Trabalho Prático 1 – Otimização Não Linear

2024/01 Turma S00
Professores Michel Bessani e João Pedro A.F. Campos

João Antônio Viola
Matrícula 202027040

Rafael Camacho de Oliveira
Matrícula 2020027040

***Resumo*—No presente trabalho, será feito um estudo sobre a otimização não linear por meio da resolução e aprofundamento em um problema relacionado à área. O objetivo do trabalho é alcançar a solução do problema, passando pelas diferentes etapas de resolução, e aprofundando em suas nuances. Serão utilizadas ferramentas computacionais para gerar a representação gráfica em duas e três dimensões, as restrições e o ponto ótimo do problema, além de métodos matemáticos como as condições de Karush-Kuhn-Tucker e o método de multiplicadores de Lagrange. Serão apresentaremos os resultados obtidos e, por fim, uma interpretação destes.**

***Palavras-chave*—Otimização, Linearidade, Não-Linearidade**

I. Introdução

Esse trabalho prático da disciplina de Otimização Não Linear tem como objetivo explicitar a resolução de um problema de programação matemática com ênfase nos métodos utilizados, e, em seguida, analisar os resultados obtidos em cada etapa do processo. Serão utilizadas ferramentas matemáticas computacionais, assim como técnicas de cálculo manuais derivados do cálculo de várias variáveis e da geometria analítica e álgebra linear.

O problema resolvido tem como fundamentação uma contextualização em uma empresa de transporte de combustíveis, onde deseja-se minimizar o custo de armazenamento dos diferentes tipos de combustíveis existentes. Os parâmetros do problema são: os custos c_i, em dólares do armazenamento e os custos c_i^m da manutenção de cada tipo de combustível.

As variáveis de entrada do sistema são os dois tipos de combustíveis $[x_1 \; e \; x_2]$ em pés cúbicos [ft³].

É apresentada a função $f(x)$ que relaciona o custo aos parâmetros e variáveis:

$$\mathrm{f}(\mathbf{x}) = \sum_{i=1}^{N} c_i(x_i + 10)^2 + \; c_i^m(x_i + 5) \quad (1)$$

Assim, podemos construir a função objetivo do problema de minimização, que se trata da busca pelos valores de $\boldsymbol{x}$ que geram uma imagem na função f(x) que seja o menor valor possível, ou ponto ótimo, representado por $\mathbf{x}^*$. Sendo assim, a função objetivo do problema pode ser escrita como:

$$\boldsymbol{x}^* = arg\,\min_x f(x) = \sum_{i=1}^{N} c_i(x_i + 10)^2 + \; c_i^m(x_i + 5) \quad (2)$$

No entanto, o armazenamento dos combustíveis deve sempre respeitar um volume mínimo V_{min} armazenado de 50ft³. Na formulação matemática, isso é apresentado na forma de uma restrição de desigualdade:

$$g(x) = \sum_{i=1}^{N} x_i \; \geq V_{min} \quad , \quad (3)$$

$$V_{min} = 50\mathrm{ft}^3 \quad (4)$$

$g(x)$ garante que a soma dos valores de $x_1 \; e \; x_2$ seja sempre maior ou igual ao valor mínimo de 50 ft³, o que, graficamente, como iremos analisar nos próximos passos, representa uma divisão na superfície da função objetivo, separando-a em uma região factível e outra infactível.

O problema então fornece valores reais para os parâmetros, explicitados na tabela a seguir:

Tabela 1: Parâmetros do problema.

Combustível (i)	c_i (\$)	c_i^m (\$)
1	2.3	0.47
2	4.27	0.26

II. Fundamentação teórica

Primeiramente o problema foi solucionado de forma analítica utilizando cálculo de várias variáveis. Essa etapa envolveu aplicação e análise de métodos de otimização como condições necessárias para a obtenção de um ponto ótimo, como o uso de derivadas parciais e estabelecimento de uma variável de folga para as restrições sejam satisfeitas simultaneamente. Em seguida as condições de Karush-Kuhn-Tucker (KKT) foram aplicadas para obter a solução final. As condições de KKT são cruciais em problemas de otimização não linear sujeitos a restrições, pois oferecem um conjunto de condições necessárias e suficientes para a otimalidade de um problema.

Em seguida o problema foi solucionado novamente, porém de forma computacional. Foi criada uma aplicação em python utilizando uma série de bibliotecas para auxílio da resolução do problema, como numpy para resolução de problemas matemáticos e álgebra linear, matplotlib para criação dos gráficos e modelos 3D e scipy.optimize para encontrar o ponto mínimo da função objetivo utilizando SLSQP (Sequential Least Squares Programming) levando em consideração as restrições propostas. Desse modo, a função objetivo e as restrições foram modeladas e em seguida o gráfico para plotar os resultados foi criado para que então o processo de otimização fosse iniciado com um valor chute de X0 = [25, 25] para que o algoritmo do scipy.optimize encontre o ponto mínimo e os valores de X correspondentes a ele respeitando as restrições propostas e esses valores descritos no gráfico assim como a função objetivo e restrições.

III. MATERIÁIS E MÉTODOS

A resolução do problema será feita de forma computacional, por meio de gráficos, e de forma manual com a aplicação dos métodos de multiplicadores de Lagrange, condições necessárias de primeira ordem e condições suficientes de segunda ordem.

O primeiro passo da resolução é a caracterização da função objetivo, que está explicitada na equação (1) na introdução do trabalho.

$$\boldsymbol{x}^* = arg \min_x f(x) = \sum_{i=1}^{N} c_i(x_i + 10)^2 + c_i^m(x_i + 5)$$

Com a função objetivo, podemos encontrar a solução para o problema irrestrito, isto é, sem considerar a restrição de desigualdade $g(x)$. Sabemos que o ponto ótimo do problema irrestrito sempre estará, necessariamente, em um ponto de mínimo local, para o caso de um problema de minimização. Dessa forma, nossa primeira busca será pelos pontos candidatos a serem mínimos. Encontraremos estes verificando os pontos da função objetivo que tenham os valores de suas respectivas derivadas valham zero, pois assim sabemos que nesses pontos a função mudou de direção. Para acharmos esses pontos, chamados pontos críticos, fazemos uso do vetor gradiente $\nabla f(x)$, que é a soma de todas as derivadas parciais da função $f(x)$.

Primeiramente, substituindo os valores dos parâmetros na função objetivo para os dois tipos de combustível obtemos:

$$\min_{x\ irrestrito} f(x) = 2.3(x_1 + 10)^2 + 0.47(x_1 + 5) + 4.27(x_2 + 10)^2 + 0.26(x_2 + 5)$$

(5)

Então, podemos calcular o gradiente e igualá-lo a zero:

$$\nabla f(x) = \begin{pmatrix} 4.6(x_1 + 10) + 0.47 \\ 8.54(x_2 + 10) + 0.26 \end{pmatrix} = 0$$

(6)

E por fim encontrar os valores de $x_1\ e\ x_2$ que formam o ponto crítico:

$$\nabla f(x) = \begin{pmatrix} -10.10 \\ -10.03 \end{pmatrix}$$

(7)

Esse resultado é chamado de Condição Necessária de Primeira Ordem (CNPO). No entanto, a informação de que o gradiente no ponto vale zero não nos permite afirmar que este é o ponto ótimo do problema irrestrito, uma vez que pode ser tanto um ponto de mínimo quanto um máximo ou um ponto de inflexão. Para verificarmos se o ponto é um mínimo, precisaremos verificar a Condição Suficiente de Segunda Ordem (CSSO).

Isso é feito analisando as segundas derivadas parciais da função objetivo, que indicará onde a função é côncava, convexa e onde tem um ponto de inflexão.

Com as derivadas de segunda ordem, podemos construir a matriz Hessiana da função $f(x)$ sendo:

$$H\left[f(x,y)\right] = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\,\partial y} \\ \\ \frac{\partial^2 f}{\partial y\,\partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \text{, Portanto } H = \begin{bmatrix} 4.6 & 0 \\ 0 & 8.54 \end{bmatrix}$$

Para o caso de H ser definida positiva, o que pode ser verificado por meio de seus autovalores, podemos afirmar que a função $f(x)$ é uma função estritamente convexa. Como os autovalores em questão são facilmente encontrados e sendo positivos (λ1 = 4.6 e λ2 = 8.54), sabemos que a função é convexa e, portanto, que o ponto encontrado é um mínimo.

$$\boldsymbol{x}^*_{irrestrito} = \begin{pmatrix} -10.10 \\ -10.03 \end{pmatrix} \text{ e } f(\boldsymbol{x}^*_{irrestrito}) = -3.68$$

Agora é necessário verificar se o ponto encontrado respeita a restrição de desigualdade (3)

$$g(x) = \sum_{i=1}^{N} x_i \geq 50$$

Isso é feito substituindo os valores de $x_1\ e\ x_2$:

$$g(x) = x_1 + x_2 = -10.10 + (-10.03)$$

$$g(x) = -20.13$$

(4)

Como o valor de $g(x)$ deve ser maior ou igual a 50, o ponto ótimo do problema irrestrito viola a restrição, sendo assim infactível. Com isso sabemos que a restrição $g(x)$ estará ativa. Isso significa que o ponto ótimo restrito está, necessariamente, sobre $g(x)$, no ponto onde a restrição tangencia a curva de nível onde a função objetivo tem o menor valor.

A busca pelo ótimo restrito então se torna a busca pelo ponto sobre a restrição onde o gradiente $\nabla f(x)$ tem mesma direção e sentidos opostos ao gradiente da restrição $\nabla g(x)$. Para isso, faremos uso das Condições de Karush-Kuhn-Tucker (ou KKT) para encontrarmos o novo ponto ótimo que respeite a restrição e seja, portanto, factível.

Para fins de formalidade, iremos inverter a restrição $g(x)$ tornando-a:

$$g_1(x) = -g(x) = \sum_{i=1}^{N} -x_i \leq -50$$

(5)

Podendo então defini-la como:

$$g_1(x) = -x_1 - x_2 + 50 \leq 0$$

(6)

A condição necessária de primeira ordem (CNPO) de um problema restrito, ou KKT, são um conjunto de condições necessárias para que um ponto seja uma solução ótima para um problema de otimização com restrições.

Esse método surgiu a partir de uma generalização do método dos multiplicadores de Lagrange, de forma a incluir restrições de desigualdade múltiplas, além de restrições de igualdade. A técnica dos multiplicadores de Lagrange permite encontrar o máximo ou o mínimo de uma função multivariável, quando há alguma restrição sobre os valores de entrada que podem ser usados.

Definimos, primeiramente, a função Lagrangiana de um problema geral com restrições de igualdade $h_i(x)$ e de desigualdade $g_i(x)$:

$$L(x,\lambda,\beta) := f(x) + \beta[g(x)+z^2] + \lambda h(x) = 0$$

$$= f(x) + \sum_{i=1}^{q} \beta_i[g_i(x)+z_i^2] + \sum_{i=1}^{p} \lambda_i h_i(x)$$

(7)

Sendo β_i e λ_i os respectivos multiplicadores de Lagrange das restrições e z_i^2 a variável de folga complementar.

No caso do problema estudado, temos apenas uma restrição de desigualdade $g(x)$:

$$L(x,\beta,z^2) = f(x) + \beta[g_1(x)+z^2] = 0$$

(8)

Com isso, podemos escrever as Condições de KKT, no ponto ótimo $\boldsymbol{x}*$:

$$\nabla L(x,\beta,z^2) = \nabla f(x) + \beta \nabla g_1(x) = 0$$

(9)

$$2\beta z = 0$$

(10)

$$g_1(x) + Z^2 = 0(9)$$

(11)

Podemos juntar as duas últimas equações para eliminar a variável de folga complementar:

$$\beta^* g_1(x) = 0$$

(12)

E como $g_1(x) \leq 0$, sabemos que $\beta^* \geq 0$. Se a restrição $g_1(x)$ está ativa, isto é, se o ponto ótimo está sobre ela, então $g_1(x) = 0$. Portanto, $\beta^* > 0$.

Resolvendo, por fim, o problema utilizando as condições KKT, temos:

$$L(x,\beta) = f(x) + \beta g_1(x) = 0$$

(13)

$$L(x,\beta) = 2.3(x_1+10)^2 + 0.47(x_1+5) + 4.27(x_2+10)^2 + 0.26(x_2+5) + \beta[-x_1 - x_2 + 50 + z^2] = 0$$

(14)

$$\nabla L(x,\beta) = \begin{pmatrix} 4.6(x_1+10)+0.47-\beta \\ 8.54(x_2+10)+0.26-\beta \end{pmatrix} = 0$$

(15)

Isolando x_1 e x_2 , podemos substituir e igualar as duas equações para encontrar o valor de β:

$$x_1 = \frac{-(46.47-\beta)}{4.6} \quad \text{e} \quad x_2 = \frac{-(85.66-\beta)}{8.54}$$

Como $x_1 + x_2 - 50 = 0$, substituindo as expressões acima, obtemos $\beta = 209.672$.

Assim, aplicando o valor de β na equação (15), obtemos os valores de x_1 e x_2

$$x_1 = 35.48$$

$$x_2 = 14.52$$

Como provado anteriormente, a função $f(x)$ é convexa. Dessa forma, as condições KKT se tornam não apenas necessárias, mas também suficientes para tomarmos o ponto ótimo restrito encontrado como um mínimo.

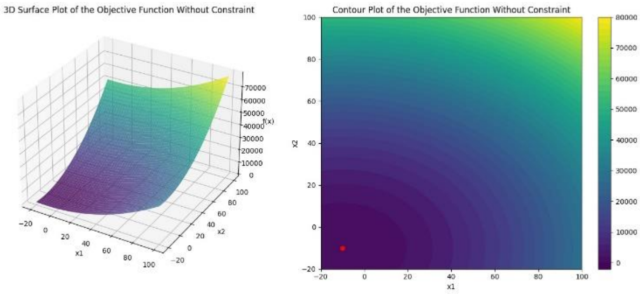
O valor da função objetivo neste ponto é:

$$f(x^*) = 7348.75$$

Tais resultados serão agora comparados à solução gráfica feita computacionalmente, explorados na sessão a seguir.

IV. Resultados

O resultado encontrado para o problema irrestrito foi ilustrado graficamente abaixo:

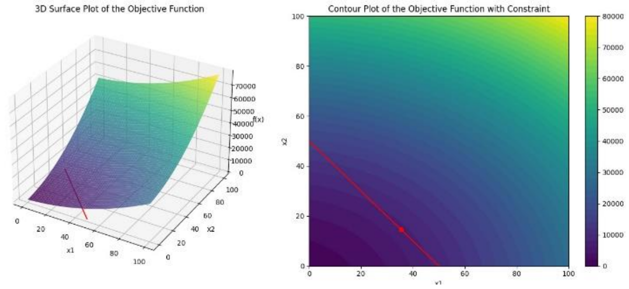


$$\boldsymbol{x}^*_{irrestrito} = \begin{pmatrix} -10.10 \\ -10.03 \end{pmatrix} \text{ e}$$

$$f(\boldsymbol{x}^*_{irrestrito}) = -3.68$$

Tal resultado não representa um valor factível como solução para um problema real, já que as variáveis de entrada x_1 e x_2 são quantidades de volume de combustível, não podendo, portanto, assumir valores negativos. Por isso é aplicada a restrição de desigualdade $g(x)$ que delimita um volume mínimo que deve ser armazenado, que graficamente se representa como uma reta que separa a curva em uma região factível e outra infactível.

A solução do problema restrito então assume a forma:



Onde

$$\boldsymbol{x}^*_{restrito} = \begin{pmatrix} 35.48 \\ 14.52 \end{pmatrix} \quad e\ f(\boldsymbol{x}^*_{restrito}) = 7348.75$$

Neste trabalho prático, exploramos a resolução de um problema de otimização não linear na área de transporte de combustíveis com o objetivo de minimizar custos de armazenamento. Utilizamos métodos matemáticos como as condições de Karush-Kuhn-Tucker (KKT) e o método de multiplicadores de Lagrange para abordar o problema, ambos sob a perspectiva teórica e computacional.

Os resultados computacionais confirmaram a solução analítica, onde o ponto ótimo restrito foi encontrado considerando a restrição de volume mínimo necessário. O ponto ótimo restrito obtido foi (x1 ,x2)=(35.48, 14.52) com um custo associado de $7348.75, evidenciando que a abordagem de otimização proposta foi eficaz.

A aplicação das condições de KKT provou ser não apenas necessária, mas também suficiente para confirmar que o ponto encontrado é de fato um mínimo local, devido à natureza convexa da função objetivo. Essa análise demonstra a robustez do método KKT para lidar com problemas complexos que incluem restrições não lineares e múltiplas variáveis.

V. Conclusão

Desse modo, pode-se concluir que,

Referências bibliográficas

[1] FG Teixeira. OTIMIZAÇÃO COM RESTRIÇÕES: TEOREMA DE KARUSH-KUHN-TUCKER, UFMG, abril, 2012

[2] .Singiresu S. Rao, Engineering Optimization: Theory and Practice, Wiley, 4th ed., 2009.