Algoritmos de Otimização Não-lineares Sem Restrições

Por que estudar?

Algoritmos para solução de problemas com restrições baseiam-se em algoritmos para solução de problemas sem restrições (métodos de penalização, obtenção da direção de busca, etc...);

Princípio Básico: Problema de n variáveis ➡ Problema de uma variável

$$\mathbf{x} = \mathbf{x}_0 + \alpha\mathbf{s}_0 \;\Rightarrow\; \mathrm{f}(\mathbf{x}) = \mathrm{f}(\mathbf{x}_0 + \alpha\mathbf{s}_0) = \mathrm{f}(\alpha) \;\Rightarrow \text{Encontrar } \alpha$$

Procedimento Típico

- Encontrar $\mathbf{s}_0$ em $\mathbf{x}_0$ que reduza função objetivo
- Encontrar α na direção de $\mathbf{s} \Rightarrow \mathbf{x}_0 + \alpha\mathbf{s}_0$ (busca unidimensional)
- Verificar convergência: se satisfeita $\Rightarrow \mathbf{x} = \mathbf{x}^*$, pare, caso contrário $\Rightarrow \mathbf{x}_{i+1} = \mathbf{x}$ e $\mathrm{i} = \mathrm{i} + 1$

Algoritmos de Otimização Não-lineares Sem Restrições

Classificação

- Ordem Zero ➡ utiliza somente o valor da função objetivo. Usados quando a função não é diferenciável ou quando a função é altamente não-linear, sendo difícil obter as derivadas de forma precisa.
  Ex.: Método das Direções Conjugadas de Powell
- Primeira Ordem ➡ utiliza o valor e gradiente da função
  Ex.: Método "Steepest Descent" e Método dos Gradientes Conjugados
- Segunda Ordem ➡ utiliza o valor, gradiente e matriz Hessiana da função.
  Ex.: Método de Newton e Quasi-Newton

Métodos de Ordem Zero

Usados quando o valor da função é obtido com precisão pobre ➡ valores de derivadas (ou gradientes) não são confiáveis.

Método das Direções Conjugadas de “Powell”

Baseado na minimização de uma função quadrática considerando direções *Q*-conjugadas linearmente independentes. Função quadrática:

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{b}^T\mathbf{x} + \mathbf{c}$$

Direções *Q*-conjugadas: $\mathbf{s}_i^T\mathbf{Q}\mathbf{s}_j = 0 \quad \text{para } i \neq j$

“Se f for minimizada ao longo de cada direção **s**, então o mínimo de f será encontrado no (ou antes) do n$^{\text{ésimo}}$ passo independentemente do ponto inicial, dado que erros de arredondamento não sejam acumulados.”

Powell ➡ método conveniente de gerar direções *Q*-conjugadas linearmente independentes. Se direções geradas são linearmente dependentes ➡ não converge para o mínimo.

Método das Direções Conjugadas de Powell

Procedimento:

1. Minimize f ao longo das direções coordenadas (busca univariada), iniciando em $\mathbf{x}_0^k$ e gerando os pontos $\mathbf{x}_1^k, \ldots, \mathbf{x}_n^k$ onde k é o número do ciclo;
2. Após encerrar a busca univariada encontre o índice m correspondente à direção em que a função f apresenta o maior decréscimo indo de $\mathbf{x}_{m-1}^k$ para $\mathbf{x}_m^k$;
3. Calcule a direção “padrão” $\mathbf{s}_p^k = \mathbf{x}_n^k - \mathbf{x}_0^k$ e determine o valor de α que minimize f tal que: $\mathbf{x} = \mathbf{x}_0^k + \alpha\mathbf{s}_p^k$
4. Se

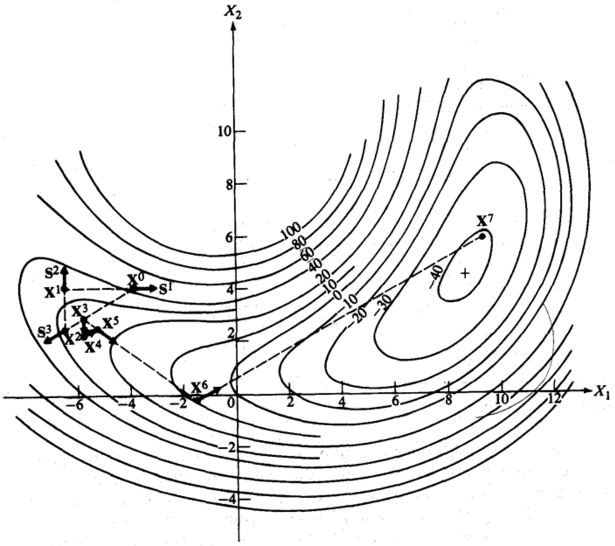
$$|\alpha| < \left[\frac{f(\mathbf{x}_0^k) - f(\mathbf{x}_0^{k+1})}{f(\mathbf{x}_{m-1}^k) - f(\mathbf{x}_m^k)}\right]^{\frac{1}{2}}$$

   então utilize as mesmas direções para a próxima busca univariada. Se a equação NÃO for satisfeita, então substitua a m$^{\text{ésima}}$ direção pela direção padrão $\mathbf{s}_p^k$

5. Comece a nova busca univariada com as direções obtidas no passo 4, e repita os passos 2, 3, e 4 até que a convergência seja atingida, ou seja: $\left\|\mathbf{x}_{k+1} - \mathbf{x}^k\right\| \leq \varepsilon$

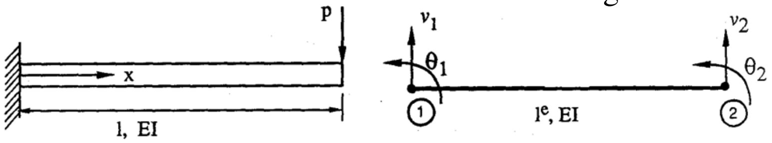
Método das Direções Conjugadas de Powell

Interpretação do Método de Direções Conjugadas de "Powell":



Exemplo

Problema de estudo: Determinação da máxima deflexão e rotação da extremidade de uma viga através da minimização da sua energia potencial total modelada usando um único elemento finito de viga cúbico.



Formulação do elemento:

$$v(\xi) = \left[\left(1-3\xi^2+2\xi^3\right) \quad l\left(\xi-2\xi^2+\xi^3\right) \quad \left(3\xi^2-2\xi^3\right) \quad l\left(-\xi^2+\xi^3\right)\right] \begin{Bmatrix} v_1 \\ \theta_1 \\ v_2 \\ \theta_2 \end{Bmatrix}$$

onde: $\xi = x/l$. A energia potencial correspondente do modelo de viga é dado por:

$$\Pi = \frac{EI}{2l^3} \int_0^l \left(\frac{d^2 v}{d\xi^2} \right)^2 d\xi + p v_2$$

Como a viga está engastada em $\xi = 0 \Rightarrow v_1 = \theta_1 = 0$, substituindo em v(ξ):

$$\Pi = \frac{EI}{2l^3}\left(12 v_2{}^2 + 2\theta_2{}^2 l^2 - 12 v_2 \theta_2 l\right) + p v_2$$ e definindo: $f = \frac{2\Pi l^3}{EI}$; $x_1 = v_2$; $x_2 = \theta_2 l \Rightarrow$

$$\boxed{f = 12 x_1{}^2 + 4 x_2{}^2 - 12 x_1 x_2 + 2 x_1}$$

Exemplo

$$\boxed{\text{Min} \quad f = 12x_1^{\,2} - 12x_1x_2 + 4x_2^{\,2} + 2x_1}$$

$$\mathbf{x}_0^1 = (-1,-2)^T \ \text{ e } \ f(\mathbf{x}_0^1) = 2$$

$$\left(\mathbf{x}^* = (-1/3,-1/2)^T \ \text{ solução exata } \right)$$

Solução:

$$\mathbf{s}_1^1 = (1,0)^T \Rightarrow \mathbf{x}_1^1 = \begin{Bmatrix} -1 \\ -2 \end{Bmatrix} + \alpha \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} = \begin{Bmatrix} -1+\alpha \\ -2 \end{Bmatrix} \Rightarrow f(\alpha) = 12(-1+\alpha)^2 + 4(-2)^2 - 12(-1+\alpha)(-2) +$$

$$+ 2(-1+\alpha) \Rightarrow \text{Min } f \Rightarrow \alpha = \text{-}1/12 \Rightarrow \mathbf{x}_1^1 = \begin{Bmatrix} -13/12 \\ -2 \end{Bmatrix} \text{e } f(\mathbf{x}_1^1) = 1{,}9166; \mathbf{s}_2^1 = (0,1)^T \Rightarrow \mathbf{x}_2^1 = \begin{Bmatrix} -13/12 \\ -2 \end{Bmatrix} +$$

$$+ \alpha \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -13/12 \\ -2+\alpha \end{Bmatrix} \Rightarrow \text{Min } f \Rightarrow \alpha = 3/8 \Rightarrow \mathbf{x}_2^1 = \begin{Bmatrix} -13/12 \\ -13/8 \end{Bmatrix} \text{ e } f(\mathbf{x}_2^1) = 1{,}3541 \Rightarrow \mathbf{s}_p^1 = \mathbf{x}_2^1 \text{ - } \mathbf{x}_0^1 = \begin{Bmatrix} -1/12 \\ 3/8 \end{Bmatrix} \Rightarrow$$

$$\Rightarrow \mathbf{x}_0^2 = \begin{Bmatrix} -1 \\ -2 \end{Bmatrix} + \alpha \begin{Bmatrix} -1/12 \\ 3/8 \end{Bmatrix} = \begin{Bmatrix} -1-\alpha/12 \\ -2+3\alpha/8 \end{Bmatrix} \Rightarrow \text{Min } f \Rightarrow \alpha = 40/49 \Rightarrow \mathbf{x}_0^2 = \begin{Bmatrix} -157/147 \\ -83/49 \end{Bmatrix} \text{ e } f(\mathbf{x}_0^2) = 1{,}31972$$

$$|\alpha| = \frac{40}{49} < \left[\frac{2 - 1{,}31972}{1{,}9166 - 1{,}3541} \right]^{\frac{1}{2}}$$

Ciclo	x_1	x_2	f
0	−1.0	−2.0	2.0
1	−1.083334	−2.0	1.916667
1	−1.083334	−1.625	1.354167
2	−0.895834	−1.625	0.9322967
2	−0.895834	−1.34375	0.6158854
2	−0.33334	−0.499999	−0.333333

Métodos de Primeira Ordem

Utilizam os valores da função objetivo e de suas derivadas (gradientes) em relação às variáveis de projeto.

Métodos de primeira ordem possuem uma taxa de convergência linear ou superlinear, ou seja, considerando:

$$\left\| \mathbf{x}_{k+1} - \mathbf{x}^* \right\| \leq c_k \left\| \mathbf{x}_k - \mathbf{x}^* \right\|^p$$

a seqüência $\mathbf{x}_k$ apresenta:

- convergência linear para $\mathbf{x}^*$, se p=1 e c_k é constante.
- convergência superlinear para $\mathbf{x}^*$, se p=1 e c_k é uma seqüência que converge para zero.

Métodos de Primeira Ordem

Método "Steepest Descent"

$$\min_{\mathbf{s}} \quad \nabla f^T \mathbf{s} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} s_i$$

$$\text{tal que} \quad \mathbf{s}^T\mathbf{s} = 1$$

$$L(\mathbf{s},\lambda) = \nabla f^T \mathbf{s} + \lambda(\mathbf{s}^T\mathbf{s} - 1) \Rightarrow \frac{\partial L}{\partial \mathbf{s}} = \nabla f + 2\lambda\mathbf{s} = 0 \Rightarrow$$

$$\Rightarrow \mathbf{s} = -\frac{\nabla f}{2\lambda} \Rightarrow \mathbf{s}^T\mathbf{s} = 1 \Rightarrow \|\nabla f\|^2 = 4\lambda^2 \Rightarrow 2\lambda = \|\nabla f\| \Rightarrow$$

Solução: $\boxed{\mathbf{s} = -\frac{\nabla f}{\|\nabla f\|}}$ onde $\| \;\|$ é a norma Euclideana

Assim: $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha\mathbf{s}$ ➡ otimização (busca) unidimensional

Se a função f é quadrática: $f = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c \Rightarrow \alpha^* = -\frac{(\mathbf{x}_k{}^T\mathbf{Q} + \mathbf{b}^T)\mathbf{s}}{\mathbf{s}^T\mathbf{Q}\mathbf{s}}$

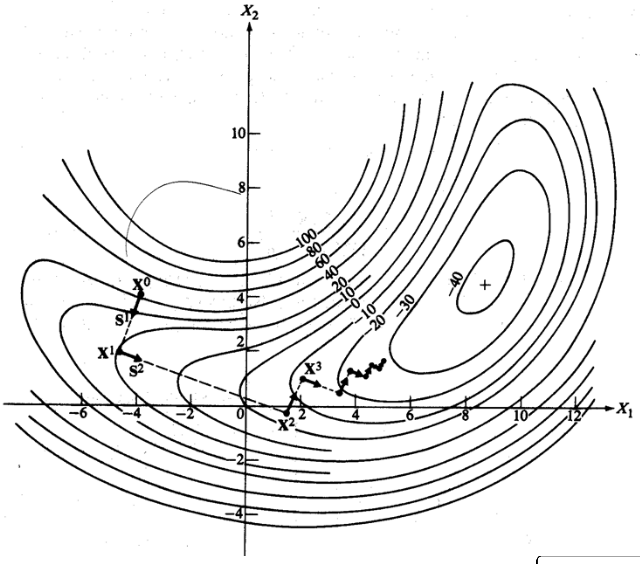
simétrica

Desempenho do Método ➡ Número de Condicionamento (CN) da matriz **Q**:

$CN = \frac{\lambda_{max}}{\lambda_{min}}$ se CN é alto ➡ progresso lento e padrão "zig-zag" (fenômeno "hemstitching")

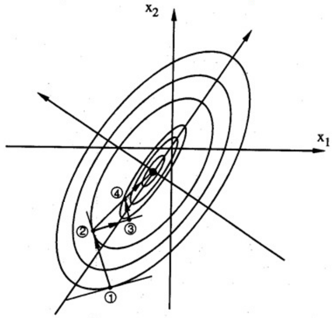
Métodos de Primeira Ordem

Interpretação do Método "Steepest Descent":



Métodos de Primeira Ordem

Agora, vejamos o Exemplo: $f = 12x_1^2 - 12x_1x_2 + 4x_2^2 + 2x_1$

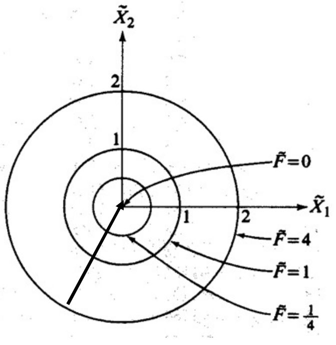


Mas, considerando a transformação de variáveis:

$$y_1 = \left(x_1 - \frac{1}{2} x_2 \right);\ y_2 = \frac{x_2}{\sqrt{12}} \Rightarrow$$

$$\Rightarrow f(y_1, y_2) = y_1^2 + y_2^2 + \frac{1}{6}\left(y_1 + \sqrt{3} y_2\right)$$

$(CN = 1)$



Métodos de Primeira Ordem

Assim, o método apresenta uma taxa linear de convergência sem a transformação de variáveis. Como é difícil encontrar a transformação correta ➡ Método dos Gradientes Conjugados

Método dos Gradientes Conjugados ("Fletcher-Reeves")

Direções $\mathbf{s}_{k+1}$ e $\mathbf{s}_k$ são Q-conjugadas: $\Rightarrow \mathbf{s}^T_{k-1}\mathbf{Q}\mathbf{s}_k = 0$

Algoritmo:

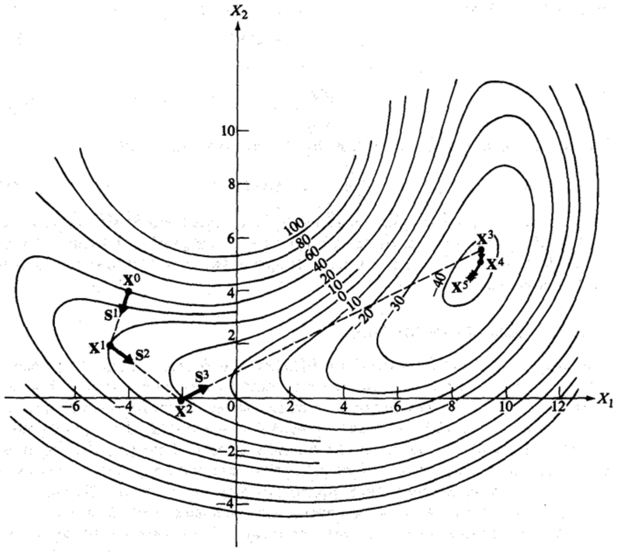
1. Calcular $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_{k+1}\mathbf{s}_k$ onde α_{k+1} é determinado tal que $\frac{df(\alpha_{k+1})}{d\alpha_{k+1}} = 0$
2. $\mathbf{s}_k = \mathbf{g}_k = -\nabla f(\mathbf{x}_k)$ se $k = 0$ e $\mathbf{s}_k = \mathbf{g}_k + \beta_k \mathbf{s}_{k-1}$ se $k > 0$ com

$$\beta_k = \frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}} \quad \text{e} \quad \mathbf{g}_k = -\nabla f(\mathbf{x}_k)$$

3. Se $\|\mathbf{g}_{k+1}\|$ ou $|f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)|$ é suficientemente pequeno, pare.
   Caso Contrário:
4. Se k<n vá para 1, ou recomece.

Métodos de Primeira Ordem

Interpretação do Método dos Gradientes Conjugados:



Exemplo

$$\boxed{\text{Min} \quad f = 12x_1^{\ 2} - 12x_1x_2 + 4x_2^{\ 2} + 2x_1}$$

Solução:

$$\mathbf{x}_0^{\ T} = (-1,-2) \Rightarrow \nabla f(\mathbf{x}_0) = \begin{Bmatrix} 24x_1 - 12x_2 + 2 \\ 8x_2 - 12x_1 \end{Bmatrix}_{\mathbf{x}=\mathbf{x}_0} \Rightarrow \mathbf{s}_0 = -\nabla f(\mathbf{x}_0) = \begin{Bmatrix} -2 \\ 4 \end{Bmatrix} \Rightarrow \mathbf{x}_1 = \begin{Bmatrix} -1 \\ -2 \end{Bmatrix} + \alpha_1 \begin{Bmatrix} -2 \\ 4 \end{Bmatrix} \Rightarrow$$

$$\Rightarrow f(\alpha_1) = 12(-1-2\alpha_1)^2 + 4(-1+4\alpha_1^{\ 2})^2 - 12(-1-2\alpha_1)(-2+4\alpha_1) + 2(-1-2\alpha_1) \Rightarrow \frac{df}{d\alpha_1} = 0 \Rightarrow$$

$$\Rightarrow \alpha_1 = 0{,}048077 \Rightarrow \mathbf{x}_1 = \begin{Bmatrix} -1{,}0961 \\ -1{,}8077 \end{Bmatrix} \text{ e } \nabla f(\mathbf{x}_1) = \begin{Bmatrix} -2{,}6154 \\ -1{,}3077 \end{Bmatrix} \Rightarrow \mathbf{s}_1 = -\nabla f(\mathbf{x}_1) + \beta_1 \mathbf{s}_0 \Rightarrow$$

$$\Rightarrow \beta_1 = \frac{(-2.6154)^2 + (-1.3077)^2}{(-2)^2 + (4)^2} = 0{,}4275 \Rightarrow \mathbf{s}_1 = \begin{Bmatrix} -2{,}6154 \\ -1{,}3077 \end{Bmatrix} + 0{,}4275 \begin{Bmatrix} -2 \\ 4 \end{Bmatrix} = \begin{Bmatrix} 1{,}76036 \\ 3{,}0178 \end{Bmatrix} \text{ e}$$

$$\mathbf{x}_2 = \begin{Bmatrix} -1{,}0961 \\ -1{,}8077 \end{Bmatrix} + \alpha_2 \begin{Bmatrix} 1{,}76036 \\ 3{,}0178 \end{Bmatrix} \Rightarrow \frac{df(\alpha_2)}{d\alpha_2} = 0 \Rightarrow \alpha_2 = 0.4334 \Rightarrow \mathbf{x}_2 = \mathbf{x}^* = \begin{Bmatrix} -0{,}3334 \\ -0{,}5 \end{Bmatrix}$$

$$\text{e} \quad \nabla f(\mathbf{x}_2) = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \Rightarrow \begin{Bmatrix} -2 \\ 4 \end{Bmatrix} \begin{bmatrix} 24 & -12 \\ -12 & 8 \end{bmatrix} \begin{Bmatrix} 1{,}76036 \\ 3{,}0178 \end{Bmatrix} \cong 0$$

Métodos de Segunda Ordem

Método de Newton

A direção **s** é obtida resolvendo-se o problema:

$$\boxed{\begin{array}{ll}\underset{\mathbf{s}}{\text{Min}} & \nabla f^T \mathbf{s} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} s_i \\ \text{tal que} & \mathbf{s}^T\mathbf{Q}\mathbf{s} = 1\end{array}}$$

Solução: $L(\mathbf{s},\lambda) = \nabla f^T\mathbf{s} + \lambda(\mathbf{s}^T\mathbf{Q}\mathbf{s} - 1) \Rightarrow \frac{\partial L}{\partial \mathbf{s}} = \nabla f + 2\lambda\mathbf{Q}\mathbf{s} = 0 \Rightarrow$

$$\Rightarrow \mathbf{s} = -\frac{\mathbf{Q}^{-1}\nabla f}{2\lambda} \Rightarrow \mathbf{s} = -\mathbf{Q}^{-1}\nabla f \Rightarrow \mathbf{Q}\mathbf{s} = -\nabla f$$

deve ser positiva-definida

$$\Rightarrow \mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_{k+1}\mathbf{s}_k$$

$\mathbf{Q}\mathbf{s} = -\nabla f$

- n^3 operações na solução
- $\frac{n(n+1)}{2}$ elementos devem ser calculados para montar Q

➡ Método de Quasi-Newton

Métodos de Segunda Ordem

Método de Quasi-Newton

$$\underbrace{\nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)}_{\mathbf{y}_k} \cong \underbrace{\mathbf{Q}}_{\mathbf{A}_k}(\underbrace{\mathbf{x}_{k+1} - \mathbf{x}_k}_{\mathbf{p}_k}) \Rightarrow \mathbf{y}_k = \mathbf{A}_k\mathbf{p}_k \;\text{ e }\; \boxed{\mathbf{B}_{k+1}\mathbf{y}_k = \mathbf{p}_k}$$

aproximação

relação da secante ou Quasi-Newton

$\mathbf{B}_{k+1}$ é a aproximação de $\mathbf{Q}^{-1}$. Eventualmente: $\mathbf{B}_{k+1}\mathbf{A}_k = \mathbf{I}$

$\mathbf{A}_k$ e $\mathbf{B}_k$ devem permanecer positivo-definidos. Após obter $\mathbf{B}_k$:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_{k+1}\mathbf{s}_k \;\text{ e }\; \mathbf{s}_k = -\mathbf{B}_k\nabla f(\mathbf{x}_k)$$

Aproximação BFGS para $\mathbf{B}_k$: $\mathbf{B}_{k+1} = \left[\mathbf{I} - \frac{\mathbf{p}_k\mathbf{y}_k^T}{\mathbf{p}_k^T\mathbf{y}_k}\right]\mathbf{B}_k\left[\mathbf{I} - \frac{\mathbf{y}_k^T\mathbf{p}_k}{\mathbf{p}_k^T\mathbf{y}_k}\right] + \frac{\mathbf{p}_k\mathbf{p}_k^T}{\mathbf{p}_k^T\mathbf{y}_k}$

Exemplo

$$\boxed{\text{Min} \quad f = 12x_1^{\,2} - 12x_1x_2 + 4x_2^{\,2} + 2x_1} \quad \mathbf{x}_0^{\,T} = (-1,-2); \quad \text{Solução:}$$

Iniciamos com a direção do "Steepest Descent" e $\mathbf{B}_0$=**I** (positiva definida):

$$\mathbf{x}_1 = \begin{Bmatrix} -1.0961 \\ -1.8077 \end{Bmatrix} \text{ e } \nabla f(\mathbf{x}_1) = \begin{Bmatrix} -2.6154 \\ -1.3077 \end{Bmatrix} \Rightarrow \mathbf{p}_0 = \begin{Bmatrix} -1.0961 \\ -1.8077 \end{Bmatrix} - \begin{Bmatrix} -1 \\ -2 \end{Bmatrix} = \begin{Bmatrix} -0.0961 \\ 0.1923 \end{Bmatrix} \text{ e como } \mathbf{s}_0 =$$

$$-\mathbf{B}_0 \nabla f(\mathbf{x}_0) = -\mathbf{I}\nabla f(\mathbf{x}_0) = \begin{Bmatrix} 2 \\ -4 \end{Bmatrix} \Rightarrow \mathbf{y}_0 = \begin{Bmatrix} -2.6154 \\ -1.3077 \end{Bmatrix} - \begin{Bmatrix} 2 \\ -4 \end{Bmatrix} = \begin{Bmatrix} -4.6154 \\ 2.6923 \end{Bmatrix} \Rightarrow \mathbf{p}_0\mathbf{y}_0^{\,T} = 0.96127 \text{ e}$$

$$\mathbf{p}_0^{\,T}\mathbf{y}_0 = \begin{Bmatrix} -0.0961 \\ 0.1923 \end{Bmatrix} \{-4.6154 \quad 2.6923\} = \begin{bmatrix} 0.44354 & -0.25873 \\ -0.88754 & 0.51773 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \mathbf{B}_1 = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{0.96127} \begin{bmatrix} 0.44354 & -0.25873 \\ -0.88754 & 0.51773 \end{bmatrix} \right) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{0.96127} \begin{bmatrix} 0.44354 & -0.88754 \\ -0.25873 & 0.51773 \end{bmatrix} \right)$$

$$+ \frac{1}{0.96127} \begin{bmatrix} 0.00923 & -0.01848 \\ -0.01848 & 0.03698 \end{bmatrix} = \begin{bmatrix} 0.37213 & 0.60225 \\ 0.60225 & 1.10385 \end{bmatrix} \Rightarrow \mathbf{s}_1 = -\begin{bmatrix} 0.37213 & 0.60225 \\ 0.60225 & 1.10385 \end{bmatrix} \begin{Bmatrix} -2.6154 \\ -1.3077 \end{Bmatrix} =$$

$$= \begin{Bmatrix} 1.7608 \\ 3.0186 \end{Bmatrix} \Rightarrow \mathbf{x}_2 = \begin{Bmatrix} -1.0961 \\ -1.8077 \end{Bmatrix} + \alpha_2 \begin{Bmatrix} 1.7608 \\ 3.0186 \end{Bmatrix} \Rightarrow \mathbf{x}_2 = \begin{Bmatrix} -0.3333 \\ -0.5 \end{Bmatrix} \text{ e } \nabla f(\mathbf{x}_2) \cong \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \text{ (convergência)}$$

$$\text{Verifica-se que:} \quad \mathbf{B}_2 = \begin{bmatrix} 0.1667 & 0.25 \\ 0.25 & 0.5 \end{bmatrix} = \mathbf{Q}^{-1}$$

Utilização do SCILAB

Rotina SCILAB para resolver problemas sem restrição: **OPTIM**

[f,[xopt,[gradopt,[work]]]]=optim(costf,[contr],x0,['algo'],[df0,[mem]], [work],[stop],['in']) resolve o problema:

$$\boxed{\begin{array}{l} \text{Min}\ \ \text{costf}(\mathbf{x}) \\ \ \ \mathbf{x} \\ \text{tal que } \mathbf{b}_i \le \mathbf{x} \le \mathbf{b}_s \end{array}}$$

- costf - função a ser minimizada
- x_{opt} - vetor da solução ótima;
- x_0 - chute inicial de x;
- f - valor ótimo da função;
- contr - 'b', $\mathbf{b}_i$, $\mathbf{b}_s$;
- algo - 'qn' (Quasi-Newton); 'gc' (gradientes conjugados); 'nd' (não-diferenciável - método de ordem zero);

Utilização do SCILAB

- mem - número de variáveis usadas na aproximação da matriz Hessiana;
- stop - ‘ar’ (palavra chave); ‘nap’ (máximo número de chamadas da função); ‘iter’ (máximo número de iterações permitida); ‘epsg’ (valor de corte na norma do gradiente); ‘epsf’ (valor de corte na variação de costf); ‘epsx’ (vetor com valores de corte para **x**);
- gradopt - gradiente de costf é fornecido;
- work - vetor para ‘restart’;

Algoritmos de Otimização Não-lineares Sem Restrições

Algumas considerações:

•Solução de Sistemas de Equações Lineares e Não-lineares:

Condição necessária na minimização da energia potencial ➡ Gradiente da energia é nulo ➡ $\nabla f(\mathbf{x}) = g(\mathbf{x}) = 0$

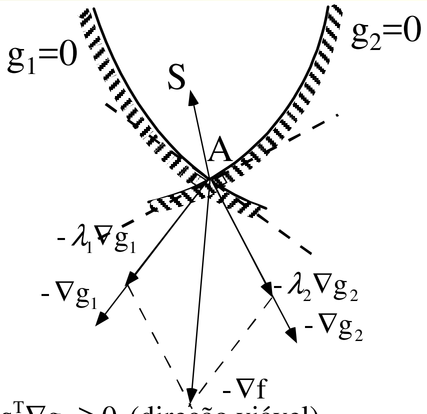
Problema equivalente a: Min $f = \frac{1}{2} g^T g$ ou seja:

$$\mathbf{Ax} = \mathbf{b} \;\equiv\; \text{Min } \frac{1}{2}(\mathbf{Ax}-\mathbf{b})^T(\mathbf{Ax}-\mathbf{b})$$

Na análise não-linear estrutural são obtidos as posições de equilíbrio estáveis e instáveis (se o método não convergir para um mínimo local)

- Algoritmos Probabilísticos:
  - * buscam o mínimo global
  - * ferramentas poderosas para problemas com variáveis discretas
  - * utilizam processo de busca randômica guiada por decisões probabilísticas.

Condições KKT de Optimalidade



$\mathbf{s}^T\nabla g_j \geq 0$ (direção viável)

$\mathbf{s}^T\nabla f < 0$ (direção útil)

$$\text{Como } \mathbf{s}^T\nabla f = \sum_{j=1}^{ng} \lambda_j \mathbf{s}^T\nabla g_j \Rightarrow$$

$\Rightarrow \mathbf{s}^T\nabla f < 0$ é impossível se $\lambda_j \geq 0$

$\mathbf{x}^*$ é mínimo local se:

$$L(\mathbf{x},\lambda) = f(\mathbf{x}) - \sum_{j=1}^{n_g} \lambda_j (g_j - t_j{}^2) \Rightarrow$$

1. $\mathbf{x}^*$ é viável $\Rightarrow g_j(\mathbf{x}^*) \leq 0 \quad j = 1, n_g$

2. $\nabla f(\mathbf{x}^*) - \sum_{j=1}^{n_g} \lambda^*{}_j \nabla g_j(\mathbf{x}^*) = 0;$

   $\lambda^*{}_j \geq 0;$

3. $\lambda^*{}_j g_j(\mathbf{x}^*) = 0 \Rightarrow \therefore$ se $\lambda_j = 0 \Rightarrow$

   $\Rightarrow g_j$ não está ativo

Caso particular:

$$\underbrace{\mathbf{s}^T\nabla f}_{\perp \text{ gradiente}} = \underbrace{\mathbf{s}^T\nabla g_j}_{\textbf{tangente à restrição}} = 0 \Rightarrow \text{altas derivadas}$$

Condições KKT de Optimalidade

KKT são condições necessárias porém não suficientes para optimalidade

KKT é condição suficiente quando:

- se $n_{\text{restrições ativas}} = n_{\text{variáveis}}$:

$$\underbrace{\mathbf{s}^T\overbrace{\nabla g_j}^{\textbf{n direções linearmente independentes}}}_{\textbf{n equações linearmente independentes}} = 0 \quad \text{única solução: } \mathbf{s} = 0$$

Obs.: restrições ativas com gradientes linearmente independentes.

- se $n_{\text{restrições ativas}} \neq n_{\text{variáveis}}$: condição de suficiência necessita de altas derivadas ➡ matriz Hessiana deve ser positiva-definida:

$$\boxed{\mathbf{s}^T(\nabla^2 L)\mathbf{s} > 0 \quad \text{para todo } \mathbf{s} \text{ tal que } \mathbf{s}^T\nabla g_j = 0 \ \ (\text{restrições ativas com } \lambda_j > 0)}$$

- Problemas Convexos

Exemplo

$$\underset{x_1,x_2}{\text{Min}}\ f=-x_1^3-2x_2^2+10x_1-6-2x_2^3$$

tal que
$$g_1=10-x_1x_2\geq 0$$
$$g_2=x_1\geq 0$$
$$g_3=10-x_2\geq 0$$

Solução: Condições KKT:

$$L=-x_1^3-2x_2^2+10x_1-6-2x_2^3-\lambda_1(10-x_1x_2)-\lambda_2x_1+$$
$$-\lambda_3(10-x_2)$$
$$\frac{\partial f}{\partial x_1}+\lambda_1\frac{\partial g_1}{\partial x_1}+\lambda_2\frac{\partial g_2}{\partial x_1}=-3x_1^2+10+\lambda_1x_2-\lambda_2=0$$
$$\frac{\partial f}{\partial x_2}+\lambda_1\frac{\partial g_1}{\partial x_2}+\lambda_3\frac{\partial g_3}{\partial x_2}=-4x_2-6x_2^2+\lambda_1x_1-\lambda_3=0$$

Casos possíveis a serem analisados:

	λ_1	λ_2	λ_3	
Caso 1	0	0	0	
Caso 2	$\neq 0$	0	$\neq 0$	
Caso 3	$\neq 0$	0	0	
Caso 4	0	$\neq 0$	0	
Caso 5	0	0	$\neq 0$	
Caso 6	0	$\neq 0$	$\neq 0$	
Caso 7	$\neq 0$	$\neq 0$	$\neq 0$	→ Não é possível

$$n_{\substack{\text{restrições}\\ \text{ativas}}}\neq n_{\text{variáveis}}$$?

Matrix Hessiana:

$$\frac{\partial^2 L}{\partial x_1{}^2}=-6x_1;\quad \frac{\partial^2 L}{\partial x_2{}^2}=-4-12x_2;$$
$$\frac{\partial^2 L}{\partial x_1\partial x_2}=-\lambda_1=\frac{\partial^2 L}{\partial x_2\partial x_1}\quad\Rightarrow\nabla^2L=\begin{bmatrix}-6x_1 & -\lambda_1\\ -\lambda_1 & -4-12x_2\end{bmatrix}$$

Exemplo

Caso 1: $x_1=1{,}826;\ x_2=0\Rightarrow f=6{,}17\Rightarrow\nabla^2L=\begin{bmatrix}-6x_1 & \lambda_1\\ \lambda_1 & -4-12x_2\end{bmatrix}$ **negativa definida** ➡ **Ponto de máximo**

Caso 2: $x_1=1;\ x_2=10;\ \lambda_1=-0{,}7;\ \lambda_3=639{,}3$ **Nem mínimo, nem máximo**

Caso 3: $\begin{cases}-3x_1^2+10+\lambda_1x_2=0\\ -4x_2-6x_2^2+\lambda_1x_1=0\\ x_1x_2=10\end{cases}\Rightarrow\begin{matrix}x_1=3{,}847\\ x_2=2{,}599\\ \lambda_1=13{,}24\\ f=-73{,}08\end{matrix}\Rightarrow\nabla^2L=\begin{bmatrix}-23{,}08 & 13{,}24\\ 13{,}24 & -35{,}19\end{bmatrix}$ **negativa definida**

Caso 4: $\begin{cases}-3x_1^2+10-\lambda_2=0\\ -4x_2-6x_2^2=0\end{cases}\Rightarrow\begin{matrix}x_1=0 & x_2=0 & \lambda_2=10 & f=-6\\ x_1=0 & x_2=-2/3 & \lambda_2=10 & f=-6{,}99\end{matrix}$ **$\Rightarrow s^T\nabla^2Ls<0$ Não é mínimo**

Caso 5: $\begin{cases}-3x_1^2+10=0\\ -4x_2-6x_2^2-\lambda_3=0\end{cases}\Rightarrow x_1=1{,}826\quad x_2=10\quad\lambda_3=640\quad f=-2194\Rightarrow\nabla^2L$ **negativa definida**

Caso 6: $x_1=0;\ x_2=10;\ \lambda_2=10;\ \lambda_3=640;\ f=-2206$

KKT é satisfeita e n° de restrições ativas = n° de variáveis de projeto

Cálculo dos Multiplicadores de Lagrange

Importante para verificação das condições KKT da solução ótima.

Seja a condição necessária de optimalidade: $\nabla f - \mathbf{N}\lambda = 0$

onde: $n_{ij} = \dfrac{\partial g_j}{\partial x_i}$ $j = 1,...,r$ (restrição ativas)

$i = 1,...,n$ (número de variáveis de projeto)

$r < n \Rightarrow n^\circ$ de incógnitas $(\lambda) < n^\circ$ de equações

Solução: Mínimos Quadrados

$$u = \mathbf{N}\lambda - \nabla f \Rightarrow \text{Min } \|u\|^2 = (\mathbf{N}\lambda - \nabla f)^T(\mathbf{N}\lambda - \nabla f) \Rightarrow$$

$$\Rightarrow 2\mathbf{N}^T\mathbf{N}\lambda - 2\mathbf{N}^T\nabla f = 0 \Rightarrow \boxed{\lambda = \left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{N}^T\nabla f}$$

Se a condição de estacionaridade é satisfeita a solução será exata, assim:

$$\mathbf{P}\nabla f = 0, \quad \text{onde} \quad \mathbf{P} = \mathbf{I} - \mathbf{N}\left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{N}^T$$

P - Matriz de Projeção

Exemplo

$$\underset{x_1,x_2,x_3}{\text{Min}} \quad f = x_1 + x_2 + x_3$$

tal que
$$g_1 = 8 - x_1^2 - x_2^2 \geq 0$$
$$g_2 = x_3 - 4 \geq 0$$
$$g_3 = x_2 + 8 \geq 0$$

Checar se (-2,-2,4) é mínimo local:

$$g_1 = 0; \quad g_2 = 0; \quad g_3 = 6 \Rightarrow$$

$$\left.\begin{array}{l} \dfrac{\partial g_1}{\partial x_1} = -2x_1 = 4; \dfrac{\partial g_1}{\partial x_2} = -2x_2 = 4; \dfrac{\partial g_1}{\partial x_3} = 0 \\ \dfrac{\partial g_2}{\partial x_1} = 0; \ \dfrac{\partial g_2}{\partial x_2} = 0; \ \dfrac{\partial g_2}{\partial x_3} = 1 \end{array}\right\} \Rightarrow \mathbf{N} = \begin{bmatrix} 4 & 0 \\ 4 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial x_3} = 1 \Rightarrow \nabla f = \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix}; \ \mathbf{N}^T\mathbf{N} = \begin{bmatrix} 32 & 0 \\ 0 & 1 \end{bmatrix}; \ \mathbf{N}^T\nabla f = \begin{Bmatrix} 8 \\ 1 \end{Bmatrix} \Rightarrow \lambda = \left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{N}^T\nabla f = \begin{Bmatrix} 1/4 \\ 4 \end{Bmatrix}$$

$\mathbf{P}\nabla f = \left[\mathbf{I} - \mathbf{N}\left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{N}^T\right]\nabla f = 0$ e todos os multiplicadores de Lagrange são positivos, mas como $r \neq n$ ➡ matriz Hessiana deve ser obtida para verificar KKT

Note que:

$$\nabla f - \mathbf{N}\lambda = 0 \Rightarrow \begin{cases} \left.\begin{array}{l} 1 - 4\lambda_1 = 0 \\ 1 - 4\lambda_1 = 0 \end{array}\right\} \textbf{linearmente dependentes} \\ 1 - \lambda_2 = 0 \end{cases}$$

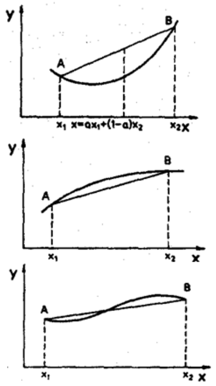
Problemas Convexos

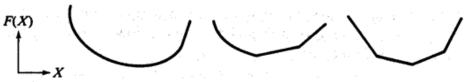
Problema covexo ➡ Função objetivo e domínio viável são convexos

Função Convexa

$$f[\alpha \mathbf{x}_2 + (1-\alpha)\mathbf{x}_1] \le \alpha f(\mathbf{x}_2) + (1-\alpha) f(\mathbf{x}_1), \quad 0 < \alpha < 1$$

Exemplos:





Funções Convexas

Funções Côncavas

Funções nem convexas nem côncavas

Função Convexa ➡ Matriz Hessiana é positiva semi-definida;

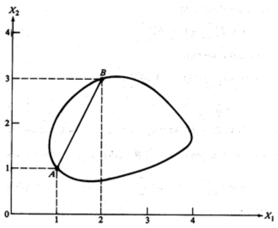
Função Côncava ➡ Matriz Hessiana é negativa semi-definida;

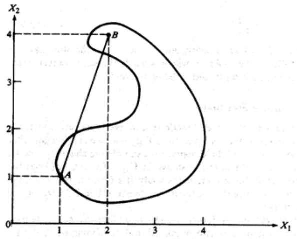
Função Linear é convexa.

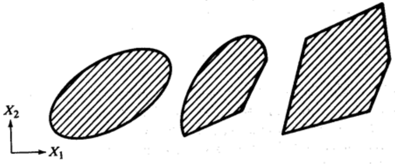
Problemas Convexos

Domínio Convexo

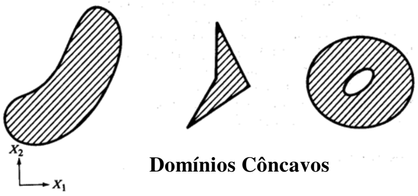
Seja: $\mathbf{w} = \alpha \mathbf{x}_2 + (1-\alpha)\mathbf{x}_1, \quad 0 < \alpha < 1$ Se $\mathbf{w} \subset$ domínio para $\forall \mathbf{x}_1, \mathbf{x}_2 \Rightarrow$ domínio é convexo







Domínios Convexos

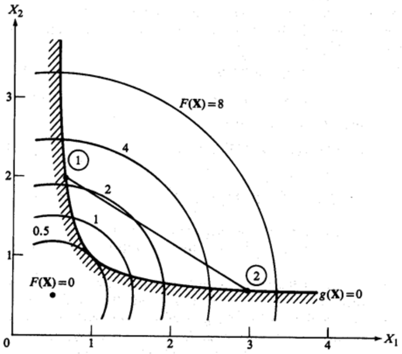


Domínios Côncavos

$$\text{Domínio é convexo se} \begin{cases} \text{Todas as restrições de inegualdade são côncavas} \\ \text{Todas as restrições de igualdade são lineares} \end{cases}$$

Problemas Convexos

Exemplo de Problema Convexo:



Vantagem dos problemas convexos:

- Ótimo Global
- KKT são também condições suficientes

A maior parte dos problemas são não convexos ➡ aproximados por uma série de problemas convexos, por exemplo, programação linear.

Exemplo

Verificar se o problema é convexo:

$$\underset{x_1,x_2}{\text{Min}} \quad f = 3x_1 + \sqrt{3}x_2$$

tal que

$$g_1 = 3 - \frac{18}{x_1} - \frac{6\sqrt{3}}{x_2} \geq 0$$

$$g_2 = x_1 - 5{,}73 \geq 0$$

$$g_3 = x_2 - 7{,}17 \geq 0$$

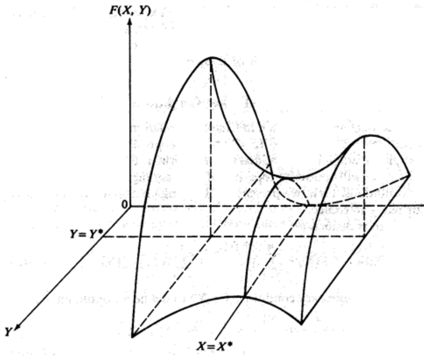
$f, g_2, g_3 \Rightarrow$ Funções Lineares

Função g_1: Matriz Hessiana vale $\mathbf{A}_1 = \begin{bmatrix} -36/x_1^3 & 0 \\ 0 & -12\sqrt{3}/x_2^3 \end{bmatrix}$

Assim a matriz Hessiana é negativa definida ➡ g_1 é côncava

Portanto como a função objetivo é linear (convexa) e as restrições são lineares ou côncavas (domínio convexo) ➡ problema é convexo.
Assim, KKT são condições suficientes para o mínimo.

Ponto Sela ("Saddle Point")



Uma função F(**x**,**y**) possue ponto "sela" se:

$$F(\mathbf{x}^*,\mathbf{y}) \le F(\mathbf{x}^*,\mathbf{y}^*) \le F(\mathbf{x},\mathbf{y}^*)$$

$$F(\mathbf{x}^*,\mathbf{y}^*)\begin{cases}\text{mínimo em }\mathbf{x}\\ \text{máximo em }\mathbf{y}\end{cases}$$

No caso da função Lagrangeana: $L(\mathbf{x}^*,\lambda^*)$

$$\begin{cases}\nabla_x L(\mathbf{x}^*,\lambda^*)=0\\ \nabla_\lambda L(\mathbf{x}^*,\lambda^*)=0\end{cases} \quad \text{e} \quad \begin{cases}\nabla_\lambda L(\mathbf{x}^*,\lambda)\le 0\\ \nabla_x L(\mathbf{x},\lambda^*)\ge 0\end{cases}$$

Solução do Ponto Sela ➡ Problema Min Max:

$$L(\lambda)=\min_{\mathbf{x}} L(\mathbf{x},\lambda) \Rightarrow \max_{\lambda} L(\lambda) = \max_{\lambda}\min_{\mathbf{x}} L(\mathbf{x},\lambda)$$

ou:

$$L(\mathbf{x})=\max_{\lambda} L(\mathbf{x},\lambda) \Rightarrow \min_{\mathbf{x}} L(\mathbf{x}) = \min_{\mathbf{x}}\max_{\lambda} L(\mathbf{x},\lambda)$$

$$\boxed{\max_{\lambda}\min_{\mathbf{x}} L(\mathbf{x},\lambda) \equiv \min_{\mathbf{x}}\max_{\lambda} L(\mathbf{x},\lambda)}$$

Exemplo

Ex.:

$$\begin{aligned}&\min_{x}\ \frac{1}{x}\\ &\text{tal que }\ x-1\le 0\\ &\qquad\qquad x\ge 0\end{aligned}$$

O Lagrangeano do problema vale:

$$L(x,\lambda)=\frac{1}{x}+\lambda(x-1)$$

Calculando o mínimo em x: $\nabla_x L(x,\lambda)=0 \Rightarrow -\frac{1}{x^2}+\lambda=0 \Rightarrow$

$$\Rightarrow x=\pm\frac{1}{\sqrt{\lambda}}\,(\text{mas } x\ge 0) \Rightarrow x=\frac{1}{\sqrt{\lambda}} \Rightarrow L(\lambda)=\sqrt{\lambda}+\lambda\left(\frac{1}{\sqrt{\lambda}}-1\right)=2\sqrt{\lambda}-\lambda$$

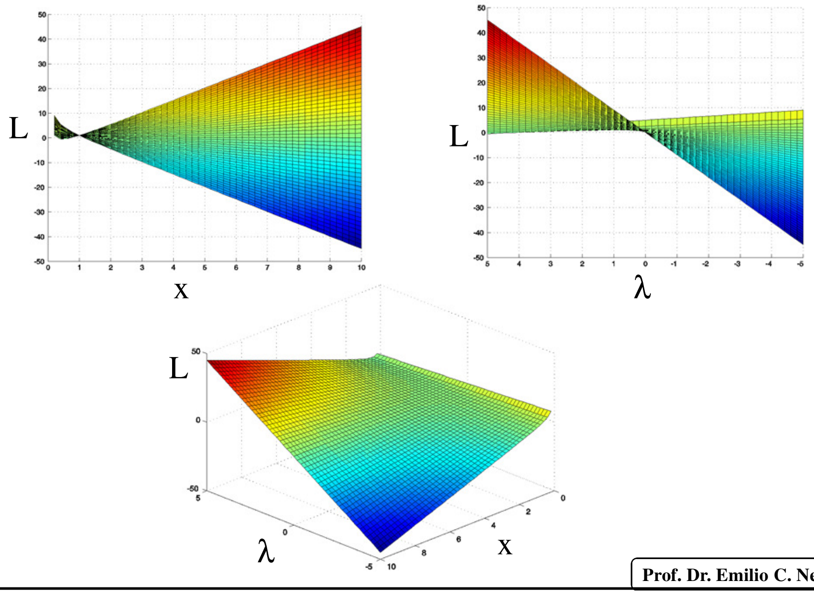
Calculando o máximo em λ: $\nabla_\lambda L(\lambda)=\frac{1}{\sqrt{\lambda}}-1 \Rightarrow \lambda^*=1 \Rightarrow x^*=1$

Ponto "Sela": $\boxed{\therefore (x^*,\lambda^*)=(1,1)}$

Os gradientes devem ser contínuos em **x** e λ , caso contrário, outra abordagem de solução para o problema Min Max deve ser utilizada.

Exemplo

Plotagem da função Lagrangeana:



Dualidade

“Primal”

$$\underset{\mathbf{x}}{\text{Min}} \quad f_p = \mathbf{c}^T\mathbf{x}$$
$$\text{tal que} \quad \mathbf{Ax} \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}$$

m restrições
n variáveis de projeto
m>n

$$\equiv$$

$$\boxed{f_{p\,\min} = f_{D\,\max}}$$

“Dual”

$$\underset{\lambda}{\text{Max}} \quad f_D = -\lambda^T\mathbf{b}$$
$$\text{tal que} \quad \mathbf{A}^T\lambda \geq -\mathbf{c}, \quad \lambda \geq \mathbf{0}$$

n restrições
m variáveis de projeto
m>n

Demonstração:

$\mathbf{x}^*$ **é mínimo**

$$L(\mathbf{x},\lambda) = \mathbf{c}^T\mathbf{x} + \lambda^T\{\mathbf{Ax} - \mathbf{b}\} \Rightarrow \nabla_{\mathbf{x}} L(\mathbf{x},\lambda) = \mathbf{c} + \mathbf{A}^T\lambda \boxed{\geq 0} \Rightarrow \mathbf{cx}^* + \lambda^T\mathbf{Ax}^* = 0 \Rightarrow$$

$L(\lambda) = -\lambda^T\mathbf{b}$ deve ser maximizada ➡ problema “Dual” e: $f(\mathbf{x}^*) = L(\lambda^*)$

Se: $\lambda_j \neq 0 \Rightarrow j^{\text{ésima}}$ restrição é ativa no problema “primal”
$\lambda_j = 0 \Rightarrow j^{\text{ésima}}$ restrição é inativa no problema “primal”

Exemplo

$$\min_{x_1,x_2} \; f = -4x_1 - x_2 + 50$$

tal que

$$\begin{aligned} &x_1 - x_2 \le 2\\ &x_1 + 2x_2 \le 8\\ &x_1 + x_2 \le 10\\ &-5x_1 + x_2 \le 5\\ &x_1 \ge 0;\;\; x_2 \ge 0\end{aligned}$$

➡ $\mathbf{c} = \begin{Bmatrix}-4\\-1\end{Bmatrix};\ \mathbf{A} = \begin{bmatrix}1 & -1\\ 1 & 2\\ 1 & 1\\ -5 & 1\end{bmatrix};\ \mathbf{b} = \begin{Bmatrix}2\\8\\10\\5\end{Bmatrix}$ ➡

➡

$$\max_{\lambda_1,\lambda_2,\lambda_3,\lambda_4} \; L = -2\lambda_1 - 8\lambda_2 - 10\lambda_3 - 5\lambda_4 + 50$$

tal que

$$\begin{aligned} &\lambda_1 + \lambda_2 + \lambda_3 - 5\lambda_4 \ge 4\\ &-\lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4 \ge 1\\ &\lambda_i \ge 0;\;\; i = 1,4\end{aligned}$$

➡ $\lambda^* = \begin{Bmatrix}7/3\\5/3\\0\\0\end{Bmatrix} \Rightarrow \boxed{L(\lambda^*) = 32}$ ➡

➡

$$\begin{cases} \nabla_x L(\mathbf{x}^*, \lambda^*) = \mathbf{c} + \mathbf{A}^T\lambda = 0 \Rightarrow \begin{Bmatrix}-4\\-1\end{Bmatrix} + \begin{bmatrix}1 & 1 & 1 & -5\\ -1 & 2 & 1 & 1\end{bmatrix}\begin{Bmatrix}7/3\\5/3\\0\\0\end{Bmatrix} = \begin{Bmatrix}0\\0\end{Bmatrix} \Rightarrow \mathbf{x}^* = \begin{Bmatrix}4\\2\end{Bmatrix} \Rightarrow \boxed{f(\mathbf{x}^*) = 32}\\ x^*_1 - x^*_2 = 2; \qquad x^*_1 + 2x^*_2 = 8 \end{cases}$$

Programação Quadrática

Um problema de Programação Quadrática é definido como:

$$\min_{\mathbf{x}} \; F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}\mathbf{x}$$

tal que $\mathbf{B}\mathbf{x} \le \mathbf{c}$

$\mathbf{x} \ge \mathbf{0}$

Matriz Hessiana

Suponha
n variáveis de projeto
e m restrições

O Lagrangeano do problema é igual à:

$$L(\mathbf{x},\lambda) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}\mathbf{x} + \lambda^T(\mathbf{B}\mathbf{x} - \mathbf{c}) \Rightarrow \nabla_x L(\mathbf{x},\lambda) = \mathbf{A}\mathbf{x} + \mathbf{b} + \mathbf{B}^T\lambda = 0 \Rightarrow$$

$$\Rightarrow \mathbf{x} = -\mathbf{A}^{-1}(\mathbf{b} + \mathbf{B}^T\lambda) \Rightarrow L(\lambda) = -\frac{1}{2}\lambda^T\mathbf{D}\lambda - \mathbf{d}.\lambda - \frac{1}{2}\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b} \longrightarrow \text{Termo constante}$$

onde: $\mathbf{D} = \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T$ e $\mathbf{d} = \mathbf{c} + \mathbf{B}\mathbf{A}^{-1}\mathbf{b}$

Assim resolve-se o novo problema usando o conceito da dualidade:

$$\max_{\lambda} \; L(\lambda) = -\frac{1}{2}\lambda^T\mathbf{D}\lambda - \mathbf{d}.\lambda$$

$$(\lambda \ge \mathbf{0})$$

m variáveis de projeto
com simples restrições

Utilização do SCILAB

Rotina SCILAB para resolver programação quadrática: **QUAPRO**

[x,lagr,f]=quapro(Q,p,C,b,ci,cs,mi,x0) resolve o problema:

$$\underset{\mathbf{x}}{\text{Min}}\ f = \frac{1}{2}\mathbf{x}^T\overset{\text{matriz simétrica (nxn)}}{\overset{\uparrow}{\mathbf{Q}}}\mathbf{x}+\mathbf{p}^T\mathbf{x}$$

$$\text{tal que}\quad \underset{\searrow(m_i\times n)}{\mathbf{C}_1}\mathbf{x}=\mathbf{b}_1$$

$$\underset{\searrow(m_d\times n)}{\mathbf{C}_2}\mathbf{x}\le\mathbf{b}_2$$

$$\mathbf{c}_i\le\mathbf{x}\le\mathbf{c}_s$$

$$\underset{\searrow(m_i+m_d)\times n}{\mathbf{C}}=[\mathbf{C}_1,\mathbf{C}_2]$$

$$\underset{\searrow 1\times(m_i+m_d)}{\mathbf{b}}=[\mathbf{b}_1,\mathbf{b}_2]$$

- x - vetor da solução ótima;
- lagr - vetor de multiplicadores de Lagrange ($n+m_i+m_d$). Se o valor de um multiplicador é zero, a restrição correspondente está inativa, caso contrário está ativa na ponto ótimo;
- f - valor ótimo da função;
- m_i - número de restrições de igualdade;
- x_0 - chute inicial de x; ou "x_0=v" o cálculo é iniciado num vértice do domínio; ou "x_0=g" o valor inicial é arbitrário;

Algoritmos de Otimização Não-lineares Com Restrições

Classificação

- Métodos Diretos ➡ verificam as restrições após seguirem o procedimento do algoritmo de otimização sem restrição. Ex.: Método das Projeções do Gradientes e dos Gradientes Reduzidos e Método das Direções Viáveis
- Métodos Indiretos ➡ transformam o problema com restrição num problema sem restrição. Ex.: Métodos de Penalização e Método de Lagrange Aumentado

Métodos Diretos

Método das Projeções do Gradientes e dos Gradientes Reduzidos

1) Inicialmente consideremos restrições lineares:

$$\min_{\mathbf{x}} \ f(\mathbf{x})$$
$$\text{tal que } g_j = \mathbf{a}_j^T\mathbf{x} - \mathbf{b}_j \geq 0 \quad j = 1,\ldots,n_g$$

Considerando *r* restrições ativas:

$\mathbf{g}_a = \mathbf{N}^T\mathbf{x} - \mathbf{b} = 0$ onde: $n_{ij} = a_{ij}$

Hipótese básica do método: $\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha\mathbf{s}$

Caminha na fronteira das restrições

x se encontra no subespaço tangente às restrições ativas, assim:

$$\mathbf{N}^T\mathbf{x}_{i+1} = \mathbf{N}^T\mathbf{x}_i + \alpha\mathbf{N}^T\mathbf{s} \Rightarrow \mathbf{b} = \mathbf{b} + \alpha\mathbf{N}^T\mathbf{s} \Rightarrow \mathbf{N}^T\mathbf{s} = 0$$

A direção **s** é obtida através da solução do problema:

$$\min_{\mathbf{s}} \ \mathbf{s}^T\nabla f$$
$$\text{tal que } \mathbf{N}^T\mathbf{s} = 0, \quad \mathbf{s}^T\mathbf{s} = 1$$

$$L = \mathbf{s}^T\nabla f - \mathbf{s}^T\mathbf{N}\lambda - \mu\left(\mathbf{s}^T\mathbf{s} - 1\right) \Rightarrow \frac{\partial L}{\partial \mathbf{s}} = \nabla f - \mathbf{N}\lambda - 2\mu\mathbf{s} = 0 \Rightarrow$$

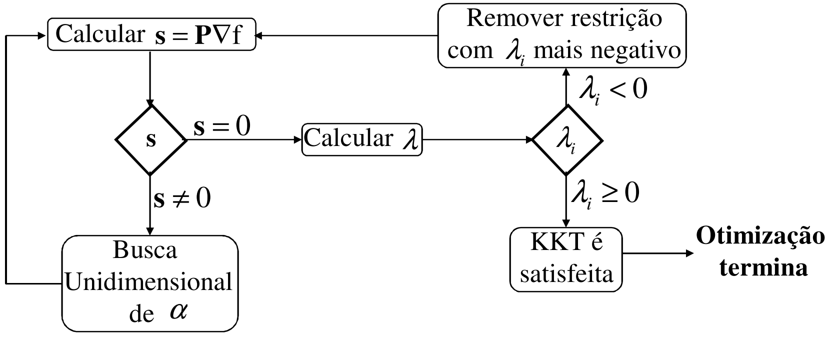
$$\Rightarrow \mathbf{N}^T\nabla f - \mathbf{N}^T\mathbf{N}\lambda = 0 \Rightarrow \lambda = \left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{N}^T\nabla f \Rightarrow$$

$$\Rightarrow s = \frac{1}{2\mu}\left[\mathbf{I} - \mathbf{N}\left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{N}^T\right]\nabla f = \frac{1}{2\mu}\mathbf{P}\nabla f$$

onde para **w** arbitrário, **Pw** é o subespaço tangente às restrições ativas, pois: $\mathbf{N}^T\mathbf{P}\mathbf{w} = \mathbf{N}^T[\mathbf{I} - \mathbf{N}\left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{N}^T]\mathbf{w} = 0$

Método das Projeções do Gradientes e dos Gradientes Reduzidos

Fluxograma do método:



Verificação da violação das restrições:

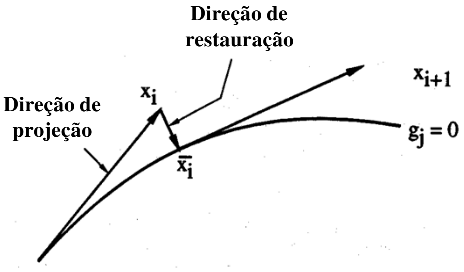
$$g_j = \mathbf{a}_j^T\left(\mathbf{x}_i + \alpha\mathbf{s}\right) - b_j \geq 0 \Rightarrow \alpha \leq -\frac{\left(\mathbf{a}_j^T\mathbf{x}_i - b_j\right)}{\mathbf{a}_j^T\mathbf{s}} = -\frac{g_j(\mathbf{x}_i)}{\mathbf{a}_j^T\mathbf{s}}$$

Válido se $\mathbf{a}_j^T\mathbf{s} < 0$, caso contrário não há valor limite superior de α devido à j-ésima restrição. O valor limite de α vale:

$$\overline{\alpha} = \min_{\substack{\alpha_j > 0 \\ restr.ativas}} \alpha_j$$

Método das Projeções do Gradientes e dos Gradientes Reduzidos

2) Considerando restrições não-lineares
Lineariza-se a restrição em torno de x_i ➡ busca unidimensional move para fora da fronteira da restrição ➡ correção



Correção:

$$g_j \approx g_j(\mathbf{x}_i) + \nabla g_j^T(\bar{\mathbf{x}}_i - \mathbf{x}_i)$$

Verifica-se que:

$$\bar{\mathbf{x}}_i - \mathbf{x}_i = -\mathbf{N}\left(\mathbf{N}^T\mathbf{N}\right)^{-1}\boxed{\mathbf{g}_a(\mathbf{x}_i)} \rightarrow \text{restrições ativas}$$

α é determinado especificando uma redução γ na função objetivo:

$$\boxed{f(\mathbf{x}_i) - f(\mathbf{x}_{i+1}) \cong \nabla f^T\left(\mathbf{x}_i - \mathbf{x}_{i+1}\right)} \approx \gamma f(\mathbf{x}_i) \Rightarrow \alpha^* = -\frac{\gamma f(\mathbf{x}_i)}{\mathbf{s}^T\nabla f} \Rightarrow$$

$$\Rightarrow \boxed{\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha^*\mathbf{s} - \mathbf{N}\left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{g}_a}$$

Exemplo

$$\underset{x_1, x_2}{\text{Min}} \quad f = 3x_1 + \sqrt{3}x_2$$

tal que

$$g_1 = 3 - \frac{18}{x_1} - \frac{6\sqrt{3}}{x_2} \geq 0$$

$$g_2 = x_1 - 5.73 \geq 0$$

$$g_3 = x_2 - 7.17 \geq 0$$

$$\mathbf{x}_0 = (11{,}61;7{,}17) \Rightarrow g_1 = 0,\ g_3 = 0 \Rightarrow f = 47{,}25 \Rightarrow$$

$$\Rightarrow \nabla f = \begin{Bmatrix} 3 \\ \sqrt{3} \end{Bmatrix};\ \nabla g_1 = \begin{Bmatrix} 0{,}1335 \\ 0{,}2021 \end{Bmatrix};\ \nabla g_3 = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix};\ N = \begin{bmatrix} 0{,}1335 & 0 \\ 0{,}2021 & 1 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \mathbf{P} = 0 \Rightarrow \mathbf{s} = 0 \Rightarrow \lambda = \begin{Bmatrix} 22{,}47 \\ -2{,}798 \end{Bmatrix} \Rightarrow$$

elimina-se g_3, assim:

$$N = \begin{bmatrix} 0{,}1335 \\ 0{,}2021 \end{bmatrix} \Rightarrow \mathbf{P} = \begin{bmatrix} 0{,}6962 & -0{,}46 \\ -0{,}46 & 0{,}3036 \end{bmatrix} \Rightarrow \mathbf{s} = -\mathbf{P}\nabla f = \begin{Bmatrix} -1{,}29 \\ 0{,}854 \end{Bmatrix}$$

Para 5% de redução na função objetivo (γ=0,05): $\alpha^* = \dfrac{0{,}05 \times 47{,}25}{\begin{bmatrix} -1{,}29 & 0{,}854 \end{bmatrix}\begin{Bmatrix} 3 \\ \sqrt{3} \end{Bmatrix}} = 0{,}988$

Como não há violação da restrição ➡ não há correção:

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha^*\mathbf{s} = \begin{Bmatrix} 11{,}61 \\ 7{,}17 \end{Bmatrix} + 0{,}988\begin{Bmatrix} -1{,}29 \\ 0{,}854 \end{Bmatrix} = \begin{Bmatrix} 10{,}34 \\ 8{,}01 \end{Bmatrix} \Rightarrow f(\mathbf{x}_1) = 44{,}89;$$

$g(\mathbf{x}_1) = -0{,}0382 \Rightarrow$ Restrição g_1 ativa ➡ $\mathbf{N} = \nabla g_1 = \begin{Bmatrix} 0{,}1684 \\ 0{,}1620 \end{Bmatrix} \Rightarrow$

$$\Rightarrow \mathbf{P} = \begin{bmatrix} 0{,}4806 & -0{,}4996 \\ -0{,}4996 & 0{,}5194 \end{bmatrix} \Rightarrow$$

Exemplo

$$\Rightarrow \mathbf{s} = -\mathbf{P}\nabla \mathrm{f} = \begin{Bmatrix} -0{,}5764 \\ 0{,}5991 \end{Bmatrix};\ \gamma = 0{,}025 \Rightarrow \alpha = -\frac{0{,}025 \times 44{,}89}{\begin{bmatrix} -0{,}5764 & 0{,}599 \end{bmatrix} \begin{Bmatrix} 3 \\ \sqrt{3} \end{Bmatrix}} = 1{,}62$$

Como g_1 é violada ➡ há correção

$$\mathbf{g}_a = -0{,}0382 \Rightarrow -\mathbf{N}\left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{g}_a = \begin{Bmatrix} 0{,}118 \\ 0{,}113 \end{Bmatrix} \Rightarrow \mathbf{x}_2 = \mathbf{x}_1 + \alpha^* \mathbf{s} - \mathbf{N}\left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{g}_a = \begin{Bmatrix} 10{,}34 \\ 8{,}01 \end{Bmatrix} +$$

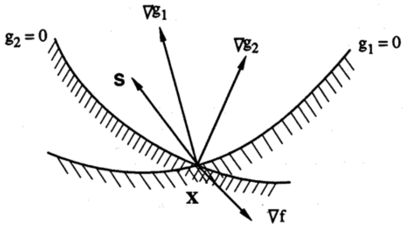
$$-1{,}62\begin{Bmatrix} 0{,}576 \\ -0{,}599 \end{Bmatrix} + \begin{Bmatrix} 0{,}118 \\ 0{,}113 \end{Bmatrix} = \begin{Bmatrix} 9{,}52 \\ 9{,}10 \end{Bmatrix} \Rightarrow \mathrm{f}(\mathbf{x}_2) = 44{,}32;\ \mathrm{g}_1(\mathbf{x}_2) = \text{-}0{,}0328$$

$$\mathbf{x}^* = \begin{Bmatrix} 9{,}464 \\ 9{,}464 \end{Bmatrix} \Rightarrow \mathrm{f}(\mathbf{x}^*) = 44{,}78 \qquad \text{(Solução Ótima)}$$

Métodos Diretos

Método das Direções Viáveis

Procura se afastar da fronteira das restrições. Inicia-se na fronteira do domínio viável.



restrições ativas

Direção viável: $\mathbf{s}^T\boxed{\nabla g_j} > 0$

Direção útil: $\mathbf{s}^T\nabla f < 0$

Assim, a direção **s** é obtida através da solução do problema de otimização:

Programação Linear

$$\boxed{\begin{array}{ll} \underset{\mathbf{s}}{\text{Min}} & \beta \\ \text{tal que} & -\mathbf{s}^T\nabla g_j + \theta_j \beta \le 0 \\ & \mathbf{s}^T\nabla f + \beta \le 0 \qquad \theta_j \ge 0 \\ & |s_i| \le 1 \end{array}}$$

Se $\beta_{max} = 0 \Rightarrow$ condições KKT são satisfeitas.

Exemplo

$$\underset{x_1,\,x_2}{\text{Min}} \quad f = 3x_1 + \sqrt{3}x_2$$

$$\text{tal que} \quad g_1 = 3 - \frac{18}{x_1} - \frac{6\sqrt{3}}{x_2} \geq 0$$

$$g_2 = x_1 - 5.73 \geq 0$$

$$g_3 = x_2 - 7.17 \geq 0$$

$$\mathbf{x}_0^T = (11{,}61;7{,}17) \Rightarrow g_1 = 0,\ g_3 = 0 \Rightarrow f = 47{,}25 \Rightarrow$$

$$\Rightarrow \nabla f = \begin{Bmatrix} 3 \\ \sqrt{3} \end{Bmatrix};\ \nabla g_1 = \begin{Bmatrix} 0{,}1335 \\ 0{,}2021 \end{Bmatrix};\ \nabla g_3 = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$$

Selecionando $\theta_1 = \theta_2 = 1$,tem-se:

$$\underset{s_1,s_2}{\text{Min}} \quad \beta$$

$$\text{tal que} \quad -0{,}1335s_1 - 0{,}2021s_2 + \beta \leq 0$$

$$-s_2 + \beta \leq 0$$

$$3s_1 + \sqrt{3}s_2 + \beta \leq 0$$

$$-1 \leq s_1 \leq 1$$

$$-1 \leq s_2 \leq 1$$

Solução: $s_1 = -0{,}6172;\ s_2 = 1 \Rightarrow$

Busca unidimensional:

$$\Rightarrow \mathbf{x}_1 = \begin{Bmatrix} 11{,}61 \\ 7{,}17 \end{Bmatrix} + \alpha \begin{Bmatrix} -0{,}6172 \\ 1 \end{Bmatrix}$$

α é limitado por $g_1 \Rightarrow \alpha < 5{,}385 \Rightarrow$

$$\Rightarrow \alpha = 5{,}385 \Rightarrow \mathbf{x}_1 = \begin{Bmatrix} 8{,}29 \\ 12{,}56 \end{Bmatrix} \Rightarrow f(\mathbf{x}_1) = 46{,}62$$

Próxima iteração (somente g_1 ativo):

$$\nabla g_1 = \begin{Bmatrix} -\dfrac{18}{x_1^2} \\ -\dfrac{6\sqrt{3}}{x_2^2} \end{Bmatrix} = \begin{Bmatrix} 0{,}2619 \\ 0{,}0659 \end{Bmatrix};$$

$$\nabla f = \begin{Bmatrix} 3 \\ 3\sqrt{3} \end{Bmatrix}$$

Exemplo

$$\underset{s_1,s_2}{\text{Min}} \quad \beta$$

$$\text{tal que} \quad -0{,}2619s_1 - 0{,}0659s_2 + \beta \leq 0$$

$$3s_1 + \sqrt{3}s_2 + \beta \leq 0$$

$$-1 \leq s_1 \leq 1$$

$$-1 \leq s_2 \leq 1$$

Solução: $s_1 = 0{,}5512;\ s_2 = -1 \Rightarrow$

Busca unidimensional:

$$\Rightarrow \mathbf{x}_2 = \begin{Bmatrix} 8{,}29 \\ 12{,}56 \end{Bmatrix} + \alpha \begin{Bmatrix} 0{,}5512 \\ -1 \end{Bmatrix}$$

α é limitado por $g_1 \Rightarrow \alpha < 4{,}957 \Rightarrow$

$$\Rightarrow \alpha = 4{,}957 \Rightarrow \mathbf{x}_2 = \begin{Bmatrix} 11{,}02 \\ 7{,}60 \end{Bmatrix} \Rightarrow f(\mathbf{x}_2) = 46{,}22$$

$$\mathbf{x}^* = \begin{Bmatrix} 9{,}464 \\ 9{,}464 \end{Bmatrix} \Rightarrow f(\mathbf{x}^*) = 44{,}78$$

(Solução Ótima)