

SOLUÇÃO EFICIENTE DE PROBLEMAS DE PROGRAMAÇÃO LINEAR DE GRANDE PORTE

GUILHERME GENTILE DE AGUIAR[†] - RA: 105020

RESUMO. Nas últimas décadas vários métodos vêm sendo propostos para a resolução de um problema de programação linear, em especial os métodos de pontos interiores. A pesquisa sobre métodos de resolução de problemas de programação linear ainda continua, visto que com o desenvolvimento da computação, e com a necessidade real, cada vez mais é necessário resolver problemas maiores em menor tempo.

Na década de 1990, Dantzig divulgou um algoritmo proposto por Von Neumann em 1948. Do algoritmo de Neumann foi proposto um algoritmo de ajustamento pelo par ótimo que trabalha com duas variáveis. Em cada iteração do algoritmo é resolvido um subproblema, o que gera o maior esforço computacional do algoritmo. A partir disso, foi proposto um algoritmo que trabalha com p variáveis, igualmente o maior esforço computacional vêm de resolver um subproblema. Ao resolvermos o subproblema por métodos de pontos interiores, o grande esforço vêm de resolver sistemas lineares.

Para resolvermos um sistema linear por métodos diretos é necessário a fatoração de uma matriz, no caso de ordem $p + 1$, o que possui complexidade $O((p + 1)^3)$, e em seguida são resolvidos dois sistemas lineares mais simples, os quais estão associados a matrizes triangulares, que possuem complexidade $O((p+1)^2)$. O que é um custo muito baixo, se considerarmos que um problema de programação linear de grande porte.

1. INTRODUÇÃO

A otimização está presente no nosso dia a dia. Sempre que precisamos tomar uma decisão, procuramos escolher entre as várias alternativas aquela que, naquele momento, nos dê maior satisfação. Nos últimos anos, a matemática vem estudando como encontrar essa solução que dê nos essa maior satisfação.

Na matemática, a área que estuda problemas de otimização é geralmente chamada de Programação Matemática. Esta denominação identifica uma ampla classe de problemas. O nome programação foi empregado porque os militares se referem ao planejamento de atividades como “programa”. Boa parte dos acontecimentos que culminaram com a criação desta importante área da matemática, se deram durante a Segunda Guerra Mundial. De fato, George B. Dantzig usou o termo Programação Linear (mais especificamente “Programming in a linear Structure”, mais tarde resumido para Linear Programming e generalizado como Mathematical Programming) para analisar um problema de planejamento para a força aérea americana. Com a disseminação do uso do computador, programação passou a ser entendido como a codificação de um algoritmo em uma determinada linguagem e às vezes a Programação Matemática é confundida com programação de computadores. Linguagens de programação e computadores são muito usados no estudo de problemas de otimização, mas a Programação Matemática é muito mais do que a codificação de um algoritmo.

[†]Graduando de Matemática Aplicada e Computacional da Unicamp e Bolsista de Iniciação Científica do CNPq. Email: aguiar.guig@gmail.com.

No nosso caso queremos estudar os problemas de programação linear, que consiste em otimizar uma função linear sujeita à um conjunto de restrições lineares. Há muitos métodos, ou programas, que resolvem esse problema. Um desses métodos é o método Simplex, desenvolvido por George B. Dantzig. O método Simplex tornou-se uma ferramenta fundamental em programação linear durante muitos anos. No entanto, este método possui convergência exponencial no número de variáveis, apesar de nem em todos os problemas tal complexidade ocorrer. Com o desenvolvimento dos computadores há uma busca por métodos que tenham melhor convergência.

Em 1984 Karmakar apresentou um algoritmo para programação linear com complexidade polinomial. A partir da publicação desse algoritmo início o desenvolvimento de outros algoritmos, chamados de algoritmos de pontos interiores. Nesse trabalho, vamos estudar sobre um método de pontos interiores.

O método estudado é baseado no algoritmo de ajustamento pelo par ótimo. O algoritmo de ajustamento pelo par ótimo trabalha com duas variáveis em cada iteração. O método é apresentado por Jair Silva em sua tese de doutorado [1] trabalha com p variáveis a cada iteração.

2. O PROBLEMA DE PROGRAMAÇÃO LINEAR

O problema de programação linear (PPL) consiste em minimizar uma função linear, chamada de função objetivo, sujeita a um conjunto de restrições lineares. A seguir temos um problema de programação linear representado na forma padrão:

$$\begin{array}{ll} \text{minimizar} & c^t x \\ \text{s.a.} & Ax = b \\ & x \geq 0 \end{array} \quad (2.1)$$

onde $A \in \mathbb{R}^{n \times m}$, $c \in \mathbb{R}^n$, $x \in \mathbb{R}^n$ e $b \in \mathbb{R}^m$. O problema escrito nessa forma é chamado de problema de programação linear na forma padrão. Geralmente os problemas não aparecem nessa forma, mas todos podem ser reduzidos à esta forma.

Para a resolução do problema há, hoje, vários métodos. Um dos mais populares é o método Simplex proposto por Dantzig na década de 1940. Durante muitos anos o método Simplex foi a ferramenta fundamental do estudo de programação linear. No entanto, este método possui convergência exponencial no número de variáveis, apesar de nem em todos os problemas tal complexidade ocorrer.

Em 1979 Khachian apresentou o método dos elipsóides que, apesar da convergência polinomial, não se mostrou prático. Em 1984, Karkarmar apresentou um novo algoritmo para programação linear, também com complexidade polinomial. A partir disso, iniciou-se uma nova linha de pesquisa conhecida como métodos de pontos interiores.

Os métodos de pontos interiores podem ser classificados como primal, dual e primal-dual dependendo em que espaço estão sendo realizados as iterações. O problema (2.1) traz a forma primal do problema de programação linear. A forma dual associada ao problema (2.1) é dada por:

$$\begin{array}{ll} \text{maximizar} & b^t y \\ \text{s.a.} & A^t y + z = 0 \\ & z \geq 0 \end{array} \quad (2.2)$$

onde $y \in \mathbb{R}^m$, $z \in \mathbb{R}^n$ e A , b , c como definido no problema primal.

Os problemas (2.1) e (2.2) juntos são chamados de par primal-dual. As condições de otimalidade de primeira ordem (Karush - Kush - Tucker) dos problemas (2.1) e

(2.2) são dadas por:

$$\begin{aligned} Ax - b &= 0 \\ A^t y + z - c &= 0 \\ XZe &= 0 \\ (x, z) &\geq 0 \end{aligned} \tag{2.3}$$

onde $X = \text{diag}(x)$, $Z = \text{diag}(z)$ e $e \in \mathbb{R}^n$ é o vetor com todas as coordenadas iguais a um.

Se (x, y, z) for uma solução de (2.3) então x é uma solução ótima de (2.1) e (y, z) é uma solução ótima de (2.2). Um ponto (x, y, z) é factível se ele satisfaz as restrições do conjunto de restrições de problema primal e dual e o ponto é dito interior se $(x, z) > 0$. O gap de um problema de programação linear é definido como a diferença entre os valores das funções objetivos do problema primal e dual, ou seja $\gamma = c^t x - b^t y$.

Alternativamente, os métodos de pontos interiores podem ser classificados em três categorias: afim-escala, redução de potencial e trajetória central. A classe de métodos que apresentou as melhores propriedades práticas e teóricas são os métodos primais-duais pertencente a categoria trajetória central. O método aqui estudado pertencente a essa classe. A seguir, vamos falar um pouco mais sobre esses métodos.

A maioria dos métodos de pontos interiores utiliza o método de Newton em suas iterações. O método de Newton é uma generalização do método de Newton para encontrar zeros de funções, que consiste em: dado um ponto x^k , devemos encontrar uma direção Δx^k . Ao encontrarmos a direção Δx^k , atualizamos o valor de x , tal que $x^{k+1} = x^k + \Delta x^k$. Os métodos primais-duais de pontos interiores podem ser vistos como aplicações do método de Newton para calcular aproximações da solução de uma sequência de sistemas não lineares (2.3), perturbados por um parâmetro μ , usado para relaxar as restrições de não-negatividades. Assim, temos:

$$F_\mu(x, y, z) = \begin{bmatrix} Ax - b \\ A^t y + z - c = 0 \\ -XZe + \mu e = 0 \end{bmatrix} = 0, \quad (x, z) \geq 0. \tag{2.4}$$

Se $\mu = 0$, então os problemas (2.3) e (2.4) são equivalentes. Assim, a solução do problema (2.4) se aproxima da solução de (2.3) conforme $\mu \rightarrow 0$.

Um método primal-dual obtém uma solução aproximada para o problema gerando uma sequência de pontos (x_k, y_k, z_k) satisfazendo (2.4) para todo $\mu > 0$ é denominado trajetória central. A cada iteração do método é aplicado um passo do método de Newton para resolver o sistema (2.4), com um dado parâmetro μ_k . A direção de Newton, $(\Delta x^k, \Delta y^k, \Delta z^k)$, é obtida da solução de um sistema de equações lineares.

A cada iteração do método é necessário resolver um sistema de equações lineares para determinar a direção de Newton. Quando métodos diretos são utilizados para resolver o sistema linear é feita uma fatoração de uma matriz e em seguida são resolvidos dois sistemas lineares mais simples, os quais estão associados a matrizes triangulares. O processo de fatoração de uma matriz de ordem n apresenta complexidade $O(n^3)$ e a solução de sistemas triangulares pode ser realizada com complexidade $O(n^2)$.

3. O ALGORITMO DE VON NEUMANN

O algoritmo de Von Neumann foi proposto em 1948 por Von Neumann à Dantzig, que o divulgou na década de 1990. O algoritmo possui propriedades interessantes, como simplicidade e convergência inicial rápida, porém ele não é muito prático para resolver problemas lineares. Consideremos o problema de encontrar uma solução

factível para o conjunto de restrições linear:

$$\begin{aligned} Px &= 0 \\ e^t x &= 1 \\ x &> 0 \end{aligned} \quad (3.1)$$

onde $P \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$ e $e \in \mathbb{R}^n$ é o vetor com todas as coordenadas igual a um, e as colunas de P tem norma um, isto é $\|P_j\| = 1$, para $j = 1, \dots, n$. Geometricamente as colunas P_j podem ser vistas como pontos sobre a hipersfera m -dimensional com raio unitário e centro na origem. O problema pode ser descrito então como atribuir ponderações x_j não negativas as colunas P_j de modo que depois de reescalado seu centro de gravidade seja a origem. Problemas de programação linear podem ser colocados na forma (3.1).

O algoritmo de Von Neumann trabalha na resolução do problema (3.1). A figura 1 descreve como método trabalha em cada iteração. Primeiramente ele encontra a coluna P_s de P que forma o maior ângulo com o resíduo b^{k-1} , e então o próximo resíduo é a projeção da origem no segmento de reta ligando b^{k-1} a P_s . A seguir trazemos o algoritmo:

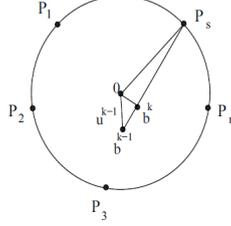


FIGURA 1. Ilustração do algoritmo de Von Neumann

Algoritmo de Von Neumann

Dado: $x^0 \geq 0$, com $e^t x^0 = 1$.

Calcule $b^0 = Px^0$.

Para $k = 1, 2, 3 \dots$ **faça**

[1] Calcule:

$$s^+ = \operatorname{argmin}_{j=1, \dots, n} P_j^t b^{k-1},$$

$$v_{k-1} = P_{s^+}^t b^{k-1}.$$

[2] **Se** $v_{k-1} > 0$, **então PARE**; o problema (3.1) é infactível.

[3] Calcule

$$u^{k-1} = \|b^{k-1}\|,$$

$$\lambda = \frac{1 - v_{k-1}}{(u^{k-1})^2 - 2v_{k-1} + 1}.$$

[4] Atualize

$$b^k = \lambda b^{k-1} + (1 - \lambda) P_{s^+},$$

$$x^k = \lambda x^{k-1} + (1 - \lambda) e_{s^+},$$

onde e_{s^+} é o vetor da base canônica com 1 na s^+ -ésima coordenada

$k = k + 1$

FIM.

Como dito anteriormente, esse algoritmo não é muito prático para resolver problemas lineares, então foram criados outros algoritmos à partir desse algoritmo. Um deles é o algoritmo de ajustamento pelo par ótimo. Este algoritmo identifica os vetores P_{s^+} e P_{s^-} que tem o maior e o menor ângulo com o vetor b^{k-1} , respectivamente. Em seguida ele encontra $x_{s^+}^k$, $x_{s^-}^k$ e λ , onde $x_j^k = \lambda x_{k-1}^j$ para todo $j \neq s^+$

e $j \neq s^-$, que minimiza a distancia de b^k à origem. A seguir trazemos o algoritmo de ajustamento pelo par ótimo:

Algoritmo de ajustamento pelo par ótimo

Dado: $x^0 \geq 0$, com $e^t x^0 = 1$.

Calcule $b^0 = Px^0$.

Para $k = 1, 2, 3 \dots$ **faça**

[1] Calcule:

$$s^+ = \operatorname{argmin}_{j=1, \dots, n} \{P_j^t b^{k-1}\},$$

$$s^- = \operatorname{argmax}_{j=1, \dots, n} \{P_j^t b^{k-1} | x_j > 0\},$$

$$v_{k-1} = P_{s^+}^t b^{k-1}.$$

[2] **Se** $v_{k-1} > 0$, **então PARE**; o problema (3.1) é infactível.

[3] Resolva o problema:

$$\begin{aligned} & \text{minimizar } \|\lambda_0(b^{k-1} - x_{s^+}^{k-1}P_{s^+} - x_{s^-}^{k-1}P_{s^-}) + \lambda_1P_{s^+} + \lambda_2P_{s^-}\|^2 \\ & \text{s.a.} \quad \lambda_0(1 - x_{s^+}^{k-1} - x_{s^-}^{k-1}) + \lambda_1 + \lambda_2 = 1, \\ & \quad \lambda_0 \geq 0, \\ & \quad \lambda_1 \geq 0, \\ & \quad \lambda_2 \geq 0. \end{aligned} \tag{3.2}$$

[4] Atualize

$$b^k = \lambda_0(b^{k-1} - x_{s^+}^{k-1}P_{s^+} - x_{s^-}^{k-1}P_{s^-}) + \lambda_1P_{s^+} + \lambda_2P_{s^-},$$

$$u^k = \|b^k\|,$$

$$x_j^k = \begin{cases} \lambda_0 x_j^{k-1}, & j \neq s^+ \text{ e } j \neq s^-, \\ \lambda_1 & j = s^+, \\ \lambda_2 & j = s^-. \end{cases}$$

$$k = k + 1.$$

FIM.

O maior esforço do algoritmo é a resolução do subproblema (3.2), que deve ser resolvido a cada iteração.

4. O MÉTODO ESTUDADO

O algoritmo de ajustamento pelo par ótimo trabalha com duas variáveis em cada iteração. O método proposto por Jair Silva em [1] trabalha com p variáveis. A idéia central utilizada no algoritmo para duas variáveis é resolver o subproblema (3.2). Esse subproblema pode ser generalizado se usarmos qualquer quantidade de colunas e assim dar importância a quantas variáveis desejarmos. A escolha de p é livre, aqui não discutiremos sobre essa escolha, e apenas sobre a generalização do algoritmo para p variáveis.

O algoritmo de ajustamento ótimo para p coordenadas, começa indentificando as s_1 colunas que fazem o maior ângulo com o vetor b^{k-1} , em seguida ele encontra as s_2 colunas que fazem o menor ângulo com o vetor b^{k-1} , onde $s_1 + s_2 = p$. Depois é resolvido o subproblema de otimização, atualizando o residuo e o ponto corrente.

Algoritmo de ajustamento ótimo para p coordenadas

Dado: $x^0 \geq 0$, com $e^t x^0 = 1$.

Calcule $b^0 = Px^0$.

Para $k = 1, 2, 3 \dots$ **faça**

[1] Calcule:

$$\{P_{\eta_1^+}, P_{\eta_2^+}, \dots, P_{\eta_{s_1}^+}\} \text{ que fazem o maior ângulo com o vetor } b^{k-1}.$$

$\{P_{\eta_1^-}, P_{\eta_2^-}, \dots, P_{\eta_{s_1}^-}\}$ que fazem o menor ângulo com o vetor b^{k-1} e tal que $x_i^{k-1} > 0$, $i = \eta_1, \eta_2, \dots, \eta_{s_1}^-$, onde $s_1 + s_2 = p$.

$$v_{k-1} = \max_{i=1, \dots, s_1} P_{\eta_i^+}^t b^{k-1}.$$

- [2] Se $v_{k-1} > 0$, **então PARE**; o problema (3.1) é infactível.
 [3] Resolva o problema:

$$\begin{aligned} & \text{minimizar } \|\lambda_0 \left(b^{k-1} - \sum_{i=1}^{s_1} x_{\eta_i^+}^{k-1} P_{\eta_i^+} - \sum_{j=1}^{s_2} x_{\eta_j^-}^{k-1} P_{\eta_j^-} \right) + \sum_{i=1}^{s_1} \lambda_{\eta_i^+} P_{\eta_i^+} + \sum_{j=1}^{s_2} \lambda_{\eta_j^-} P_{\eta_j^-}\|^2 \\ & \text{s.a. } \lambda_0 \left(1 - \sum_{i=1}^{s_1} x_{\eta_i^+}^{k-1} - \sum_{j=1}^{s_2} x_{\eta_j^-}^{k-1} \right) + \sum_{i=1}^{s_1} \lambda_{\eta_i^+} + \sum_{j=1}^{s_2} \lambda_{\eta_j^-} = 1, \\ & \lambda_{\eta_i^+} \geq 0, \text{ para } i = 1, \dots, s_1, \\ & \lambda_{\eta_j^-} \geq 0, \text{ para } j = 1, \dots, s_2. \end{aligned} \tag{4.1}$$

- [4] Atualize

$$\begin{aligned} b^k &= \lambda_0 \left(b^{k-1} - \sum_{i=1}^{s_1} x_{\eta_i^+}^{k-1} P_{\eta_i^+} - \sum_{j=1}^{s_2} x_{\eta_j^-}^{k-1} P_{\eta_j^-} \right) + \sum_{i=1}^{s_1} \lambda_{\eta_i^+} P_{\eta_i^+} + \sum_{j=1}^{s_2} \lambda_{\eta_j^-} P_{\eta_j^-}, \\ u^k &= \|b^k\|, \\ x_j^k &= \begin{cases} \lambda_0 x_j^{k-1}, & j \notin \{\eta_1^+, \dots, \eta_{s_1}^+, \eta_1^-, \dots, \eta_{s_2}^-\}, \\ \lambda_1 & j = \eta_i^+, i = 1, \dots, s_1, \\ \lambda_2 & j = \eta_i^-, i = 1, \dots, s_2. \end{cases} \\ k &= k + 1. \end{aligned}$$

FIM.

Em cada iteração do algoritmo de ajustamento ótimo para p coordenadas, é necessário resolver o subproblema (4.1). Este subproblema é resolvido encontrando uma solução no octante positivo, sujeito a factibilidade para um sistema linear de ordem no máximo $(p+1) \times (p+1)$. Uma forma de resolver este subproblema é verificar todos os possíveis casos de soluções factíveis. Contudo, ao resolver o subproblema desta forma o número de casos possíveis de solução aumenta exponencialmente com o valor de p , mais precisamente $2^{p+1} + 1$.

Com a finalidade de contornar o problema do crescimento exponencial do número de casos com o valor de p , podemos abordar o subproblema (4.1) de outra forma e resolvê-lo aplicando métodos de pontos interiores. A grande vantagem de usar métodos de pontos interiores é que em cada iteração do método de pontos interiores surge sistemas lineares, cujo o custo é a fatoração de uma matriz de ordem $(p+1) \times (p+1)$.

O custo computacional de fatorar uma matriz de ordem 10×10 ou 100×100 , por exemplo, não muito diferente se considerarmos que o problema de programação linear a ser resolvido é de grande porte. Já verificar $2^{10} - 1$ ou $2^{100} - 1$ casos possíveis representam custos bem diferentes.

Em [1] é mostrada a convergência do método. O método apresenta um desempenho superior em relação ao algoritmo de Von Neumann. Também é mostrado que se $p_2 \geq p_1$ o algoritmo de ajustamento ótimo para p_2 coordenadas possui desempenho superior em relação ao algoritmo de ajustamento ótimo para p_1 coordenadas.

Também, em [1] é mostrado resultados da implementação do algoritmo de ajustamento ótimo para p coordenadas e desempenho foi muito bom. Durante esse semestre, além do estudo da teoria do método, houve um reconhecimento do algoritmo implementado na linguagem de programação C. Para o próximo semestre o objetivo é melhorar a implementação.

5. CONCLUSÃO

O estudo realizado foi sobre o algoritmo de ajustamento ótimo para p coordenadas para programação linear. Para isso foi feito um estudo sobre programação linear, métodos de pontos interiores e do Algoritmo de Von Neumann, o qual o algoritmo se baseia.

Em cada iteração do algoritmo de ajustamento ótimo para p coordenadas é resolvido um subproblema verificando as condições de KKT. A proposta é resolver as equações de KKT do subproblema por métodos de pontos interiores, o que é vantajoso.

A grande vantagem de usar métodos de pontos interiores é que em cada iteração do método de pontos interiores surge sistemas lineares. Quando métodos diretos são utilizados para resolver o sistema linear é feita uma fatoração de uma matriz e em seguida são resolvidos dois sistemas lineares mais simples, os quais estão associados a matrizes triangulares. O processo de fatoração de uma matriz de ordem n apresenta complexidade $O(n^3)$ e a solução de sistemas triangulares pode ser realizada com complexidade $O(n^2)$. Considerando que temos um problema de grande porte, e que os sistemas lineares a serem resolvidos possuem ordem $p + 1$ o custo é proporcionalmente pequeno.

A implementação do algoritmo apresentou um desempenho muito bom, para o próximo semestre o objetivo é melhorar essa implementação.

REFERÊNCIAS

- [1] SILVA, JAIR *Uma família de algoritmos de programação linear baseada no algoritmo de Von Neumann*, Matemática Aplicada, Unicamp, Campinas, SP, 2009.