Otimização ELE037
Introdução

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

Universidade Federal de Minas Gerais
Departamento de Engenharia Elétrica

Sumário

Introdução

Assunto abordado:

- O problema de otimização;
- Otimização de funções matemáticas simples (2 variáveis);
- Caracterização de diferentes tipos de funções;
- Caracterização de diferentes estratégias de otimização;
- Princípio de funcionamento dos métodos de otimização.

A *Otimização* representa um conjunto de ferramentas capazes de determinar as melhores configurações possíveis para a construção ou o funcionamento de sistemas de interesse.

A mesma teoria é aplicada a diferentes contextos:

- Projeto de circuitos, antenas, motores (Eng. Elétrica);
- Controle de processos industriais (Eng. CA);
- Política eficiente de vacinação (Mat. Computacional);
- Otimização do tráfego de informação em redes (Cientista da Computação);
- etc.

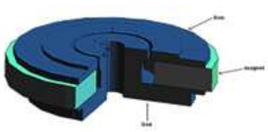
Como modelar um problema de otimização?

Formalmente, um problema de otimização pode ser definido por:

$$\boldsymbol{x}^* = \arg \min_{\boldsymbol{x}} f(\boldsymbol{x})$$
$$\text{sujeito a: } \begin{cases} g_i(\boldsymbol{x}) \leq 0; i = 1, \ldots, p \\ h_j(\boldsymbol{x}) = 0; j = 1, \ldots, q \end{cases} \tag{1}$$

Obs.: variáveis em negrito são vetores; as demais são escalares.

Seja o problema de otimização de um alto-falante como ilustrado a seguir.



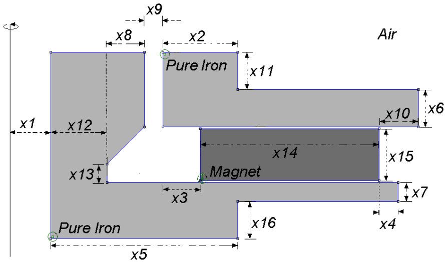
Objetivo:

- Minimizar o volume do alto-falante e atender um valor mínimo de densidade de fluxo magnético **B** no entreferro.

Matematicamente:

$$\begin{cases} \text{min: } f(\boldsymbol{x}) = \mathit{volume} \\ \text{sujeito a: } g_1(\boldsymbol{x}) : |\mathbf{B}| \geq \mathbf{B}_{min} \end{cases}$$

Modelo 2D do alto-falante:



Questões práticas:

- Como calcular o volume do alto-falante usando $x_1, \ldots, x_{16}$?
- Quais os limites de $x_1, \ldots, x_{16}$?
- Quais materiais serão usados?
- Como calcular **B**?

O vetor $\boldsymbol{x}$ é o *vetor de variáveis de otimização*.

O processo de otimização busca especificar os valores destas variáveis.

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{2}$$

O vetor $\boldsymbol{x}$ possui um significado concreto/físico?

O vetor $\boldsymbol{x}$ é composto de variáveis reais ou discretas?

As técnicas de otimização são as mesmas para qualquer representação das variáveis de decisão?

A *função objetivo* $f(\cdot)$, ou *função custo*:

- Representa um índice específico do sistema, cujo valor, por convenção, queremos minimizar para alcançarmos o desempenho ótimo.

Qual a possível função custo do projeto do alto-falante?

Qual a possível função custo do projeto de um motor?

Ex. funções de custo: consumo de combustível; ruído; probabilidade de defeitos, etc.

Como tratar problemas de maximização?

Neste caso basta minimizarmos a função que se deseja maximizar multiplicada por -1.

Maximizar a função $p(\boldsymbol{x})$ é o mesmo que minimizar $f(\boldsymbol{x}) = -p(\boldsymbol{x})$.

O vetor $\boldsymbol{x}$ que minimiza $f(\cdot)$ é também o vetor que maximiza $p(\cdot)$.

Note que $f(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}$.

As variáveis de otimização $\boldsymbol{x}$ são reais?

No exemplo anterior, quantas possibilidades de construção existem para o alto-falante?

Na solução ótima de um problema de minimização:

> O vetor ótimo $\boldsymbol{x}^*$ é igual ao **argumento** da função $f(\cdot)$ que faz com que essa função atinja seu **mínimo** valor.

Como encontrar $\boldsymbol{x}^*$?

- As restrições significam o conjunto dos requisitos que o resultado do projeto deve atender para ser admissível enquanto solução.

- Restrições de desigualdade:

$$g_i(\boldsymbol{x}) \leq 0 \tag{3}$$

- Restrições de igualdade:

$$h_j(\boldsymbol{x}) = 0 \tag{4}$$

Restrições de natureza não-técnica:

São fisicamente implementáveis, porém violam certos padrões.

Exemplo:

- Objetivo: Projeto de um automóvel de baixo custo;
- Restrição: Emissão de gases poluentes abaixo de um limiar estabelecido por lei.

$g_i(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^p$: $g_i(\boldsymbol{x}) \leq 0, i = 1, \ldots, p$

$h_j(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}^q$: $h_j(\boldsymbol{x}) = 0, j = 1, \ldots, q$

Em relação às restrições, definimos ainda a seguinte nomenclatura:

- **Região factível**: Conjunto dos pontos do espaço $\Re^n$ que satisfazem, simultaneamente, a todas as restrições;
- **Região infactível**: Conjunto dos pontos do espaço $\Re^n$ que violam pelo menos uma das restrições do problema;
- **Ponto factível**: Ponto pertencente à *região factível*;
- **Ponto infactível**: Ponto pertencente à *região infactível*;
- **Restrição violada**: Cada uma das componentes do vetor $g_i(\boldsymbol{x})$ que apresentar valor positivo, ou cada uma das componentes do vetor $h_j(\boldsymbol{x})$ que apresentar valor não-nulo será chamada de *restrição violada* no ponto $\boldsymbol{x}$.

Regras de acesso à informação:

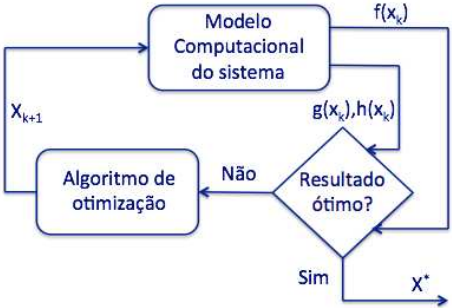
- Não conhecemos expressões matemáticas explícitas que representem a função objetivo $f(\cdot)$ e as funções de restrição $g_i(\cdot)$ e $h_j(\cdot)$;
- Temos, entretanto, a possibilidade de descobrir quanto valem as funções objetivo e de restrição em qualquer ponto do espaço de variáveis de otimização. Essa é a única informação que conseguiremos adquirir, ao longo do processo de otimização, para nos guiar em direção à solução desejada.

Por que nem sempre conhecemos $f(\boldsymbol{x})$?

Ex.: Qual a função objetivo do problema do alto-falante? E a função de restrição?

Os métodos de otimização serão comparados entre si de acordo com os critérios:

- Número de avaliações da função objetivo e das restrições;
- Quanto menos avaliações forem necessárias, melhor será considerado o método;
- Precisão e robustez: quanto mais a solução fornecida pelo método se aproximar da solução exata do problema, melhor será considerado o método.



Modelo Computacional do sistema
f(x_k)
X_k+1
g(x_k),h(x_k)
Algoritmo de otimização
Não
Resultado ótimo?
Sim
X*

Embora $f(\boldsymbol{x})$ seja do tipo *caixa-preta*, ela é bem definida.

O processo de otimização utiliza informações locais da superfície de $f(\boldsymbol{x})$.

Problema de minimização de uma função objetivo sem restrição:

$$\boldsymbol{x}^* = \arg \min_{\boldsymbol{x}} f(\boldsymbol{x}) \tag{5}$$

Para viabilizar a representação gráfica do problema, supõe-se $\boldsymbol{x} \in \Re^2$.

Consideremos a seguinte função não-linear $f(\boldsymbol{x})$:

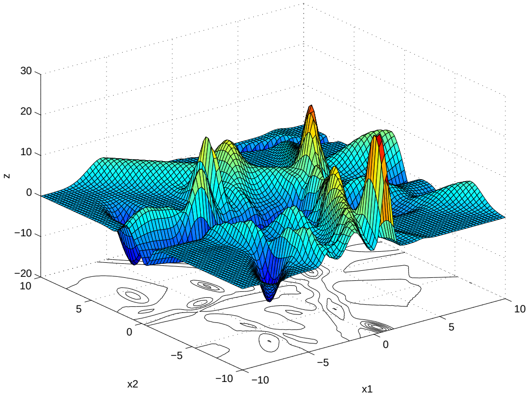


Figura: Superfície que representa o gráfico de uma função não-linear $f(\boldsymbol{x})$ de duas variáveis reais.

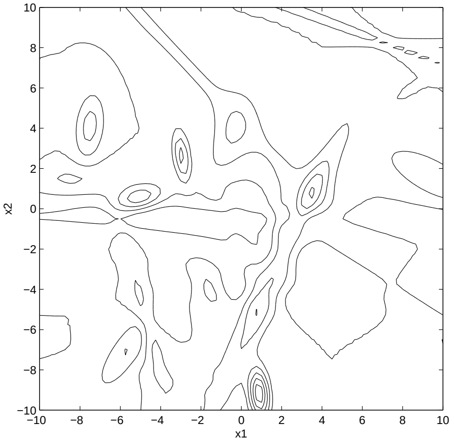


Figura: Gráfico de *curvas de nível* de $f(\boldsymbol{x})$.

Metáfora para a solução do problema de otimização:

- Um aluno é lançado de pára-quedas sobre um ponto qualquer da superfície de $f(\boldsymbol{x})$;
- O objetivo do aluno é encontrar o ponto mais baixo de $f(\boldsymbol{x})$, i.e. o *ponto de mínimo*, com o menor número possível de "passos";
- Deverá caminhar com uma venda cobrindo seus olhos, sem poder "olhar" para a superfície;
- A única informação que ele pode utilizar é a altura do ponto no qual estiver "pisando";
- Pode, entretanto, se "lembrar" das alturas dos pontos em que já tiver pisado;
- Esta informação pode ser utilizada para tomar a decisão de "para onde caminhar".

A metáfora descrita anteriormente ilustra bem o que é o problema de otimização.

Construir os chamados **métodos de otimização** corresponde, dentro dessa metáfora, a formular as **estratégias** a serem utilizadas pelo "aluno" em sua **busca pelo ponto de mínimo**.

Que tipo de **estratégia de otimização** utilizar?

Esta escolha depende das características da superfície de $f(\boldsymbol{x})$:

- Diferenciabilidade: diferenciável ou não-diferenciável
- Modalidade: unimodal ou multimodal
- Convexidade: convexa, quasi-convexa, não-convexa
- Linearidade: linear ou não-linear
- Escala: uni-escala ou multi-escala

Estas características são discutidas ao longo da apresentação das diferentes estratégias de otimização.

Vamos considerar a função quadrática definida por:

$$f(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{x}_0)' Q (\boldsymbol{x} - \boldsymbol{x}_0)$$

$$Q = \begin{bmatrix} 2 & 0.3 \\ 0.3 & 1 \end{bmatrix} \qquad \boldsymbol{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{6}$$

cujo gráfico e curvas de nível são ilustrados a seguir.

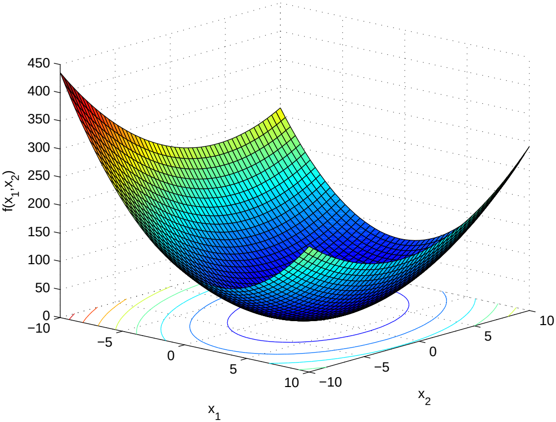


Figura: Superfície que representa o gráfico de uma função quadrática $f(\boldsymbol{x})$ de duas variáveis reais.

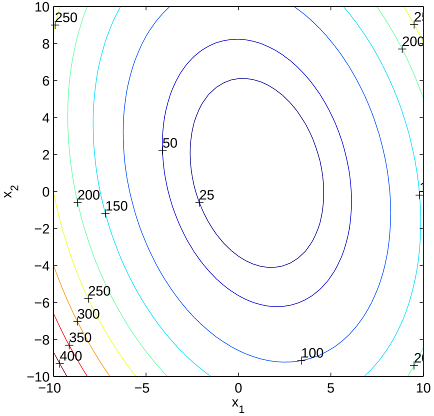


Figura: Gráfico de *curvas de nível* da mesma função quadrática de duas variáveis reais, $f(\boldsymbol{x})$.

Método do Gradiente (*um método de direção de busca*)

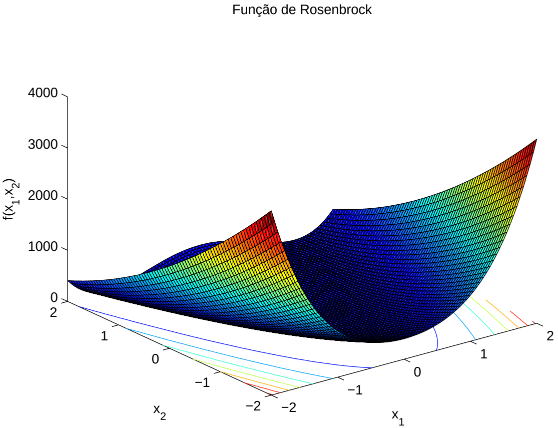
O aluno:

- Passo 1: colhe amostras locais da função e determina em qual direção a função decresce mais rapidamente (usa aproximação numérica do **gradiente** da função).
- Passo 2: caminha na direção encontrada enquanto a função decrescer.
- Passo 3: decide se pára (caso considere que esteja suficientemente próximo do ponto de mínimo da função) ou se continua a busca, retornando ao Passo 1.

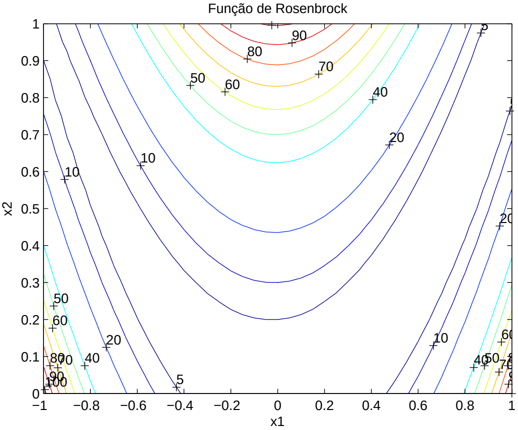
Requisitos exigidos pelas Estratégias de Direção de Busca

- A função deve ser **diferenciável** (a aproximação numérica do gradiente da função contém informação significativa sobre a forma como a função varia nas vizinhanças do ponto em que tiver sido estimado).
- A função deve ser **unimodal** (possuir **um único mínimo global**, no interior de **uma única bacia de atração**).

Consideremos uma função um pouco mais complexa.
Função de Rosenbrock (diferenciável e unimodal; $\boldsymbol{x}^* = [1 \;\; 1]'$).



O *Método do Gradiente* é eficiente para resolver este problema?
Existem outros *Métodos de Direção de Busca* mais eficientes?



Consideremos agora uma função $f(\boldsymbol{x})$ com as seguintes características:

- Ainda **unimodal**;
- Porém **não-diferenciável**.

Qual a dificuldade encontrada por técnicas de Direções de Busca?

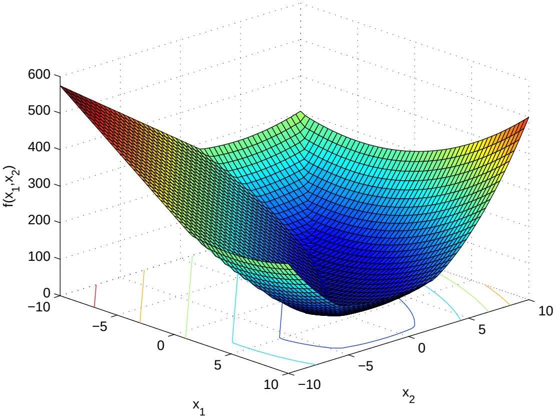


Figura: Superfície que representa o gráfico de uma função unimodal não diferenciável $f(\boldsymbol{x})$ de duas variáveis.

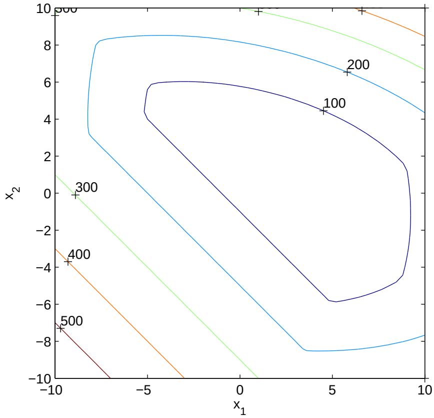
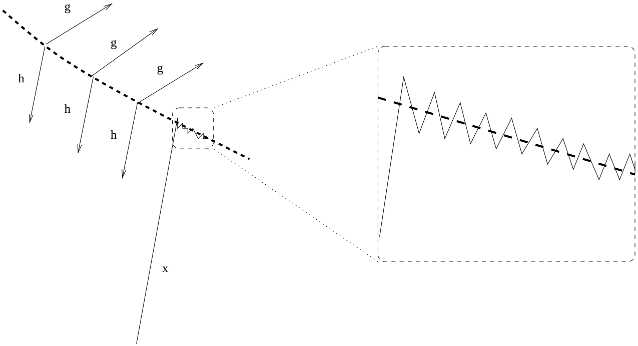


Figura: Gráfico de *curvas de nível* da função unimodal não diferenciável de duas variáveis reais, $f(\boldsymbol{x})$.

- Como tratar problemas unimodais e não-diferenciáveis usando Métodos de Direção de Busca?
- A não-diferenciabilidade ocorre em todos os pontos da função?
- Como saber, numericamente, se a função não é diferenciável em um ponto específico?
- Quando identificada a não-diferenciabilidade em um ponto, o que fazer?
- **A otimização por esses métodos pode se tornar inviável**.
- **Esta dificuldade é intrínseca a toda a família dos métodos de direção de busca**.

Dificuldade encontrada pelos Métodos de Direção de Busca em funções não-diferenciáveis:



- Funções não-diferenciáveis são muito comuns em problemas práticos.
- Uma nova família de métodos é formulada para tratar tais problemas.
- Métodos de Exclusão de Regiões:
  - Unimodalidade da função;
  - Diferenciabilidade não é exigida;
  - **Convexidade da função**.

Por que a função $f(\boldsymbol{x})$ precisa ser **convexa**?

- Uma curva de nível de uma função convexa sempre delimita uma região convexa em seu interior.
- O vetor gradiente é sempre perpendicular à curva de nível que passa pelo ponto onde o vetor foi calculado.
- A reta perpendicular ao vetor gradiente que passa no ponto onde esse vetor foi calculado é tangente à curva de nível.
- Devido à convexidade da região no interior da curva de nível, esta região sempre fica inteiramente localizada em apenas um dos lados dessa reta tangente.

Funcionamento do Método de Exclusão de Regiões

Partindo de um ponto inicial sobre o espaço de busca:

- Passo 1: adquire-se informação local, e faz-se uma **estimativa do gradiente** da função objetivo nesse ponto.
- Passo 2: com base no gradiente, identifica-se qual é a **reta tangente à curva de nível que passa pelo ponto atual** (todo o semi-plano que contém o vetor gradiente é descartado).
- Passo 3: move-se para algum ponto no interior da região que ainda não foi descartada.
- Passo 4: caso o novo ponto esteja suficientemente próximo do mínimo da função, o processo termina. Do contrário, retorna-se ao Passo 1.

Algumas observações:

- A convergência do método ocorre devido a **diminuição sistemática da região** que contém $\boldsymbol{x}*$.
- Como a região "viável" diminui, o novo ponto **tende** a aproximar-se cada vez mais de $\boldsymbol{x}*$.
- A não-diferenciabilidade não impede a convergência do método.

Qual o possível critério de parada do método?

Qual a relação entre a **velocidade de convergência** no início e no final do processo de otimização?

Otimização sem Restrição
Estratégia de exclusão de regiões

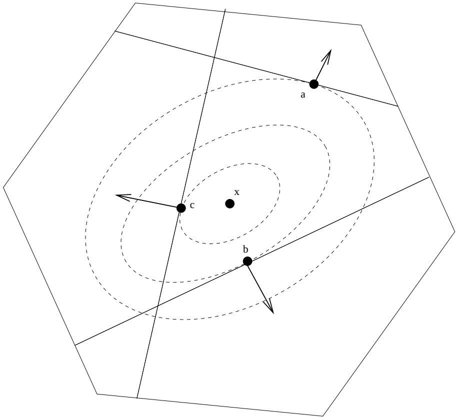


Figura: Iterações de um método de exclusão de regiões, mostradas sobre as curvas de nível de uma função cujo mínimo exato é $\boldsymbol{x}^*$.

- Grande parte das funções objetivo que queremos otimizar na prática, infelizmente, **não são unimodais**.

- Note que **uma função multimodal também não é convexa**. Mas pode apresentar regiões convexas.

- Por consequência, tanto as estratégias de **direção de busca** quanto as estratégias de **exclusão de regiões** *irão falhar* em sua otimização.

- Caso $f(\boldsymbol{x})$ seja multimodal, o resultado encontrado por qualquer uma das técnicas estudadas sempre será o **mínimo local** associado à bacia de atração onde a busca tiver sido iniciada.

- Para determinar o mínimo global, todas as bacias devem ser investigadas.

- As **estratégias de populações** evoluem um **conjunto de soluções candidatas** em paralelo, e estas cooperam entre si.

Considere a função $f(\boldsymbol{x})$ multimodal:

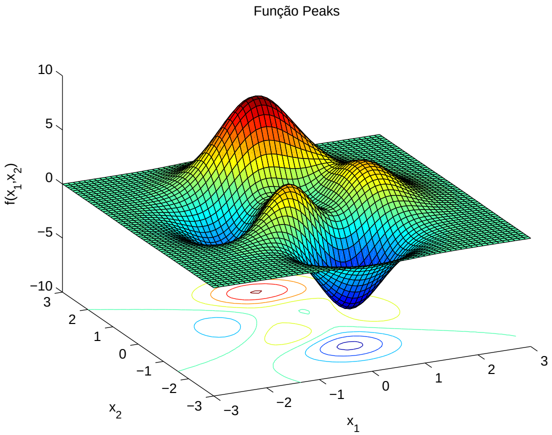


Figura: Função multimodal

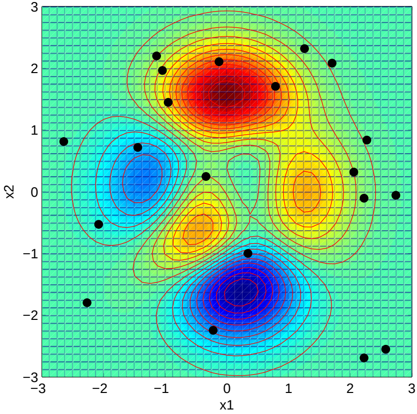


Figura: Curvas de nível da função multimodal

Vários Otimizadores:

- Passo 1: encontram-se distribuídos no espaço de busca, e colhem **informações locais** da função.
- Passo 2: comunicam entre si, e **trocam informações** a respeito dos valores da função objetivo em cada ponto.
- Passo 3: um pequeno sub-grupo de Otimizadores, que estiver nas melhores regiões, fica parado.
- Passo 4: os demais se deslocam, com movimentos que simultaneamente: (i) os façam se **aproximarem daqueles melhor localizados**; e (ii) os façam **explorarem outras regiões** ainda não investigadas.
- Passo 5: cada Otimizador avalia a função no ponto para onde foi.
- Passo 6: o algoritmo pára ou retorna ao Passo 2.

Algoritmos baseados em Populações

- Vantagem (não exigem unimodalidade, convexidade ou diferenciabilidade).
- Limitação (custo computacional elevado).
- Quando são úteis?
- Métodos híbridos:
  - São viáveis? Por que? (ex.: custo computacional)
  - Quando podem ser empregados? (ex.: funções multi-escala)
  - Quais requisitos devem ser considerados? (ex.: convexidade)

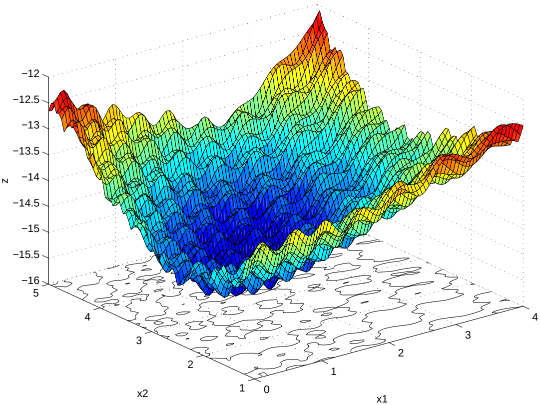


Figura: Superfície que representa o gráfico de uma função multimodal $f(\boldsymbol{x})$ de duas variáveis que apresenta a característica de *múltiplas escalas*.

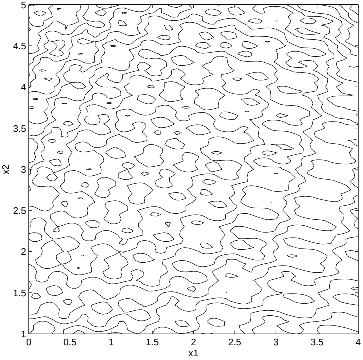


Figura: Superfície que representa o gráfico de uma função multimodal $f(\boldsymbol{x})$ de duas variáveis que apresenta a característica de *múltiplas escalas*.

Suponha agora um problema de otimização com **restrições de desigualdade**:

$$\begin{aligned} \boldsymbol{x}^* = \ & \arg\ \min_{\boldsymbol{x}}\ f(\boldsymbol{x}) \\ & \text{Sujeito a: } \left\{\ g_i(\boldsymbol{x}) \leq 0;\ i = 1, \ldots, p \right. \end{aligned} \tag{7}$$

O ponto de ótimo $\boldsymbol{x}^*$ deve satisfazer às p desigualdades:

$$\begin{aligned} g_1(\boldsymbol{x}^*) &\leq 0 \\ g_2(\boldsymbol{x}^*) &\leq 0 \\ &\vdots \\ g_p(\boldsymbol{x}^*) &\leq 0 \end{aligned} \tag{8}$$

Consideremos inicialmente uma única restrição de desigualdade:

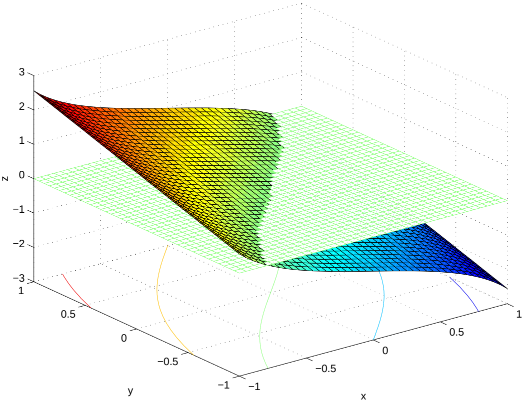
$$g_1(\boldsymbol{x}) \leq 0 \tag{9}$$

- Admitamos que a função $g_1(\cdot)$ seja contínua;
- Se verdadeiro, essa função nunca muda “bruscamente” de valor;
- Logo, se existem...
  - $\mathcal{P}_1 \subset \mathbb{R}^n$ em que $g_1(\cdot) > 0$; e
  - $\mathcal{N}_1 \subset \mathbb{R}^n$ em que $g_1(\cdot) < 0$;
  - Então, deve existir $\mathcal{G}_1 \subset \mathbb{R}^n$ em que $g_1(\cdot) = 0$;
  - Além disso, $\mathcal{G}_1$ separa $\mathcal{P}_1$ de $\mathcal{N}_1$.

Matematicamente:

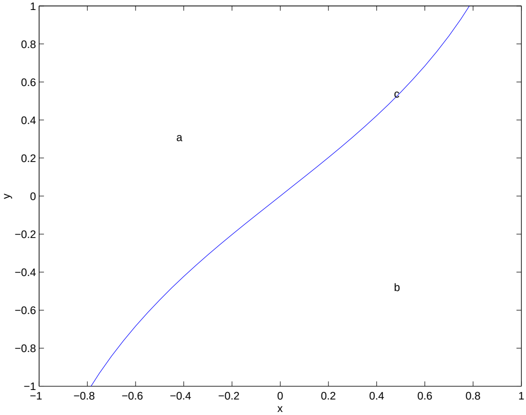
$$
\begin{aligned}
\mathcal{P}_1 &\triangleq \{\boldsymbol{x} \mid g_1(\boldsymbol{x}) > 0\} \\
\mathcal{N}_1 &\triangleq \{\boldsymbol{x} \mid g_1(\boldsymbol{x}) < 0\} \\
\mathcal{G}_1 &\triangleq \{\boldsymbol{x} \mid g_1(\boldsymbol{x}) = 0\}
\end{aligned}
\tag{10}
$$

Exemplo de restrição: superfície $z = g_1(\boldsymbol{x})$



Exemplo de restrição: curva de nível onde $g_1(\boldsymbol{x}) = 0$

- $a = \mathcal{P}_1$; $b = \mathcal{N}_1$; $c = \mathcal{G}_1$

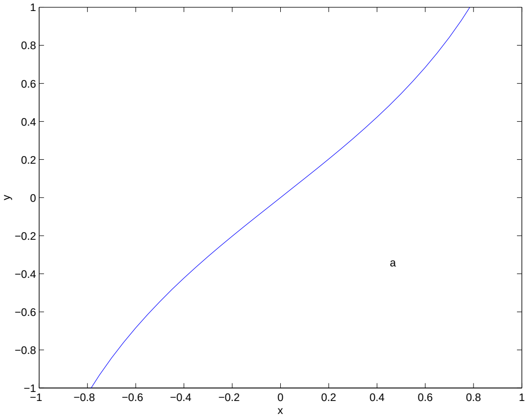


A exigência de que $g_1(\boldsymbol{x}^*) \leq 0$ implica que:

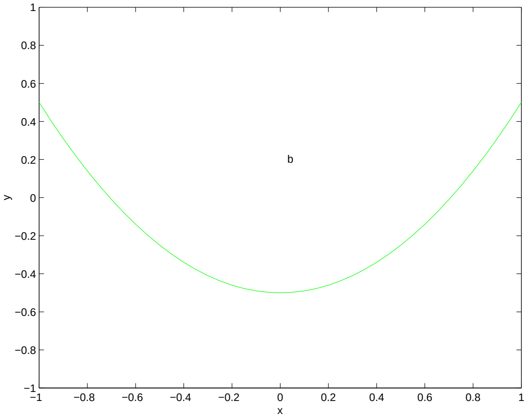
- $\mathcal{P}_1$ representa o conjunto/região inviável/infactível;
- $\mathcal{N}_1$ e $\mathcal{G}_1$ são conjuntos/regiões viáveis/factíveis;
- O *conjunto factível* ou a *região factível* $\mathcal{F}_1$ é a união de $\mathcal{G}_1$ e $\mathcal{N}_1$:

$$\mathcal{F}_1 = \mathcal{G}_1 \cup \mathcal{N}_1 \tag{11}$$

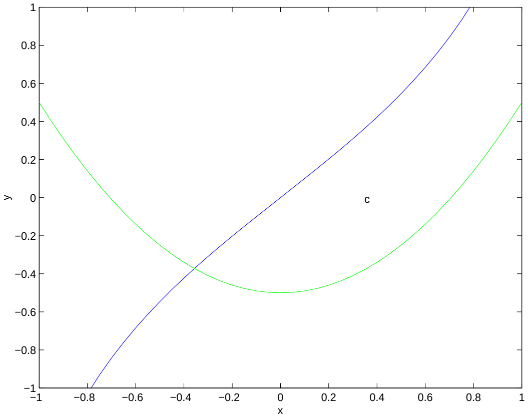
Interpretação para problemas com várias restrições: $a = \mathcal{F}_1$



Interpretação para problemas com várias restrições: $b = \mathcal{F}_2$



Interpretação para problemas com várias restrições: $c = \mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2$



- Os métodos discutidos são aplicáveis a problemas restritos?

- O ótimo restrito deverá pertencer a $\mathcal{F}$.

- Como garantir que $\boldsymbol{x}^* \in \mathcal{F}$?

- **Como adaptar os métodos conhecidos**, de maneira simples e eficiente, para resolver problemas restritos?

- Solução Geral: **Método de Barreiras** ou **Método de Penalidades**.
- A ideia é transformar a função-objetivo em uma **função pseudo-objetivo**.
- Barreiras: impedem a existência de soluções em regiões infactíveis.
- Penalidades: penalizam soluções localizadas em regiões inviáveis.
- A função pseudo-objetivo deve ser semelhante à função original na região viável.

Em termos matemáticos, o problema de otimização original:

$$\begin{aligned} \boldsymbol{x}^* = & \ \arg \min_{\boldsymbol{x}} \ f(\boldsymbol{x}) \\ & \text{Sujeito a: } \{g_i(\boldsymbol{x}) \leq 0 \end{aligned} \tag{12}$$

é transformado no problema:

$$\boldsymbol{x}^* = \ \arg \min_{\boldsymbol{x}} \ f(\boldsymbol{x}) + F(\boldsymbol{x}) \tag{13}$$

A função $F(\cdot)$ deve ser muito pequena (ou zero) no interior da região factível, de tal forma que $f(\cdot)$ seja muito parecida com $f(\cdot) + F(\cdot)$ em qualquer ponto deste espaço.

- O valor de $F(\boldsymbol{x})$ cresce subtamente perto da fronteira.
- Pré-requisito: **a solução inicial deve ser factível**.

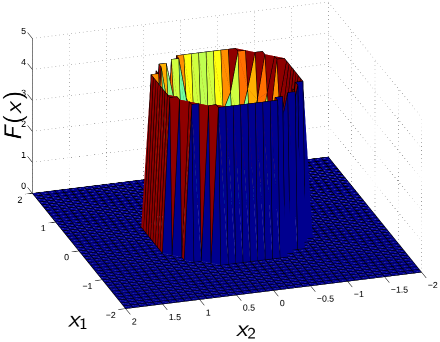


Figura: Ilustração de uma função de barreira.

- O valor de $F(\boldsymbol{x})$ cresce a medida que se afasta da região viável.
- **A solução inicial pode ser factível ou não**.

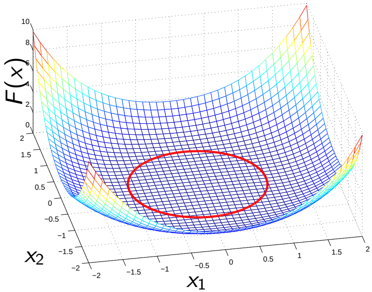


Figura: A função de penalidade é igual a zero no interior da região factível, e cresce rapidamente à medida em que o ponto se afasta dessa região.

- O problema restrito é então tratado como irrestrito.
- Logo, os métodos discutidos podem ser aplicados na otimização dos problemas transformados.
- Qual método utilizar: Direção de Busca, Exclusão de Regiões ou Método de Populações?
- Barreiras ou Penalidades?

Exclusão de Regiões e tratamento de $g_i(\boldsymbol{x}) \leq 0$ por
Composição pelo Máximo

- Dado um ponto $\boldsymbol{x}$, identifica-se a restrição mais violada:

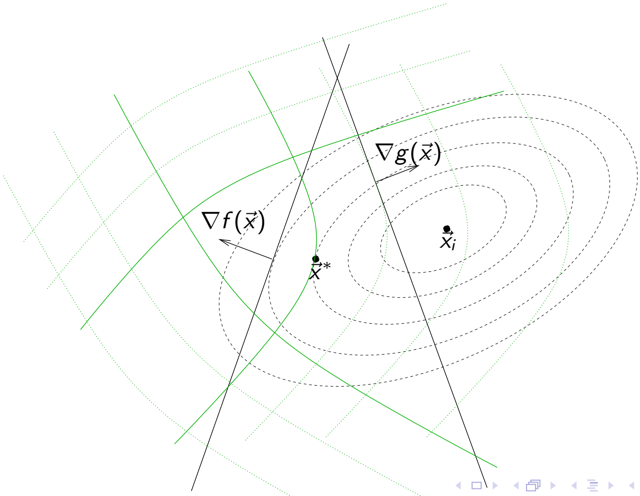
$$G(\boldsymbol{x}) = \max(g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), \ldots, g_p(\boldsymbol{x})) \tag{14}$$

- Aplica-se exclusão de região à função $G(\cdot)$.
- Caso $\boldsymbol{x}$ seja viável, aplica-se exclusão de região à função $f(\cdot)$.

Exclusão de Regiões e tratamento de $g_i(\boldsymbol{x}) \leq 0$ por
Composição pelo Máximo

- A sequência de pontos $\boldsymbol{x}$ pode oscilar entre viável e inviável.
- No entanto, a região factível diminui iterativamente.
- A solução final termina arbitrariamente próxima do ótimo.

Exclusão de Regiões e **Composição pelo Máximo**



Consideremos agora o problema de otimização com *restrições de igualdade*:

$$
\begin{aligned}
& \mathbf{x}^* = \arg \min_{\mathbf{x}} \; f(\mathbf{x}) \\
& \text{Sujeito a: } \{ h_j(\mathbf{x}) = 0 \; , \quad j = 1, \ldots, q
\end{aligned}
\tag{15}
$$

O ótimo $\mathbf{x}^*$ deve satisfazer às q equações:

$$\begin{aligned} h_1(\mathbf{x}^*) &= 0 \\ h_2(\mathbf{x}^*) &= 0 \\ &\vdots \\ h_q(\mathbf{x}^*) &= 0 \end{aligned} \tag{16}$$

- A região factível é descrita por um conjunto de pontos que satisfazem $h(\boldsymbol{x}) = 0$.
- Esta região viável é uma superfície de dimensão $n - 1$.
- No caso de q restrições de igualdade, o conjunto factível corresponde à interseção de todas as superfícies $h_j(\boldsymbol{x}) = 0$.

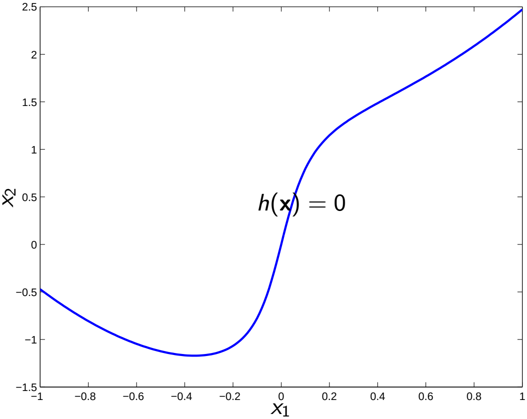


Figura: A linha corresponde ao lugar geométrico dos pontos que satisfazem $h(\mathbf{x}) = 0$. Essa linha é a região factível de um problema de otimização com essa restrição.

- Como tratar as restrições de igualdade?
- Quais das técnicas discutidas são viáveis neste caso?
- Por que?
- As técnicas de barreiras e composição pelo máximo dependem da existência de pontos *interiores* à região factível do problema.
- A técnica de penalidades é a única viável nestes problemas.

Suponha funções objetivo e de restrições lineares.
O problema agora é chamado de otimização linear:

$$
\begin{aligned}
\boldsymbol{x}^* &= \arg\min_{\boldsymbol{x}} \boldsymbol{c}'\boldsymbol{x} \\
\text{Sujeito a: } &\left\{ A\boldsymbol{x} \leq \boldsymbol{b} \right.
\end{aligned} \tag{17}
$$

De outra forma:

$$
f(\boldsymbol{x}) = c_1x_1 + c_2x_2 + \ldots + c_nx_n \tag{18}
$$

e o conjunto de restrições corresponde às m desigualdades:

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \leq b_1 \\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \leq b_2 \\
&\vdots \\
&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \leq b_m
\end{aligned} \tag{19}
$$

Importância da otimização linear

- Problemas práticos podem ser modelados via funções lineares;
- Podem ser resolvidos mais rapidamente que problemas não-lineares:
  - Para o mesmo número de variáveis; e
  - Mesmo número de restrições.
- Algoritmos especializados são capazes de lidar com problemas muito grandes (centenas de variáveis de decisão).

- As curvas de nível de uma função linear são retas paralelas.
- Não exite mínimo local irrestrito de um problema linear.

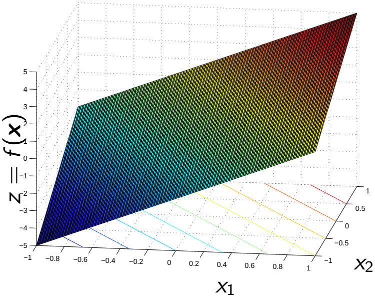


Figura: Superfície correspondente à função objetivo linear $f(\boldsymbol{x}) = \boldsymbol{c}'\boldsymbol{x}$.

- A presença de restrições de desigualdade definem um poliedro factível.

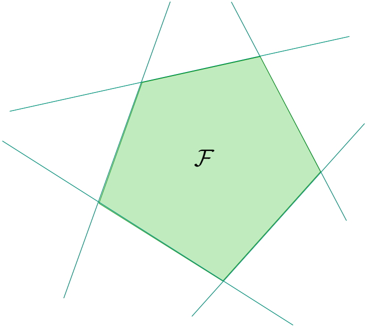


Figura: Região factível $\mathcal{F}$ correspondente às restrições lineares de desigualdade.

- Curvas de nível de uma função linear e região factível:

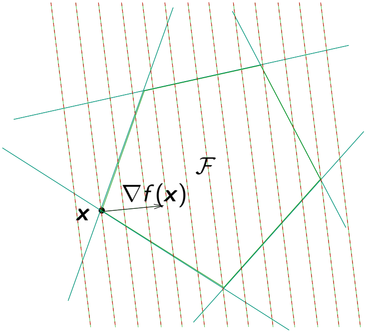


Figura: O vetor gradiente, $\nabla f(\boldsymbol{x})$, é constante em todo o espaço.

- O mínimo da função linear nunca estará no interior da região factível.
- O mínino deverá estar na fronteira, e sobre algum vértice.
- O conjunto ótimo pode ser toda uma fronteira.
- Como otimizar eficientemente tais funções?
  - Método Simplex.
  - Métodos de Pontos Interiores.