UNIVERSIDADE FEDERAL DO CEARÁ

PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM E MÉTODOS QUANTITATIVOS

OTIMIZAÇÃO NÃO-LINEAR

PROJETO COMPUTACIONAL III

ISMAEL FERNANDES BRITO

FORTALEZA - CE

2021

1 Fundamentos Teóricos dos Métodos

1.1 Otimização Restrita

Na matemática, otimizar algo significa minimizar ou maximizar uma função, com ou sem restrição nas suas variáveis. Denotaremos o problema de otimização da seguinte forma:

$$\begin{array}{ll} min & f(x) \\ sujeito\ a & c_i(x) = 0, i \in \mathcal{E}, \\ & c_i(x) \leqslant 0, i \in \mathcal{I}, \\ & x \in \mathbb{R}^n \end{array}$$

onde a função f e as funções c_i são reais e suaves no conjunto $\mathbb{R}^n$, e os índices $\mathcal{E}$ e $\mathcal{I}$ são finitos e disjuntos.

Podemos reescrever o problema anterior de forma compacta

$$\begin{array}{ll} min & f(x), \\ x \in \Omega & \end{array}$$

denotamos por Ω o conjunto de busca do problema onde tentamos minimizar f. Neste caso, temos $\Omega = \{x \in \mathbb{R} \mid c_i(x) = 0, i \in \mathcal{E}\ e\ c_i(x) \leqslant 0, i \in \mathcal{I}\}$, o qual chamaremos de conjunto viável. Uma solução viável que minimiza (ou maximiza) a função objetivo é chamada de uma solução ótima. Por convenção, a forma padrão de um problema de otimização é definida em termos de minimização. Geralmente, a menos que tanto a função objetivo quanto a região viável sejam convexas em um problema de minimização, pode haver alguns mínimos locais.

1.2 Método de Barreira

O método de barreira é utilizado para resolução de problemas com restrições de desigualdade, cujo interior é não vazio. Pode ser visto como um caso particular do método das penalidades, mas diferencia-se deste por exigir uma penalização interna, ou seja, por trabalhar no interior da região factível. Ao trabalhar no interior dessa região, os fatores de barreira impedem que os pontos saiam da região factível. Com isso, parte-se de um ponto factível e geram-se novos pontos factíveis. Uma das vantagens desse método é a obtenção de, pelo menos uma solução factível, caso ocorra uma parada prematura do método.

Vamos considerar o problema

$$\begin{array}{lc} min & f_0(x) \\ s.a & f_i(x) \leqslant 0,\ i = 1, 2, \cdots, m \end{array}$$

onde $f_i : \mathbb{R}^n \to \mathbb{R}$ para $i = 0, 1, \cdots, m$ são funções com derivadas segundas contínuas.

Supomos que existe um ponto interior, isto é, um ponto x tal que $f(x) < 0$.

Seguindo as estratégias de pontos interiores, acrescentamos variáveis de folga e uma penalidade logarítmica, obtendo o seguinte problema de barreira associado ao problema.

$$\begin{array}{lc} min & f_0(x) - \mu \sum_{i=1}^{m} log\ s_i \\ s.a & f(x) + s = 0 \\ & s > 0, \end{array}$$

em que $\mu > 0$ é o parâmetro de barreira e os vetores das variáveis de folga $s = (s_1, s_2, \cdots, s_m)^T$ são positivos.

O método de barreira para resolver o problema de desigualdade é estabelecido abaixo.

Algoritmo 1: Método de Barreira

Dados: Dado o problema restrito, construa a função auxiliar de barreira.

Faça $k = 0$, obtenha uma estimativa inicial $x_0 \in \mathbb{R}^n, \mu^0 > 0$ e o parâmetro de atualização da barreira $\beta > 1$:

1 enquanto $|\mu B(x)| > \epsilon$ **faça**

2 Resolva o problema barreira por um método de otimização irrestrita com μ^k com valor fixo.

3 Atualize o parâmetro de barreira por alguma heurística, $\mu^{k+1} = \frac{\mu^k}{\beta}$.

4 Determine: x^{k+1}.

5 fim

6 x^k é a solução para o problema restrito.

Exemplo Seja o problema

$$\begin{array}{cc} min & f(x) = x, \\ sujeito\ a & g(x) = -x + 3 \leqslant 0. \end{array}$$

Solução

Usando a função barreira, tem-se:

$$f(x) + b(x, u) = x + u\frac{1}{x-3}$$

Assim:

$$\frac{d}{dx}\left(x + u\frac{1}{x-3}\right) = 1 - u\frac{1}{(x-3)^2} = 0 \to x^* = 3 + \sqrt{u}$$

Com $u \to 0^+,\ temos\ x^* \to 3^+$.

1.3 Método de Penalidade

Métodos de penalidade são aplicados a problemas de programação não linear com restrições. A base dos métodos de penalidade é transformar o problema restrito original em uma sequência de problemas sem restrições, derivados do inicial, acrescentando um termo de penalidade à função objetivo, possibilitando sua resolução pelos métodos conhecidos para este tipo de problema. Este termo penaliza, ou seja, aumenta, o valor da função objetivo quando as restrições não forem satisfeitas.

As técnicas ou estratégias utilizadas em cada método dependem da quantidade de informação disponível e possível de ser utilizada, da maior ou menor eficiência e robustez do algoritmo a utilizar, da facilidade de implementação, e das especificidades do próprio problema.

consideramos uma substituição do problema de otimização com restrições, pelo problema de minimizar uma função que é formada pela função objetivo original do problema, e um termo adicional positivo quando a solução atual viola a restrição, ou é nula, caso contrário.

Definimos a função $Q(x; \mu)$, com $x \in \Omega$ e $\mu > 0$ como

$$Q(x; \mu) = f(x) + \mu P(x),$$

onde $\mu > 0$ é o parâmetro de penalidade e $P(x)$ é a função de penalidade escolhida. A função $P : \mathbb{R}^n \to \mathbb{R}$ deve ser contínua e satisfazer

$$P(x) = \begin{cases} 0, & se\ x \in \Omega, \\ > 0, & se\ x \notin \Omega. \end{cases}$$

Apesar desta estratégia raramente poder ser usada na prática, suas características teóricas são base para algoritmos mais sofisticados, inclusive o método de lagrangeana aumentada.

Algoritmo 2: Método de Penalidade

Dados: Dado o problema restrito, construa a função auxiliar de barreira.

Faça $k = 0$, obtenha uma estimativa inicial $x_0 \in \mathbb{R}^n, \mu^0 > 0$ e o parâmetro de atualização da barreira $\beta > 1$:

1 enquanto $|\mu B(x)| > \epsilon$ **faça**

2 Resolva o problema barreira por um método de otimização irrestrita com μ^k com valor fixo.

3 Atualize o parâmetro de barreira por alguma heurística, $\mu^{k+1} = \frac{\mu^k}{\beta}$.

4 Determine: x^{k+1}.

5 fim

6 x^k é a solução para o problema restrito.

A teoria prevê que a solução converge para a solução ótima quando $u_k \to \infty$ (*ou* $u_k \to 0^+$), porém, a hessiana da função objetivo penalizada depende de u_k e pode apresentar mal-condicionamento numérico, a variação de u_k deve ser gradual e a precisão final fica bastante limitada. Diversas modificações podem ser feitas na abordagem por penalidades, tais como, usar termos de penalidade distintos para cada restrição evitando problemas de escala e penalizar apenas a restrição mais violada.

1.4 Método Lagrangeana Aumentada

O uso de algoritmos de penalidade pode gerar sequências extremamente difíceis de calcular devido ao mal condicionamento da penalidade μ relacionada com o cálculo das restrições e das derivadas de primeira e segunda ordem envolvidas. Para evitar o surgimento destes problemas recorremos, em geral, ao uso de **multiplicadores de Lagrange** para controlar o comportamento da função fora do conjunto viável.

O próximo algoritmo para a minimização de funções restritas propõe uma transformação na função original para que seja feita a resolução de um problema irrestrito. Considere a função lagrangeana com apenas uma restrição de igualdade dada por:

$$L(x, \lambda_1) = f(x) - \lambda \cdot c(x)$$

Onde, pelas condições de KKT, encontramos pontos estacionários em:

$$\nabla L(x, \lambda) = \nabla f(x) - <\lambda \cdot \nabla c(x)> = 0$$

Dizemos que $\lambda_i \geqslant 0$ que resolve a equação acima é um multiplicador de Lagrange.

A Lagrangeana aumentada é um método que usa as propriedades da função lagrangeana para calcular com maior eficácia as novas iterações do algoritmo já que a lagrangeana estima um valor exato λ que minimiza a função sem a necessidade de tornar μ um escalar pequeno. A fórmula geral é dada por:

$$L(x, \lambda, \mu) = f(x) - \sum \lambda_i \cdot c_i(x) + \frac{1}{2 \cdot \mu} \cdot \sum c_j^2(x)$$

em que temos por objetivo usar o valor real de λ_i em vez de uma estimativa para μ que decresce a cada iteração.

O método do Lagrangeana Aumentada é uma extensão dos métodos de penalidades, porém sem o problema de tornar a função de penalidade mal-condicionada com a variação de u. As vantagens dele é que a penalidade u não precisa tender a infinito, pode ser aumentada de forma suave, a solução inicial não precisa ser factível, é possível encontrar $g_i(x) = 0$ e $h_j(x) = 0$ com precisão e os multiplicadores de Lagrange diferentes de zero identificam as restrições ativas no ponto solução.

Algoritmo 3: Método de Lagrangeana Aumentada

Dados $\mu_0 > 0, \epsilon > 0$ uma sequência não negativa $\{\epsilon_k\}$ com $\epsilon_k \to 0, x_0$ e λ^0
para $k = 0, 1, 2, \cdots$ **faça**
    Encontre x tal que $||\nabla L_A(., \lambda^k, \mu_k)|| < \epsilon_k$
    se $||\nabla L(x^k, \lambda^k)|| \leqslant \epsilon$ e $||c(x^k)|| \leqslant \epsilon$ **então**
        Retorne x^k
    fim se
    Atualizar os multiplicadores de Lagrange usando $\lambda_i^{k+1} = \lambda_i^k + \mu_k c_i(x_k)$, para $i = 1, \cdots, m$
    Escolha um novo parâmetro de penalidade $\mu_{k+1} \geqslant \mu_k$
fim para

Exemplo (Luenberger, pg. 409)

$$\begin{array}{cc} minimizar & 2x_1^2 + 2x_1x_2 + x_2^2 - 2x_2 \\ sujeito\ a & x_1 = 0 \end{array}$$

- **Lagrangeana Aumentada:**
  $L_c(x_1, x_2, \lambda) := 2x_1^2 + 2x_1x_2 + x_2^2 - 2x_2 + \lambda x_1 + \frac{1}{2}cx_1^2$
- **Mínimo Analítico:**
  $x_1(\lambda) = -\frac{(2+\lambda)}{(2+c)},\ x_2(\lambda) = \frac{(4+c+\lambda)}{(2+c)}$
- **Processo iterativo:**$(h(x) = x_1)$
  $\lambda^{k+1} = \lambda^k - \frac{c(2+\lambda^k)}{(2+c)} = \frac{2}{2+c}\lambda^k - \frac{2c}{2+c}$
  e $\lambda^k \to -2(= \lambda^*)$, qualquer que seja $c > 0(= c^*)$

2 Manual de utilização dos métodos no Octave

Os dados de entrada foram armazenados em arquivo texto padrão do Matlab/Octave com o nome de "funcao.m"que deve ser carregado na memória antes de fazer a chamada dos métodos computacionais. Neste arquivo contém todas as 10 funções pedidas que foram criadas como fcn1 até fcn10 juntamente com as suas respectivas derivadas criadas como dfcn1 até dfcn10 que serão utilizadas nos métodos que usam a informação da derivada e também os vetores de entrada $x^{(1)}$ foram criados como fcn1x1 até fcn10x10 e para o $x^{(2)}$ foram criados como fcn1x2 até fcn10x10 que serão utilizados nos métodos que usam como parâmetro a informação do ponto inicial.

Outra entrada utilizada será a informação do intervalo de incerteza (0,1) que será utilizada no método sem o uso da informação da derivada no qual será inserido como vetor linha da seguinte forma [0 0] e [1 1]. Em todos os métodos tem um parâmetro n que significa o número máximo de iterações.

Em geral a sintaxe para executar os métodos comupacionais no shell do Octave é a seguinte: metodo(parametros). A saída será o valor de xmin para o valor de x_{min}, o valor de fval para $f(x_{min})$ e a quantidade de iterações com k para o método de fibonacci. Já no método da interpolação a saída será alphak para o valor de α_k. O valor de xmin, fval e o número de iterações foram feitos utilizando uma função com comando fminunc do próprio Octave através do seguinte comando no shell dele da seguinte forma: [xmin, fval, info, output] = fminunc(funcao, ponto-inicial), por exemplo: "[xmin, fval, info, output] = fminunc(fcn1, fcn1x1)".

Já a saída do método da condição de wolfe será mostrado o valor de alphak para α_k. Por fim, todos os método foram implementados como funções e salvas com extensão ".m"que é o padrão do Octave. Os arquivos dos mesmos irão junto com este trabalho.

3 Resultados dos Experimentos

3.1 Dados Utilizados nos Testes

Dados utilizados no **Método de Barreira**:

* Primeiro problema

$$\begin{array}{ll} min & f_{11}(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\ s.a & x_1 - x_2^2 \leqslant 0 \\ & x_1^2 - x_2 \geqslant 0 \\ & \mathbf{x}^0 = [-2\ 1]^t \end{array}$$

* Segundo problema

$$\begin{array}{ll} min & f_{12}(\mathbf{x}) = x_1^2 + 0.5x_2^2 + x_3^2 + 0.5x_4^2 - x_1x_3 + x_3x_4 - x_1 - 3x_2 + x_3 - x_4 \\ s.a & 5 - x_1 - 2x_2 - x_3 - x_4 \geqslant 0 \\ & 4 - 3x_1 - x_2 - 2x_3 + x_4 \geqslant 0 \\ & x_2 + 4x_3 - 1.5 \geqslant 0 \\ & x_i \geqslant 0, \forall i \in \{1,2,3,4\} \\ & \mathbf{x}^0 = [0.5\ 0.5\ 0.5\ 0.5]^t \end{array}$$

* Terceiro problema

$$\begin{array}{ll} min & f_{13}(\mathbf{x}) = (x_1 - 10)^3 + (x_2 - 20)^3 \\ s.a & (x_1 - 5)^2 + (x_2 - 5)^2 - 100 \geqslant 0 \\ & (x_1 - 6)^2 + (x_2 - 5)^2 \geqslant 0 \\ & 82.81 - (x_1 - 6)^2 - (x_2 - 5)^2 \geqslant 0 \\ & \mathbf{x}^0 = [14.115\ 0.885]^t \end{array}$$

Dados utilizados no **Método de Penalidade e Lagrangeana Aumentada**:

* Primeiro problema

$$\begin{array}{ll} min & f_{21}(\mathbf{x}) = 0.01(x_1 - 1)^2 + (x_2 - x_1^2)^2 \\ s.a & x_1 + x_3^2 + 1 = 0 \\ & \mathbf{x}^0 = [2\ 2\ 2]^t \end{array}$$

* Segundo problema

$$\begin{array}{ll} min & f_{22}(\mathbf{x}) = (x_1 - x_2)^2 + (x_2 + x_3 - 2)^2 + (x_4 - 1)^2 + (x_5 - 1)^2 \\ s.a & x_1 + 3x_2 - 4 = 0 \\ & x_3 + x_4 - 2x_5 = 0 \\ & x_2 - x_5 = 0 \\ & \mathbf{x}^0 = [2.5\ 0.5\ 2\ \ -1\ 0.5]^t \end{array}$$

* Terceiro problema

$$\begin{array}{ll} min & f_{23}(\mathbf{x}) = (x_1 - 2)^2 + (x_2 - 1)^2 \\ s.a & 0.25x_1^2 + x_2^2 - 1 \leqslant 0 \\ & x_1 2x_2 + 1 = 0 \\ & \\ & \mathbf{x}^0 = [2\ 2]^t \end{array}$$

Dados utilizados em **Todos os métodos**:

* Primeiro problema

$$\begin{array}{ll} min & f_{31}(\mathbf{x}) = x_1x_4(x_1 + x_2 + x_3) + x_3 \\ s.a & x_1x_2x_3x_4 - 25 \geqslant 0 \\ & x_1^2 + x_2^2 + x_3^2 + x_4^2 - 40 \geqslant 0 \\ & 1 \leqslant x_i \leqslant 5, \forall i \in \{1,2,3,4\} \\ & \mathbf{x}^0 = [3\ 4\ 4\ 3]^t \end{array}$$

* Segundo problema

$$\begin{array}{ll} min & f_{32}(\mathbf{x}) = x_1 - x_2 - x_3 - x_1x_3 + x_1x_4 + x_2x_3 - x_2x_4 \\ s.a & 8 - x_1 - 2x_2 \geqslant 0 \\ & 12 - 4x_1 - x_2 \geqslant 0 \\ & 12 - 3x_1 - 4x_2 \geqslant 0 \\ & 8 - 2x_3 - x_4 \geqslant 0 \\ & 8 - x_3 - 2x_4 \geqslant 0 \\ & 5 - x_3 - x_4 \geqslant 0 \\ & x_i \geqslant 0, \forall i \in \{1,2,3,4\} \\ & \mathbf{x}^0 = [1\ 1\ 1\ 1]^t \end{array}$$

* Terceiro problema

$$\begin{array}{ll} min & f_{33}(\mathbf{x}) = (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2 + 10x_5^6 + 7x_6^2 + x_7^4 - 4x_6x_7 - 10x_6 - 8x_7 \\ s.a & 127 - x_1^2 - 3x_2^4 - x_3 - 4x_4^2 - 5x_5 \geqslant 0 \\ & 282 - 7x_1 - 3x_2 - 10x_3^2 - x_4 + x_5 \geqslant 0 \\ & 196 - 23x_1^2 - x_2^2 - 6x_6^2 + 8x_7 \geqslant 0 \\ & -4x_1^2 - x_2^2 + 3x_1x_2 - 2x_3^2 - 5x_6 + 11x_7 \geqslant 0 \\ & \mathbf{x}^0 = [1\ 2\ 0\ 4\ 0\ 1\ 1]^t \end{array}$$

Em todos os métodos implementados foi utilizado os seguintes parâmetros:

Tabela 1: Dados com os resultados dos testes

3.2 Resultados

Problema	xmin	fval
1	[1.1194 1.2508]	2.1867e-03
2	[0.1112 2.2408 -0.1852 0.5925]	-4.9676
3	—	—
4	—	—
5	[1.00 1.00 1.00 1.00]	2.9089e-17
6	—	—
7	—	—
8	—	—
9	—	—

Referências

[1] `https://pt.wikibooks.org/wiki/Otimiza%C3%A7%C3%A3o/M%C3%A9todos_de_regi%C3%A3o_de_confian%C3%A7a`

[2] John W. Eaton, David Bateman, Søren Hauberg, Rik Wehbring (2020). GNU Octave version 6.1.0 manual: a high-level interactive language for numerical computations. `https://www.gnu.org/software/octave/doc/v6.1.0/`

[3] `https://www.ime.unicamp.br/~sandra/MS629/handouts/livro28jul.pdf`

[4] `https://www.ime.unicamp.br/~friedlan/livro.pdf`

[5] Nonlinear Programming - M. S. Bazaraa & C. M. Shetty - John Wiley & Sons, 1979.

[6] J. M. Martínez, Otimização Prática Usando o Langrageano Aumentado. IMECC - UNICAMP, 2009.