OTIMIZAÇÃO NÃO LINEAR DE FLUXO EM REDE – MULTIPLICADORES DE HESTENES

LUIS ERNESTO TORRES GUARDIA
Mestrado de Engenharia de Produção
Universidade Federal Fluminense
Rua Passo da Pátria 156, São Domingos
24210-240, Niterói, R.J., Brasil
tepletg@vm.uff.br

Resumo

Neste trabalho, a solução de um problema de otimização convexa não linear e separável, sujeito a restrições lineares, é determinada. A técnica usada para obter a solução do anterior problema é o método de pontos interiores primal – dual. Como em todos os métodos de pontos interiores, aparece um sistema linear, cuja solução geralmente requer de um maior esforço computacional. Neste trabalho, dita solução é obtida usando o esquema de multiplicadores de Hestenes. Como ilustração, o método primal – dual é aplicado em modelos de fluxo em rede sem precisar armazenar as matrizes envolvidas no sistema linear, sendo assim apropriado para problemas de grande porte. Alguns experimentos numéricos são apresentados para mostrar a execução do método primal – dual e confirmam a eficiência do método para todos os casos testados.
Palavras-chave: Programação Não Linear. Método de Pontos Interiores Primal - Dual. Fluxo em Redes.

Abstract

In this work, the solution of the nonlinear optimization problem, with convex cost and subject to linear constraints, is determined. The technique to solve the above problem is the Primal-Dual Interior Point Method. The linear system that appears in the primal-dual method is solved using the Hestenes multipliers´ scheme. As an illustrated, the primal – dual method is applied to the convex minimum cost flow problem without storing the constraint matrix and in this case, the method is appropriated to solve large scale problems. The efficiency of this method is demonstrated through computational experiments on some network problems.
Keywords: Nonlinear Programming. Primal-Dual Interior Point Method. Network Flow Problem.

1. Introdução

Depois do anuncio do método de pontos interiores com complexidade polinomial, apresentado por Karmarkar [1984], para resolver problemas de programação linear de grande porte, esse método mostrou-se ser de grande sucesso e eficiente técnica de solução desses problemas. O método primal – dual foi que atraiu maior atenção entre os diferentes métodos de pontos interiores. Para maiores detalhes desse método primal – dual, pode-se consultar o livro de Wright, [1997]. Experiências computacionais e desenvolvimentos teóricos mostrarão que ele se desempenha muito melhor que os outros métodos de pontos interiores. Por isso, é natural que as maiores atenções agora sejam na aplicação desse método primal- dual para uma área geralmente mais complexa como da programação não linear com restrições não lineares, e cuja solução originalmente era determinada usando o método seqüencial de programação quadrática. Várias experiências numéricas foram realizadas para resolver esse problema de otimização usando o método seqüencial, mas a solução do problema quadrático poderia ser cara computacionalmente, em especial para problemas de grande porte. Assim, o sucesso computacional do método primal - dual em programação linear deu lugar para um substancial interesse na aplicação para a programação não linear, demonstrando que o método de pontos interiores primal - dual pode resolver esses problemas eficientemente, como atualmente está sendo provado numericamente.

Neste trabalho, apresentamos o problema de programação não linear com restrições lineares de fluxo em rede, e cuja solução, como já foi mencionado, é determinada usando o método de pontos interiores primal - dual de tal forma a explorar a estrutura da rede devido à matriz envolvida nas restrições ser esparsa. Tais problemas de otimização não linear com restrições de fluxo em rede aparecem em uma ampla variedade de aplicações, por exemplos em planejamento de rede de transporte, ver Nagurney [1984], em redes de telecomunicações , ver o trabalho de Ouorou [2000], entre outras aplicações.

O trabalho é organizado como segue. Na seção 2, o método de pontos interiores primal-dual é apresentado para resolver um problema de programação não linear com restrições lineares. Neste caso, a função não linear envolvida é uma função separável, e a solução do correspondente sistema linear, que aparece em todos os métodos de pontos interiores, é obtida usando o esquema de multiplicadores de Hestenes, ver Bonettine, *et al.*[2005], isto é, devido a que o sistema está representado por uma matriz que é simétrica indefinida. Na seção 3, apresentamos a aplicação do método primal-dual para diferentes funções não lineares e separáveis sendo as restrições lineares de conservação de fluxo em uma rede. Finalmente, algumas conclusões são dadas na seção 4.

2. Método de Pontos Interiores Primal-Dual

Consideremos o seguinte problema de otimização não linear restringido linearmente:

$$\text{Minimizar} \quad f(\mathbf{x}) \tag{1.1}$$
$$\text{sujeito a:} \quad h(\mathbf{x}) = \mathbf{0} \tag{1.2}$$
$$\mathbf{x} \geq \mathbf{0}. \tag{1.3}$$

Assume-se que a função f: $\mathbf{R}^n \rightarrow \mathbf{R}$ é convexa e com segunda derivada continua. A função h($\mathbf{x}$) é dada pelas seguintes restrições lineares:

$$h(\mathbf{x}) = \mathbf{Ax} - \mathbf{b}$$

onde $\mathbf{A}$ é uma matriz, $\mathbf{A} \in \mathbf{R}^{mxn}$, m < n, e de posto completo e $\mathbf{b} \in \mathbf{R}^m$. O vetor $\mathbf{x} \in \mathbf{R}^n$ é a variável de decisão.

Seja a função de Lagrange associada com o problema acima (1) definida por:

$$L(\mathbf{w}) = f(\mathbf{x}) + \mathbf{y}^t (\mathbf{Ax} - \mathbf{b}) - \mathbf{z}^t \mathbf{x}$$

onde $\mathbf{w} = (\mathbf{x},\mathbf{y},\mathbf{z})^t$, $\mathbf{y} \in \mathbf{R}^m$ e $\mathbf{z} \in \mathbf{R}^n$, $\mathbf{z} \geq \mathbf{0}$ são os vetores multiplicadores de Lagrange associados às restrições de igualdades e desigualdades respectivamente.

As condições de otimalidade de Karush – Kuhn – Tucker (KKT) do problema (1) são dadas pelo seguinte sistema:

$$F(\mathbf{w}) = \begin{pmatrix} \nabla_x L(x,y,z) \\ Ax - b \\ XZe \end{pmatrix} = \mathbf{0} \tag{2}$$

$$\text{e} \qquad \mathbf{x} \geq \mathbf{0} \ , \ \mathbf{z} \geq \mathbf{0},$$

sendo:

$$\nabla_x \mathbf{L}(\mathbf{x},\mathbf{y},\mathbf{z}) = \nabla f(\mathbf{x}) + \mathbf{A}^t \mathbf{y} - \mathbf{z},$$
$$\mathbf{X} = \text{diag}(x_1,x_2,...,x_n),$$
$$\mathbf{Z} = \text{diag}(z_1,z_2,...,z_n),$$

$$\mathbf{e} = (1,1,...,1)^t \in \mathbf{R}^n,$$

e ∇_x $\mathbf{L}(\mathbf{x},\mathbf{y},\mathbf{z})$ é a gradiente de $\mathbf{L}(\mathbf{x},\mathbf{y},\mathbf{z})$ em relação a variável $\mathbf{x}$ e $\nabla f(\mathbf{x})$ é a gradiente de f e $\mathbf{X}$ e $\mathbf{Z}$ são matrizes diagonais com diagonal (x_i) e (z_i) respectivamente, componentes dos vetores $\mathbf{x}$ e $\mathbf{z}$.

Para resolver o problema (1) usando o método de pontos interiores primal – dual, aplica-se o método de Newton às condições perturbadas de KKT (de barreira) as quais são dadas por:

$$F_\mu(\mathbf{w}) = \begin{pmatrix} \nabla_x L(x,y,z) \\ Ax - b \\ XZe - \mu e \end{pmatrix} = \mathbf{0} \tag{3.1}$$

$$\mathbf{x} > \mathbf{0}, \quad \mathbf{z} > \mathbf{0}, \tag{3.2}$$

sendo $\mu > 0$ o parâmetro de barreira.

As condições (3) são denominadas as condições de KKT de barreira e um ponto $\mathbf{w}(\mu) = (\mathbf{x}(\mu),\mathbf{y}(\mu),\mathbf{z}(\mu))$ que satisfaz essas condições é denominado de ponto KKT de barreira. O método de pontos interiores serve para alcançar um ponto que aproximadamente satisfaça as condições acima (3) e finalmente obter um ponto que satisfaça as condições de KKT fazendo $\mu \to 0$. Isto significa que forçamos $\mathbf{x}$ e $\mathbf{z}$ sejam estritamente positivos durante as iterações.

Para determinar um ponto KKT de barreira aproximado para um μ dado, $\mu > 0$, usamos o método de Newton. Seja a direção de Newton $\Delta\mathbf{w} = (\Delta\mathbf{x},\Delta\mathbf{y},\Delta\mathbf{z})^t$ definida pela solução do sistema de equações:

$$F'_\mu(\mathbf{w})\,\Delta\mathbf{w} = -F_\mu(\mathbf{w}) \tag{4}$$

onde

$$F'_\mu(\mathbf{w}) = \begin{pmatrix} \nabla_x^2 L(w) & \nabla h(x)^t & -I \\ \nabla h(x) & 0 & 0 \\ Z & 0 & X \end{pmatrix} = \begin{pmatrix} H & A^t & -I \\ A & 0 & 0 \\ Z & 0 & X \end{pmatrix} \tag{5}$$

sendo ∇^2_x L($\mathbf{w}$) a matriz Hessiana da função L($\mathbf{w}$) e, neste caso particular, por ser as restrições lineares temos que ∇^2_x L($\mathbf{w}$) = ∇^2 f($\mathbf{x}$) sendo ∇^2 f($\mathbf{x}$) a matriz Hessiana de f(x). Podemos colocar neste caso ∇^2 f($\mathbf{x}$) = $\mathbf{H}(\mathbf{x})$, ou simplesmente $\mathbf{H}$, uma matriz diagonal se f é uma função separável.

Assim, a relação (4) pode ser escrita como:

$$\begin{pmatrix} H & A^t & -I \\ A & 0 & 0 \\ Z & 0 & X \end{pmatrix} . \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = - \begin{pmatrix} \nabla_X L(x,y,z) \\ Ax - b \\ XZe - \mu e \end{pmatrix} \tag{6}$$

Sejam as seguintes expressões:

$$\xi_c = \nabla_x L(\mathbf{w}) = \nabla f(\mathbf{x}) + \mathbf{A}^t\mathbf{y} - \mathbf{z},$$
$$\xi_b = \mathbf{Ax} - \mathbf{b},$$
$$\xi_\mu = \mathbf{XZe} - \mu\mathbf{e},$$

e são denominadas de folgas de viabilidades.

Existem varias formas de determinar a solução do sistema dada na relação anterior (6), digamos que podemos expressar a solução da seguinte forma:

$$\{ \mathbf{A}[\, \mathbf{H} + \mathbf{X}^{-1}\mathbf{Z}\,]^{-1}\, \mathbf{A}^{t} \}\boldsymbol{\Delta}\mathbf{y} = - \mathbf{A}[\, \mathbf{H} + \mathbf{X}^{-1}\mathbf{Z}\,]^{-1} (\, \xi_{c} + \mathbf{X}^{-1}\xi_{\mu}) + \xi_{b}. \qquad (7.1)$$

$$\Delta\mathbf{x} = - [\, \mathbf{H} + \mathbf{X}^{-1}\, \mathbf{Z}\,]^{-1} (\, \mathbf{A}^{t}\, \Delta\mathbf{y} + \xi_{c} + \mathbf{X}^{-1}\xi_{\mu}). \qquad (7.2)$$

$$\Delta\mathbf{z} = - \mathbf{X}^{-1} (\, \xi_{\mu} + \mathbf{Z}\, \Delta\mathbf{x}\,). \qquad (7.3)$$

e neste caso, primeiro determinamos a solução de (7.1), $\Delta\mathbf{y}$ **,** usando qualquer processo iterativo, digamos o método do gradiente conjugado, ou um método que determina a matriz inversa de $\{\mathbf{A}[\, \mathbf{H} + \mathbf{X}^{-1}\mathbf{Z}\,]^{-1}\mathbf{A}^{t}\}$, para depois calcular $\Delta\mathbf{x}$ e $\Delta\mathbf{z}$ respectivamente. Este último procedimento foi apresentado no trabalho de Torres [2005].

Uma outra forma solução do sistema dado em (6) é a seguinte. Usando a última linha do sistema (6), podemos eliminar a variável $\Delta\mathbf{z}$, como em (7.3), a qual é dada por:

$$\Delta\mathbf{z} = - \mathbf{X}^{-1} (\, \xi_{\mu} + \mathbf{Z}\, \Delta\mathbf{x}\,), \qquad (8)$$

e temos então o seguinte sistema, denominado de aumentado, simétrico e indefinido das variáveis ($\Delta\mathbf{x}$, $\Delta\mathbf{y}$) dado por:

$$\begin{bmatrix} H + X^{-1}Z & A^{t} \\ A & 0 \end{bmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = - \begin{pmatrix} \xi_{c} + X^{-1}\xi_{\mu} \\ \xi_{b} \end{pmatrix} \qquad (9)$$

Seja $\Delta\mathbf{w}_{k} = (\Delta\mathbf{x}_{k},\Delta\mathbf{y}_{k},\Delta\mathbf{z}_{k})^{t}$ a solução do sistema (6), obtida já seja usando (7.1)-(7.3), ou como neste trabalho, usando(8) – (9), então uma nova iteração é realizada da seguinte forma:

$$\mathbf{w}_{k+1} = \mathbf{w}_{k} + \alpha_{k}\, \Delta\mathbf{w}_{k}, \qquad (10)$$

onde α_k é o tamanho de passo determinado por um procedimento de busca de linha.

No ponto $\mathbf{w}_{k} = (\, \mathbf{x}_{k}, \mathbf{y}_{k}, \mathbf{z}_{k}\,)^{t}$ calculamos o máximo passo permitido à fronteira da região viável dado por:

$$\alpha_{kmax} = \min\{\ \min_{i}\{-(x_{k})_{i} / (\Delta x_{k})_{i}\ \ /\ \ (\Delta x_{k})_{i}<0\},\ \min_{i}\{ -(z_{k})_{i}/(\Delta z_{k})_{i}\ \ / (\Delta z_{k})_{i}<0\}\ \}.$$

Para ter robustez do algoritmo primal – dual, um passo na iteração seguinte é dado por:

$$\alpha_{k} = \alpha'_{k}\, \beta^{j} \quad , \qquad \alpha'_{k} = \min\, \{\, \gamma\alpha_{kmax},\, 1\, \},$$

sendo $\gamma \in (0,1)$ e $\beta \in (0,1)$ escalares fixos, e $j = j_k$ e o menor inteiro não negativo tal que:

$$\phi(\mathbf{w}_{k} + \alpha_{k}\Delta\mathbf{w}_{k}, \mu_{k}) - \phi(\mathbf{w}_{k}, \mu_{k}) \leq \tau\, \alpha_{k}\nabla\phi(\mathbf{w}_{k}, \mu_{k})^{t}\, \Delta\mathbf{w}_{k},$$

onde $\tau \in (0,1)$, e sendo $\phi(\mathbf{w},\mu)$ a função de mérito definida por:

$$\phi(\mathbf{w},\mu) = \tfrac{1}{2}\ \|\, F_{\mu}(\mathbf{w})\, \|^{2}.$$

Apresentamos a seguir, um resumo do método de pontos interiores primal-dual, ver El-Bakry, *et al.* [1996].

Consideremos um ponto interior inicial dado $\mathbf{w}_{0} = (\mathbf{x}_{0},\mathbf{y}_{0},\mathbf{z}_{0})$ tal que $(\mathbf{x}_{0},\mathbf{z}_{0}) > 0$, $\tau \in (0,1)$, $\beta \in (0,1)$, e $\gamma \in (0,1)$, n o número de variáveis.

Para k=0,1,2,... até a convergência, realizar os seguintes passos:

Passo 1. Fazer $\mu_k = \sigma_k(\mathbf{x}_k{}^t\mathbf{z}_k)/n$, $\sigma_k \in (0,1)$.
Passo 2. Resolver o seguinte sistema linear para determinar $\Delta\mathbf{w}_k$ usando (8) e (9):

$$F'\mu_k(\mathbf{w}_k)\,\Delta\mathbf{w}_k = -\,F\mu_k(\mathbf{w}_k).$$

Passo 3. Calcular α_k usando a função de mérito $\phi(\mathbf{w},\mu)$.
Passo 4. Fazer $\mathbf{w}_{k+1} = \mathbf{w_k} + \alpha_k\,\Delta\mathbf{w}_k$.

Geralmente o maior esforço computacional requerido no método de pontos interiores primal – dual é a resolução do sistema no **passo 2**, que se reduz ao sistema linear de equações simétrico e indefinido, também denominado de sistema aumentado, dado em (9), que por simplicidade podemos escrever na forma:

$$\begin{bmatrix} D & A^t \\ A & 0 \end{bmatrix}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix} = \begin{pmatrix}\psi_1 \\ \psi_2\end{pmatrix} \tag{11}$$

onde $\mathbf{\Psi_1} = -(\xi_\mathbf{c} + \mathbf{X}^{-1}\xi_\mu)$, $\mathbf{\Psi_2} = -\,\xi_\mathbf{b}$ e $\mathbf{D} = \nabla^2 f(\mathbf{x}) + \mathbf{X}^{-1}\mathbf{Z}$.

Esse sistema aumentado aparece, como é citado em Golub, *et al.* [2001], em diferentes aplicações tais como na otimização restringida, no método de elementos finitos para resolver a equação de Navier-Stokes, no problema de mínimo quadrado, etc.

Existem vários métodos para resolver o sistema simétrico (11), entre eles podemos mencionar os trabalhos de Golub, *et al*. [2001], Bai e Li [2003], de Durazzi e Ruggiero [2003], de Bergamaschi *et al*. [2004] entre outros mais. Neste trabalho apresentamos um método seguindo o trabalho de Bonettini *et al*. [2005] que usam um esquema de multiplicadores de Hestenes. Assim, a solução do sistema simétrico e indefinido (11) é determinada da seguinte forma:

$$(\mathbf{D} + \chi\,\mathbf{A^t}\,\mathbf{A}\,)\,\Delta\mathbf{x_k} = -\mathbf{A}^t\,\Delta\mathbf{y_k} + \Psi_1 + \chi\,\mathbf{A}^t\,\Psi_2. \tag{12}$$

$$\Delta\mathbf{y_{k+1}} = \Delta\mathbf{y_{\,k}} + \chi(\mathbf{A}\,\Delta\mathbf{x_{\,k}} - \Psi_2\,), \tag{13}$$

onde χ é um parâmetro de penalidade e positivo.

Bonettini *et al.* [2005] usam um teorema para afirmar que existe um parâmetro positivo χ^* tal que para todo $\chi > \chi^*$, a matriz $(\mathbf{D} + \chi\mathbf{A^t}\,\mathbf{A})$ é definida positiva.

Com o anterior resultado, podemos primeiro determinar a solução do sistema (12), $\Delta\mathbf{x_k}$, e depois determinar $\Delta\mathbf{y_{k+1}}$ usando a relação (13). Existem vários métodos para determinar $\Delta\mathbf{x_k}$. Por exemplo, Bonettini *et al.* [2005] aplicam o método de decomposição de Cholesky. Neste trabalho, por sua eficiência computacional, usamos o método do Benzi *et al*. [2000], tal que :

$$(\mathbf{D} + \chi\,\mathbf{A^t}\,\mathbf{A}\,)^{-1} \cong \mathbf{Z}\,\mathbf{P}^{-1}\,\mathbf{Z}^t, \tag{14}$$

sendo **Z** uma matriz triangular superior com elementos na diagonal igual a 1's, e **P** uma matriz diagonal.

A seguir apresentamos o método de decomposição de Benzi *et al*. [2000], para obter a matriz inversa de $(\mathbf{D} + \chi\,\mathbf{A^t}\,\mathbf{A}\,)$ de dimensão n, onde e_i é o i-ésimo vetor unitário base:

(1) Seja $z_i^{(0)} = e_i$ $(1 \le i \le n)$
(2) Para i = 1,2,...,n fazer
(3) $\quad v_i = (\mathbf{D} + \chi\mathbf{A}^t\,\mathbf{A}\,)\,z_i^{(i-1)}$
(4) $\quad$ Para j = i,i+1,…, n fazer
(5) $\quad\quad p_j^{(i-1)} = v_i^T\,z_j^{(i-1)}$

(6) final
(7) se i = n ir a (12)
(8) Para j = i+1,..., n fazer
(9) $z_j^{(i)} = z_j^{(i-1)} - (p_j^{(i-1)} / p_i^{(i-1)})\, z_i^{(i-1)}$
(10) final
(11) final
(12) Seja $z_i = z_i^{(i-1)}$ e $p_i = p_i^{(i-1)}$, para $1 \le i \le n$. Retornar
$\mathbf{Z} = [\, z_1, z_2, \ldots, z_n \,]$ e $\mathbf{P} = \mathrm{diag}\,(p_1, p_2, \ldots, p_n)$.

Neste trabalho, calcula-se de uma vez ($\Delta\mathbf{x}$, $\Delta\mathbf{y}$) para depois calcular $\Delta\mathbf{z}$. Alguns autores preferem usar este último esquema devido à robustez do método.

3. Experiência Numérica

Consideremos o problema não linear de fluxo em redes para um único produto com custos separáveis e não lineares associados nos arcos. Este problema é definido por um grafo direcionado G=(N,*A*) onde N é o conjunto de nós com $\| N \|$=n e A é o conjunto de arcos com $\| A \|$=m. Para cada nó i $\in$ N, um escalar s_i é associado, onde s_i é a oferta/procura nó i, dependendo se seu valor é maior/menor que zero. Para cada arco (i,j) $\in A$, uma função convexa, com segunda derivada continua f_{ij}; R$\to$R é dada. Seja x_{ij} o número de unidades de fluxo através do arco (i,j) $\in A$, do nó i ao nó j.

O problema de fluxo em rede, com função não linear e separável nos arcos e no esquema de fluxo-arco, é dado por:

$$\min\ f(\mathbf{x}) = \sum_{(i,j)\in A} f_{ij}(x_{ij}) \tag{15.1}$$

$$\text{sujeito a:} \quad \sum_{\{j:(i,j)\in A\}} x_{ij} - \sum_{\{j:(j,i)\in A\}} x_{ji} = s_i, \quad i \in N, \tag{15.2}$$

$$x_{ij} \ge 0, \ (i,j) \in A. \tag{15.3}$$

A função objetiva (15.1) é uma função de custo não linear e separável. A equação (15.2) é denominada de restrições de conservação de fluxo na rede. É suposto que a matriz que formam essas restrições é de posto completo, e para isso, elimina-se uma dessas restrições. A restrição (15.3) é a condição de os fluxos sejam não negativos.

Na literatura existente para resolver o problema não linear de fluxo de rede (15) podemos mencionar o trabalho de Marín [1995], de Bertsekas *et al.* [1997], de Ourou [2002], de Beraldi *et al.* [2001], de Garcia *et al.* [2003], entre outros trabalhos.

O método de pontos interiores primal – dual foi inteiramente codificado em linguagem FORTRAN com dupla precisão. Todos os experimentos foram realizados em um microcomputador PC Duron com 512 Mb de RAM e 1600 MHZ de freqüência rodando a plataforma Windows XP.

As restrições (15.2) que formam a matriz **A** de (1.2) são armazenadas em dois vetores. A componente de primeiro vetor representa a origem do arco e a componente correspondente do segundo vetor representa o destino do arco.

A seqüência do parâmetro de barreira $\{\mu_k\}$, no passo 1 do método primal – dual, deve convergir a zero tão rápido quanto seja possível. Várias regras foram estudadas para decrescer o parâmetro μ. Neste trabalho, usamos a estratégia adotada por Luksan , *et al.* [2005] que segundo os autores desenvolve-se relativamente bem na prática. Essa regra é dada a seguir:

$$\mu = \sigma\, \mathbf{x}^t \mathbf{z} / n$$

onde $\sigma = 0{,}1 \min\,(0{,}05(1-\rho)/\rho,\ 2)^3$

e $$\rho = \frac{\min(x_i z_i)}{x^t z / n}$$

Ao aplicar o método primal – dual para determinar a matriz inversa da matriz **R** onde **R** = (**D** + χ $\mathbf{A}^t$ **A**), o método de Benzi *et al.* [2000] usa diretamente a matriz **R** sem precisar armazenar dita matriz explicitamente nem a matriz **A**, portanto o método em estudo pode ser aplicado para problemas de grande porte.

A característica do método de pontos interiores é que o ponto inicial dado é qualquer ponto não necessariamente viável e essa característica é usada nesta experiência computacional.

Para ilustrar o método de pontos interiores primal – dual para resolver o problema não linear (15) de fluxo em rede, consideramos a rede da figura 1, p. 197 do artigo de García *et al.* [2003] que consiste de 22 nós e 120 arcos, e a rede da figura 1, p. 476 e da figura 3, p. 479 do trabalho de Nagurney [1984], que consiste de uma rede de 20 nós e 28 arcos e de uma rede de 40 nós e 66 arcos respectivamente.

As funções de custo não lineares tradicionalmente estudadas na literatura de equilíbrio de tráfego ou de telecomunicações são:

i) $f_{ij}(x_{ij}) = x_{ij} \ln (x_{ij})$ para todo $(i,j) \in A$.

ii) $f_{ij}(x_{ij}) = x_{ij} / (c_{ij} - x_{ij})$, $x_{ij} < c_{ij}$, $c_{ij} > 0$, para todo $(i,j) \in A$.

Figura 1. Redes Menor Porte.

Função Objetivo	Valor função objetivo	Número iterações	Rede
$f_{ij}(x_{ij})=x_{ij}\ln(x_{ij})$	441.111507	7	Nagur28
	3766.938968	8	Nagur66
	460.517019	8	Garcia et al.
$f_{ij}(x_{ij})=x_{ij}/(c_{ij}-x_{ij})$	1.995356	11	Nagur28
	14.820400	11	Nagur66

Nestes experimentos, os parâmetros adotados são: σ = 0,05; β = 0,5; γ = 0,99995; e τ = 0,0001.

O tempo (em segundos) para determinar a solução do problema não linear em todos os casos é desprezível.

A rede de 20 nós e 28 arcos e usando as funções (i) e (ii) também foram resolvidas aplicando o pacote computacional **LINGO**. Os resultados obtidos por esse último pacote são iguais com os resultados obtidos pelo método de pontos interiores primal-dual. Para redes maiores, não foi possível o **LINGO** proporcionar uma solução.

Com a experiência computacional favorável nos casos anteriores, analisamos algumas redes de grande porte, baseada na rede básica da figura 1 do trabalho de Nagurney, e assim temos novas redes estendidas. Para isso, implementamos um programa de tal modo que gere redes de qualquer dimensão. Assim, temos redes de 510 nós e 960 arcos e 1020 nós e 1970 arcos, e usamos as mesmas funções de custos acima mencionadas.

Os resultados computacionais, incluindo informações adicionais para estas redes estendidas, são as seguintes:

Figura 2. Redes Grande Porte.

$f(x_{ij}) = x_{ij}\ln(x_{ij})$

Número variáveis	Número de nós	Valor função objetivo	μ	Tempo (segundos)	Número iterações
960	510	4056462,8236430	0,00000000001	139,00	14
1970	1020	4323152,6244214	0,0000000254	959,28	10

$f(x_{ij}) = x_{ij}/(c_{ij} - x_{ij})$

Número variáveis	Número de nós	Valor função objetivo	μ	Tempo (segundos)	Número iterações
960	510	342,9080185	0,00000000001	303,30	29
1970	1020	287,1903669	0,0000000171	3312,03	23

Vale a pena mencionar que a mesma eficiência computacional é alcançada quando não é usada a função de mérito $\phi(\mathbf{w},\mu)$, isto é, neste caso usamos $\alpha_k = \alpha'_k$,

4.Conclusões

Este artigo apresenta o problema não linear de fluxo em rede que pela sua estrutura é um problema de grande porte. O método de pontos interiores primal – dual é apresentado para resolver o anterior problema, de tal forma a explorar a estrutura da rede sem precisar armazenar as matrizes **A** nem ($\mathbf{ADA}^t$), onde $\mathbf{D=(H+X^{-1}Z)^{-1}}$, $\mathbf{H} = \nabla^2 f(\mathbf{x})$, todas elas envolvidas no problema não linear. Isto é confirmado pela aplicação do esquema de multiplicadores de Hestenes ao sistema linear simétrico indefinido que igualmente não armazena a matriz $\mathbf{R} = (\mathbf{D}+\chi\, \mathbf{A}^t \mathbf{A})$. Os resultados numéricos realizados em algumas redes de fluxo de diferentes dimensões confirmam a eficiência do método primal – dual.

5. Referências Bibliográficas

[1] Bai, Z. e G. Li, (2003), "Restrictively Preconditioned Conjugate Gradient Methods for Systems of Linear Equations", IMA Journal of Numerical Analysis, v. 23, p. 561 – 580.

[2] Benzi, M., J. Cullum e M. Tuma, (2000), " Robust Approximate Inverse Preconditioning fo r the Conjugate Gradient Method", SIAM Journal on Scientific Computing, v. 22, no. 4, p. 1318 - 1332.

[3] Beraldi, P., F. Guerriero e R. Musmanno, (2001), "Parallel Algorithms for Solving the Convex Minimum Cost Flow Problem", Computational Optimization and Applications, v. 18. no. 2, p. 175 – 190.

[4] Bergamaschi, L., J. Gondzio e G. Zilli, (2004), " Preconditioning Indefinite Systems in Interior Point Methods for Optimization ", Computational Optimization and Applications, v. 28, p. 149 – 171.

[5] Bertsekas, D., L. Polymenakos e P. Tseng, (1997), "An ∈-Relaxations Method for Convex Network Optimization Problems ", SIAM Journal on Optimization, v. 7, p. 853 – 870.

[6] Bonettini, S., E. Galligani e V. Ruggiero, (2005), "An Inexact Newton Method Combined with Hestenes Multipliers' scheme for the Solution of Karush – Kuhn – Tucker Systems", Applied Mathematics and Computation, v. 168, p. 651-676.

[7] El-Bakry, A., R. Tapia, T. Tsuchiya e Y. Zhang, (1996), “On the Formulation and Theory of Newton Interior -Point Method for Nonlinear Programming”, Journal of Optimization Theory and Applications, v. 89, p. 507 – 541.
[8] Durazzi, C. e V. Ruggiero, (2003), “ Indefinitely Preconditioned Conjugate Gradient Method for Large Sparse Equality and Inequality Constrained Quadratic Problems”, Numerical Linear Algebra with Applications, v. 10, p. 673 – 688.
[9] García, R., A. Marín e M. Patriksson, (2003), “ Column Generation Algorithms for Nonlinear Optimization, I: Convergence Analysis”, Optimization, v. 52, no . 2, p. 171 – 200.
[10] Golub, G., X. Wu e J. Yuan, 2001, “SOR - Like Methods for Augmented Systems”, BIT, v. 41, no. 1, p. 71-85.
[11] Karmarkar, N., (1984), “A Polynomial - Time Algorithm for Linear Programming”, Combinatorica, v. 4, p. 373 – 395.
[12] Luksan, L., Matonoba, C. E Vlcek, J., “Interior Point Methods for Large - Scale Nonlinear Programming”, Optimization Methods and Software, v. 20, n. 4-5, p. 569 – 582.
[13] Marín, A., (1995), “ Restricted Simplicial Decomposition with Side Constraints” , Networks, v. 26, p. 199 – 215.
[14] Nagurney , A. (1984), “ Comparative Tests of Multimodal Traffic Equilibrium Methods ”, Transportation Research, v. 18B, no. 6, p. 469- 485.
[15] Ouorou, A., (2000), “A Primal – Dual Algorithm for Monotropic Programming and its Application to Network Optimization”, Computational Optimization and Applications, v. 15, p.125 – 143.
[16] Torres, L., (2005), “Método Primal – Dual para Otimização Não Linear Separável”, VIII Simpósio de Pesquisa Operacional e Logística da Marinha, 16 e 17 de Agosto de 2005, Rio de Janeiro.
[17] Wright, S., (1997), Primal – Dual Interior – Point Methods, SIAM, Philadelphia, PA.