EEL 5102-47: Métodos Numéricos de Otimização I

# Teoria de Otimização com Restrições

Prof.: Erlon Cristian Finardi, D. Eng.
erlon.finardi@ufsc.br

LabPlan

Laboratório de Planejamento de Sistemas de Energia Elétrica
Centro Tecnológico – Departamento de Engenharia Elétrica
Tel. +55 (48) 3721.9731/9933 – Fax +55 (48) 3721.7538
Homepage: htto://www.labplan.ufsc.br

# Formulação Geral

$$\min_{x \in \Re^n} \; f(x)$$

$$\text{sujeito a:} \begin{cases} c_i(x) = 0, & i \in \varepsilon \\ c_i(x) \geq 0, & i \in I \end{cases}$$

- $f$ e $c_i$ são funções suaves definidas em um subconjunto que pertence ao $\Re^n$
- $f$ é a função objetivo, enquanto $c_i$, $i \in \varepsilon$ são as restrições de igualdade e $c_i$, $i \in I$ são as restrições de desigualdade
- Abaixo é definido um conjunto viável $\Omega$ que contém todos os pontos $x$ que atendem as restrições

$$\Omega = \{x \mid c_i(x) = 0, \; i \in \varepsilon; \quad c_i(x) \geq 0, \; i \in I\}$$

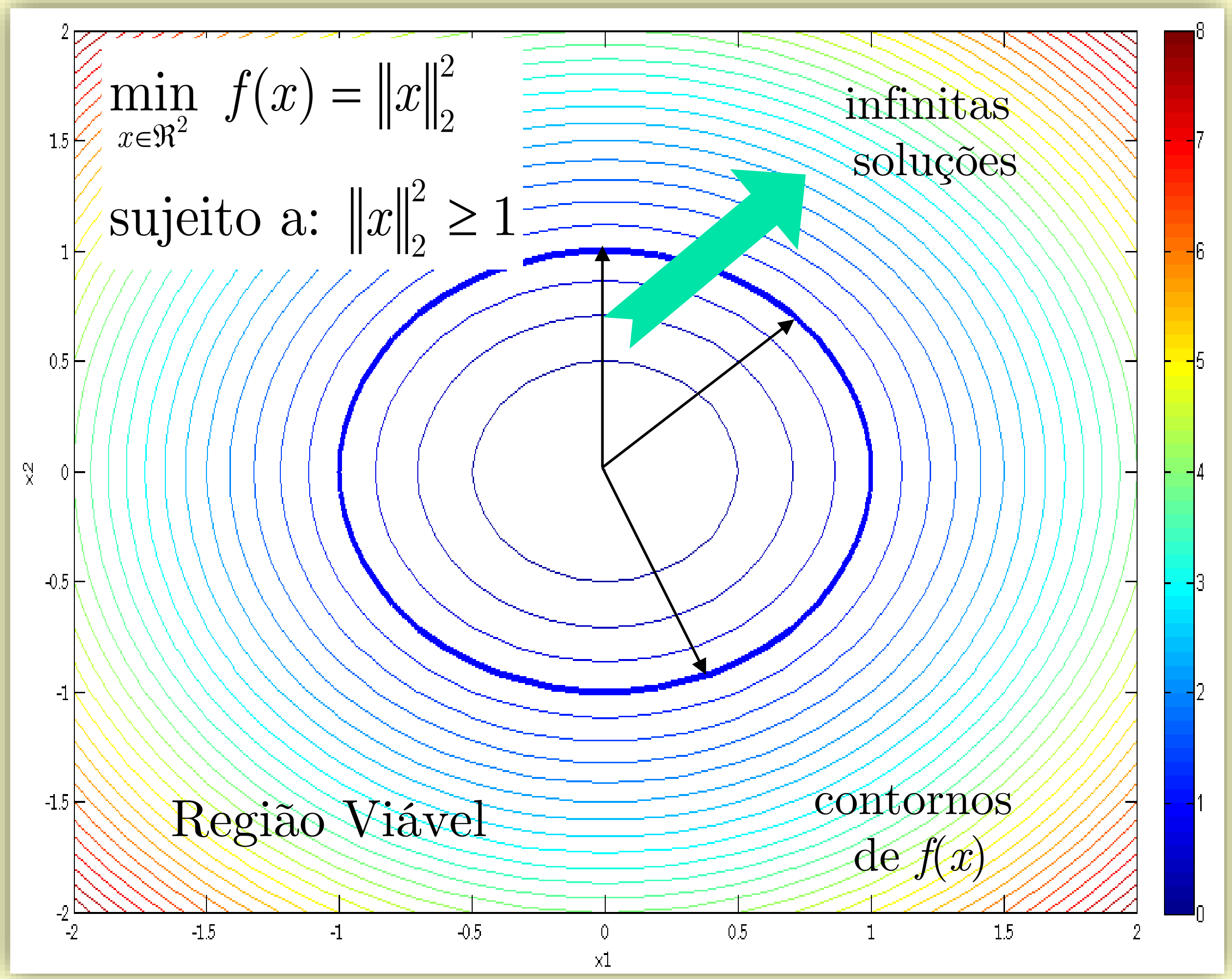
- Deste modo é possível escrever o problema acima na forma compacta

$$\min_{x \in \Omega} \; f(x)$$

# Caracterizando as Soluções

- Encontrar o mínimo global é uma tarefa árdua em problemas irrestritos
- Adição das restrições pode excluir muitos mínimos locais e facilitar a busca pelo global
- Contudo, as restrições também podem tornar essa tarefa mais complicada
- As restrições podem produzir um grande número de soluções que não formam um conjunto fechado
- Embora, no caso geral, as restrições alterem a solução, em alguns casos elas não tem efeito na determinação de um mínimo

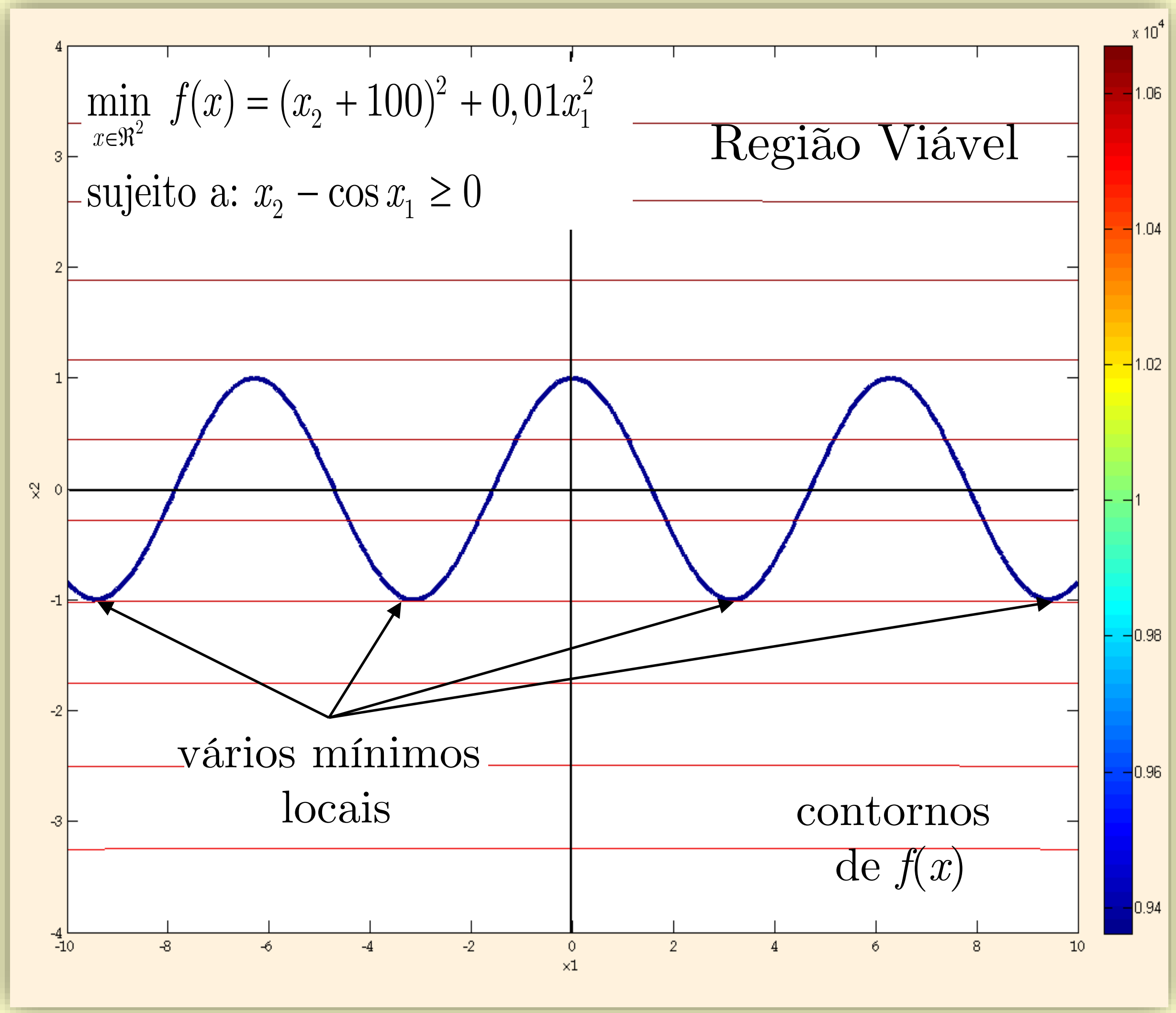
# Vários Mínimos Locais




- Sem a restrição é um problema quadrático estritamente convexo com um único minimizador em $x = 0$

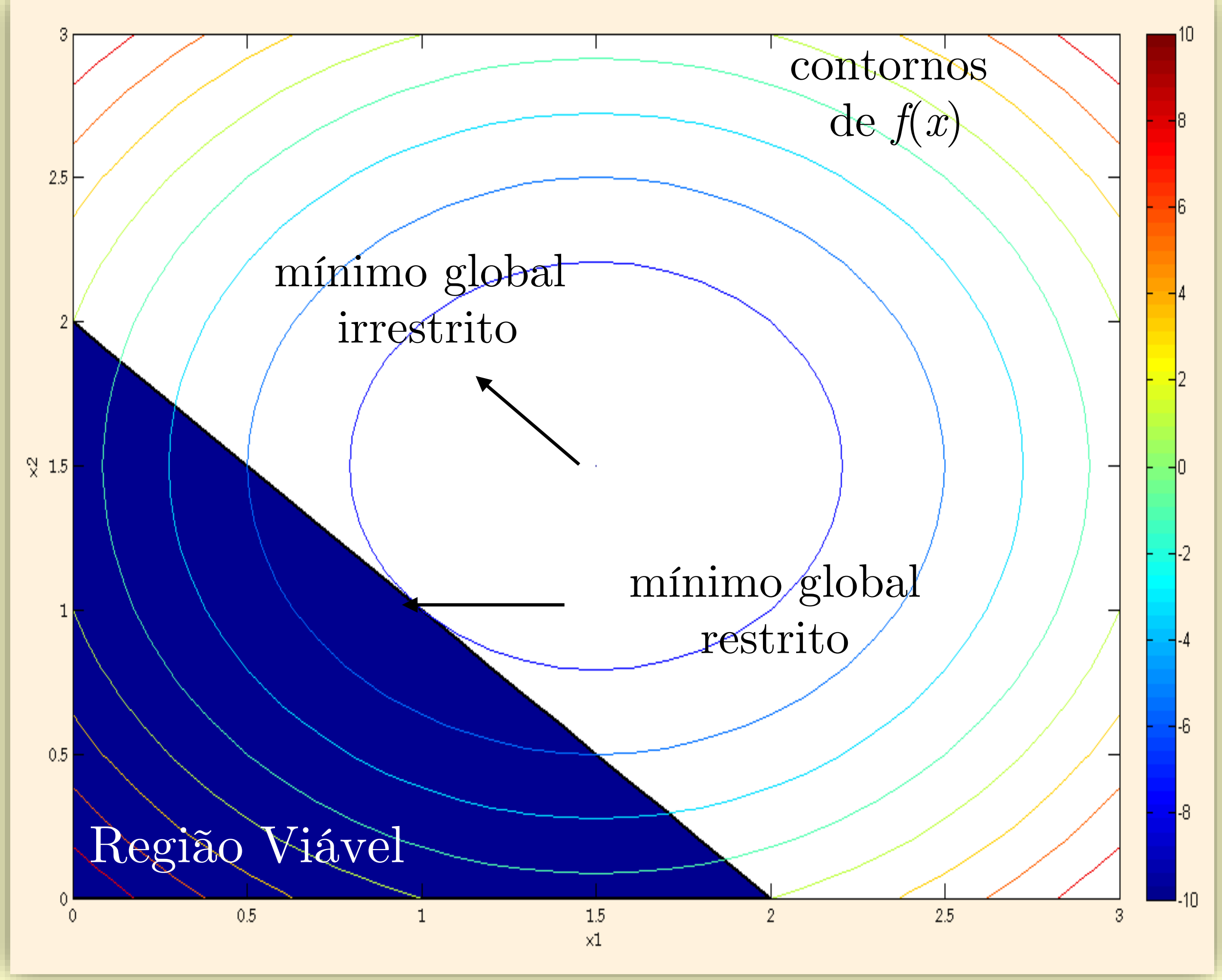
- Com a restrição, $\forall x$ tal $||x||_2 = 1$ resolve o problema existindo, portanto, infinitos mínimos globais

# Vários Mínimos Locais




- Sem a restrição o único minimizador em $x = (0,-100)$
- Com a restrição, $\forall$ vetor $(x_1,x_2) = (k\pi,\text{-}1)$ tal que $k=\pm1,\ \pm3,\quad 5,...$ resolve o problema

# Único Minimizador

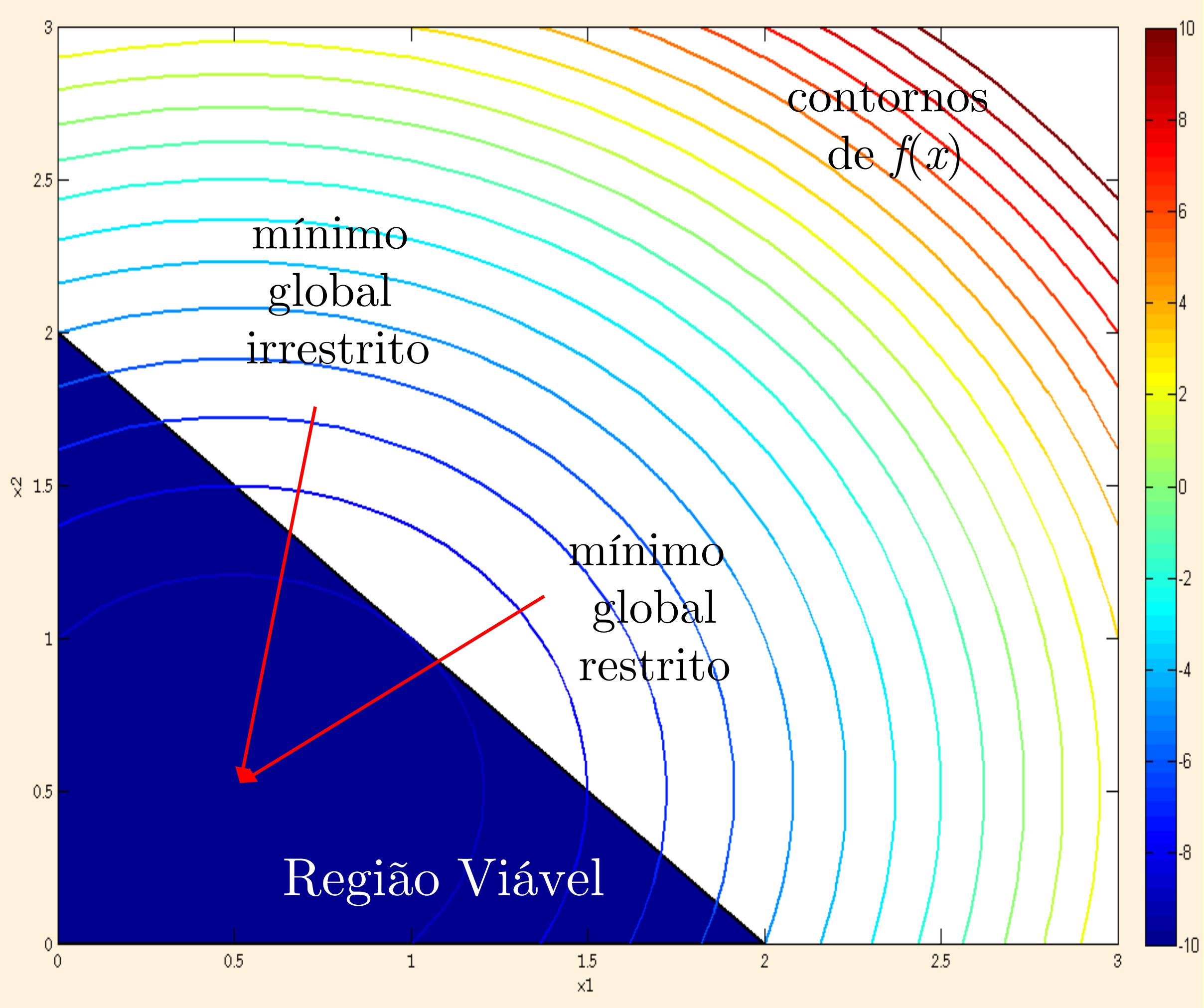



$$\min_{x \in \Re^2} \; f(x) = (x_1 - 1,5)^2 + (x_2 - 1,5)^2$$

$$\text{sujeito a: } \begin{cases} x_1 + x_2 - 2 \leq 0 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

- Se a primeira restrição for ignorada, o ponto ótimo é $(1{,}5 \; 1{,}5)^T$ e $f(x) = 0$
- O ponto ótimo restrito é $(1{,}1)^T$ e $f(x) = 0{,}5$

# Restrições Não Interferem na Solução




$$\min_{x \in \Re^2} \; f(x) = (x_1 - 0,5)^2 + (x_2 - 0,5)^2$$

$$\text{sujeito a:} \begin{cases} x_1 + x_2 - 2 \le 0 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

- O ponto ótimo deste problema pode ser obtido ignorando-se todas as restrições do problema $x = (0{,}5 \;\; 0{,}5)^T$ e $f(x) = 0$

# Definindo Soluções Locais

- Simples extensões do caso irrestrito, exceto que agora deve-se levar em consideração os pontos viáveis na vizinhança de $x^*$
  - Um vetor $x^*$ é uma solução local se $x^* \in \Omega$ e existe uma vizinhança $V$ de $x^*$ tal que $f(x) \geq f(x^*)$ para $x \in \mathrm{V} \cap \Omega$
  - Um vetor $x^*$ é uma solução local estrita (solução local forte) se $x^* \in \Omega$ e existe uma vizinhança $V$ de $x^*$ tal que $f(x) > f(x^*)$ para todo $x \in \mathrm{V} \cap \Omega$ com $x \neq x^*$

# Problemas Somente com Restrições de Igualdade

# Considerações Iniciais

- Introduzir os princípios básicos que caracterizam as soluções dos problemas com restrições de igualdade

- Formulação

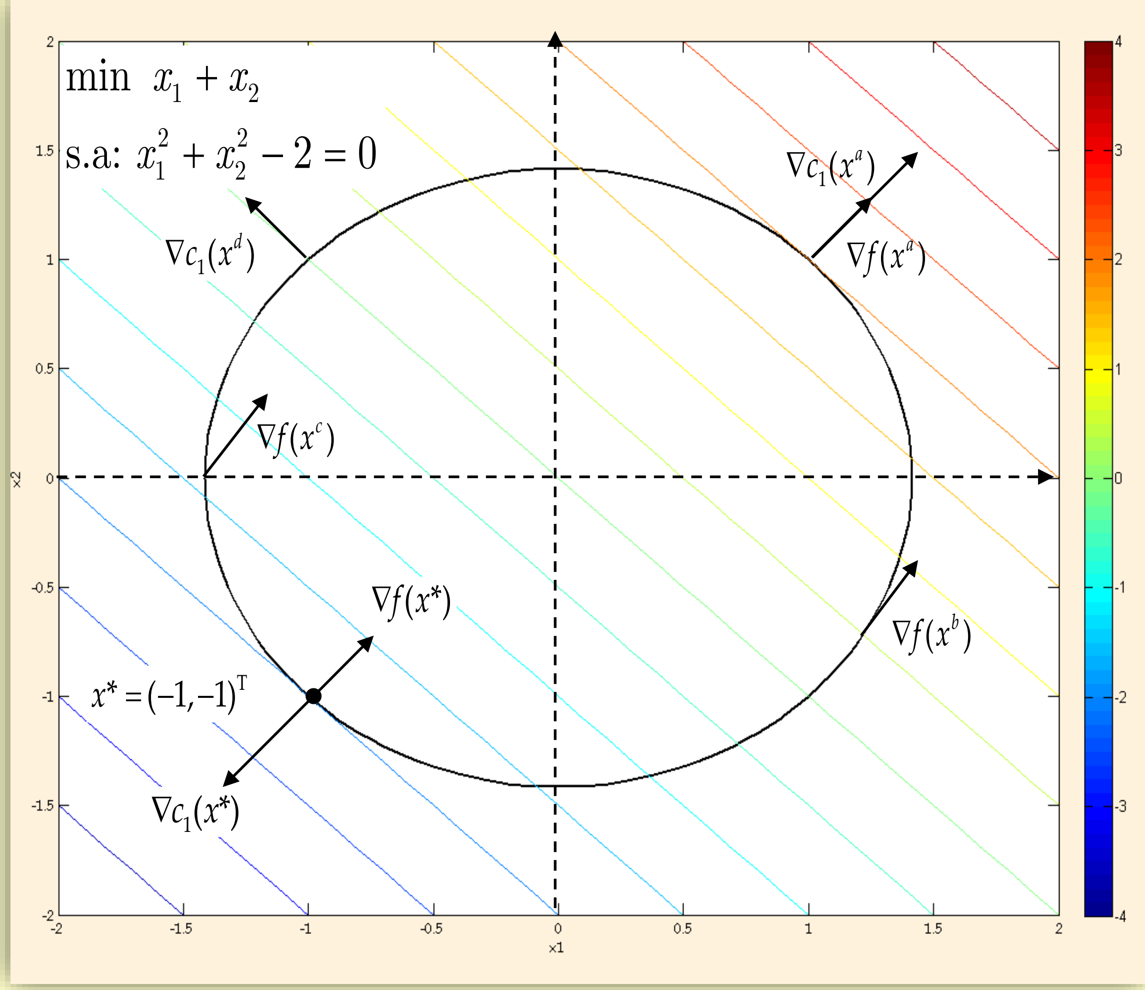
$$\min_{x \in \Re^n} \; f(x)$$

$$\text{sujeito a: } c_i(x) = 0 \quad i \in \varepsilon \text{ e } \; I = \phi$$

- Uma restrição não linear

- Uma restrição linear

# Uma Restrição Não Linear




- $\varepsilon = \{1\}$ e $I = \phi$
- De qualquer ponto em $\Omega$, tal que $x \neq x^*$, é possível realizar um movimento viável enquanto $f$ é decrementada
- Em $x^*$ os gradientes da restrição e da função objetivo são paralelos
- Existe um escalar $\lambda^*$ tal que

  $$\nabla f(x^*) = \lambda^* \nabla c_1(x^*)$$

# Caracterização da Solução

1) Para tentar manter a viabilidade deve-se respeitar $c_1(x+d)=0$

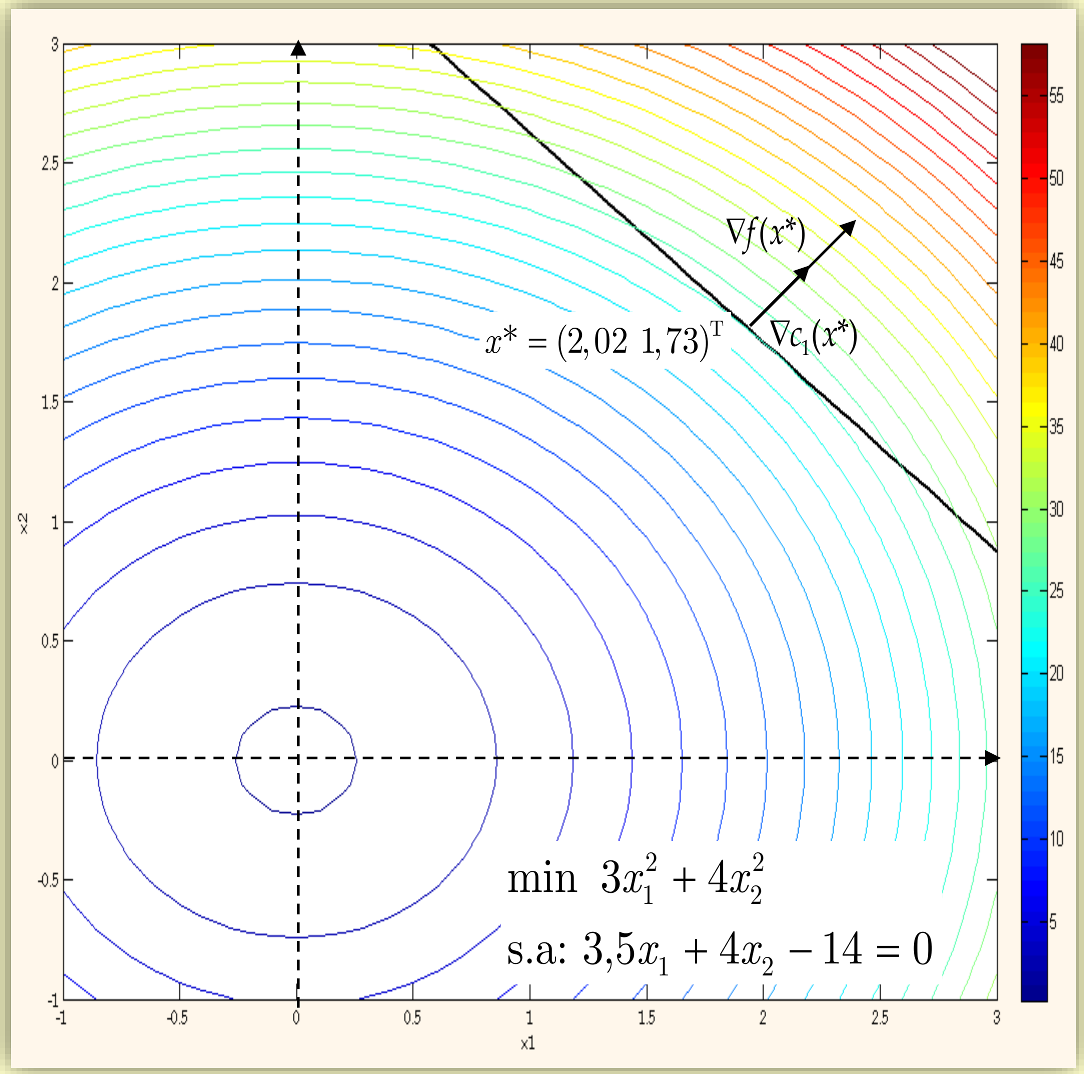
$$0 = c_1(x+d) \approx c_1(x) + \nabla c_1(x)^T d = \nabla c_1(x)^T d$$

2) De forma análoga, para uma direção de descida produzir um decréscimo em $f$ então

$$0 > f(x+d) - f(x) \approx \nabla f(x)^T d$$

- Se existe um $d$ que satisfaz 1) e 2) então é possível ter melhorias
- Em consequência, a condição necessária de otimalidade é que não exista nenhuma direção $d$ que satisfaça 1) e 2) simultaneamente
- Com base no exemplo, tem-se que tal direção não pode existir se $\nabla f(x)$ e $\nabla c_1(x)$ forem paralelos
- A condição $\nabla f(x)=\lambda\nabla c_1(x)$ não é suficiente para encontrar a solução ótima. No exemplo, essa condição é atendida para o ponto $x=(1,1)^T$ → ponto de máximo, em que $\lambda=0{,}5$

# Uma Restrição Linear




- Condição necessária é atendida para um único valor de $\lambda = -3{,}46$
- Em problemas com restrições de igualdade não é possível colocar como condição adicional que $\lambda > 0$!
- O ponto encontrado é um (único) mínimo global

# Função Lagrangiana

- Problema restrito é analisado sob ponto de vista de um problema irrestrito
- Sendo a função Lagrangiana definida por: $L(x,\lambda) = f(x) - \lambda c_1(x)$
- O ponto estacionário é calculado da seguinte maneira:

$$\nabla_x L(x,\lambda) = 0 = \nabla_x f(x) - \lambda \nabla_x c_1(x) \qquad \nabla_\lambda L(x,\lambda) = 0 = -c_1(x)$$

- É possível buscar a solução de um problema com restrições de igualdade calculando os pontos estacionários da função Lagrangiana

# Problemas com Restrições de Desigualdade

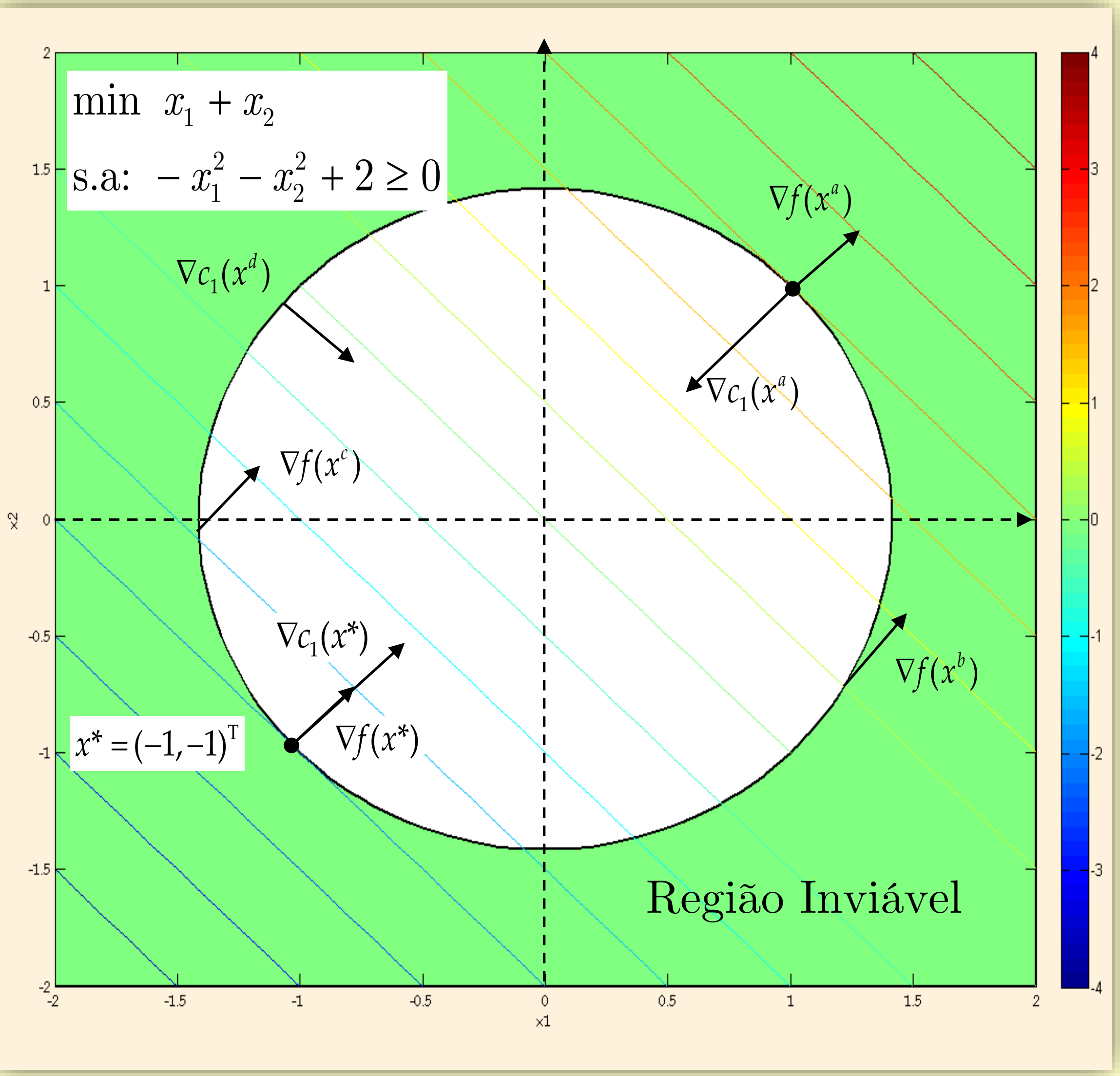
# Considerações Iniciais

- Introduzir os princípios básicos que caracterizam as soluções dos problemas com restrições de desigualdade
- Formulação

$$\min_{x \in \Re^n} \; f(x)$$
$$\text{sujeito a: } c_i(x) \geq 0 \quad i \in I$$

- Em um ponto viável $x$ a $i$–ésima restrição de desigualdade, tal que $i \in I$, é dita estar ativa se $c_i(x) = 0$ e inativa se $c_i(x) > 0$

# Uma Restrição Não Linear




- O vetor gradiente da restrição, nos pontos onde a mesma está ativa, sempre aponta para o interior da região viável
- A condição $\nabla f(x) = \lambda \nabla c_1(x)$ é atendida nos pontos
  - $x = (1,1)^T$ com $\lambda = -0{,}5$
  - $x = -(1,1)^T$ com $\lambda = 0{,}5$
- Porém, o sinal de $\lambda$ é um importante parâmetro para encontrar uma solução

# Caracterização da Solução

1) Para manter a viabilidade deve-se respeitar: $0 \leq c_1(x+d) \approx c_1(x) + \nabla c_1(x)^T d \rightarrow c_1(x) + \nabla c_1(x)^T d \geq 0$

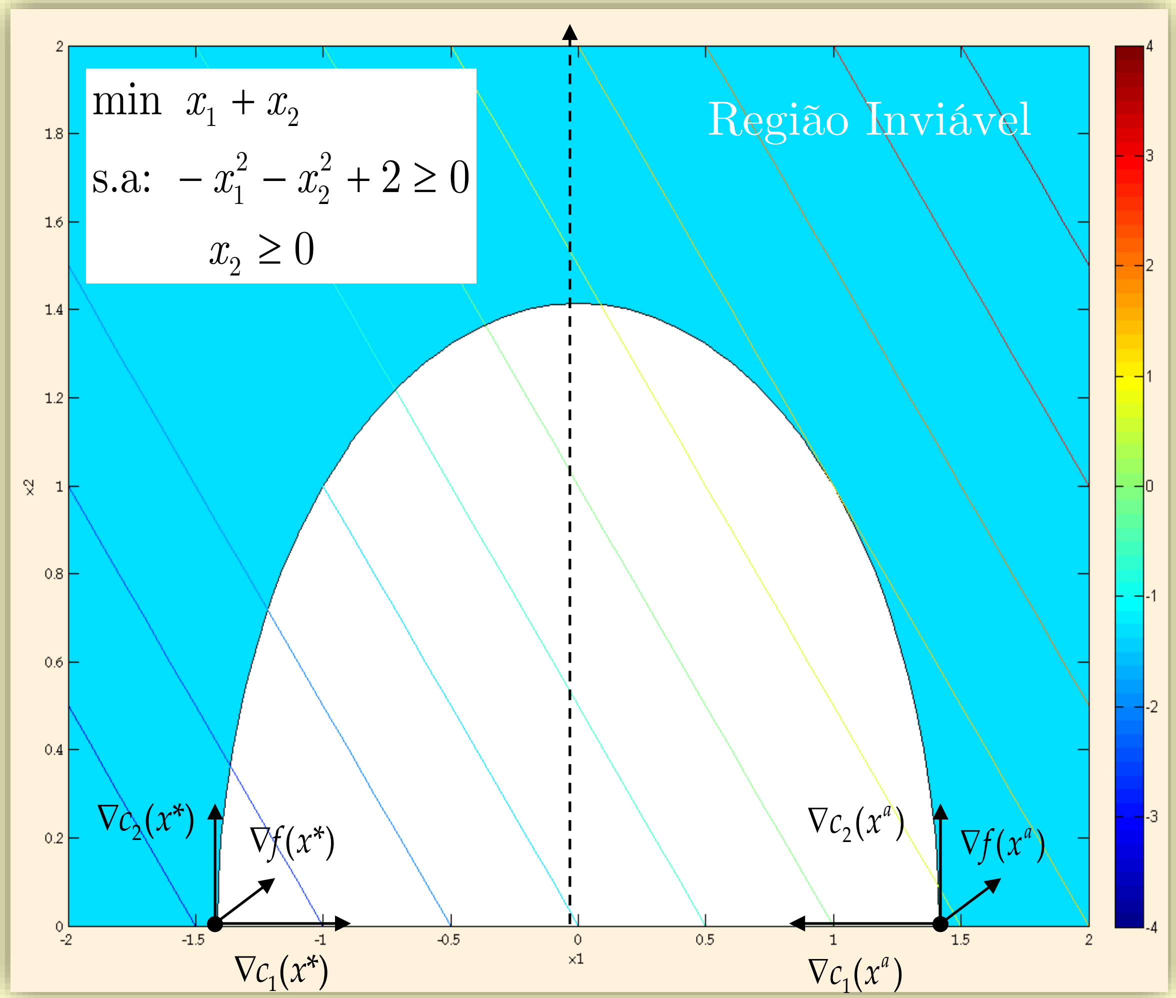
2) Direção de descida: $\nabla f(x)^T d < 0$

- Para encontrar $d$ que satisfaz 1) e 2) simultaneamente é necessário analisar dois casos:

- $x$ é um ponto tal que $c_1(x) > 0$. Neste caso $\forall d$ atende 1), dado que seja suficientemente pequeno. A única situação onde não é possível atender 1) e 2) é com $\nabla f(x)=0$

- $x$ é um ponto tal que $c_1(x) = 0$. Assim 1) e 2) tornam-se $\nabla f(x)^{\mathrm{T}} d < 0$ e $\nabla c_1(x)^{\mathrm{T}} d \geq 0$. A única situação onde essas duas regiões não tem nenhum ponto em comum é quando $\nabla f(x)$ e $\nabla c_1(x)$ apontam na mesma direção, isto é, quando $\nabla f(x) = \lambda \nabla c_1(x)$, com $\lambda \geq 0$.

# Duas Restrições




$x^* = (-\sqrt{2}, 0)^T$

- $\nabla c_i(x^*)^T d \geq 0,\ i = 1,2$, são satisfeitas com $d$ no quadrante definido por $\nabla c_1(x^*)$ e $\nabla c_2(x^*)$

- Vetores neste quadrante atendem a $\nabla f(x^*)^T d \geq 0$

$$\Downarrow$$

$$\nabla f(x^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \nabla c_1(x^*) = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}, \nabla c_2(x^*) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\Downarrow$$

$$\lambda^* = \begin{bmatrix} \frac{1}{2\sqrt{2}} \\ 1 \end{bmatrix}$$

# Não Linearidade Acentuada

(1)

$$\min\ x_1 + x_2$$
$$\text{s.a: } x_1^2 + x_2^2 - 2 = 0$$

(2)

$$\min\ x_1 + x_2$$
$$\text{s.a: } (x_1^2 + x_2^2 - 2)^2 = 0$$

- Os dois problemas possuem o ponto ótimo em $x = -(1,1)^T$
- Em (2) tem-se que $\nabla c_1(x) = 0$ para todos pontos viáveis e $\nabla f(x) = \lambda \nabla c_1(x)$ não pode ser assegurado
- **Definição**: Dado um ponto $x^*$ e um conjunto $A(x^*)$, diz-se que as restrições estão qualificadas se $\{\nabla c_i(x),\ i \in A(x^*)\}$ são linearmente independentes
- Se a condição acima for verificada, nenhum dos elementos de $\{\nabla c_i(x),\ i \in A(x^*)\}$ é nulo. Ainda, permite expressar as condições de otimalidade para um problema de programação não linear qualquer

# Condições de Otimalidade de Primeira Ordem

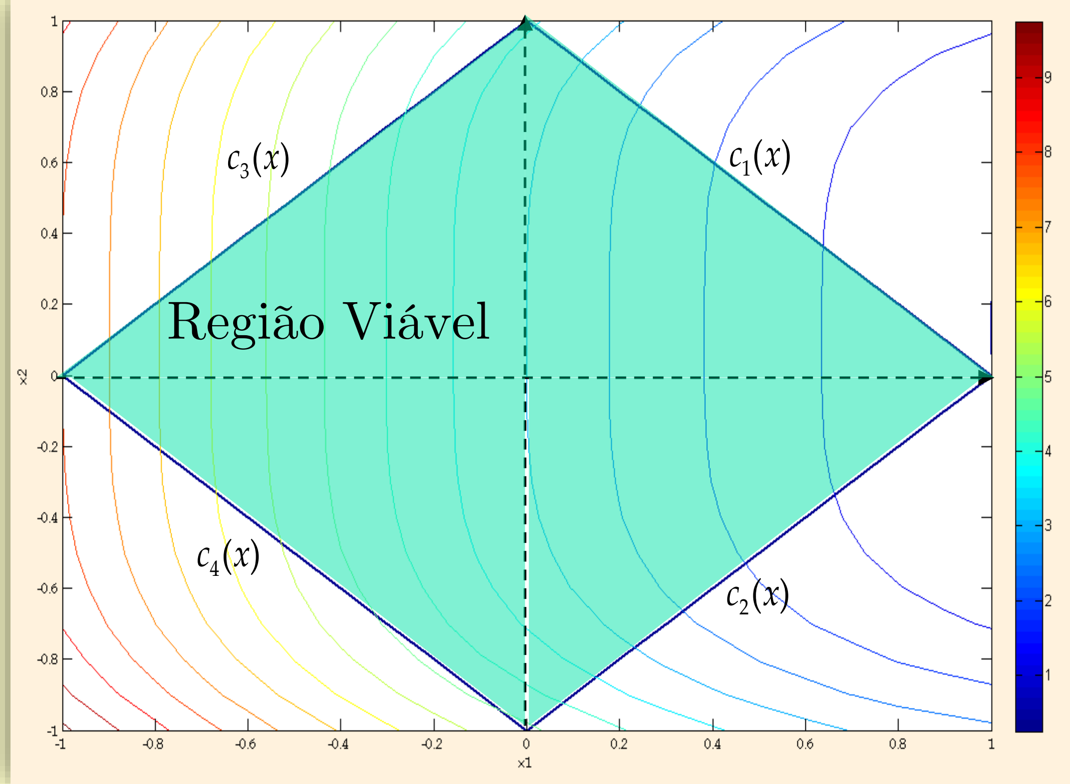
# Condições Necessárias de Primeira Ordem

- Seja $x^*$ um mínimo local e que as restrições ativas em $x^*$ estão qualificadas. Assim, existe um vetor $\lambda^*$, $i \in \varepsilon \cup I$ , tal que as seguintes condições são satisfeitas em $(x^*, \lambda^*)$

  - $\nabla_x L(x^*, \lambda^*) = 0$
  - $c_i(x^*) = 0 \qquad i \in \varepsilon$
  - $c_i(x^*) \geq 0 \qquad i \in I$
  - $\lambda^*_i \geq 0 \qquad i \in I$
  - $\lambda^*_i \, c_i(x^*) = 0 \qquad i \in \varepsilon \cup I$

**Condições de Karush-Kuhn-Tucker (KKT)**

# Condições de KKT - Ilustração




$$\min \; f(x) = \left(x_1 - \frac{3}{2}\right)^2 + \left(x_2 - \frac{1}{8}\right)^4$$

$$\text{s.a:} \begin{bmatrix} c_1(x): 1 - x_1 - x_2 \geq 0 \\ c_2(x): 1 - x_1 + x_2 \geq 0 \\ c_3(x): 1 + x_1 - x_2 \geq 0 \\ c_4(x): 1 + x_1 + x_2 \geq 0 \end{bmatrix}$$

$$x^* = (1,0) \rightarrow \nabla f(x^*) = \begin{bmatrix} -1 \\ -\frac{1}{2} \end{bmatrix}, \nabla c_1(x^*) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \nabla c_2(x^*) = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightarrow \lambda^* = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 & 0 \end{bmatrix}^T$$

# Exemplo

- Escrever as condições de KKT para o problema abaixo

$$\text{minimize } f(x) = -x_1^3 + x_2^3 - 2x_1x_3^2$$
$$\text{sujeito a:} \quad 2x_1 + x_2^2 + x_3 - 5 = 0$$
$$5x_1^2 - x_2^2 - x_3 \geq 2$$
$$x_1, x_2, x_3 \geq 0$$

$$\nabla f(x) - \lambda_1 \nabla c_1(x) - \lambda_2 \nabla c_2(x) - \lambda_3 \nabla c_2(x) - \lambda_4 \nabla c_4(x)$$
$$-\lambda_5 \nabla c_5(x) = 0$$
$$c_1(x) = 0, \quad c_j(x) \geq 0, j = 2, \dots, 5$$
$$\lambda_j \geq 0, j = 2, \dots, 5, \quad \lambda_j c_j(x) = 0, j = 2, \dots, 5$$

$$\begin{bmatrix} -3x_1^2 - 2x_3^2 \\ 3x_2^2 \\ -4x_1x_3 \end{bmatrix} - \lambda_1 \begin{bmatrix} 2 \\ 2x_2 \\ 1 \end{bmatrix} - \lambda_2 \begin{bmatrix} 10x_1 \\ -2x_2 \\ -1 \end{bmatrix} - \lambda_3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \lambda_4 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \lambda_5 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$$

$$2x_1 + x_2^2 + x_3 - 5 = 0$$
$$5x_1^2 - x_2^2 - x_3 \geq 2$$
$$x_1, x_2, x_3 \geq 0$$

$$\lambda_2 \geq 0, \lambda_3 \geq 0, \lambda_4 \geq 0, \lambda_5 \geq 0$$
$$\lambda_2 (5x_1^2 - x_2^2 - x_3 - 2) = 0$$
$$\lambda_3 x_1 = 0, \lambda_3 x_2 = 0, \lambda_5 x_2 = 0$$

# Exemplo

- Verifique se as condições de KKT são atendidas em $x = [1\ 0\ 3]^t$

$$\begin{aligned} &\text{minimize } f(x) = -x_1^3 + x_2^3 - 2x_1x_3^2 \\ &\text{sujeito a:} \quad 2x_1 + x_2^2 + x_3 - 5 = 0 \\ &\qquad\qquad\quad 5x_1^2 - x_2^2 - x_3 \geq 2 \\ &\qquad\qquad\quad x_1, x_2, x_3 \geq 0 \end{aligned}$$

$$\nabla f(x) - \lambda_1 \nabla c_1(x) - \lambda_2 \nabla c_2(x) - \lambda_4 \nabla c_4(x) = 0$$

$$\begin{bmatrix} -3x_1^2 - 2x_3^2 \\ 3x_2^2 \\ -4x_1x_3 \end{bmatrix} - \lambda_1 \begin{bmatrix} 2 \\ 2x_2 \\ 1 \end{bmatrix} - \lambda_2 \begin{bmatrix} 10x_1 \\ -2x_2 \\ -1 \end{bmatrix} - \lambda_4 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 0$$

$$\begin{bmatrix} -21 \\ 0 \\ -12 \end{bmatrix} - \lambda_1 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \lambda_2 \begin{bmatrix} 10 \\ 0 \\ -1 \end{bmatrix} - \lambda_4 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 0 \rightarrow \begin{cases} -2\lambda_1 - 10\lambda_2 = 21 \\ \lambda_4 = 0 \\ -\lambda_1 + \lambda_2 = 12 \end{cases} \rightarrow \begin{cases} \lambda_1 = -11,75 \\ \lambda_2 = 0,25 \\ \lambda_3 = \lambda_4 = \lambda_5 = 0 \end{cases}$$

# Exemplo

- Outros pontos que atendem as condições de KKT
  - ✓ **Combinar as possibilidades das restrições de desigualdade estarem ativas**

$$
\begin{aligned}
&-3x_1^2 - 2x_3^2 - 2\lambda_1 - 10\lambda_2 x_1 - \lambda_3 = 0\\
&3x_2^2 - 2\lambda_1 x_2 + 2\lambda_2 x_2 - \lambda_4 = 0\\
&-4x_1 x_3 - \lambda_1 + \lambda_2 - \lambda_5 = 0
\end{aligned}
\qquad
\begin{aligned}
&2x_1 + x_2^2 + x_3 - 5 = 0\\
&5x_1^2 - x_2^2 - x_3 \geq 2\\
&x_1, x_2, x_3 \geq 0
\end{aligned}
\qquad
\begin{aligned}
&\lambda_2 \geq 0, \lambda_3 \geq 0, \lambda_4 \geq 0, \lambda_5 \geq 0\\
&\lambda_2(5x_1^2 - x_2^2 - x_3 - 2) = 0\\
&\lambda_3 x_1 = 0, \lambda_3 x_2 = 0, \lambda_5 x_2 = 0
\end{aligned}
$$

- Todas as desigualdades ativas: $x_1,x_2,x_3=0$ inviável em $c_1(x)$ e $c_2(x)$
- Somente $c_5(x)$ não ativa: $x_1,x_2=0$, por $c_1(x)$ $x_3=5$ e viola $c_2(x)$
- Somente $c_4(x)$ não ativa: $x_1,x_3=0$, por $c_1(x)$ $x_2=\pm\sqrt{5}$ e viola $c_2(x)$
- **Somente $c_3(x)$ não ativa: $x_2,x_3=0$, por $c_1(x)$ $x_1=2{,}5$ e não viola $c_2(x)$**
- E assim por diante...

# Condições de Otimalidade de Segunda Ordem

# Introdução... (1)

- Condições de KKT indicam como as derivadas primeiras de $f(x)$ e das restrições ativas estão relacionadas em $x^*$

- Quando as condições de KKT são verificadas, um movimento de qualquer vetor $w \in F_1$ faz com que $w^{\mathrm{T}}\nabla f(x^*)>0$ ou $w^{\mathrm{T}}\nabla f(x^*)=0$, onde

$$F_1 = \left\{ \alpha d \mid \alpha > 0, \begin{array}{ll} d^T \nabla c_i(x^*) = 0, & \forall i \in \varepsilon \\ d^T \nabla c_i(x^*) \geq 0, & \forall i \in A(x^*) \cap I \end{array} \right.$$

- Para direções $w \in F_1$ tal que $w^{\mathrm{T}}\nabla f(x^*)=0$, não é possível determinar da informação de primeira ordem se um movimento nesta direção irá incrementar ou diminuir a função objetivo

# Introdução... (2)

- Dado $F_1$ e algum multiplicador de Lagrange $\lambda^*$ que satisfaz as condições de KKT, define-se um subconjunto $F_2(\lambda^*)$ de $F_1$ por

$$w \in F_2(\lambda^*) \Leftrightarrow \begin{cases} \nabla c_i(x^*)^{\mathrm{T}} w = 0, & \forall i \in \varepsilon, \\ \nabla c_i(x^*)^{\mathrm{T}} w = 0, & \forall i \in A(x^*) \cap I \text{ com } \lambda_i^* > 0, \\ \nabla c_i(x^*)^{\mathrm{T}} w \geq 0, & \forall i \in A(x^*) \cap I \text{ com } \lambda_i^* = 0. \end{cases}$$

- Portanto, $F_2(\lambda^*)$ contém direções de $F_1$ que não informam, com base nas derivadas primeira, se $f(x^*)$ irá aumentar ou diminuir

# Condições Necessárias

- Envolvem as derivadas de segunda ordem: se $x^*$ é uma solução local, então a curvatura da função Lagrangiana nas direções de $F_2(\lambda^*)$ devem ser não negativas

- **Teorema:** Seja $x^*$ é uma solução local e que os gradientes das restrições ativas sejam linearmente independentes. Seja $\lambda^*$ um vetor de multiplicadores de Lagrange tal que as condições de KKT são verificadas e seja $F_2(\lambda^*)$ conforme definido anteriormente. Então,

$$w^{\mathrm{T}} \nabla_{xx} L(x^*, \lambda^*) w \geq 0, \quad \text{para } \forall w \in F_2(\lambda^*).$$

# Condições Suficientes

- São condições em $f$ e $c_i$ que asseguram que $x^*$ é uma solução local – em oposição às condições necessárias que assumem que $x^*$ é uma solução e deduzem propriedades em $f$ e $c_i$

- Teorema: Suponha que para $x^*$ viável $\in \Re^n$ existe um vetor $\lambda^*$, $i \in \varepsilon \cup I$ , tal que as condições de KKT são satisfeitas. Suponha também que

$$w^{\mathrm{T}} \nabla_{xx} L(x^*, \lambda^*) w > 0, \quad \text{para } \forall w \in F_2(\lambda^*), w \neq 0.$$

Então $x^*$ é um mínimo local

# Exemplo 1

$$\min\ f(x) = x_1 + x_2 \quad \text{s.a: } c_1(x) = 2 - x_1^2 - x_2^2 \geq 0$$

$$L(x,\lambda) = x_1 + x_2 - \lambda(2 - x_1^2 - x_2^2)$$

- É fácil verificar que as condições de KKT são satisfeitas em $x^* = -(1,1)^T$ com $\lambda^* = 1/2$.

  A hessiana de $L(x^*,\lambda^*)$ é

$$\nabla_{xx} L(x^*,\lambda^*) = \begin{bmatrix} 2\lambda^* & 0 \\ 0 & 2\lambda^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- Essa matriz é definida positiva para qualquer $w \neq 0$, e portanto, certamente satisfaz as condições do teorema mostrado no slide anterior

- Assim $x^*$ é uma solução estrita do PNL – de fato, é uma solução global dado que o PNL é convexo

# Exemplo 2

$$\min\ f(x) = x_1 + x_2 \quad \text{s.a:}\ c_1(x) = x_1^2 + x_2^2 - 2 = 0$$

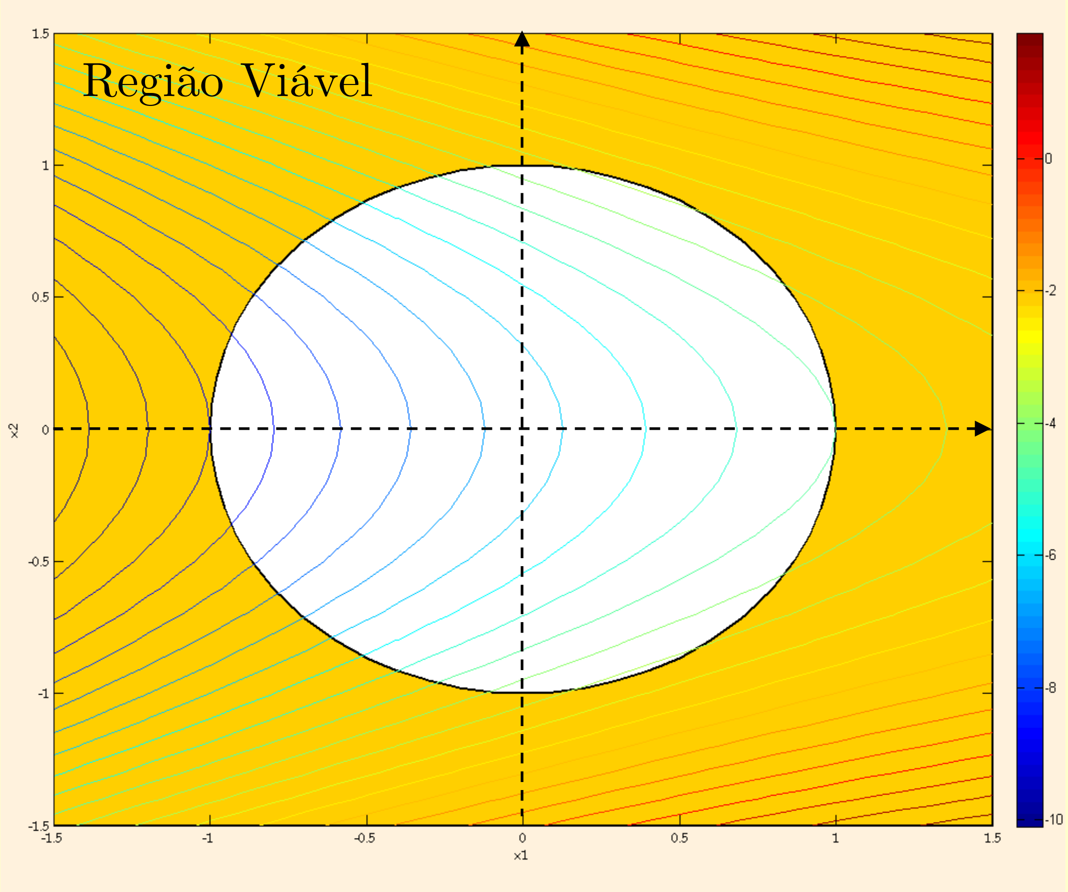
$$L(x,\lambda) = x_1 + x_2 - \lambda(x_1^2 + x_2^2 - 2)$$

- Condições de KKT são satisfeitas em $x_a = -(1,1)^T$ com $\lambda_a^* = -1/2$ e em $x_b = (1,1)^T$ com $\lambda_b^* = 1/2$. As hessianas são

$$\nabla_{xx} L(x_a,\lambda_a) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ e } \nabla_{xx} L(x_b,\lambda_b) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

- $x_a$ é uma solução estrita do PNL – De fato, é uma solução global, pois $x_b$ não atende as condições de segunda ordem e não existem outros pontos que atendem as condições de primeira ordem
- Adicionalmente, a região viável é limitada

# Exemplo 3... (1)




$$\min \quad -0,1(x_1-4)^2 + x_2^2$$
$$\text{s.a:} \quad x_1^2 + x_2^2 - 1 \geq 0$$

$$\nabla_x L(x,\lambda) = \begin{bmatrix} -0,2(x_1-4) - 2\lambda x_1 \\ 2x_2 - 2\lambda x_2 \end{bmatrix}, \nabla_{xx}^2 L(x,\lambda) = \begin{bmatrix} -0,2-2\lambda & 0 \\ 0 & 2-2\lambda \end{bmatrix}$$

- $x^* = (1,0)^T$ atende KKT em $\lambda^* = 0,3$
- Note que $\nabla c(x^*) = (2,0)^T$, então:

$$= w^T \nabla_{xx}^2 L(x^*,\lambda^*) w$$
$$= \begin{bmatrix} 0 \\ w_2 \end{bmatrix}^T \begin{bmatrix} -0,8 & 0 \\ 0 & 1,4 \end{bmatrix} \begin{bmatrix} 0 \\ w_2 \end{bmatrix}$$
$$= 1,4 w_2^2 > 0$$

- $x^* = (1,0)^T$ é um mínimo local

# Exemplo 3... (2)

- $x^* = (1,0)^T$ é um mínimo global? Uma vez que o problema é não convexo deve-se verificar se existem outros pontos que atendem as condições de 1° e 2° ordem

- Condições de KKT→ $\nabla_x L = \begin{bmatrix} -0,2(x_1 - 4) - 2\lambda x_1 \\ 2x_2 - 2\lambda x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \nabla_\lambda L = x_1^2 + x_2^2 - 1 = 0$

- De $\nabla_x L=0$ (segunda equação) tem-se que $\lambda=1$ ou $x_2=0$. Com $x_2=0$ chega-se ao ponto testado no slide anterior. Com $\lambda=1$ obtém-se da primeira equação de $\nabla_x L$ que $x_1=0,3636$ e, aplicando-se na equação $\nabla_y L$ tem-se $x_2 = \pm 0,9315$

- Portanto dois pontos adicionais atendem as condições de KKT
  - $x_a=(0,3636 \quad 0,9315)^T$ com $\lambda=1$
  - $x_b=(0,3636 \quad -0,9315)^T$ com $\lambda=1$

# Exemplo 3... (3)

- Considerando $x_a$=(0,3636  0,9315)$^T$ com $\lambda$=1, tem-se

$$\mapsto \nabla c(x_a) = \begin{bmatrix} 0,7273 \\ 1,8631 \end{bmatrix} \rightarrow \nabla c(x_a)^T w = 0 \rightarrow w_1 = -2,5617 w_2$$

$$\mapsto w^T \nabla^2_{xx} L(x_a, \lambda) w$$

$$\mapsto \begin{bmatrix} -2,5617 w_2 \\ w_2 \end{bmatrix}^T \begin{bmatrix} -2,2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -2,5617 w_2 \\ w_2 \end{bmatrix} = -14,4371 w_2^2 < 0$$

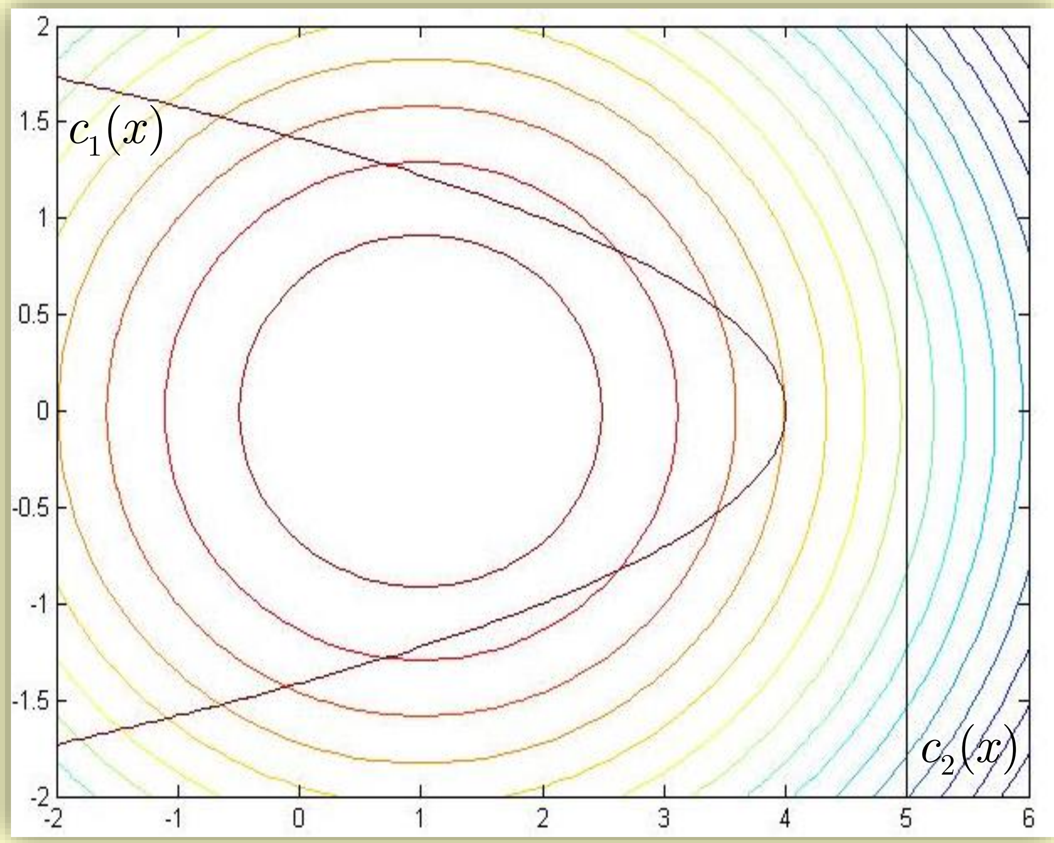
- Portanto, $x_a$ não é um ponto de mínimo, pois a função Lagrangiana tem curvatura negativa
- Uma vez que neste problema a curvatura não depende do valor das variáveis primais, então $x_b$ também não representa uma solução local
- Embora $x^* = (1,0)^T$ seja um mínimo local, a $f(x)$ não é limitada inferiormente na região viável e, portanto, não existe solução global

# Exemplo 4... (1)

$$\max_{x\in\Re^2} f(x) = \frac{3}{4}(x_1-1)^2 + 2x_2^2$$
$$\text{s.a}: c_1(x): x_1 + 2x_2^2 - 4 = 0$$
$$c_2(x): x_1 \le 5$$

$\Rightarrow$

$$\min_{x\in\Re^2} f(x) = -\frac{3}{4}(x_1-1)^2 - 2x_2^2$$
$$\text{s.a}: c_1(x): x_1 + 2x_2^2 - 4 = 0$$
$$c_2(x): -x_1 + 5 \ge 0$$



# Exemplo 4... (2)

- Somente $c_1(x)$ ativa

$$\nabla_x L = \begin{bmatrix} -1,5(x_1 - 1) \\ -4x_2 \end{bmatrix} - \lambda \begin{bmatrix} 1 \\ 4x_2 \end{bmatrix} = 0 \Rightarrow \begin{array}{ll} -1,5x_1 + 1,5 - \lambda = 0 & (1) \\ -4x_2 - 4\lambda x_2 = 0 & (2) \\ x_1 + 2x_2^2 - 4 = 0 & (3) \end{array}$$

- De (2) tem-se $x_2 = 0$ ou $\lambda = -1$.

- Com $x_2 = 0$, pode-se obter por (3) $x_1 = 4$. Note que $c_2(x) > 0$. De (1) tem-se que $\lambda = -4{,}5$. Assim, $x = (4{,}0)^T$ atende KKT

- Considerando $x = (4{,}0)^T$ e $\lambda = -4{,}5$ as condições de segunda ordem são

$$\begin{bmatrix} 1 & 4 \cdot 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0, \quad \begin{bmatrix} 0 & w_2 \end{bmatrix} \begin{bmatrix} -1,5 & 0 \\ 0 & 14 \end{bmatrix} \begin{bmatrix} 0 \\ w_2 \end{bmatrix} > 0, \quad 14w_2^2 > 0$$

- Assim, $x = (4{,}0)^T$ é um ponto de mínimo local estrito

# Exemplo 4... (3)

- ...continuando...

$$\nabla_x L = \begin{bmatrix} -1,5(x_1 - 1) \\ -4x_2 \end{bmatrix} - \lambda \begin{bmatrix} 1 \\ 4x_2 \end{bmatrix} = 0 \quad \Rightarrow \quad \begin{aligned} &-1,5x_1 + 1,5 - \lambda = 0 && (1) \\ &-4x_2 - 4\lambda x_2 = 0 && (2) \\ &x_1 + 2x_2^2 - 4 = 0 && (3) \end{aligned}$$

- Com $\lambda = -1$, pode-se obter por (1) que $x_1 = 5/3$. Note que $c_2(x) > 0$. De (3) tem-se que $x_2 = \pm 1{,}0801$.

- Considerando $x = (5/3\ 1{,}0801)^T$ e $\lambda = -1$ as condições de segunda ordem são

$$\begin{bmatrix} 1 & 4,32 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0, \quad \begin{bmatrix} -4,32w_2 & w_2 \end{bmatrix} \begin{bmatrix} -1,5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -4,32w_2 \\ w_2 \end{bmatrix} > 0, \quad -28w_2^2 < 0$$

- Assim, $x = (5/3\ 1{,}0801)^T$ não é um ponto de mínimo local estrito

# Exemplo 4... (4)

- Considerando $x = (5/3\ -1{,}0801)^T$ e $\lambda = -1$

$$\begin{bmatrix} 1 & -4,32 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0, \quad \begin{bmatrix} 4,32w_2 & w_2 \end{bmatrix} \begin{bmatrix} -1,5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 4,32w_2 \\ w_2 \end{bmatrix} > 0, \quad -28w_2^2 < 0$$

- Assim, $x = (5/3\ \ -1{,}0801)^T$ não é um ponto de mínimo local estrito

- $x^* = (4{,}0)^T$ é um mínimo global?
  - ✓ Embora seja o único ponto de KKT que atende a segunda ordem, deve-se verificar se o valor de $f(x)$ na região viável é limitada inferiormente
  - ✓ Note que $f(x^*) = -6{,}75$. Pode-se fixar $x_1 = -4$ em $c_1(x)$. Assim $x_2 = 2$ e $f(-4{,}2) = -14{,}75 < f(x^*)$
  - ✓ Portanto, $x^*$ não é um mínimo global

# OBRIGADO!


## Prof. Erlon Cristian Finardi

erlon.finardi@ufsc.br