Programação Não-Linear Restrita: Fundamentos de Algoritmos

Eduardo Camponogara

Departamento de Automação e Sistemas
Universidade Federal de Santa Catarina

DAS-9011: Métodos de Otimização

Introdução

Categorizando Algoritmos de Otimização

Eliminação de Variáveis

Dificuldades

Sumário

Introdução

Categorizando Algoritmos de Otimização

Eliminação de Variáveis

Dificuldades

Introdução

Agenda

Aqui iniciamos a discussão de algoritmos para a solução do problema geral de otimização sob restrições.

$$
\begin{aligned}
P: \quad & \textit{Minimize} \quad && f(x) \\
& x \in \mathbb{R}^n \\
& \textit{Sujeito a}: \\
& && c_i(x) = 0 \qquad i \in \mathcal{E} \\
& && c_i(x) \leq 0 \qquad i \in \mathcal{I}
\end{aligned}
\tag{1}
$$

Onde:

- f e c_i são funções suaves com valores reais; e
- $\mathcal{I}$ e $\mathcal{E}$ são conjuntos de índices finitos para as restrições de desigualdade e igualdade, respectivamente.

Introdução

Estudos Anteriores

- Desenvolvemos condições necessárias para solução ótima local (*Karush-Kuhn-Tucker Conditions*).
- Algoritmos eficientes levam em conta propriedades e a estrutura das funções que definem o problema.
- Algoritmos especializados foram desenvolvidos para os problemas:
  a) Programação Linear
  b) Programação Quadrática
  c) Programação Não-Linear

Introdução

- Algoritmos para otimização não-linear são iterativos:
- Algoritmos geram uma seqüência x^k de aproximações convergente a uma solução ótima local x^*.
- Algoritmos utilizam informações sobre a solução corrente x^k, bem como as derivadas de f e c_i, $i \in \mathcal{E} \cup \mathcal{I}$, no processo de obtenção do próximo iterando, x^{k+1}.

Introdução

É recomendável o estudo inicial de um problema na busca de simplificações:

a) um exame das restrições pode indicar que a região factível é vazia;

b) a função objetivo pode ser ilimitada; e

c) identificar quais restrições estarão ativas na solução ótima, permitindo o emprego direto das condições KKT.

Introdução

Um aspecto relevante durante a modelagem e pertinente à escolha de um algoritmo se refere aos tipos de restrições:

Restrições Duras

- Aquelas que devem ser satisfeitas para que os resultados sejam significativos
- Por exemplo, a restrição de que $x \geq 0$ para que uma outra restrição tal como $\sqrt{x} \leq 2$ tenha significado
- As restrições duras podem até mesmo não ser definidas fora da região de factibilidade

Restrições Flexíveis Aquelas que são definidas mesmo na região infactível.

Sumário

Introdução

Categorizando Algoritmos de Otimização

Eliminação de Variáveis

Dificuldades

O Método da Função Penalidade

- Podemos atacar o problema não-linear por meio da combinação da função objetivo e das restrições em uma função penalidade,
- E resolvendo uma seqüência de problemas irrestritos.
- Se apenas restrições de igualdade estão presentes, isto é ($\mathcal{I} = \phi$), então a função de penalidade pode tomar a forma:

$$f(x) + \frac{1}{2\mu} \sum_{i \in \mathcal{E}} c_i^2(x) \tag{2}$$

onde $\mu > 0$ é chamado de parâmetro de penalidade.

O Método da Função Penalidade

- Se apenas restrições de igualdade estão presentes, isto é ($\mathcal{I} = \phi$), então a função de penalidade pode tomar a forma:

$$f(x) + \frac{1}{2\mu} \sum_{i \in \mathcal{E}} c_i^2(x) \tag{3}$$

onde $\mu > 0$ é chamado de parâmetro de penalidade.

- Minimizar a função acima para uma série decrescente de valores de μ, até que uma solução do problema original seja encontrada.

O Método de Barreiras

- Adicionar à função objetivo, termos que são insignificantes quando x está seguramente no interior da região factível
- Mas que tendem a $+\infty$ quando x se aproxima da fronteira da região factível.
- Se apenas desigualdades estão presentes, a função de barreira logarítmica tem forma:

$$f(x) - \mu \sum_{i \in \mathcal{I}} \log c_i(x) \tag{4}$$

onde $\mu > 0$ é dito parâmetro de barreira.

O Método de Barreiras

- Se apenas desigualdades estão presentes, a função de barreira logarítmica tem forma:

$$f(x) - \mu \sum_{i \in \mathcal{I}} \log c_i(x) \tag{5}$$

onde $\mu > 0$ é dito parâmetro de barreira.

- O minimizador da função (4) tende para uma solução do problema original à medida que μ se aproxima de zero
- Encontrar um minimizador aproximado para uma seqüência decrescente de valores de μ.

Lagrangeano Aumentado

- Combinar as propriedades da função Lagrangeana e da penalidade quadrática:

$$\mathcal{L}(x, y) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x), \quad \lambda_i \geq 0, \ i \in \mathcal{I} \qquad (6)$$

- Quando o problema (2) possui apenas restrições de igualdade, a função Lagrangeana aumentada toma a forma:

$$\mathcal{L}(x, \lambda, \mu) = f(x) - \sum_{i \in \mathcal{E}} \lambda_i c_i(x) + \frac{1}{2\mu} \sum_{i \in \mathcal{E}} c_i^2(x) \qquad (7)$$

Lagrangeano Aumentado

- Métodos baseados na função (7) procedem fixando λ a alguma estimativa dos multiplicadores de Lagrange ótimos e μ a um valor positivo,
- depois encontram um valor x que aproximadamente minimize $\mathcal{L}(x, \lambda, \mu)$.
- O iterando x é utilizado para recalcular o valor de λ, permitindo que μ talvez seja decrescido.

Programação Quadrática Seqüencial

- Modelar o problema não-linear como um subproblema quadrático a cada iteração e definir uma direção de busca como solução deste subproblema.
- No caso de existirem apenas restrições de igualdade, definimos a direção de busca p_k a cada iterando (x_k, λ_k) como a solução do problema abaixo.

$$\begin{aligned} \underset{p}{\textit{Minimize}} \quad & \tfrac{1}{2} p^T W_k p + \nabla f_k^T p \\ \textit{Sujeito a}: \quad & \\ & A_k p + c_k = 0 \end{aligned} \tag{8}$$

Programação Quadrática Seqüencial

- O objetivo do problema quadrático é uma aproximação da função Lagrangeana e as restrições são uma linearização das restrições do problema não-linear.
- O iterando seguinte é obtido por meio de uma busca ao longo da direção p_k, através da solução do problema quadrático, até que uma certa função mérito seja decrescida.
- Programação Quadrática Seqüencial é considerado um dos métodos mais eficientes na prática.

Sumário

Introdução

Categorizando Algoritmos de Otimização

Eliminação de Variáveis

Dificuldades

Eliminação de Variáveis

- É natural buscar eliminar restrições de maneira a se obter um problema irrestrito,
- Ou pelo menos eliminar algumas restrições e se obter um problema simplificado.
- Ilustraremos o procedimento de eliminação de restrições em um exemplo simples:

$$\begin{aligned} &\textit{Minimize} && f(x) = f(x_1, x_2, x_3, x_4) \\ &\textit{Sujeito a}: && \\ & && x_1 + x_3^2 - x_3 x_4 = 0 \\ & && -x_2 + x_4 + x_3 = 0 \end{aligned} \tag{9}$$

Programação Quadrática Seqüencial

- Não há risco algum em definirmos:

$$\begin{cases} x_1 = x_3 x_4 - x_3^2 \\ x_2 = x_4 + x_3^2 \end{cases} \tag{10}$$

- E trabalhamos com o problema irrestrito:

$$\underset{x_3, x_4}{\textit{Minimize}} \quad f(x_3 x_4 - x_3^2, x_4 + x_3^2, x_3, x_4) \tag{11}$$

- Este problema pode ser resolvido com os métodos de otimização irrestrita.

Sumário

Introdução

Categorizando Algoritmos de Otimização

Eliminação de Variáveis

Dificuldades

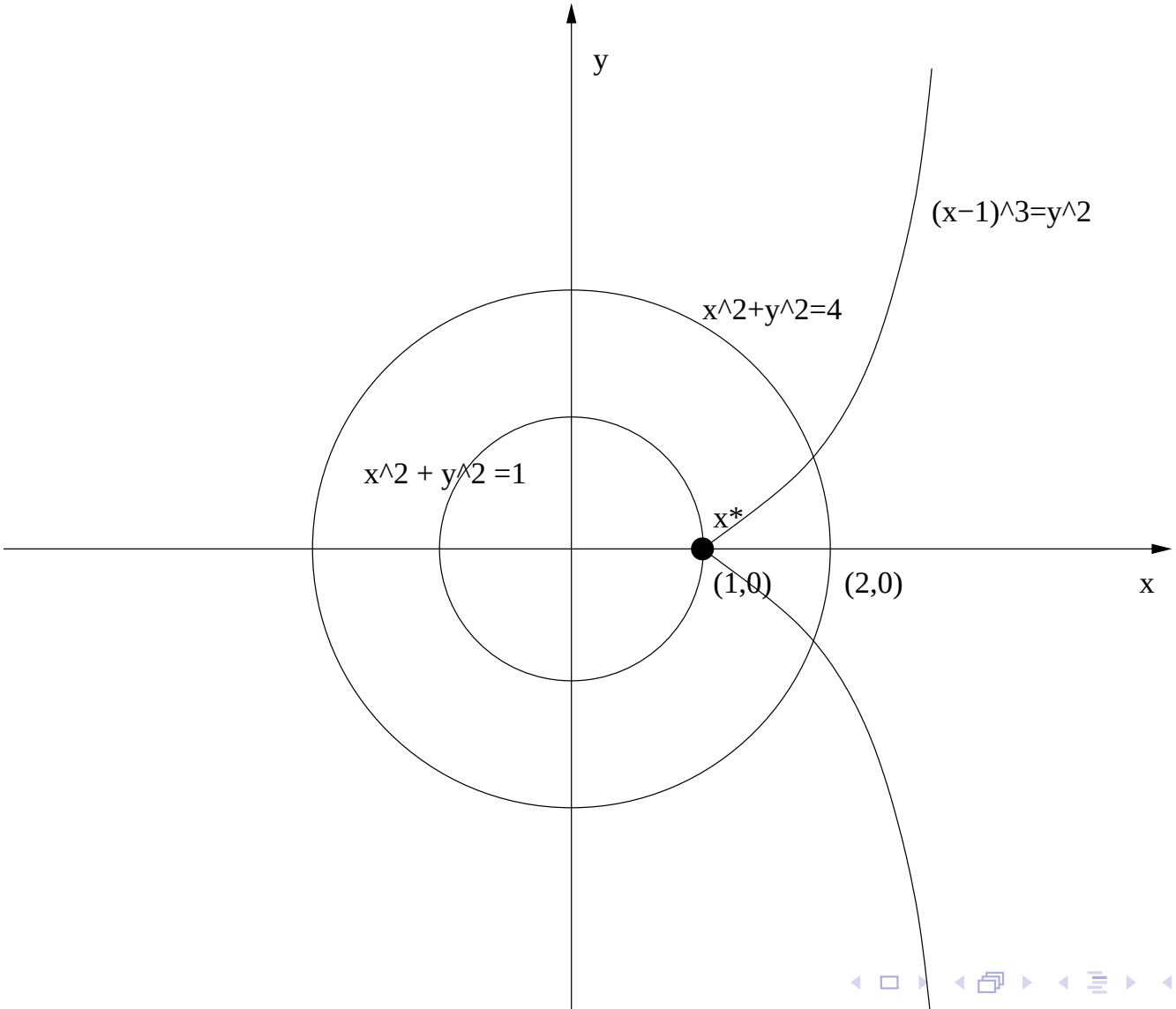
Problemas Que Podem Surgir

- Considere o problema:

$$\begin{aligned} &\textit{Minimize} && x^2 + y^2 \\ &\textit{Sujeito a}: && \\ & && (x-1)^3 = y^2 \end{aligned} \tag{12}$$

- A Figura abaixo mostra as curvas de nível para a função objetivo e a restrição.
- A mesma mostra que a solução ótima é $(x, y) = (1, 0)$.

Problemas Que Podem Surgir



Problemas Que Podem Surgir

- eliminamo a variável y e resolvemos o problema irrestrito:

$$\textit{Mininize } h(x) = x^2 + (x-1)^3 \tag{13}$$

- Claramente $h(x) \to -\infty$ se $x \to -\infty$.
- **O que aconteceu?**

Problemas Que Podem Surgir

- Ignoramos que a restrição $(x-1)^3 = y^2$ implicitamente impõe a restrição $x \geq 1$.
- Note que esta restrição está ativa na solução ótima.
- **Conclusão:** a eliminação de restrições não-lineares pode resultar em erros que são difíceis de se identificar.

Eliminação de Restrições Lineares

- Considere a minimização de uma função não-linear sujeita a um conjunto de igualdades lineares.

$$\begin{aligned} &\textit{Minimize} && f(x) \\ &\textit{Sujeito a}: && \\ & && Ax = b \end{aligned} \tag{14}$$

- Onde $A \in \mathbb{R}^{m \times n}$ e $m \leq n$.

Eliminação de Restrições Lineares

- Suponha que $rank(A) = m$.
- Sob essas condições, podemos encontrar um subconjunto de m colunas de A que são linearmente independentes.
- Podemos reunir essas colunas linearmente independentes na matriz B, definindo uma matriz $P \in \mathbb{R}^{n \times n}$.
- P é uma matriz de permutação que move as colunas linearmente independentes para as primeiras colunas de A.

Eliminação de Restrições Lineares

- Ou seja,

$$AP = \left[\begin{array}{c|c} B & N \end{array} \right] \tag{15}$$

- Onde N denota as $n - m$ colunas restantes.

Eliminação de Restrições Lineares

- Definimos os subvetores $x_B \in \mathbb{R}^m$ e $x_N \in \mathbb{R}^{n-m}$ de forma que:

$$P^T x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} \tag{16}$$

- E dizemos que x_B são as variáveis básicas e x_N as variáveis não-básicas.

Eliminação de Restrições Lineares

Notando que $PP^T = I$, podemos reescrever a restrição $Ax = b$ como:

$$\begin{aligned} b &= Ax \\ &= APP^T x \\ &= (AP)(P^T x) \\ &= \left[\begin{array}{c|c} B & N \end{array}\right] \left[\begin{array}{c} x_B \\ x_N \end{array}\right] \\ &= Bx_B + Nx_N \end{aligned} \tag{17}$$

Eliminação de Restrições Lineares

- A partir da fórmula acima, deduzimos que as variáveis básicas são dadas por:

$$x_B = B^{-1}b - B^{-1}Nx_N \tag{18}$$

- Podemos então computar um ponto factível para as restrições $Ax = b$ escolhendo qualquer valor de x_N, e então definindo x_B de acordo com a fórmula acima.
- O problema não-linear restrito é portanto equivalente ao problema irrestrito que segue abaixo:

$$Min \quad f(x_N) = f(P \begin{bmatrix} B^{-1}b - B^{-1}Nx_N \\ x_N \end{bmatrix}) \tag{19}$$

Eliminação de Restrições Lineares

Considere o problema:

$$\begin{aligned} &\textit{Minimize} \quad \sin(x_1 + x_2) + x_3^2 + \tfrac{1}{3}(x_4 + x_5^4 + \tfrac{x_6}{2}) \\ &\textit{Sujeito a}: \\ &\qquad 8x_1 - 6x_2 + x_3 + 9x_4 + 4x_5 = 6 \\ &\qquad 3x_1 + 2x_2 - x_4 + 6x_5 + 4x_6 = -4 \end{aligned} \tag{20}$$

Eliminação de Restrições Lineares

Definimos a matriz P de forma a reordenar os componentes de x como $x^T = (x_3, x_6, x_1, x_2, x_4, x_5)^T$, obtemos:

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{21}$$

Eliminação de Restrições Lineares

- E também:

$$AP = \begin{bmatrix} 1 & 0 & 8 & -6 & 9 & 4 \\ 0 & 4 & 3 & 2 & -1 & 6 \end{bmatrix} \tag{22}$$

- A matriz base B é diagonal, portanto fácil de se inverter.

Eliminação de Restrições Lineares

Com base em (18), obtemos:

$$\begin{bmatrix} x_3 \\ x_6 \end{bmatrix} = -\begin{bmatrix} 8 & -6 & 9 & 4 \\ \frac{3}{4} & \frac{1}{2} & \frac{-1}{4} & \frac{3}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 6 \\ -1 \end{bmatrix} \qquad (23)$$

Eliminação de Restrições Lineares

Substituindo-se x_3 e x_6 em (20), obtemos o seguinte problema irrestrito:

$$\begin{aligned} \textit{Minimize} \quad & f(x_1, x_2, x_4, x_5) = \operatorname{sin}(x_1 + x_2) \\ & +(8x_1 - 6x_2 + 9x_4 + 4x_5 - 6)^2 \\ x_1, x_2, x_4, x_5 \quad & +\tfrac{1}{3}\left[x_4 + x_5^4 - \left(\tfrac{1}{2} + \tfrac{3}{8}x_1 + \tfrac{1}{4}x_2 - \tfrac{1}{8}x_4 + \tfrac{3}{4}x_5\right)\right] \end{aligned}$$

Os Efeitos das Desigualdades

- Eliminação de variáveis não é sempre recomendada na presença de igualdades e desigualdades.
- Considere o problema:

$$
\begin{aligned}
&\textit{Minimize} && \sin(x_1 + x_2) + x_3^2 + \tfrac{1}{3}(x_4 + x_5^2 + \tfrac{x_6}{2}) \\
&\textit{Sujeito a}: && \\
& && 8x_1 - 6x_2 + x_3 + 9x_4 + 4x_5 = 6 \qquad (24) \\
& && 3x_1 + 2x_2 - x_4 + 6x_5 + 4x_6 = -4 \\
& && x_1, x_2, x_3, x_4, x_5, x_6 \geq 0
\end{aligned}
$$

Os Efeitos das Desigualdades

- Após eliminarmos as variáveis x_3 e x_6, conforme desenvolvimento anteiro, obteremos o seguinte conjunto de restrições:

$$\begin{array}{rcl} (x_1, x_2, x_4, x_5) & \geq & 0 \\ 8x_1 - 6x_2 + 9x_4 + 4x_5 & \leq & 6 \\ \frac{3}{4}x_1 + \frac{1}{2}x_2 - \frac{1}{4}x_4 + \frac{3}{2}x_5 & \leq & -1 \end{array} \tag{25}$$

- Portanto, o custo de eliminarmos as igualdades foi tornar as desigualdades muito mais complicadas.

Programação Não-Linear Restrita: Fundamentos de Algoritmos

- Fim!
- Obrigado pela presença