.55$.
- A solução analítica encontrada usando o multiplicador de Lagrange indica que é necessário multiplicar o $\nabla g_1(\cdot)$ por $\beta_1 = 0.55$ para que o somatório de $\nabla f(\cdot)$ e $\nabla g_1(\cdot)$ seja zero no ponto solução.

Condições Analíticas

Problemas com restrição de desigualdade: condições de 1a ordem

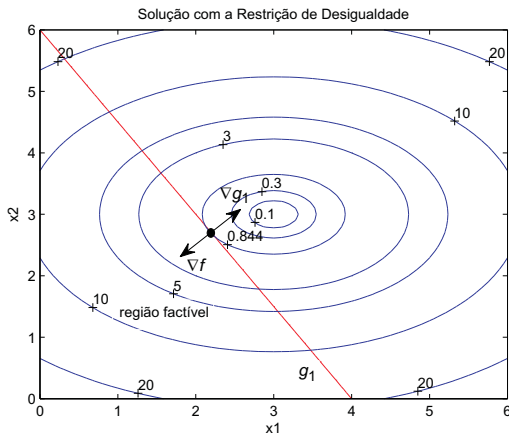


Figura: Solução gráfica do problema exemplo – restrição de desigualdade.

Condições Analíticas

Problemas com restrição de desigualdade: condições de 2a ordem

Condições de segunda ordem:

$$\Delta \bar{f}(\cdot) = \bar{f}(\mathbf{x}^* + \Delta \mathbf{x}) - \bar{f}(\mathbf{x}^*) = \nabla \bar{f}(\mathbf{x}^*)^T \Delta \mathbf{x} + \frac{1}{2} \Delta \mathbf{x}^T [\bar{H}(\mathbf{x}^*)] \Delta \mathbf{x} > 0 \quad (54)$$

$$\nabla g_1(\cdot)^T \Delta \mathbf{x} = 0 \quad (55)$$

Condições Analíticas

Problemas com restrição de desigualdade: condições de 2a ordem

Considerando um problema com duas variáveis e uma restrição de desigualdade, tem-se $(\nabla \bar{f}(\mathbf{x}^*) = 0)$:

$$\Delta \bar{f}(\cdot) = \frac{1}{2} \left[\frac{\partial^2 \bar{f}(\cdot)}{\partial x_1^2} (\Delta x_1)^2 + 2 \frac{\partial^2 \bar{f}(\cdot)}{\partial x_1 \partial x_2} (\Delta x_1)(\Delta x_2) + \frac{\partial^2 \bar{f}(\cdot)}{\partial x_2^2} (\Delta x_2)^2 \right] > 0 \quad (56)$$

Rearranjando os termos de resulta em:

$$\Delta \bar{f}(\cdot) = \frac{1}{2} \left[\frac{\partial^2 \bar{f}(\cdot)}{\partial x_1^2} \left(\frac{\Delta x_1}{\Delta x_2} \right)^2 + 2 \frac{\partial^2 \bar{f}(\cdot)}{\partial x_1 \partial x_2} \left(\frac{\Delta x_1}{\Delta x_2} \right) + \frac{\partial^2 \bar{f}(\cdot)}{\partial x_2^2} \right] (\Delta x_2)^2 > 0 \quad (57)$$

Condições Analíticas

Problemas com restrição de desigualdade: condições de 2a ordem

A equação (55) pode ser expressa por:

$$\frac{\Delta x_1}{\Delta x_2} = -\frac{\partial g_1(\cdot)/\partial x_2}{\partial g_1(\cdot)/\partial x_1} \quad (58)$$

Substituindo a equação (58) na equação (57) obtém-se:

$$\Delta \bar{f}(\cdot) = \frac{1}{2} \left[\frac{\partial^2 \bar{f}(\cdot)}{\partial x_1^2} \left(\frac{\partial g_1(\cdot)/\partial x_2}{\partial g_1(\cdot)/\partial x_1} \right)^2 - 2 \frac{\partial^2 \bar{f}(\cdot)}{\partial x_1 \partial x_2} \left(\frac{\partial g_1(\cdot)/\partial x_2}{\partial g_1(\cdot)/\partial x_1} \right) + \frac{\partial^2 \bar{f}(\cdot)}{\partial x_2^2} \right] (\Delta x_2)^2 \quad (59)$$

que representa a condição de segunda ordem, ou condição suficiente.

No problema exemplo, $[\cdot] = 44/9 > 0$.

Condições Analíticas

Problemas com restrição de desigualdade: condições de 1a e 2a ordem

Otimização Restrita – Desigualdade: Sejam $f(\cdot) \in C^2$ e $g_i(\cdot) \in C^2$, $i = 1, \dots, p$, e $\mathbf{x}^*$ tal que $g_i(\mathbf{x}^*) \leq 0$, $i = 1, \dots, p$. Se existem multiplicadores $\beta_1, \beta_2, \dots, \beta_p$ tais que

❶ $\beta_i \geq 0$, $i = 1, \dots, p$

❷ $\beta_i g_i(\mathbf{x}^*) = 0$, $i = 1, \dots, p$

❸ $\nabla f(\mathbf{x}^*) + \sum_{i=1}^p \beta_i \nabla g_i(\mathbf{x}^*) = 0$

❹ $H(\mathbf{x}^*) + \sum_{i=1}^p \beta_i G_i(\mathbf{x}^*) > 0$ sobre

$$M = \{\mathbf{y} \in \mathbb{R}^n : \nabla g_i(\mathbf{x}^*)' \mathbf{y} = 0, i \in I(\mathbf{x}^*)\},$$

$$I(\mathbf{x}^*) = \{i : g_i(\mathbf{x}^*) = 0, \beta_i > 0\}$$

são simultaneamente satisfeitos, então $\mathbf{x}^*$ é um mínimo local estrito de $f(\cdot)$ sobre $g_i(\mathbf{x}) \leq 0$, $i = 1, \dots, p$.

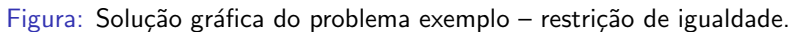
Condições Analíticas

Problemas com restrição de igualdade

O problema sujeito à restrição de igualdade pode ser definido como:

$$\begin{aligned} \mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) &= (x_1 - 3)^2 + 2(x_2 - 3)^2 \\ \text{sujeito a: } \left\{ \begin{array}{l} h_1(\mathbf{x}) : x_1 + x_2 = 5 \\ 0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6 \end{array} \right. & \end{aligned} \quad (60)$$

Problemas com restrição de igualdade



Condições Analíticas

Problemas com restrição de igualdade

- No ponto solução, $\nabla f(\cdot)$ está na mesma direção e no sentido oposto do vetor $\nabla h(\cdot)$.
- Examinando outros pontos factíveis, pode-se afirmar que essa relação entre $\nabla f(\cdot)$ e $\nabla h(\cdot)$ só é possível em $\mathbf{x}^*$.
- Existe em $\mathbf{x}^*$ uma relação proporcional entre $\nabla f(\cdot)$ e $\nabla h(\cdot)$.
- Representando a **constante de proporcionalidade** por λ_1 :

$$\nabla f(\cdot) = -\lambda_1 \nabla h_1(\cdot) \quad \text{ou} \quad \nabla f(\cdot) + \lambda_1 \nabla h_1(\cdot) = 0 \quad (61)$$

Condições Analíticas

Problemas com restrição de igualdade

- A equação (61) pode ser obtida usando o método de Lagrange.
- No método de Lagrange o problema original (60) é transformado com a introdução de uma função Lagrangeana $\bar{f}$.

$$\text{minimize } \bar{f}(x_1, x_2, \lambda_1) = f(x_1, x_2) + \lambda_1 h_1(x_1, x_2) \quad (62)$$

- Neste caso, o problema restrito passa a ser expresso por:

$$\text{minimize } \bar{f}(\mathbf{x}, \lambda_1) = (x_1 - 3)^2 + 2(x_2 - 3)^2 + \lambda_1(x_1 + x_2 - 5)$$

$$\text{sujeito a: } \begin{cases} h_1(x_1, x_2) : x_1 + x_2 = 5 \\ 0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6 \end{cases} \quad (63)$$

sendo λ_1 um multiplicador de Lagrange.

Condições Analíticas

Problemas com restrição de igualdade: condições de 1ª ordem

- As condições necessárias de primeira ordem são obtidas considerando $\bar{f}(\cdot)$ como uma função irrestrita das variáveis x_1 , x_2 e λ_1 .

$$\frac{\partial \bar{f}(\cdot)}{\partial x_1} = \frac{\partial f(\cdot)}{\partial x_1} + \lambda_1 \frac{\partial h_1(\cdot)}{\partial x_1} = 0 \quad (64)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial x_2} = \frac{\partial f(\cdot)}{\partial x_2} + \lambda_1 \frac{\partial h_1(\cdot)}{\partial x_2} = 0 \quad (65)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial \lambda_1} = h_1(\cdot) = 0 \quad (66)$$

Condições Analíticas

Problemas com restrição de igualdade: condições de 1ª ordem

Em linhas gerais:

$$\nabla f(\cdot) + \lambda_1 \nabla h_1(\cdot) = 0 \quad (67)$$

$$h_1(\cdot) = 0 \quad (68)$$

A generalização para q restrições de igualdade é direta.

Condições Analíticas

Problemas com restrição de igualdade: condições de 1ª ordem

- Aplicando as condições de 1ª ordem ao problema (63), obtém-se:

$$\frac{\partial \bar{f}(\cdot)}{\partial x_1} = 2x_1 - 6 + \lambda_1 = 0 \quad (69)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial x_2} = 4x_2 - 12 + \lambda_1 = 0 \quad (70)$$

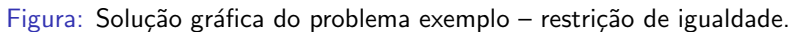
$$h_1(\cdot) = x_1 + x_2 = 5 \quad (71)$$

Condições Analíticas

Problemas com restrição de igualdade: condições de 1ª ordem

- A solução do sistema de equações (69)–(71) é $x_1^* = 2,33$, $x_2^* = 2,67$ e $\lambda_1^* = 1,34$.
- A solução analítica encontrada coincide com a solução geométrica obtida por inspeção.
- A verificação das condições de 2ª ordem, ou suficientes, é deixada para o leitor.

Problemas com restrição de igualdade: condições de 1ª ordem



Condições Analíticas

Problemas com restrição de igualdade: condições de 1ª e 2ª ordem

Otimização Restrita – Igualdade: Sejam $f(\cdot) \in C^2$ e $h_j(\cdot) \in C^2$, $j = 1, \dots, q$ e $\mathbf{x}^*$ tal que $h_j(\mathbf{x}^*) = 0$, $j = 1, \dots, q$. Se existem multiplicadores $\lambda_1, \lambda_2, \dots, \lambda_q$ tais que

$$\textcircled{1} \quad \nabla f(\mathbf{x}^*) + \sum_{j=1}^q \lambda_j \nabla h_j(\mathbf{x}^*) = 0$$

$$\textcircled{2} \quad H(\mathbf{x}^*) + \sum_{j=1}^q \lambda_j H_j(\mathbf{x}^*) > 0 \text{ sobre}$$

$$M = \{\mathbf{y} \in \mathbb{R}^n : \nabla h_j(\mathbf{x}^*)' \mathbf{y} = 0, j = 1, \dots, q\}$$

são simultaneamente satisfeitos, então $\mathbf{x}^*$ é um mínimo local estrito de $f(\cdot)$ sujeito a $h_j(\mathbf{x}) = 0$, $j = 1, \dots, q$.

O **problema geral de otimização** é definido incluindo simultaneamente as restrições de desigualdade e igualdade:

$$\begin{aligned} \mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x}) &= (x_1 - 3)^2 + 2(x_2 - 3)^2 \\ \text{sujeito a: } \left\{ \begin{array}{l} g_1(\mathbf{x}) : 3x_1 + 2x_2 \leq 12 \\ h_1(\mathbf{x}) : x_1 + x_2 = 5 \\ 0 \leq x_1 \leq 6; \quad 0 \leq x_2 \leq 6 \end{array} \right. & \quad (72) \end{aligned}$$

Usando o método de multiplicadores de Lagrange:

$$\begin{aligned} \mathbf{x}^* = \arg \min_{\mathbf{x}} \bar{f}(\mathbf{x}, \beta_1, \lambda_1, z_1^2) &= f(\mathbf{x}) + \beta_1[g_1(\mathbf{x}) + z_1^2] + \lambda_1 h_1(\mathbf{x}) \\ \text{sujeito a: } \begin{cases} g_1(\mathbf{x}) + z_1^2 = 0 \\ h_1(\mathbf{x}) = 0 \\ x_1^{\min} \leq x_1 \leq x_1^{\max} \text{ e } x_2^{\min} \leq x_2 \leq x_2^{\max} \end{cases} \end{aligned} \quad (73)$$

Condições Analíticas

O Problema Geral de Otimização

Considerando a função Lagrangeana $\bar{f}(\cdot)$ como uma função irrestrita, as **condições de primeira ordem** para este caso podem ser obtidas por:

$$\frac{\partial \bar{f}(\cdot)}{\partial x_1} = \frac{\partial f(\cdot)}{\partial x_1} + \beta_1 \frac{\partial g_1(\cdot)}{\partial x_1} + \lambda_1 \frac{\partial h_1(\cdot)}{\partial x_1} = 0 \quad (74)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial x_2} = \frac{\partial f(\cdot)}{\partial x_2} + \beta_1 \frac{\partial g_1(\cdot)}{\partial x_2} + \lambda_1 \frac{\partial h_1(\cdot)}{\partial x_2} = 0 \quad (75)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial z_1} = 2\beta_1 z_1 = 0 \quad (76)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial \beta_1} = g_1(\cdot) + z_1^2 = 0 \quad (77)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial \lambda_1} = h_1(\cdot) = 0 \quad (78)$$

Condições Analíticas

O Problema Geral de Otimização

As equações (74) a (78) podem ser reduzidas a quatro expressões:

$$\frac{\partial \bar{f}(\cdot)}{\partial x_1} = \frac{\partial f(\cdot)}{\partial x_1} + \beta_1 \frac{\partial g_1(\cdot)}{\partial x_1} + \lambda_1 \frac{\partial h_1(\cdot)}{\partial x_1} = 0 \quad (79)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial x_2} = \frac{\partial f(\cdot)}{\partial x_2} + \beta_1 \frac{\partial g_1(\cdot)}{\partial x_2} + \lambda_1 \frac{\partial h_1(\cdot)}{\partial x_2} = 0 \quad (80)$$

$$\beta_1 g_1(\cdot) = 0 \quad (81)$$

$$h_1(\cdot) = 0 \quad (82)$$

A equação (81) requer que duas possibilidades sejam testadas:

❶ Caso a: $\beta_1 = 0$ [$g_1 < 0$]

❷ Caso b: $\beta_1 \neq 0$ [$g_1 = 0$]

As condições de 1a ordem podem ser resumidas da seguinte maneira:

$$\nabla f(\cdot) + \beta_1 \nabla g_1(\cdot) + \lambda_1 \nabla h_1(\cdot) = 0 \quad (83)$$

$$\beta_1 g_1(\cdot) = 0 \quad (84)$$

$$h_1(\cdot) = 0 \quad (85)$$

A generalização para p restrições de desigualdade e q restrições de igualdade é direta.

No problema exemplo, as **condições de primeira ordem** resultam em:

$$\frac{\partial \bar{f}(\cdot)}{\partial x_1} = 2x_1 - 6 + \lambda_1 + 3\beta_1 = 0 \quad (86)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial x_2} = 4x_2 - 12 + \lambda_1 + 2\beta_1 = 0 \quad (87)$$

$$\beta_1 g_1(\cdot) = \beta_1(3x_1 + 2x_2 - 12) = 0 \quad (88)$$

$$\frac{\partial \bar{f}(\cdot)}{\partial \lambda_1} = x_1 + x_2 - 5 = 0 \quad (89)$$

Condições Analíticas

O Problema Geral de Otimização

Duas soluções devem ser examinadas:

- 1 Caso a: ($\beta_1 = 0$ e $[g_1 < 0]$). Obtém-se $x_1 = 7/3$, $x_2 = 8/3$ e $\lambda_1 = 4/3$, porém $g_1(\cdot) = 1/3 > 0$.
- 2 Caso b: ($\beta_1 \neq 0$ [$g_1 = 0$]). Obtém-se $x_1 = 2$, $x_2 = 3$, $\beta_1 = 2$ e $\lambda_1 = -4$.

Observe que no ponto solução ($x_1 = 2$; $x_2 = 3$), verifica-se que $\nabla f(\cdot) + \beta_1 g_1(\cdot) + \lambda_1 h_1(\cdot) = 0$.

Condições Analíticas

O Problema Geral de Otimização

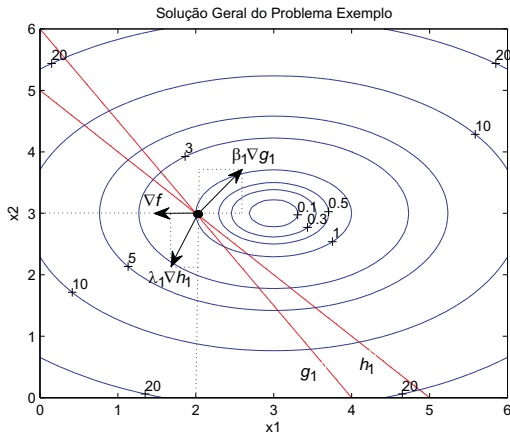


Figura: Solução analítica do caso geral.

Condições Necessárias de Karush-Kuhn-Tucker: Seja $\mathbf{x}^*$ um ponto regular das restrições do problema de otimização:

$$\begin{aligned} \mathbf{x}^* &= \arg \min_{\mathbf{x}} f(\mathbf{x}) \\ \text{sujeito a } &\begin{cases} g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, p \\ h_j(\mathbf{x}) = 0, \quad j = 1, \dots, q \end{cases} \end{aligned} \quad (90)$$

sendo que $f(\cdot), g(\cdot), h(\cdot) \in C^1$.

Condições Analíticas

Condições de Karush-Kuhn-Tucker

Para $\mathbf{x}^*$ ser um ótimo local do problema, deve existir um conjunto de multiplicadores de KKT $\beta_i^* \in \mathbb{R}^p$ com $\beta_i^* \geq 0$ e $\lambda_j^* \in \mathbb{R}^q$ tal que:

$$\nabla f(\mathbf{x}^*) + \sum_{j=1}^q \lambda_j^* \nabla h_j(\mathbf{x}^*) + \sum_{i=1}^p \beta_i^* \nabla g_i(\mathbf{x}^*) = 0$$

$$\beta_i^* g_i(\mathbf{x}^*) = 0 \quad \text{e} \quad \beta_i \geq 0 \quad \forall i = 1, \dots, p \quad (91)$$

$$h_j(\mathbf{x}) = 0 \quad \forall j = 1, \dots, q$$