O MÉTODO DE PONTOS INTERIORES NO PLANEJAMENTO DA RADIOTERAPIA

Andréa Camila dos Santos Martins

Dissertação apresentada à Universidade Estadual Paulista "Júlio de Mesquita Filho" para a obtenção do título de Mestre em Biometria.

BOTUCATU
São Paulo - Brasil
Fevereiro - 2011

O MÉTODO DE PONTOS INTERIORES NO PLANEJAMENTO DA RADIOTERAPIA

Andréa Camila dos Santos Martins

Orientadora: Profa. Dra. **Helenice de Oliveira Florentino Silva**

Dissertação apresentada à Universidade Estadual Paulista "Júlio de Mesquita Filho" para a obtenção do título de Mestre em Biometria.

BOTUCATU
São Paulo - Brasil
Fevereiro - 2011

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação
Divisão Técnica de Biblioteca e Documentação - Campus de Botucatu - UNESP
Bibliotecária responsável: *Sulamita Selma Clemente Colnago – CRB 8/4716*

Martins, Andréa Camila dos Santos.
O método de pontos interiores no planejamento da radioterapia / Andréa Camila dos Santos Martins. - Botucatu, 2011

Dissertação (mestrado) – Instituto de Biociências de Botucatu, Universidade Estadual Paulista, 2011
Orientador: Helenice de Oliveira Florentino Silva
Capes: 10104003

1. Radioterapia - Matemática. 2. Biometria.

Palavras-chave: Otimização; Pontos interiores; Radioterapia

Dedicatória

Dedico essa dissertação ao meu esposo, meus pais, meus irmãos e especialmente à minha avó Tereza.

Agradecimentos

A Deus, por todas as graças recebidas e pela força concedida para completar mais essa etapa da minha vida.

Ao meu marido Thiago por estar sempre do meu lado nas horas de dificuldade e por acreditar e apoiar os meus sonhos.

Aos meus pais Carlos e Solange por sempre incentivarem a realizar os meus objetivos.

Aos meus irmãos Hellen, Carlos Jr e Ana Elisa por estarmos sempre juntos.

Aos familiares e amigos sempre presentes e indispensáveis em minha vida.

À Profa.Dra. Helenice pelo voto de confiança depositado em mim, por todas as coisas que me ensinou para que o melhor seja feito.

Ao Prof.Dr. Antônio Roberto Balbo pelo incentivo para continuar os estudos.

As amizades feitas no mestrado e a ajuda recebida durante todo esse processo.

Aos professores do Departamento de Bioestatística - IBB - UNESP, por todo conhecimento científico compartilhado.

Aos funcionários do Departamento de Bioestatística - IBB - UNESP, pela boa convivência que tivemos.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pelo suporte financeiro parcial para a realização desse projeto de pesquisa.

Se o conhecimento pode criar problemas, não é através da ignorância que podemos solucioná-los.

(Isaac Asimov)

Sumário

Página

LISTA DE FIGURAS **viii**

LISTA DE TABELAS **x**

RESUMO **xi**

SUMMARY **xii**

1 INTRODUÇÃO **1**

2 TRATAMENTO POR RADIOTERAPIA **3**

2.1 Formulação do modelo matemático 5

2.2 Modelo de programação linear 8

3 PROBLEMA DE PROGRAMAÇÃO LINEAR **11**

3.1 Problemas de otimização 11

3.2 Problema de programação linear 12

3.3 Resolução de um problema de programação linear 13

3.3.1 O método Simplex 13

3.3.2 O método de Pontos Interiores 14

3.4 Algoritmos afins de pontos interiores 15

3.4.1 Método Primal-Afim 15

3.4.2 Método Dual-Afim 27

3.4.3 Método Primal-Dual 35

3.4.4 Implementação do algoritmo Primal-Dual 45

4 APLICAÇÃO DA PROGRAMAÇÃO LINEAR NO PLANEJAMENTO ÓTIMO DA RADIOTERAPIA **53**

5 CONCLUSÃO **62**

ANEXOS **63**

REFERÊNCIAS BIBLIOGRÁFICAS **70**

APÊNDICE **74**

Lista de Figuras

Página

1 Imagem de tomografia computadorizada com estruturas de interesse selecionadas - Fonte: Viana et al. (2008) 6

2 Estrutura da matriz de deposição de dose A - Fonte: Viana (2010) 7

3 Esboço geométrico da transformação afim no espaço bidimensional - Fonte: Fang & Puthenpura (1993) 19

4 Esboço da projeção do vetor direção do método Primal-Afim no espaço bidimensional - Fonte: Fang & Puthenpura (1993) 21

5 Soluções obtidas pelo algoritmo Primal-Afim 27

6 Significado geométrico do comprimento do passo para o método Primal-Dual - Fonte: Fang & Puthenpura (1993) 44

7 Comparação das soluções obtidas pelo algoritmo Primal-Afim e Primal-Dual . 52

8 Fluxograma do planejamento otimizado considerando a localização das regiões de interesse 54

9 Imagem hipotética, simulando um tumor interior a uma região crítica 55

10 Divisão por pixels da imagem hipotética 56

11 Representação da imagem simulada com os subfeixes de radiação 57

12 Distribuição da porcentagem de dose por pixel com os resultados obtidos pelo aplicativo linprog do *software* Matlab 59

13 Distribuição da porcentagem de dose por pixel ao longo da imagem 11 com os resultados obtidos pelo aplicativo linprog do *software* Matlab 60

14 Distribuição da porcentagem de dose por pixel com os resultados obtidos pelo programa desenvolvido em linguagem C++ relativo ao MPDPI 60

15 Distribuição da porcentagem de dose por pixel ao longo da imagem 11 com os resultados obtidos pelo programa desenvolvido em linguagem C++ relativo ao MPDPI 61

Lista de Tabelas

Página

1 Resultados obtidos pelo algoritmo Primal-Afim 26

2 Resultados obtidos pelo algoritmo Dual-Afim 34

3 Resultados obtidos pelo algoritmo Primal-Dual 51

4 Resultados parciais da função objetivo minimizada no planejamento otimizado 58

O MÉTODO DE PONTOS INTERIORES NO PLANEJAMENTO DA RADIOTERAPIA

Autora: ANDRÉA CAMILA DOS SANTOS MARTINS
Orientadora: Profa. Dra. HELENICE DE OLIVEIRA FLORENTINO SILVA

RESUMO

Um tratamento do câncer por radioterapia tem como objetivo a eliminação das células do tumor e preservação das células saudáveis, obtendo assim uma melhor homogeneização da dose administrada e menor possibilidade de complicações clínicas durante o tratamento. O sucesso do tratamento depende de um bom planejamento. Para um planejamento ótimo, técnicas matemáticas estão sendo utilizadas com o objetivo de maximizar a radiação no tumor e minimizar a radiação nas regiões vizinhas, com isto modelos de programação linear têm sido ótimas ferramentas para auxiliar a construção dos planos de tratamento por radioterapia. Assim, este trabalho visa: estudar os principais conceitos envolvidos no planejamento do tratamento do câncer por radioterapia; estudar modelos de programação linear (PL) aplicados ao planejamento ótimo; fazer um amplo estudo sobre a técnica de pontos interiores para PL e apresentar uma aplicação desta técnica para resolução de um problema de planejamento ótimo para o tratamento do câncer por radioterapia.

THE INTERIOR POINTS METHOD IN RADIOTHERAPY PLANNING

Author: ANDRÉA CAMILA DOS SANTOS MARTINS
Adviser: Profa. Dra. HELENICE DE OLIVEIRA FLORENTINO SILVA

SUMMARY

A cancer treatment by radiotherapy aims to eliminate tumor cells and preservation of healthy cells, thus getting a better homogenization of the administered dose and fewer chances of complications during treatment. Treatment success depends on good planning. For an optimal planning, mathematical techniques are being used in order to maximize radiation at tumor and minimize radiation in the surrounding regions, thus linear programming models has been great tools to assist the construction of treatment plans for radiation therapy. Thus, this work aims: studying the key concepts involved in planning the treatment of cancer by radiotherapy; study the models the linear programming (PL) applied to optimal planning; make a broad study on the technique of interior point for PL and present an enforcement of this technique for solving a problem of optimal planning for cancer treatment by radiotherapy.

1 INTRODUÇÃO

O tratamento do câncer tem como finalidade a interrupção do crescimento e reprodução das células cancerosas. As modalidades de tratamento utilizadas são a radioterapia, a cirurgia e a quimioterapia. A radioterapia é um método capaz de destruir células tumorais, empregando feixe de radiações ionizantes. Após a indicação para esse tipo de tratamento, uma dose pré-calculada de radiação é aplicada no paciente, em um determinado tempo, a um volume de tecido que engloba o tumor, buscando erradicar todas as células tumorais, com o menor dano possível às células normais circunvizinhas. Um planejamento da radioterapia é considerado ótimo quando todos os parâmetros envolvidos, sejam eles físicos ou biológicos, foram investigados e adequados para cada paciente. Com a exigência de planejamentos otimizados surgiram novas modalidades de tratamento e de aparelhos. As imagens dos exames de tomografia computadorizada também começaram a ser utilizadas pelos sistemas de planejamento, por simularem uma interação entre o paciente e os feixes de radiação durante as seções de radioterapia. Devido à existência de uma grande variedade de algoritmos para reconstrução tomográfica e cálculo de dose absorvida na radioterapia, diferentes aplicações envolvendo técnicas matemáticas foram propostas (Viana, 2010).

Do ponto de vista matemático, o desafio é emitir uma dosagem de radiação no tumor suficiente para a sua eliminação e simultaneamente, minimizar a radiação nas regiões vizinhas, que por muitas vezes são críticas, evitando assim complicações. Muitos autores Sonderman & Abrahamson (1985), Rosen et al. (1990) e Shepard et al. (1999) tem trabalhado com técnicas matemáticas para auxiliar o planejamento por radioterapia, dentre estas técnicas a programação linear tem obtido sucesso, pois apresenta resultados promissores quando comparados aos resultados dosimétricos (Men et al., 2007).

A programação linear é a área da matemática que estuda a modelagem matemática e as técnicas de resolução do problema de programação linear (PPL). A mode-

lagem matemática de um problema de otimização parte do conhecimento de um problema real onde o objetivo é maximizar ou minimizar um dado comportamento. Os métodos exatos existentes para a resolução deste PPL são: o Método Simplex e o Método de Pontos Interiores. De acordo com Cid (2003), o método de pontos interiores é o mais adequado para resolver o problema de planejamento da radioterapia, principalmente aqueles modelos que se baseiam na formulação de Holder (2003), pois o problema apresenta uma estrutura matricial bem definida, facilitando o uso desta técnica e obtendo um melhor desempenho computacional.

O objetivo deste trabalho é discutir a técnica de pontos interiores para programação linear e sua aplicação na construção de planejamentos ótimos da radioterapia para tratamento do câncer, pois o método de pontos interiores atende as características dos modelos de programação linear desta área (Cid, 2003). Optou-se pela utilização do modelo linear proposto por Holder (2003), devido a grande importância e ao número de citações que esta formulação tem recebido (Zhang et al., 2004), (Lim et al., 2008) e (Rocha & Dias, 2009).

O trabalho está organizado da seguinte forma: o capítulo 02 apresenta os principais conceitos envolvidos no planejamento do tratamento do câncer por radioterapia e também uma formulação matemática de um problema de programação linear para auxílio neste planejamento. O capítulo 03 apresenta a técnica de pontos interiores para a programação linear visando a resolução do modelo apresentado no capítulo 02. O capítulo 04 apresenta a aplicação do modelo de programação linear para a resolução de um problema de planejamento ótimo da radioterapia e os resultados computacionais obtidos nesta aplicação. No final, o capítulo 05 apresenta as conclusões deste trabalho.

2 TRATAMENTO POR RADIOTERAPIA

O câncer é uma doença que tem início quando ocorre uma mutação genética no DNA da célula, onde o mecanismo de controle do crescimento normal do tecido celular é alterado. O tratamento do câncer têm como objetivo a erradicação da doença, contenção do crescimento das células, ou alívio dos sintomas. As modalidades existentes de tratamento do câncer são: cirurgia, radioterapia e quimioterapia (Lorencetti & Simonetti, 2005).

A radioterapia é o tipo de tratamento em que se utilizam os raios-X, com a finalidade de interrupção do crescimento e reprodução de células tumorais. Uma dose pré-calculada de radiação é aplicada, buscando erradicar todas as células tumorais, com o menor dano possível às células normais circunvizinhas. Pode-se dizer que grande parte dos indivíduos que tiveram um diagnóstico de câncer, realizaram seu tratamento baseado em radioterapia ou como uma combinação desta com outras modalidades de tratamento (Sultan, 2006).

Os raios-X foram descobertos em 1895 por Wilhelm Conrad Röntgen, quando estudava o fenômeno da luminescência produzida por raios catódicos em um tubo de Crookes (Arruda, 1996). Percebeu-se então que a exposição dos tecidos aos raios-X, sem a devida proteção e por tempo prolongado, causam irritações na pele, ulcerações, dermatites, sérias lesões na epiderme, morte celular, ou até a morte do indivíduo (Santos, 1995). A capacidade da radiação causar lesões e morte à nível celular motivou estudos, visando pesquisar uma aplicação eficaz no combate de doenças desse tipo. As pesquisas mostraram que a radiação provocava efeitos nocivos tanto a células sadias quanto a células cancerígenas porém, as células cancerígenas são mais radiossensíveis. A radiossensibilidade indica a regressão do tecido após a radioterapia, sendo assim as células sadias se recuperam mais facilmente dos efeitos causados pela radiação (Stoll, 1968).

A maioria dos tratamentos realizados com radioterapia utilizam fótons de

alta energia gerados por um acelerador linear e aplicados em diferentes posições angulares ao redor do paciente. O feixe de fótons incidentes pode ser conformado (colimado), de maneira a irradiar apenas o perímetro tumoral, evitando um espalhamento dos raios no paciente. Essa conformação do feixe pode ser feita com o uso de colimadores multi-lâmina, sistema automatizado acoplado a uma máquina logo na saída do feixe. Os colimadores são feitos de material de alta densidade eletrônica a fim de atenuar porções específicas do feixe principal. Sua montagem obedece a uma configuração, tal que várias barras ficam dispostas paralelamente, obedecendo a um controle automático estático para sua abertura ou fechamento, criando uma perspectiva do perímetro tumoral (Shepard et al., 1999).

Na década de 90, foi desenvolvida a radioterapia com intensidade modulada (IMRT - Intensity Modulated Radiation Therapy), onde os colimadores empregados obedecem a um regime de controle automático de abertura e fechamento dinâmico, movendo-se enquanto o feixe é incidido no paciente de acordo com os tecidos irradiados e suas respectivas doses toleradas. Desta forma, o modelo dinâmico de colimação utilizado na IMRT é capaz de produzir feixes não-homogêneos, permitindo o planejamento de distribuição de doses para regiões de difícil acesso ou de baixa tolerância à radiação, como medula espinhal e tumores cerebrais (Webb, 2003). A IMRT é caracterizada como uma modalidade de radioterapia conformacional, pois encerra o volume tumoral irradiado, onde o paciente é tratado com campos de radiação que não possuem uma fluência uniforme (número de fótons por unidade de área), pois o feixe perde certo número de fótons (atenuação) conforme interage com o sistema dinâmico de colimação multi-lâmina (Webb, 2003).

Nos dois tipos de plano de tratamento (convencional ou otimizado) o princípio funcional é baseado nos efeitos da interação da radiação com a matéria e a reação do organismo humano nesse processo.

No tratamento convencional, sem o uso de técnicas de otimização, a prática mais comum consiste em transmitir a maior radiação possível no tumor. Há duas razões para evitar este procedimento:

1. altos níveis de radiação podem conduzir uma grande soma de necrose, e o corpo humano teria dificuldade para eliminá-lo;

2. as células doentes estão distribuídas entre os tecidos saudáveis, consequentemente,

uma dose letal uniformemente distribuída na região do tumor é crucial para o sucesso do plano de tratamento.

Do ponto de vista matemático, o desafio é emitir uma alta dosagem de radiação no tumor, suficiente para sua eliminação, e simultaneamente, minimizar a radiação nas regiões vizinhas, reduzindo ao máximo as complicações nestas regiões, que são muitas vezes críticas.

O processo de planejamento ótimo por radioterapia começa após o diagnóstico, avaliação da localização, tamanho e capacidade radiobiológica do tumor. Posteriormente é prescrita pelo médico uma dose para cada tipo de tecido: tumoral, crítico e saudável, de tal forma que apresente condições propícias para se obter a menor possibilidade de complicações clínicas durante o tratamento. Para o tratamento não deve-se levar em conta somente a distribuição de dose profunda no eixo central, pois não é suficiente para caracterizar o feixe de radiação. Por isso, é feita a curva de isodose, que são linhas que representam pontos de mesma dose.

A proteção de órgãos vitais do campo de radiação é uma das maiores preocupações na radioterapia. Para conter isso, tempo e esforços consideráveis são gastos na modelagem dos campos para proteger não somente órgãos críticos, mas para evitar igualmente a irradiação desnecessária do tecido normal circunvizinho. Em tratamentos mais modernos, os planos de tratamento são baseados em técnicas matemáticas feitos com auxílio de um software. Para o cálculo matemático, são utilizados dados clínicos do tumor para cada paciente tratado. Estes cálculos podem ser auxiliados por técnicas matemáticas de otimização, que são atualmente os mais utilizados na literatura.

2.1 Formulação do modelo matemático

Quando o câncer é diagnosticado e há a indicação médica pelo tratamento por radioterapia, são realizados vários exames no paciente com a finalidade de conhecer a localização, forma e volume do tumor, bem como os tecidos críticos presentes na região a ser tratada. Geralmente os exames realizados são: a Tomografia Computadorizada e a Ressonância Magnética. As imagens obtidas por esses exames são fundidas, resultando em uma única imagem com mais informações, auxiliando assim uma maior precisão sobre os

dados do tumor. A dose mínima a ser aplicda no tumor é prescrita, assim como as doses máximas que os tecidos saudáveis e críticos podem receber.

A partir desta nova imagem é realizada a localização das estruturas de interesse (tumor, tecido crítico e tecido saudável) e feito um mapeamento dividindo-a em segmentos chamados pixels. A figura 1 é uma imagem de tomografia computadorizada, onde as estruturas de interesse selecionadas representam um tumor localizado na próstata e os tecidos críticos são as cabeças de fêmur, o reto e a bexiga. Os demais tecidos são considerados saudáveis.

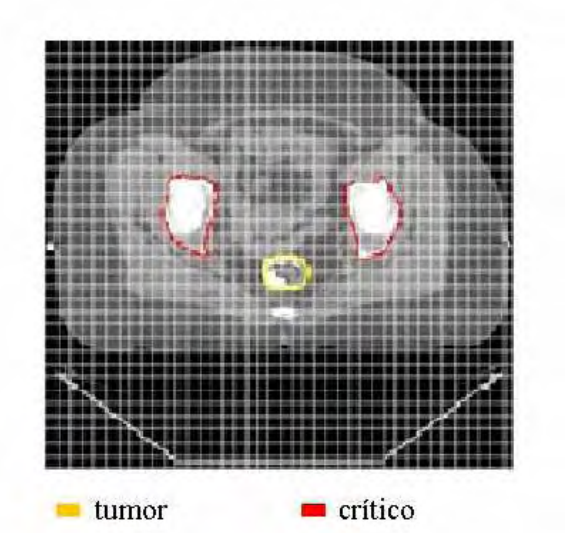

Figura 1: Imagem de tomografia computadorizada com estruturas de interesse selecionadas - Fonte: Viana et al. (2008)

O número de ângulos é fixado de forma que sejam considerados os k ângulos de localização para emissão dos feixes principais de radiação contendo em cada ângulo η subfeixes. Os sistemas de tratamento modernos são capazes de realizar combinações entre estes subfeixes de modo a utilizá-los ao longo de um arco completo, fazendo com que o planejamento utilize $k.\eta$ subfeixes. Esta decomposição dos feixes de radiação é baseada nos fundamentos teóricos e práticos da radioterapia com intensidade modulada (IMRT).

Feito a delimitação das estruturas, o plano de tratamento deve ser elaborado de forma a administrar corretamente a dose de radiação no tumor, buscando assim a

homogeneidade do plano de tratamento, de acordo com a determinação da prescrição de dose das regiões de interesse.

Sendo $x_{(a,i)}$ a dose ao longo do i-ésimo subfeixe ($i = 1, 2, ..., n$) do a-ésimo ângulo ($a = 1, 2, ..., k$), a matriz de deposição de dose é da ordem $m \times k\eta$, formada por p linhas e (a, i) colunas.

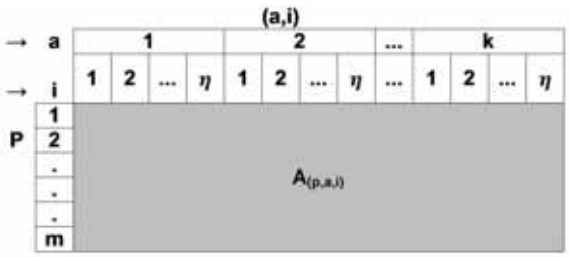

Figura 2: Estrutura da matriz de deposição de dose A - Fonte: Viana (2010)

Define-se assim a matriz $A_{(p,a,i)}$ como:

$$A_{(p,a,i)} = Se^{-\mu\ d_{(p,a,i)}}$$

Onde S a área geométrica do pixel p que recebe a dose $x_{(a,i)}$ e $d_{(p,a,i)}$ a distância entre o pixel p e a fonte posicionada no ângulo a emitido pelo subfeixe i.

O fator $e^{-\mu\ d_{(p,a,i)}}$ mede a atenuação da radiação sobre o tecido, e os valores deste coeficiente de atenuação (μ) dependem da energia do raio. O coeficiente de atenuação mede a característica do tecido com relação a sua densidade que são grandezas proporcionais, isto é, quanto menor a densidade do tecido, menor o coeficiente de atenuação.

Seja, m_t, m_c e m_g o número de pixels dos tecidos tumoral, crítico e saudável, respectivamente, o número total de pixels é dado por $m = m_t + m_c + m_g$.

Assim, sem perda de generalidade, a matriz de deposição de dose A pode ser representada pelas sub-matrizes $A_t \in \Re^{m_t \times n}$, $A_c \in \Re^{m_c \times n}$ e $A_g \in \Re^{m_g \times n}$, que são matrizes dos respectivos tecidos tumoral, crítico e saudável.

$$A = \begin{bmatrix} A_t \\ A_c \\ A_g \end{bmatrix}$$

Sendo x o vetor de dose cujos elementos são $x_{(a,i)}$, dados por:

$$x^T = [x_{(1,1)}, x_{(1,2)}, ..., x_{(1,n)}, x_{(2,1)}, ..., x_{(2,n)}, ..., x_{(k,1)}, ..., x_{(k,n)}]$$

A dose de radiação total para o pixel p é dada pela p-ésima componente $(p = 1, 2, ..., m)$ e obtida pela multiplicação: Ax.

No tratamento por radioterapia é prescrito um limitante de dose para cada tipo de tecido (tumoral, crítico e saudável). Devido à divisão em pixels, a prescrição da dose é dada em forma de vetor com a seguinte notação:

- u_t representa o limite superior de radiação no tumor ($u_t \in \Re^{m_t}$);
- l_t representa o limite inferior de radiação no tumor ($l_t \in \Re^{m_t}$);
- u_c representa o limite superior de radiação no tecido crítico ($u_c \in \Re^{m_c}$);
- u_g representa o limite superior de radiação no tecido saudável ($u_g \in \Re^{m_g}$)

Obviamente, temos que $0 < l_t \leq u_t$, $u_c \geq 0$ e $u_g \geq 0$. O limite superior e inferior de dose para os pixels do tumor são obtidos através de metas estabelecidas (t_g). Os valores de u_{t_i} e l_{t_i} são geralmente $(1 + \varepsilon)t_g$ e $(1 - \varepsilon)t_g$, respectivamente, onde ε é a porcentagem da variação para a doseagem do tumor, que é usualmente usado na faixa de 2% à 15% (Holder, 2003).

A partir destas informações o modelo de programação linear será abordado.

2.2 Modelo de programação linear

O modelo proposto por Holder (2003), objetiva alcançar 3 metas:

1. minimizar a deficiência de dose no tumor;
2. minimizar a quantidade média de radiação que o tecido crítico está recebendo acima da dose prescrita;
3. minimizar a quantidade média de radiação que o tecido saudável está recebendo acima da dose prescrita.

O modelo é dado por:

$$\text{minimizar } w\, l^T t + u_c^T c + u_g^T g \tag{1}$$

$$\text{sujeito a } \quad l_t - Lt \leq A_T x \leq u_t \tag{2}$$

$$A_C\, x \leq u_c + U_C\, c \tag{3}$$

$$A_G\, x \leq u_g + U_G\, g \tag{4}$$

$$0 \leq Lt \leq l_t \tag{5}$$

$$-u_c \leq U_C\, c \tag{6}$$

$$U_G\, g \geq 0 \tag{7}$$

$$x \geq 0 \tag{8}$$

O modelo consiste em determinar x que é o vetor de doses dos sub-raios que entram na imagem para alcançar os pixels mapeados, $t \in \Re^{m_T}$, $c \in \Re^{m_C}$, $g \in \Re^{m_G}$ que são, respectivamente, o vetor tumoral, crítico e saudável, que otimizam a função objetivo representada pela soma de três metas no tratamento: $w\, l^T t + u_c^T c + u_g^T g$, onde:

- $l^T\, t$ mede o quanto falta para que o plano encontrado consiga aplicar a dose mínima na região do tumor;
- $u_c^T\, c$ mede a quantidade de radiação acima da prescrita recebida pela região crítica;
- $u_g^T\, g$ mede a quantidade de radiação acima da prescrita nos demais tecidos saudáveis.

O escalar positivo w pondera a importância da formulação de um plano que obtenha a dose mínima na região do tumor, isto é, valores grandes de w forçam $l^T\, t$ a ser tão pequeno quanto possível. Seria desejável que houvesse um valor para um finito $w \geq 0$ tal que o valor ótimo da componente $l^T\, t$ fosse zero, o que garantiria que o tumor estaria recebendo o nível mínimo de radiação necessário para sua eliminação.

As restrições (2)-(4) são responsáveis pelo controle da deposição da dose prescrita para as regiões de interesse e são denominadas elásticas, pois seus limites podem variar de acordo com os vetores t, c e g, que correspondem as variáveis do problema.

As restrições (5)-(7) controlam a penalização ou recompensa com relação à elasticidade e restrição (8) exige que os pesos relativos sejam positivos.

As matrizes L, U_C e U_G definem como medir a elasticidade, e os vetores l, u_c e u_g controlam a penalização ou recompensa com relação à elasticidade. Valores fixos de L, U_C e U_G , l, u_c e u_g definem um conjunto de funções elásticas. Elas garantem que o conjunto de restrições seja sempre estritamente factível e que a diferença dos limites inferiores nas funções elásticas permitam incorporar diferentes objetivos de tratamento (Berman & Plemmons, 1979) e (Holder, 2003).

O modelo (1)-(8) é definido como problema de programação linear e pode ser resolvido por técnicas clássicas de otimização.

A seguir serão apresentados os métodos de programação linear utilizados para a resolução deste problema.

3 PROBLEMA DE PROGRAMAÇÃO LINEAR

3.1 Problemas de otimização

Um problema geral de otimização consiste em minimizar ou maximizar uma função, onde o seu domínio é um conjunto dado por $\{x \in \Re^n / g(x) \leq 0\}$, que poderia também ser $g(x) = 0$ ou $g(x) \geq 0$. Na programação matemática este problema é expresso na forma:

$$\begin{aligned} &\text{minimizar } f(x) \ (ou \ \ maximizar) \\ &\text{sujeito a } g(x) \leq 0 \ \ (ou \ \ g(x) = 0, \ \ g(x) \geq 0) \\ &x \in \Re^n \end{aligned} \tag{9}$$

Em que: a função $f : \Re^n \to \Re$ é chamada de função objetivo; as restrições $g : \Re^n \to \Re^p$ limitam o espaço de soluções do problema, chamadas de soluções factíveis ou soluções viáveis e x é o vetor variável de decisão.

Dependendo da natureza do problema de otimização (9), a função objetivo, bem como as restrições assumem diferentes características, necessitando de diferentes técnicas para a sua resolução.

- Se $f(x)$ e/ou $g(x)$ forem não lineares, tal que $x = (x_1, x_2, ..., x_n)$, $x_i \in \Re$, $i = 1, 2, ..., n$, tem-se um problema de programação não linear (PPNL).
- Se $f(x)$ e $g(x)$ forem lineares, tal que $x = (x_1, x_2, ..., x_n)$, $x_i \in \Re$, $i = 1, 2, ..., n$, tem-se um problema de programação linear(PPL).
- Se $f(x)$ e $g(x)$ forem lineares, tal que $x = (x_1, x_2, ..., x_n)$, x_i inteiro para todo $i \in 1, 2, ..., n$, tem-se um problema de programação linear inteira (PPLI).

- Se $f(x)$ e $g(x)$ forem lineares, tal que $x = (x_1, x_2, ..., x_n)$, x_i inteiro para algum $i \in 1, 2, ..., n$, tem-se um problema de programação linear inteira mista (PPLIM).

O foco deste trabalho são os problemas de programação linear. A seguir será abordado o modelo matemático de um PPL e discutido sua forma de resolução.

3.2 Problema de programação linear

O problema de programação linear é um problema da forma (9), onde a função objetivo e as restrições são formadas por equãções lineares.

Para a resolução de problemas de programação linear será utilizada a forma padrão dada por:

$$\begin{aligned} &\text{minimizar } c^T x \\ &\text{sujeito a } Ax = b \\ &x \geq 0 \end{aligned} \tag{10}$$

Onde A, c, b e x têm dimensões apropriadas.

A programação linear é a área da matemática que estuda a modelagem e técnicas de resolução de problemas de programação linear (PPL) da forma definida em (10). A resolução do modelo de programação linear (10) consiste em determinar o vetor x que satisfaça as restrições impostas e otimize a função objetivo. Havendo solução possível e não havendo solução ilimitada, a solução ótima é uma solução factível que otimiza a função objetivo.

A modelagem matemática de um problema de otimização parte do conhecimento do problema real onde o objetivo é maximizar ou minimizar uma quantidade que depende de uma variável de decisão e o sistema está sujeito a um conjunto de restrições que limitam a região de busca da solução do problema.

Na modelagem de um problema de programação linear devem estar bem definidos:

- O conjunto de variáveis manipuláveis: variáveis de decisão (x);

- O objetivo a ser alcançado representado pela função objetivo: função linear das variáveis de decisão;
- As restrições do problema: equações lineares representadas através das variáveis de decisão.

Modelado o problema, ou seja, colocado o problema em forma de equações matemáticas, o próximo passo é identificar qual o método de otimização deverá ser usado para resolução deste. A seguir será discutida a forma de resolução de um PPL.

3.3 Resolução de um problema de programação linear

As técnicas utilizadas para determinar numericamente a solução ótima de um problema de programação linear são baseadas em dois métodos: o método Simplex e o método de Pontos Interiores.

3.3.1 O método Simplex

O método Simplex foi apresentado por George Dantzig em 1947 como uma forma sistemática de resolução da programação linear. Hoje existem várias modificações e adaptações desse método e a grande vantagem desse algoritmo é a facilidade de implementação e a existência de vários *softwares* comerciais.

O método Simplex é um procedimento algébrico e iterativo que fornece a solução exata de qualquer problema de programação linear em um número finito de iterações. É também capaz de indicar se o problema tem solução ilimitada, se não tem solução ou se possui infinitas soluções. Este algoritmo pode ser visto como um processo combinatório que busca encontrar as colunas da matriz de restrições que induzem uma base e, portanto, uma solução básica ótima.

De forma resumida, para iniciar o método Simplex é necessária a existência de uma solução básica viável inicial. O método usa uma estratégia para verificar se a presente solução é ótima, e em caso afirmativo o problema está resolvido, caso contrário, outra solução básica viável é obtida através da redução Gauss-Jordan sob o sistema $Ax=b$.

Ou seja, há uma mudança de solução básica de forma a diminuir o valor da função objetivo e a otimalidade desta nova solução é analisada. Este procedimento é repetido até que a solução ótima seja encontrada.

Segundo Cid (2003) para a resolução do modelo (1-8) para o planejamento ótimo da radioterapia, o método de pontos interiores é o mais adequado, pois o modelo apresenta uma estrutura matricial bem definida (esparsa e de blocos), facilitando o uso desta técnica e obtendo um melhor desempenho computacional. Assim, este trabalho dará enfâse ao método de pontos interiores.

3.3.2 O método de Pontos Interiores

Desde a introdução do algoritmo de escala projetiva proposto por Karmarkar (1984), o método de pontos interiores tornou-se o método mais notável para resolver problemas de programação linear. Este pioneirismo provocou uma agitação nas atividades de pesquisa desta área, pois Karmarkar afirmou que este era mais eficiente que o método Simplex, por utilizar a estratégia de determinar uma solução ótima caminhando por pontos interiores do domínio factível, diferentemente do Simplex que a cada iteração percorre as extremidades da região factível.

Entre todas as alteracões do algoritmo original feito por Karmarkar a que mais atraiu a atenção dos pesquisadores foi a do procedimento afim. Este procedimento usa uma transformação afim simples, permitindo trabalhar com os problemas de programação linear na forma padrão, substituindo a transformação projetiva original de Karmarkar. O algoritmo proveniente desta mudança foi chamado de Primal-Afim Escala, atualmente conhecido como Primal-Afim de pontos interiores.

O algoritmo afim básico foi primeiramente apresentado por Dikin (1967), um matemático soviético. Depois, o trabalho foi independentemente redescoberto por Barnes (1986) e Vanderbei et al. (1986). Eles propuseram usar o algoritmo Primal-Afim para resolver o problema linear na forma padrão (primal). Um algoritmo similar, chamado algoritmo Dual-Afim de pontos interiores foi projetado e implementado por Adler et al. (1989) para resolver problemas lineares na forma de desigualdade (dual). Estes dois algoritmos são geralmente usuados para problemas de grande dimensão que mostram resultados

promissores e apresentam complexidade de tempo polinomial.

Posteriormente, o trabalho de Megiddo & Shub (1989) mostrou que a convergência polinomial do algoritmo afim depende da solução inicial. Uma má solução inicial, que está muito perto de um vértice do domínio factível, por exemplo, pode resultar em um longo tempo computacional, percorrendo todos os vértices da região factível. Todavia, a complexidade de tempo polinomial do algoritmo Primal-Afim e Dual-Afim pode ser restabelecida por incorporar uma função barreira logarítmica no contorno do octante positivo impedindo que uma solução fique "presa", possivelmente na fronteira. Existe também uma terceira variação do método de Karmarkar, chamado algoritmo Primal-Dual de pontos interiores, que foi apresentado e analisado por Monteiro et al. (1990) e também por Kojima et al. (1989). A prova teórica da complexidade de tempo polinomial foi realizada com êxito.

No que segue, são abordadas as variações dos métodos afim mencionados acima, mostrando as principais definições e processos iterativos. As atenções serão focalizadas nos três elementos básicos para um processo iterativo, que são:

1. Solução inicial;
2. Boa direção de busca;
3. Critério de parada.

Sendo assim, inicia-se com os conceitos básicos dos algoritmos afins de pontos interiores e apresentação dos algoritmos dos métodos Primal-Afim, Dual-Afim e Primal-Dual, segundo Fang & Puthenpura (1993).

3.4 Algoritmos afins de pontos interiores

3.4.1 Método Primal-Afim

Considerando o problema de programação linear em sua forma padrão

$$\begin{aligned} &\text{minimizar } c^T x \\ &\text{sujeito a } Ax = b \\ &x \geq 0 \end{aligned} \tag{11}$$

onde A é uma matriz $m \times n$ de posto m, $b \in \Re^m$, $x \in \Re^n$ e $c \in \Re^n$.

O domínio viável é $\mathrm{P} = \{\mathrm{x} \in \mathrm{R^n}/\mathrm{Ax} = \mathrm{b},\ \mathrm{x} \geq 0\}$ e o seu interior é definido como:

$$\mathrm{P^0} = \{\mathrm{x} \in \mathrm{R^n}/\mathrm{Ax} = \mathrm{b},\ \mathrm{x} > 0\} \tag{12}$$

Um vetor x é chamado um ponto interior viável do problema de programação linear, se $x \in P^0$. Para garantirmos assim a existência de pontos interiores será considerado que: $P^0 \neq \emptyset$.

Há diversas maneiras de encontrar uma solução interior de um problema de programação linear, dentre estas, dois mecanismos são muito usados para encontrar uma solução inicial interior factível. O primeiro é o método Big-M, que é executado mais facilmente e apropriado para a maioria das aplicações. O outro é o método Duas-Fases que têm uma execução comercial mais utilizada devido a sua estabilidade (Fang & Puthenpura, 1993). Estes dois métodos estão discutidos no Apêndice A.

Idéias básicas do algoritmo Primal-Afim

Os dois principais fundamentos observados por Karmarkar na definição de seu algoritmo são:

1. No processo iterativo, se a solução interior atual estiver próximo do centro do politopo, então faz sentido mover-se numa direção de descida relacionada ao gradiente da função objetivo para conseguir um valor de mínimo;
2. Sem grandes alterações no problema original, uma transformação apropriada pode ser aplicada ao espaço solução tal que a solução interior atual seja levada para perto do centro definido no espaço transformado.

Na formulação de Karmarkar, a estrutura especial e simples é definida:

$$\Delta = \{\mathrm{x} \in \mathrm{R^n} / \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{x_i} = 1,\ \mathrm{i} = 1, ..., \mathrm{n}\}$$

e o seu ponto central $e/n = (1/n, 1/n, ..., 1/n)^T$, esta condição foi introduzida obedecendo os princípios fundamentais 1 e 2. Quando se trabalha diretamente no problema na forma padrão, a estrutura simples adotada por Karmarkar é difícil de ser realizada, e o domínio factível poderia transformar-se num conjunto poliedral ilimitado. Toda a estrutura a ser obedecida consiste na intersecção do espaço $\{\mathrm{x} \in \mathrm{R^n} / \mathrm{Ax} = \mathrm{b}\}$ formado pelas restrições explícitas e com o conjunto definido pelo octante positivo $\{\mathrm{x} \in \mathrm{R^n} / \mathrm{x} \geq 0\}$ requerido pelas restrições não-negativas. Obviamente, o octante não-negativo não tem um ponto "central" real. Então, a partir de uma transformação pode-se posicionar um ponto e no centro ou pelo menos conseguir mantê-lo a mesma distância de cada face do octante não-negativo. Isto ocorre, pois a partir do ponto central, a distância de movimento de qualquer ponto é sempre menor que 1 e permanece no interior do octante não negativo. Conseqüentemente, podendo encontrar uma transformação apropriada que descreva uma solução interior atual ao ponto, então, analogamente ao algoritmo de transformação projetiva de Karmarkar, pode-se afirmar a seguinte estratégia:

> "Assumindo uma solução interior, aplica-se uma transformação apropriada ao espaço que caminhe na direção de descida máxima no núcleo da matriz das restrições explícitas transformadas, mas controlando o comprimento do passo até o contorno das restrições não-negativas a fim de permanecer com uma solução interior do espaço transformado. Então examina-se se a transformação inversa descreve, posteriormente, uma solução no espaço original, como uma nova solução interior. Repete-se este processo até a otimalidade ou quando as condições de parada forem satisfeitas."

Uma transformação apropriada neste caso foi denominada transformação afim. Portanto, a partir desta transformação, os pesquisadores chamaram suas variações de algoritmos afins. Quando essa transformação é aplicada diretamente ao problema primal na forma padrão, esse algoritmo é denominado Primal-Afim.

A transformação afim no octante não negativo

Seja $x^k \in R^n$ um ponto interior do octante não negativo R^n, isto é, $x_i > 0$, com i =1,.....,n, define-se uma matriz diagonal $X_k \in R^{nxn}$ por:

$$X_k = diag(x^k) \begin{bmatrix} x_1^k & 0 & \cdots & 0 \\ 0 & x_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n^k \end{bmatrix}$$

A matriz X_k é não singular e admite inversa X_k^{-1}, que também é diagonal, com $1/x_i^k$ como seu i-ésimo elemento diagonal.

A transformação afim é definida no octante não negativo R_+^n, onde:

$$y = T_k(x) = X_k^{-1}x \tag{13}$$

Essa transformação simplesmente redimensiona o i-ésimo componente de x por um número positivo x_i^k. Geometricamente, transforma pontos e segmentos de reta do espaço original para pontos e segmentos de retas no espaço transformado. A figura 3 ilustra o desenho geométrico da transformação no espaço bidimensional.

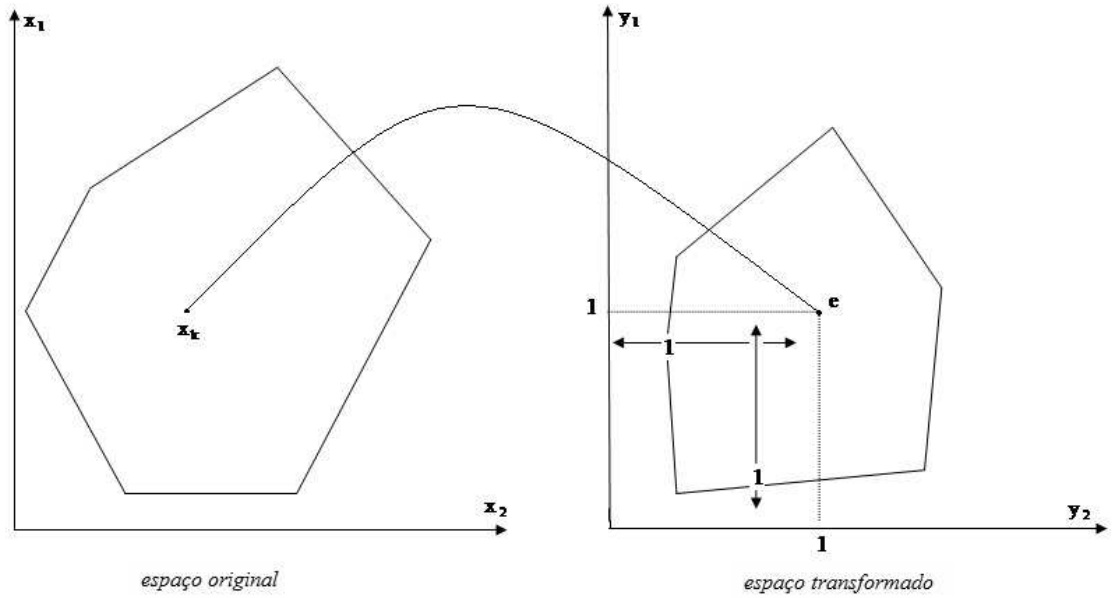

Figura 3: Esboço geométrico da transformação afim no espaço bidimensional - Fonte: Fang & Puthenpura (1993)

As seguintes propriedades para a transformação T_k podem ser verificadas:

1. T_k é uma aplicação bem definida de R^n_+ em R^n_+, se x^k for um ponto interior de R^n_+ então, y^k é um ponto interior do R^n_+.
2. $T_k(x^k) = e$.
3. $T_k(x)$ é um vértice de R^n_+ se x for um vértice.
4. $T_k(x)$ está na fronteira do octante R^n_+ se x estiver na fronteira.
5. $T_k(x)$ é um ponto interior de R^n_+ se x for um ponto interior.
6. T_k é um transformação biunívoca, admitindo assim uma transformação inversa T_k^{-1} tal que :

$$T_k^{-1}(y) = X_k\, y = x \quad para\, cada\ y \in\ R^n_+ \tag{14}$$

Dada uma solução interior x^k do problema de programação linear (11). Aplica-se a transformação afim T_k para "centralizar"sua imagem em e. Através da relação

$x = X_k y$ mostrada em (14), no espaço transformado, tem-se o seguinte problema de programação linear:

$$\begin{aligned} &\text{minimizar } (c^K)^T y \\ &\text{sujeito a } A_k y = b \\ &\qquad\qquad y \geq 0 \end{aligned} \tag{15}$$

onde: $(c^k)^T = c^T X_k$ e $A_k = AX_k$.

No problema acima, a imagem de x^k, isto é, $y^k = T_k(x^k)$ torna-se e, pois mantêm uma distância unitária da fronteira no octante não - negativo. Se mover ao longo de uma direção d_y^k, que se encontra no espaço nulo da matriz A_k, com comprimento de passo apropriado, $\alpha_k > 0$, tem-se um novo ponto: $y^{k+1} = e + \alpha_k d_y^k$, que permanece no interior do domínio factível do problema (15) e sua imagem inversa $x^{k+1} = T_k^{-1}(y^{k+1}) = X_k y^{k+1}$, torna se uma nova solução interior para o problema original (11).

Como o objetivo é minimizar o valor da função objetivo, a estratégia indicada é adotar um procedimento de máxima descida, ou seja, projetar o gradiente negativo $-c^k$ no espaço nulo da matriz A_k e criar uma direção factível d_y^k, que melhora, no caso diminui, o valor da função objetivo no espaço transformado. Para este fim, define-se primeiramente a matriz projeção no espaço nulo por:

$$P_k = I - A_k^T (A_k A_k^T)^{-1} A_k \tag{16}$$

Lembrando que: $A_k = AX_k$ e $(A_k)^T = (AX_k)^T = X_k^T A^T = X_k A^T$, pois X_k é uma matriz diagonal e portanto $X_k^T = X_k$.

Substituindo em (16), temos: $P_k = I - X_k A^T (AX_k X_k A^T)^{-1} AX_k$. Portanto a matriz de projeção do espaço fica:

$$P_k = I - X_k A^T (AX_k^2 A^T)^{-1} AX_k \tag{17}$$

Em seguida, a direção de movimento $d_y^k = P_k(-c^k)$. Como $(c^k)^T = c^T X_k$, então $c^k = X_k c$, com isso:

$$d_y^k = P_k(-c^k) = -[I - X_k A^T (AX_k^2 A^T)^{-1} AX_k] X_k c \tag{18}$$

A matriz projeção P_k é bem definida sempre que A tiver posto completo m e $x^k > 0$, assim é facilmente verificado que $AX_k d^k = 0$. A figura 4 mostra o esboço da projeção.

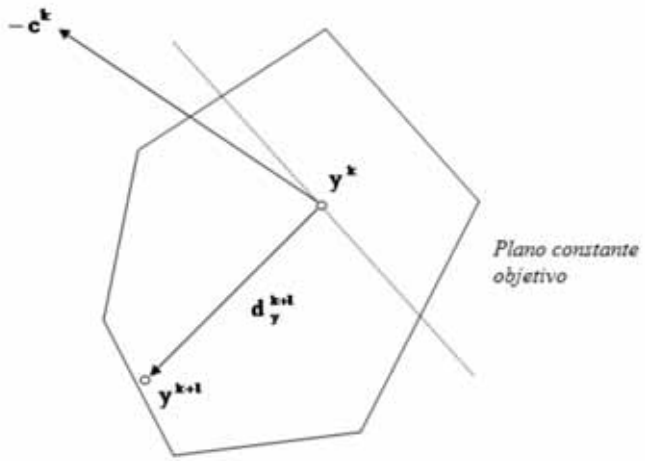

Figura 4: Esboço da projeção do vetor direção do método Primal-Afim no espaço bidimensional - Fonte: Fang & Puthenpura (1993)

Agora tem-se condição de mudar para uma nova solução interior $y^{k+1} > 0$ a partir da solução interior atual $y^k = e$ ao longo da direção d_y^k, garantindo o decréscimo do valor da função objetivo. Fazendo isso, tem-se que escolher um comprimento do passo apropriado $\alpha_k > 0$, tal que $y^{k+1} = y^k + \alpha_k d_y^k > 0$.

Observamos os possíveis valores de d_y^k :

1. Se $d_y^k \geq 0$, então α_k pode ser escolhido como qualquer número positivo, pois a solução continuará permanecendo no interior da região factível.

2. Se $(d_y^k)_i < 0$, para algum i, então $\alpha_k < \frac{1}{(-(d_y^k)_i)}$, pois $y_i^k + \alpha_k (d_y^k)_i > 0$, como $y_i^k > 0$ e $\alpha_k > 0$, tem-se $y_i^k > -\alpha_k (d_y^k)_i$ portanto $\frac{y_i^k}{(-(d_y^k)_i)} > \alpha_k$. Como $y_i^k = 1$, então $\frac{1}{(-(d_y^k)_i)} > \alpha_k$.

O valor de α_k tem que ser o mínimo possível e como não deseja-se pontos que estejam na fronteira da região factível, desta forma escolhe-se $0 < \alpha < 1$, obtendo:

$$\alpha_k = \min_i \left\{ \frac{\alpha}{-(d_y^k)_i} / (d_y^k)_i < 0 \right\} \tag{19}$$

Determinando assim um comprimento de passo apropriado, que garanta a positividade de y^{k+1}. Quando α está perto de 1, a solução atual é movida quase sempre para a fronteira mais próxima do octante positivo para definir uma nova solução interior no espaço transformado. Esta mudança de solução é ilustrada também na figura 4. O próximo passo é transformar a nova solução y^{k+1} do espaço transformado para o espaço original, obtendo assim uma melhor solução x^{k+1} para o problema (11). Isso é realizado aplicando a transformação inversa T_k^{-1} a y^{k+1}, onde: $x^{k+1} = X_k y^{k+1}$.

Sabe-se que: $d_y^k = -P_k X_k c$ e $y^{k+1} = y^k + \alpha_k d_y^k$.

Determina-se $w^k = (AX_k^2 A^T)^{-1} AX_k^2 c$ como vetor estimativa dual associado ao x^k. A partir disso tem-se:

$$x^{k+1} = x^k - \alpha_k X_k^2 [c - A^T w_k] \tag{20}$$

A direção de movimento no espaço solução original é $d_x^k = -X_k^2[c - A^T w^k]$, o comprimento do passo é α_k, $d_y^k = -X_k[c - A^T w^k]$ no espaço transformado.

Algoritmo Primal-Afim

Baseado na discussão acima apresenta-se aqui um procedimento iterativo do algoritmo Primal - Afim.

Passo 1 (inicialização): Fazer $k = 0$ e encontrar $x^0 > 0$ tal que $Ax^0 = b$.

Passo 2 (vetor estimativa dual): Calcular o vetor estimativa dual.

$$w^k = (AX_k^2A^T)^{-1}AX_k^2c$$

onde X_k é a matriz diagonal cujos elementos são as componentes de x^k.

Para facilitação do cálculo do vetor estimativa, ao invés de calcular a matriz inversa de $(AX_k^2A^T)$, é resolvido o sistema linear $(AX_k^2A^T)\ w^k = AX_k^2c$ pelo método de Cholesky.

Passo 3 (vetor custo relativo): Calcular o vetor custo reduzido.

$$r^k = c - A^Tw^k$$

Passo 4 (checando a otimalidade): Se $r^k \geq 0$ e $e^TX_kr^k \leq \varepsilon$ (um número positivo pequeno fixo), então PARE, x^k é ótimo primal e w^k é ótimo dual. Caso contrário, vá para o passo seguinte.

Passo 5 (direção de movimento): Calcular a direção.

$$d_y^k = -X_kr^k$$

Passo 6 (testando se o objetivo é ilimitado): Se $d_y^k > 0$ PARE, o problema é ilimitado. Se $d_y^k = 0$ também PARE, x^k é ótimo primal. Caso contrário vá para o Passo 7.

Passo 7 (comprimento do passo): Calcular o comprimento do passo.

$$\alpha_k = \min_i \left\{ \frac{\alpha}{-(d_y^k)_i} / (d_y^k)_i < 0 \right\} \quad onde\ 0 < \alpha < 1$$

Passo 8 (determinação de uma nova solução): Atualizar

$$x^{k+1} = x^k + \alpha_k X_k d_y^k$$

Faça $k \leftarrow k + 1$ e vá para o passo 2.

Para ilustração do algoritmo Primal-Afim será apresentado o seguinte exemplo.

Exemplo 01

Seja o seguinte problema de programação linear

$$\begin{aligned} \text{minimizar} \quad & -2x_1 + x_2 \\ \text{sujeito a} \quad & x_1 - x_2 \leq 15 \\ & x_2 \leq 15 \\ & x_1, x_2 \geq 0 \end{aligned} \tag{21}$$

Colocando o problema (21) na forma padrão, tem-se:

$$\begin{aligned} \text{minimizar} \quad & -2x_1 + x_2 \\ \text{sujeito a} \quad & x_1 - x_2 + x_3 = 15 \\ & x_2 + x_4 = 15 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

Neste caso:

$$A = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 15 \\ 15 \end{bmatrix}, \quad \text{e} \quad c = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Considerando $\alpha = 0.99$ e $\varepsilon = 10^{-3}$ e a solução inicial interior e factível como: $x^0 = \begin{bmatrix} 10 & 2 & 7 & 13 \end{bmatrix}^T$. Assim, seguindo os passos do algoritmo Primal-Afim temos descrito a primeira iteração.

Calculando a matriz diagonal, o vetor estimativa dual e o custo relativo iniciais.

$$X_0 = diag(x^0) = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 13 \end{bmatrix}, \quad w^0 = \begin{bmatrix} -1.3335 \\ -0.0077 \end{bmatrix} \quad \text{e} \quad r^0 = \begin{bmatrix} -0.6664 \\ -0.3258 \\ 1.3335 \\ 0.0077 \end{bmatrix}$$

Checando a otimalidade, percebe-se que existem algumas componentes de r^0 que são negativas e $e^T X_0 r^0 = 2.1187$, então a solução não é ótima.

Calculando a direção de movimento e o comprimento do passo tem-se:

$$d_y^0 = \begin{bmatrix} 6.6646 & 0.6516 & -9.3347 & -0.1002 \end{bmatrix}^T \quad e \quad \alpha_0 = 0.10659$$

A nova solução e o valor da função objetivo são:

$$x^1 = x^0 + \alpha_0 X_0 d_y^0 = \begin{bmatrix} 17.1039 \\ 2.1389 \\ 0.0350 \\ 12.8610 \end{bmatrix} \quad e \quad z^1 = -31.9982.$$

Como a solução encontrada ainda não é a ótima, mais iterações são realizadas. O resultado obtido pelo algoritmo Primal-Afim e a interpretação geométrica do problema dado podem ser vistos, respectivamente na tabela 1 e na figura 5.

iteração(k)	x_i^k	função objetivo
1	17.1039 2.1389 0.0350 12.8610	-31.9982
2	29.8296 14.8713 0.0417 0.1286	-44.7879
3	29.9837 14.9987 0.0149 0.0012	-44.9688
4	29.9986 14.9987 0.0001 0.0012	-44.9984
5	29.9998 14.9999 0.0001 0.00001	-44.9997
6	29.9999 14.9999 0.000001 0.00001	-44.9999
7	29.9999 14.9995 0.0006 0.0041	-44.9999
8	30.0000 14.9999 0.0000 0.0000	-45.0001

Tabela 1: Resultados obtidos pelo algoritmo Primal-Afim

A solução ótima encontrada na 8ª iteração é: $x_1^8 = 30$ e $x_2^8 = 14.9999$ e o valor da função objetivo é: $z^8 = -45.0001$.

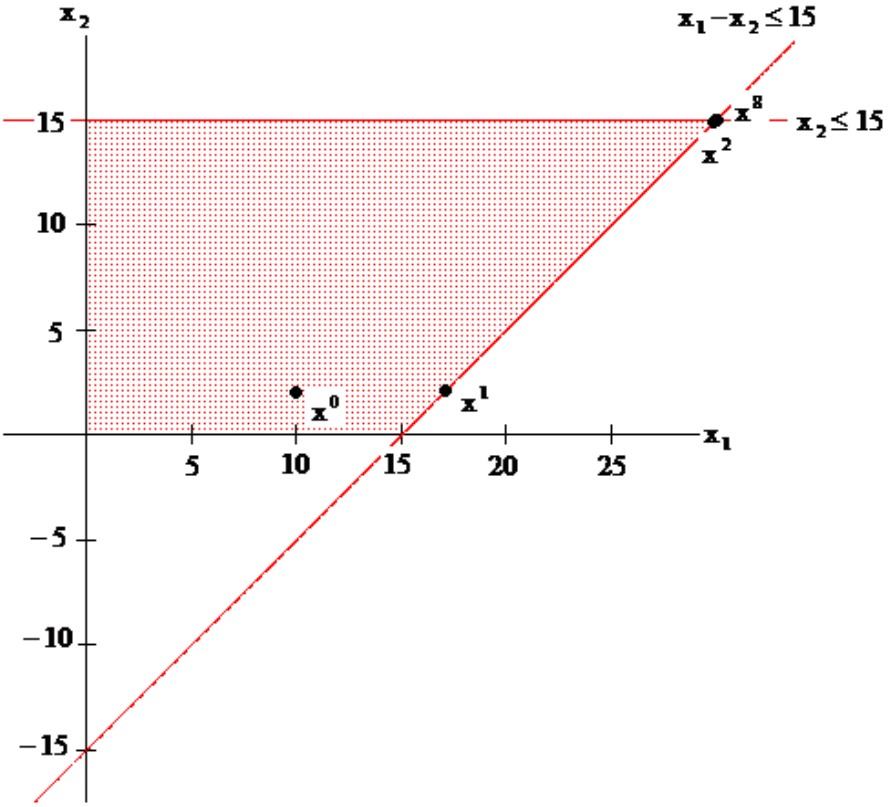

Figura 5: Soluções obtidas pelo algoritmo Primal-Afim

3.4.2 Método Dual-Afim

O problema de programação linear dual do problema (11) é:

$$\begin{aligned} &\text{maximizar } b^T w \\ &\text{sujeito a } A^T w + s = c \\ &s \geq 0 \text{ e w : irrestrito} \end{aligned} \tag{22}$$

O algoritmo Dual-Afim inicia com uma solução factível e determina medidas de passo em direções que garantem a otimização progressiva da função objetivo mantendo a factibilidade dual do problema durante o processo. O problema (22) contém variáveis irrestritas $w \in \Re^m$ e não negativas $s \in \Re^n$. Neste caso, $(w; s)$ é definida como uma solução interior factível se $A^T w + s = c$ e $s \geq 0$. Note que para as variáveis w, não tem sentido 'centralizar' uma vez que ela são irrestritas, mas as variáveis s, podem ser tratadas como as variáveis x do problema primal.

Idéias básicas do algoritmo Dual-Afim

O algoritmo Dual-Afim consiste de três procedimentos:

1. iniciar o algoritmo com uma solução interior dual factível;
2. mover a solução interior , melhorando o valor da função objetivo;
3. parar o processo quando obter uma solução ótima do problema (22).

Considerando a k-ésima iteração, tem-se uma solução interior dual $(w^k; s^k)$ tal que $A^T w^k + s^k = c$ e $s^k \geq 0$. O objetivo é encontrar uma boa direção de busca $(d_w^k; d_s^k)$ que, juntamente com um comprimento de passo $\beta_k > 0$ adequado, admita uma solução nova $(w^{k+1}; s^{k+1})$ definida por:

$$w^{k+1} = w^k + \beta_k d_w^k \tag{23}$$

$$s^{k+1} = s^k + \beta_k d_s^k \tag{24}$$

Satisfazendo:

$$A\, w^{k+1} + s^{k+1} = c \tag{25}$$

$$s^{k+1} > 0 \tag{26}$$

$$b^T w^{k+1} \geq b^T w^k \tag{27}$$

Substituindo (23) e (24) em (25), tem-se uma condição para a direção de movimento:

$$A\, d_w^k + d_s^k = 0 \tag{28}$$

Substituindo (23) em (27), com o objetivo de melhorar o valor da função objetivo, tem-se assim uma outra condição para a direção de movimento:

$$b^T d_w^k \geq 0 \tag{29}$$

Para garantir a expressão (26), um processo análogo ao método afim é aplicado. A idéia básica é 'recentralizar' s^k em $e = (1, 1, ..., 1)^T \in \Re^n$, no espaço transformado de tal forma que a distância de cada face positiva seja conhecida. Assim, define-se uma matriz mudança de escala $S_k = diag(s^k)$, também diagonal, com s_i^k como seu i-ésimo elemento na diagonal, que admite inversa S_k^{-1}. Desta forma $S_k^{-1} s^k = e$ e a cada variável s é transformada em uma nova variável $u \geq 0$ tal que:

$$u = S_k^{-1} s \tag{30}$$

$$s = S_k u \tag{31}$$

Além disso, d_u^k, que é uma direção que melhora a função objetivo no espaço transformado, tem uma direção correspondente no espaço original dada por:

$$d_s^k = S_k d_u^k \tag{32}$$

Pode-se estudar as iterações do algoritmo Dual-Afim no espaço transformado. Para determinar uma boa direção de movimento no espaço transformado determina-se que.

$$A^T d_w^k + d_s^k = 0 \text{ que implica em } A^T d_w^k + S_k d_u^k = 0$$

Multiplicando por S_k^{-1}, obte-se: $S_k^{-1} A^T d_w^k = -d_u^k$ e depois multiplicando ambos os lados por AS_k^{-1} tem-se:

$$AS_k^{-2} A^T d_w^k = -AS_k^{-1} d_u^k$$

Como a matriz A tem posto igual a m, existe $(AS_k^{-2}A)^{-1}$

$$d_w^k = -(AS_k^{-2} A^T)^{-1} AS_k^{-1} d_u^k \tag{33}$$

Definindo $Q_k = (AS_k^{-2}A^T)^{-1}AS_k^{-1}$, a equação (33) torna-se:

$$d_w^k = -Q_k d_u^k \tag{34}$$

Na equação acima vê-se que d_w^k é determinada por d_u^k. Encontrando uma direção apropriada d_u^k tal que (29) seja satisfeita, então o objetivo será conquistado. Para fazer isso, precisa-se definir:

$$d_u^k = -Q_k^T b \tag{35}$$

então: $b^T d_w^k = -b^T Q_k d_u^k = b^T Q_k Q_k^T b = \|b^T Q_k\|^2 \geq 0$

Combinando (35) e (34), tem-se:

$$d_w^k = (AS_k^{-2}A^T)^{-1}b \tag{36}$$

Conseqüentemente, da equação (28), tem-se a seguinte direção de movimento no espaço original:

$$d_s^k = -A^T d_w^k = -A^T(AS_k^{-2}A^T)^{-1}b \tag{37}$$

Dada a direção de movimento $(d_w^k; d_s^k)$ o comprimento do passo β_k é definido pela condição de não-negatividade de s^{k+1}, assim da mesma forma que no algoritmo Primal-Afim têm-se as seguintes condições:

1. Se $d_s^k = 0$, então o problema dual tem um valor objetivo constante no domínio factível e $(w^k; s^k)$ é ótimo dual;

2. Se $d_s^k > 0$, mas não pode ser o vetor nulo, então o problema (22) é ilimitado ;

3. Caso contrário:

$$\beta_k = \min_i \left\{ \frac{\alpha s_i^k}{-(d_s^k)_i} / (d_s^k)_i < 0 \right\} \quad onde \;\; 0 < \alpha < 1$$

Determina-se o vetor estimativa do algoritmo afim, como:

$$x^k = -S_k^{-2} d_s^k \tag{38}$$

então $Ax^k = AS_k^{-2}A^T d_u^k = b$. Portanto, x^k pode ser visto como "o vetor estimativa primal"para algoritmo Dual-Afim. Uma vez calculado o vetor estimativa dual

$x^k \geq 0$, determina-se uma solução factível primal com uma folga complementar que é definida por $c^T x^k - b^T w^k$. Além disso, se $c^T x^k - b^T w^k = 0$, então $(w^k; s^k)$ é ótimo dual e x^k ótimo primal. Esta informação pode ser usada para definir o critério de parada para o algoritmo Dual-Afim.

Algoritmo Dual-Afim

Baseado nas idéias básicas discutidas anteriormente define-se a seguir um processo iterativo para o algoritmo Dual-Afim.

Passo 1 (inicialização): Considerar $k = 0$ e encontrar uma solução inicial $(w^0; s^0)$ tal que $Aw^0 + s^0 = b$ e $s^0 > 0$ (solução interior).

Passo 2 (direção de translação): Calcular:

$$d_w^k = (AS_k^{-2}A^T)^{-1}b \text{ e } d_s^k = -A^T d_w^k$$

onde S_k é a matriz diagonal cujos elementos são as componentes de s_k.

Para facilitação do cálculo da direção de movimento, ao invés de calcular a matriz inversa de $(AS_k^{-2}A^T)$, é resolvido o sistema linear $(AS_k^{-2}A^T)\ d_w^k = b$ pelo método de Cholesky.

Passo 3 (testando se o objetivo é ilimitado): Se $d_s = 0$, então PARE, $(w^k; s^k)$ é otimo dual. Se $d_s^k > 0$, então PARE, o problema dual é ilimitado. Caso contrário vá para o passo seguinte.

Passo 4 (vetor estimativa primal): Calcular o vetor estimativa dual dado por:

$$x^k = -S_k^{-2} d_s^k$$

Passo 5 (testando a otimalidade): Se $x^k \geq 0$ e $c^T x^k - b^T w^k \leq \varepsilon$, onde ε é uma pequena tolerância positiva, então PARE, $(w^k; s^k)$ é otimo dual e x^k é ótimo primal. Caso contrário, vá para o passo seguinte.

Passo 6 (comprimento do passo): Calcular o comprimento do passo.

$$\beta_k = \min_i \left\{ \frac{\alpha s_i^k}{-(d_s^k)_i} / (d_s^k)_i < 0 \right\} \quad onde \quad 0 < \alpha < 1$$

Passo 7 (determinação de uma nova solução): Atualizar

$$w^{k+1} = w^k + \beta_k d_w^k$$
$$s^{k+1} = s^k + \beta_k d_s^k$$

Faça $k \leftarrow k + 1$ e vá para o passo 2.

Para ilustração do algoritmo Dual-Afim será apresentado o seguinte exemplo.

Exemplo 02

Considera-se o problema do exemplo (21), o seu dual:

$$\begin{aligned} \text{maximizar } & 15w_1 + 15w_2 \\ \text{sujeito a } & w_1 + s_1 = -2 \\ & -w_1 + w_2 + s_2 = 1 \\ & w_1 + s_3 = 0 \\ & w_2 + s_4 = 0 \\ & s_1, s_2, s_3, s_4 \geq 0 \end{aligned} \tag{39}$$

A primeira iteração será apresentada para a ilustração do algoritmo Dual-Afim na resolução do problema (21).

Considerando $\alpha = 0.99$ e $\varepsilon = 10^{-3}$ e a solução inicial interior e factível como: $w^0 = \begin{bmatrix} -3 & -3 \end{bmatrix}^T$ e $s^0 = \begin{bmatrix} 1 & 1 & 3 & 3 \end{bmatrix}^T$. Assim, seguindo os passos do

algoritmo Dual-Afim, calcula-se a matriz diagonal $S_0 = diag(s^0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$.

As direções de busca são:

$$d_w^0 = \begin{bmatrix} 23.5321 & 34.6788 \end{bmatrix}^T \quad \text{e} \quad d_s^0 = \begin{bmatrix} -23.5321 & -11.1467 & -23.5321 & -34.6788 \end{bmatrix}^T$$

Analisando os valores obtidos no cálculo da direção d_s^0, pode-se verificar que o problema não é ilimitado, então é calculado o vetor estimativa primal:

$$x^0 = -S_0^{-2} d_s^0 = \begin{bmatrix} 23.5321 \\ 11.1467 \\ 2.6146 \\ 3.8532 \end{bmatrix}$$

Checando a otimalidade, percebe-se que $x^0 \geq 0$ e a folga complementar $c^T x_0 - b^T w^0 = 54.0825 > \varepsilon$, portanto a solução não é ótima, então calcula-se o comprimento do passo: $\beta_0 = 0.0420$.

A nova solução e o valor da função objetivo são:

$$w^1 = w^0 + \beta_0 d_w^0 = \begin{bmatrix} -2.0100 \\ -1.5410 \end{bmatrix} \quad e \quad s^1 = s^0 + \beta_0 d_s^0 = \begin{bmatrix} 0.0100 \\ 0.5310 \\ 2.010 \\ 1.5410 \end{bmatrix}$$

Como a solução encontrada ainda não é a ótima, mais iterações serão realizadas. O resultado obtido pelo algoritmo Dual-Afim pode ser visto na tabela 2.

iteração(k)	w_i^k	x_i^k	função objetivo	folga complementar
1	-2.0100 -1.5410	28.4060 13.4067 0.0007 1.5932	-53.2657	9.8604
2	-2.0096 -1.0149	29.9962 14.9969 0.0006 0.0030	-45.3678	0.3722
3	-2.0000 -1.0039	29.9997 14.9997 0.0000 0.0002	-45.0607	0.0609
4	-2.0000 -1.0001	29.9999 14.9999 0.0000 0.0000	-45.0033	0.0033
5	-2.0000 -1.0000	29.9999 14.9999 0.0000 0.0000	-45.0004	0.0004

Tabela 2: Resultados obtidos pelo algoritmo Dual-Afim

A solução ótima encontrada na 5ª iteração é: $w_1^5 = -2.0000$ e $w_2^5 = -1.0000$ e o valor da função objetivo é: $z^5 = -45.0004$. A solução ótima para o problema original é: $x_1^5 = 29.9999$ e $x_2^5 = 14.9999$

3.4.3 Método Primal-Dual

O método Primal-Dual de pontos interiores é baseado na abordagem da função barreira logarítmica. A idéia de usar o método de barreira logarítmica para problemas de programação convexa foi proposto por K. R. Frisch em 1955. Em 1984, depois do algoritmo de Karmarkar ser introduzido, o método de barreira logarítmica foi reconsiderado para resolver problemas de programação linear. Gill et al. (1986) usaram este método para desenvolver a projeção de barreira logarítmica de Newton e demonstraram uma equivalência com a escala projetiva de Karmarkar. Megiddo (1989) providenciou uma análise teórica para o método de barreira logarítmica e propôs uma estrutura para o algoritmo Primal-Dual. Com base nesses conhecimentos, foi iniciado o estudo do método Primal-Dual de pontos interiores para problemas de programação linear.

Idéias básicas do algoritmo Primal-Dual

Considerando o problema linear:

$$\begin{aligned} &\text{minimizar } c^T x \\ &\text{sujeito a } Ax = b \\ &x \geq 0 \end{aligned} \tag{40}$$

e o seu dual:

$$\begin{aligned} &\text{maximizar } b^T w \\ &\text{sujeito a } A^T w + s = c \\ &s \geq 0 \text{ e w : irrestrito} \end{aligned} \tag{41}$$

Tem-se as seguintes hipóteses para o problema Primal-Dual:

(A1) O conjunto $S \equiv \{x \in \Re^n / Ax = b, x > 0\}$ é não vazio.
(A2) O conjunto $T \equiv \{(w; s) \in \Re^m \times \Re^n / A^T w + s = c, s > 0\}$ é não vazio.
(A3) A matriz de restrições A tem posto completo m.

Sob estas hipóteses, a partir do teorema da dualidade os problemas (40) e (41) tem as mesmas soluções ótimas com um valor comum. Além disso, os conjuntos das soluções ótimas de (40) e (41) são limitados. Note que, para $x > 0$ em (40), aplica-se a técnica da função barreira logarítmica e considera-se a seguinte família de problemas de programação não linear (P_μ).

$$\begin{aligned} \text{minimizar} \quad & c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{sujeito a} \quad & Ax = b \\ & x > 0 \end{aligned} \tag{42}$$

onde $\mu > 0$ é um parâmetro de penalidade ou barreira.

Quando $\mu \longrightarrow 0$, espera-se que as soluções ótimas do problema (42) convirjam para a solução ótima do problema original de programação linear (40). Com a finalidade de provar isso, observe que a função objetivo do problema (42) é uma função estritamente convexa, portanto sabe-se que tem pelo menos um mínimo global. A teoria de programação convexa implica que, se existir um mínimo global, ele está caracterizado pelas condições de Kuhn - Tucker, descritas abaixo:

$$Ax = b, \;\; x \geq 0 \qquad (factibilidade\, primal) \tag{43}$$

$$A^T w + s = c, \;\; s \geq 0 \quad (factibilidade\, dual) \tag{44}$$

$$XSe - \mu e = 0 \quad (folga\, complementar) \tag{45}$$

onde: X e S são matrizes diagonais com, respectivamente , x_i^k e s_i^k como seu i-ésimo elemento na diagonal.

Sob as hipóteses (A1) e (A2) e assumindo que o problema (40) tenha uma região factível limitada, o problema (42) é factível e assume pelo menos um ponto de mínimo $x(\mu)$, para algum $\mu > 0$. Conseqüentemente, o sistema formado pelas equações (43), (44) e (45) tem uma solução finita. Assim tem-se o seguinte lema:

Lema 01. Sob as hipóteses (A1) e (A2), ambos os problemas (42) e o sistema de equações (43), (44) e (45) tem pelo menos uma solução finita.

Esse sistema fornece as condições necessárias e suficientes para que $(w(\mu); s(\mu))$ maximize o seguinte problema $(D(\mu))$:

$$\begin{aligned} \text{maximizar} \quad & b^T w + \mu \sum_{j=1}^{n} \ln s_j \\ \text{sujeito a} \quad & A^T w + s = c \\ & s \geq 0 \quad w : irrestrito \end{aligned} \tag{46}$$

Note que a equação (45) do sistema pode ser escrita com as componentes de X e S.

$$x_j s_j = \mu \quad para\ j = 1, 2, ..., n \tag{47}$$

Portanto, quando a hipótese (A3) é imposta, x unicamente determina w a partir das equações (47) e (44), denotando assim uma solução finita $(x(\mu); w(\mu); s(\mu))$ para o sistema de equações para algum $\mu > 0$. Obviamente, $x(\mu) \in S$ e $(w(\mu); s(\mu)) \in T$. Entretanto, a folga complementar (ou gap dual) torna-se:

$$\begin{aligned} g(\mu) &= c^T x(\mu) - b^T w(\mu) \\ &= (c^T - w(\mu)^T A)\ x(\mu) \\ &= s(\mu)^T x(\mu) = n\mu \end{aligned}$$

Portanto, quando $\mu \to 0$, a folga complementar $g(\mu)$ converge para zero. Isto implica que $x(\mu)$ e $w(\mu)$ realmente convergem para a solução ótima dos problemas (40) e (41) respectivamente. Assim tem-se o seguinte lema:

Lema 02. Sob as hipóteses A1, A2 e A3, quando $\mu \to 0$, $x(\mu)$ converge para a solução ótima do problema (40) e $(w(\mu); s(\mu))$ converge para a solução ótima do problema (41).

Para $\mu > 0$, denota-se Γ a curva, ou caminho, que consiste na solução do sistema formado pelas condições de Kuhn - Tucker, isto é:

$$\Gamma = \{(x(\mu); w(\mu); s(\mu))/x(\mu); w(\mu); s(\mu) \text{ seja solução do sistema para } \mu > 0\} \tag{48}$$

Quando $\mu \to 0$, a curva Γ nos leva para um par de soluções: a solução primal x^* e a dual $(w^*; s^*)$. Deste modo, a trajetória descrita pela curva Γ serve para definir uma classe de métodos Primal-Dual de pontos interiores para programação linear. Por essa razão, o método Primal-Dual é classificado como um método de trajetória central.

Dado um ponto inicial $(x^0; w^0; s^0) \in S \times T$, o algoritmo Primal-Dual gera uma seqüência de pontos $((x^k; w^k; s^k) \in S \times T)$ e desloca-se em uma direção escolhida apropriadamente $(d_x^k; d_w^k; d_s^k)$ com um comprimento de passo β_k para cada iteração. Para calcular o desvio da curva em cada $(x^k; w^k; s^k)$, introduz-se as seguintes notações para $k = 0, 1, 2, ...$

$$\phi_i^k = x_i^k s_i^k \quad i = 1, 2, ...n \tag{49}$$

$$\phi_{\text{ave}}^k = \frac{1}{n} \sum_{i=1}^{n} \phi_i^k \tag{50}$$

$$\phi_{\min}^k = \min \ \{\phi_i^k, \ \ para \ \ i = 1, 2, ..., n\} \tag{51}$$

$$\theta^k = \frac{\phi_{\text{ave}}^k}{\phi_{\min}^k} \ \ \text{(normalização)} \tag{52}$$

Pode ser visto que $\theta^k \geq 1$ e $(x^k; w^k; s^k) \in \Gamma$ se e somente se $\theta^k = 1$. Quando o desvio θ^0, no ponto inicial $(x^0; w^0; s^0) \in S \times T$ for grande, o algoritmo do método Primal-Dual reduz não somente o gap dual, mas também o desvio. Com parâmetros devidamente

escolhidos, a seqüência de pontos $(x^k; w^k; s^k) \in S \times T$ gerados pelo algoritmo Primal-Dual satisfaz as seguintes desigualdades:

$$c^T x^{k+1} - b^T w^{k+1} = \left(1 - \frac{2}{n\theta^k}\right)(c^T x^k - b^T w^k) \tag{53}$$

$$\theta^{k+1} - 2 \leq \left(1 - \frac{1}{(n+1)}\right)(\theta^k - 2) \quad se\ \theta^k > 2 \tag{54}$$

$$\theta^{k+1} \leq 2 \quad se\ \theta^k \leq 2 \tag{55}$$

A expressão (53) assegura que o *gap* dual diminui monotonicamente. Nas outras duas desigualdades, o desvio θ^k é menor que 3 em, no máximo, $O(n \log \theta^0)$ iterações, e então, o *gap* dual converge para zero linearmente com a taxa de convergência com pelo menos $(1 - 2/(3n))$, ou seja, a convergência ocorre em tempo polinomial.

Seguem algumas etapas importantes para o algoritmo Primal-Dual.

Direção de Newton

O método de Newton é um dos métodos mais usados para encontrar as soluções de sistemas de equações não lineares, utilizando sucessivas aproximações do sistema por equações lineares. Sendo $F(z)$ uma aplicação não linear de $\Re^n$ para $\Re^n$ precisa-se encontrar uma solução ótima ($z^* \in \Re^n$) tal que $F(z^*) = 0$. Usando a expansão em série de Taylor (para $z = z^k$), obtem-se a seguinte aproximação linear:

$$F(z^k + \Delta z) \approx F(z^k) + J(z^k)\Delta z \tag{56}$$

onde $J(z^k)$ é a matriz jacobiana, seu (i,j)-ésimo elemento é dado por:

$$\left[\frac{\delta F_i(z)}{\delta z_j}\right]_{z=z^k} \tag{57}$$

e Δz é o vetor de translação. O lado esquerdo de (56) estima uma solução para $F(z) = 0$, tendo assim um sistema linear:

$$J(z^k)\Delta z = -F(z^k) \tag{58}$$

Um vetor solução para a equação (58) prevê uma iteração de Newton de z^k para z^{k+1}, com direção de Newton e um comprimento de passo unitário. Quando $J(z^*)$ é não singular e o ponto inicial z^0 "está muito perto"de z^*, o método de Newton converge quadraticamente para z^*. Mas esta taxa de convergência é apenas um comportamento "local". Geralmente para uma aplicação não-linear, se z^0 não for próximo o suficiente de z^*, a iteração de Newton pode divergir.

Nas condições de Kuhn - Tucker, assumindo um ponto $(x^k; w^k; s^k)$ para algum $\mu^k > 0$, tal que x^k e s^k sejam positivos, a direção de Newton $(d_x^k; d_w^k; d_s^k)$ é determinada pelo seguinte sistema de equações lineares:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S_k & 0 & X_k \end{bmatrix} \begin{bmatrix} d_x^k \\ d_w^k \\ d_s^k \end{bmatrix} = - \begin{bmatrix} A\,x^k - b \\ c\,A^T w^k + s^k - c \\ X_k S_k\,e - \mu^k\,e \end{bmatrix} \tag{59}$$

onde X_k e S_k são matrizes diagonais formadas respectivamente por x_i^k e s_i^k, como seus elementos diagonais.

Para simplificação do sistema acima usa-se:

$$t^k = A d_x^k \tag{60}$$

$$u^k = A^T d_w^k + d_s^k \tag{61}$$

$$v^k = S_k\,d_x^k + X_k\,d_s^k \tag{62}$$

em que t^k, u^k e v^k são denominados os resíduos primal, dual e da folga complementar, respectivamente, sendo definidos por:

$$t^k = b - Ax^k,\ u^k = c - A^T w^k - s^k \text{ e } v^k = \mu^k e - X_k S_k e.$$

Note que se $x^k \in S$ e $(w^k; s^k) \in T$, correspondentemente $t^k = 0$ e $u^k = 0$.

Para resolver as equações (59), multiplica-se ambos os lados da equação (61) por $AX_kS_k^{-1}$:

$$AX_kS_k^{-1}A^T d_w^k = AX_kS_k^{-1}u^k - AX_kS_k^{-1}d_s^k \tag{63}$$

Isolando d_s^k na equação (62):

$$d_s^k = X_k^{-1} v^k - X_k^{-1} S_k d_x^k \tag{64}$$

Nominando-se $p^k = X_k^{-1} v^k$. Substituindo (60) em (64), tem-se

$$d_s^k = p^k - X_k^{-1} S_k d_x^k$$

multiplicando $AX_k S_k^{-1}$ em ambos os lados, fica

$$\begin{aligned} AX_k S_k^{-1} d_s^k &= AX_k S_k^{-1} p^k - AX_k S_k^{-1} X_k^{-1} S_k d_x^k \\ AX_k S_k^{-1} d_s^k &= AX_k S_k^{-1} p^k - A d_x^k \\ AX_k S_k^{-1} d_s^k &= AX_k S_k^{-1} p^k - t^k \end{aligned} \tag{65}$$

e (63) em (65), tem-se:

$$\begin{aligned} AX_k S_k^{-1} A^T d_w^k &= AX_k S_k^{-1} u^k - AX_k S_k^{-1} p^k + t^k \\ AX_k S_k^{-1} A^T d_w^k &= AX_k S_k^{-1} (u^k - p^k) + t^k \\ d_w^k &= [AX_k S_k^{-1} A^T]^{-1} (AX_k S_k^{-1} (u^k - p^k) + t^k) \end{aligned} \tag{66}$$

onde $X_k S_k^{-1}$ é uma matriz diagonal definida positiva.

Uma vez que d_w^k é obtido, d_s^k e d_x^k podem ser facilmente calculados por:

$$d_s^k = u^k - A^T d_w^k \tag{67}$$

$$d_x^k = X_k S_k^{-1} [p^k - d_s^k] \tag{68}$$

Novamente, para $(x^k; w^k; s^k) \in S \times T$, as equações (66)-(68) são simplificadas por:

$$d_w^k = -[AD_k^2 A^T]^{-1} A S_k^{-1} v^k \tag{69}$$

$$d_s^k = -A^T d_w^k \tag{70}$$

$$d_x^k = S_k^{-1} [v^k - X_k d_s^k] \tag{71}$$

onde $D_k^2 = X_k S_k^{-1}$ e $D_k = diag(\sqrt{x_i^k / s_i^k})$.

É importante notar que d_w^k, d_s^k e d_x^k em (69), (70) e (71) estão relacionados. Denotando o vetor:

$$r^k(\mu) = X_k^{-1} D_k v^k(\mu) \tag{72}$$

e a matriz:

$$Q = D_k A^T (A D_k^2 A^T)^{-1} A D_k \tag{73}$$

onde Q é a matriz de projeção ortogonal no espaço coluna da matriz $D_k A^T$.

Assim as direções de busca d_w^k, d_s^k e d_x^k podem ser reescritas como:

$$d_w^k = D_k (I - Q) r^k(\mu) \tag{74}$$

$$d_s^k = -(A D_k^2 A^T) A D_k r^k(\mu) \tag{75}$$

$$d_x^k = D_k^{-1} Q r^k(\mu) \tag{76}$$

Após ter obtido uma direção de Newton para a k-ésima iteração, o algoritmo Primal-Dual determina um novo ponto de acordo com a equação seguinte:

$$x^{k+1} = x^k + \beta^k d_x^k$$
$$w^{k+1} = w^k + \beta^k d_w^k$$
$$s^{k+1} = s^k + \beta^k d_s^k$$

com um comprimento do passo β^k adequadamente escolhido na k-ésima iteração onde $x^{k+1} \in S$ e $(w^{k+1}; s^{k+1}) \in T$.

Comprimento do passo e parâmetro de penalidade

Como $(x^k; w^k; s^k) \in S \times T$ o algoritmo Primal-Dual necessita de dois parâmetros σ e τ, tais que $0 \leq \tau < \sigma < 1$, para controlarem, respectivamente, o parâmetro de penalidade μ^k e o comprimento do passo β^k na k-ésima iteração. Para o parâmetro de penalidade, relembrando as notações definidas em (49)-(52), para reduzir a folga complementar $(n\,\phi^k_{\text{ave}})$, escolhe-se um parâmetro de penalidade como:

$$\mu^k = \sigma\,\phi^k_{\text{ave}} \tag{77}$$

desta forma, as definições (60)-(62) implicam que $v^k \leq 0$.

Quanto ao comprimento do passo β^k, a escolha está relacionada com a folga complementar. Note que as equações (60)-(62) implicam $x^k_i d^k_{s_i} + s^k_i d^k_{x_i} = \mu^k - \phi^k_i$. Assim a folga complementar varia de forma quadrática em função do comprimento do passo β uma vez que:

$$\phi^k_i(\beta) \equiv (x^k_i + \beta d^k_{x_i})(s^k_i + \beta d^k_{s_i}) = \phi^k_i + \beta(u^k - \phi^k_i) + \beta^2(d^k_{x_i} d^k_{s_i}) \;\; i = 1, 2, ..., n \tag{78}$$

Além disso, uma vez que $((d^k_x)^T d^k_s) \to 0$ desconsiderando este termo e calculando a folga complementar média, tem-se que a folga complementar muda de forma linear em β, isto é:

$$\phi^k_{\text{ave}}(\beta) = (x^k + \beta d^k_x)^T (s^k + \beta d^k_s)/n = \phi^k_{\text{ave}} + \beta(u^k - \phi^k_{\text{ave}}) \tag{79}$$

Desconsiderando o termo quadrático em (78), diminuindo o valor $u^k = \sigma\phi_{\text{ave}}$ por um fator $\tau < \sigma$, pode-se definir uma função linear:

$$\psi^k(\beta) = \phi^k_{\min} + \beta(\tau\phi^k_{\text{ave}} - \phi^k_{\min}) \tag{80}$$

A função $\phi^k_i(\beta)$ pode ser convexa ou côncava dependendo do valor do produto de $d^k_{x_i} d^k_{s_i}$. Quando $d^k_{x_i} d^k_{s_i} \geq 0$, $\phi^k_i(\beta)$ é convexa e portanto a curva ϕ^k_i está acima da curva $\psi^k(\beta)$ para $0 \leq \beta \leq 1$. Quando $d^k_{x_i} d^k_{s_i} < 0$, $\phi^k_i(\beta)$ é côncava e pode cruzar com $\psi^k(\beta)$, como mostra a figura 6. A fim de controlar o parâmetro de desvio θ^k, reduzindo a folga complementar, escolhe-se:

$$\alpha^k = \max\{\overline{\beta}/\phi^k_i(\beta) \geq \psi^k(\beta),\;\; \textit{para todo}\;\; \beta \in (0, \overline{\beta}),\;\; \textit{onde}\;\; 0 < \overline{\beta} < 1\;\; e\;\; i = 1, 2, ...n\} \tag{81}$$

Então o comprimento do passo na k-ésima iteração é definida por:

$$\beta^k = \min\{1, \alpha^k\} \tag{82}$$

O significado geométrico de α^k e β^k é representado na figura 6. É claramente visível a partir da figura, que a escolha do β^k depende da escolha de $0 < \tau < 1$ para garantir a existência de α^k.

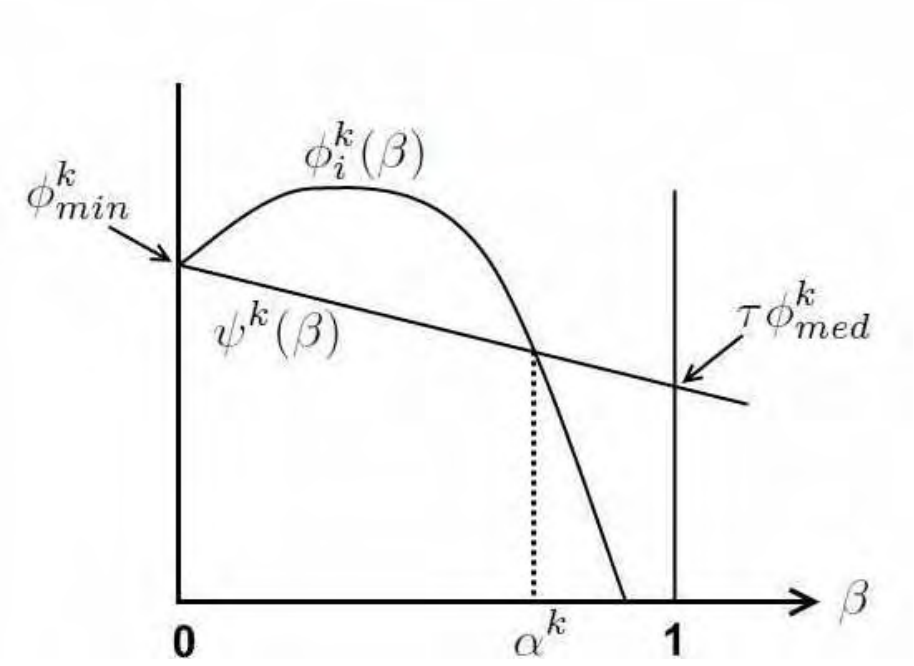

Figura 6: Significado geométrico do comprimento do passo para o método Primal-Dual - Fonte: Fang & Puthenpura (1993)

Quando $(x^k; w^k; s^k) \in S \times T$, desde que $(d_x^k; d_w^k; d_s^k)$ seja uma solução para (60)-(62) com $t^k = 0$ e $u^k = 0$, sabe-se que $Ax^{k+1} = b$ e $A^T w^{k+1} + s^{k+1} = c$. Além disso, a definição de α^k em (81) implica que $x^{k+1} > 0$ e $s^{k+1} > 0$. Em outras palavras, $(x^{k+1}; w^{k+1}; s^{k+1}) \in S \times T$.

3.4.4 Implementação do algoritmo Primal-Dual

A implementação do algoritmo Primal-Dual é uma tarefa difícil de se realizar devido a problemas numéricos e a escolha dos parâmetros de controle que alteram extremamente o desempenho do algoritmo. Muito esforço tem sido dedicado para uma versão prática de implemetação do algoritmo Primal-Dual. Sendo assim, uma versão do algoritmo Primal-Dual é apresentada, permitindo iniciá-lo com um ponto arbitrário $(x^0; w^0; s^0)$, com x^0, $s^0 > 0$.

Esta versão produz uma seqüência de iterações $(x^k; w^k; s^k)$, com x^k, $s^k > 0$, o que leva a uma solução ótima, embora as mesmas não permaneçam ao longo da curva $S \times T$.

Direção de movimento

Suponha que se esteja em um ponto $(x^k; s^k; w^k)$ para algum $\mu^k > 0$, tal que x^k, $s^k > 0$. A direção de Newton $(d_x^k; d_w^k; d_s^k)$ é determinada pelas equações (66), (67) e (68) e combinando-as tem-se:

$$d_x^k = \mu^k D_k P_k D_k X_k^{-1} e - D_k P_k D_k c + D_k^2 A^T (A D_k^2 A^T)^{-1} t^k \tag{83}$$

onde $D_k^2 = X_k S_k^{-1}$ e $P_k = I - D_k A^T (A D_k^2 A^T)^{-1} A D_k$, que é a matriz de projeção do espaço nulo da matriz AD_k. Definindo $d_{x,\mu}^k = \mu^k D_k P_k D_k X_k^{-1} e$, $d_{x,c}^k = -D_k P_k D_k c$ e $d_{x,t^k}^k = D_k^2 A^T (A D_k^2 A^T)^{-1} t^k$, então tem-se:

$$d_x^k = d_{x,\mu}^k + d_{x,c}^k + d_{x,t^k}^k \tag{84}$$

O termo de $d_{x,\mu}$ é chamado de *direção de centralização*, pois auxilia a projeção do vetor a se distanciar da fronteira da região factível primal. O termo $d_{x,c}$ é chamado de *direção redução objetiva*, pois a projeção do gradiente negativo da função objetivo transformada faz com que a função objetivo primal original seja reduzida. E o termo d_{x,t^k} é chamado de *direção de viabilidade*, uma vez que t^k é uma medida de viabilidade primal. Como $A d_{x,\mu}^k = 0$ e $A d_{x,c}^k = 0$, então estas duas direções pertencem ao espaço nulo da matriz A, bem como a factibilidade primal é influenciada apenas por d_{x,t^k}^k, o que justifica considerar o resíduo t^k na determinação da direção de busca d_x^k.

Na prática, iniciando-se com um ponto inicial arbitrário $(x^0; s^0; w^0)$ com x^0, $s^0 > 0$, o valor t^0 pode ser grande, já que x^0 pode ser infactível. Neste ponto, o esforço principal do algoritmo será em um ponto próximo possível da trajetória central. Uma vez que uma solução viável é encontrada (por exemplo, na k-ésima iteração), o algoritmo tentará manter $t^k = 0$ para todos os $k^{'} > k$, podendo perder a factibilidade devido a erros de truncamento numérico ou de arredondamento.

De forma semelhante, pode-se realizar a análise das direções de busca, d_w^k e d_s^k.

Comprimento do passo

Uma vez que a direção de movimento é obtida, pode-se avançar para um novo ponto $(x^{k+1}; w^{k+1}; s^{k+1})$ com $x^{k+1} > 0$ e $s^{k+1} > 0$. Para isso, considera-se:

$$x^{k+1} = x^k + \beta_P d_x^k \tag{85}$$

$$w^{k+1} = w^k + \beta_D d_w^k \tag{86}$$

$$s^{k+1} = s^k + \beta_D d_s^k \tag{87}$$

onde β_P e β_D são, respectivamente, o comprimento do passo nos espaços primal e dual. As exigências de não negatividade de x^{k+1} e s^{k+1} determinam a escolha do comprimento do passo β_P e β_D. Uma maneira simples de obtê-los é:

$$\beta_P = \min\left\{1, \frac{\alpha x_i^k}{-d_{x_i}^k} / (d_x^k)_i < 0\right\} \tag{88}$$

$$\beta_D = \min\left\{1, \frac{\alpha s_i^k}{-d_{s_i}^k} / (d_s^k)_i < 0\right\} \tag{89}$$

onde $\alpha < 1$, $(d_x^k)_i$ é a i-ésima componente de d_x^k, x_i^k é a i-ésima componente de x^k, $(d_s^k)_i$ é a i-ésima componente de d_s^k e s_i^k é a i-ésima componente de s^k.

Ajustando os parâmetros de penalidade e o critério de parada

Observe que a direção de busca na k-ésima iteração é determinada usando um valor fixo do parâmetro de penalidade μ^k, de forma que os passos de Newton convirjam para a trajetória central. Para que isso ocorra é necessário alguns passos para se aproximar

da trajetória central. A fim de satisfazer a folga complementar, μ^k deve tender a zero, consequentemente, em implementações práticas, o valor de μ^k é reduzido a cada iteração. Usa-se apenas um passo de Newton para cada valor dado de μ^k, mesmo que este valor permaceça exatamente na trajetória central.

A cada iteração, o próprio algoritmo sugere a redução do parâmetro μ^k. Da equação (45), temos que $\mu = s^T x/n$. Assim, na k-ésima iteração os valores de x^k e s^k nos dão uma medida razoavelmente boa para o parâmetro de penalidade. Um valor mais baixo de μ^k, por exemplo, $\sigma[(s^k)^T x^k]/n$ com $\sigma < 1$, pode acelerar a convergência do algoritmo. Algumas literatuas citam outras idéias semelhantes sobre essa escolha. No entanto, Fang & Puthenpura (1993) afirmam que a regra funciona bem para uma variedade de problemas práticos que foram resolvidos. Como os critérios de parada, podemos verificar a factibilidade primal, factibilidade dual e a folga complementar que são medidas respectivamente por t^k, u^k e v^k, tal como definido em (60)-(62).

Algoritmo Primal-Dual

Baseado nas idéias discutidas anteriormente apresenta-se a seguir um um procedimento passo-a-passo para a implementação do algoritmo Primal-Dual.

Passo 1 (inicialização): Considerar $k = 0$. Escolher uma solução inicial $(x^0; w^0; s^0)$ com $x^0 > 0$ e $s^0 > 0$, e ε_1, ε_2 e ε_3 números positivos suficientemente pequenos (tolerâncias).

Passo 2 (cálculos intermediários): Calcular:

$$\mu^k = \frac{(x^k)^T s^k}{n}$$

em que $t^k = b - Ax^k$, $u^k = c - A^T w^k - s^k$, $v^k = \mu^k e - X_k S_k e$, $p^k = X_k^{-1} v^k$, e $D_k^2 = X_k S_k^{-1}$, onde X_k e S_k são matrizes diagonais cujos elementos são, respectivamente, as componentes de x_i^k e s_i^k.

Passo 3 (checando a otimalidade): Se

$$\mu^k < \varepsilon_1, \quad \frac{\|t\|}{\|b\| + 1} < \varepsilon_2 \quad \text{e} \quad \frac{\|u\|}{\|c\| + 1} < \varepsilon_3$$

então PARE, a solução é otima. Caso contrário vá para o passo seguinte.

Passo 4 (calculando a direção de busca): Calcular:

$$d_w^k = (AD_k^2A^T)^{-1}(AD_k^2(u^k - p^k) + t^k)$$
$$d_s^k = u^k - A^T d_w^k$$
$$d_x^k = D_k^2(p^k - d_s^k)$$

Para facilitar os cálculos, ao invés de calcular a matriz inversa $(AD_k^2A^T)^{-1}$, é resolvido o sistema linear $(AD_k^2A^T)\ d_w^k = (AD_k^2(u^k - p^k) + t^k)$ pelo método de Cholesky.

Passo 5 (verificando se o objetivo é ilimitado): Se

$$t^k = 0,\ d_x^k > 0 \text{ e } c^T d_x^k < 0$$

então o problema primal (40) é ilimitado. Se

$$u^k = 0,\ d_s^k > 0 \text{ e } b^T d_w^k > 0$$

então o problema dual (41) é ilimitado. Se qualquer um desses casos acontecer, PARE. Caso contrário vá para o próximo passo.

Passo 6 (comprimento do passo): Calcular o comprimento do passo.

$$\beta_P = \min\left\{1, \frac{\alpha x_i^k}{-d_{x_i}^k} / (d_x^k)_i < 0\right\}$$
$$\beta_D = \min\left\{1, \frac{\alpha s_i^k}{-d_{s_i}^k} / (d_s^k)_i < 0\right\}$$

onde $\alpha < 1$

Passo 7 (determinação de uma nova solução): Atualize

$$x^{k+1} \leftarrow x^k + \beta_P d_x^k$$
$$w^{k+1} \leftarrow w^k + \beta_D d_w^k$$
$$s^{k+1} \leftarrow s^k + \beta_D d_s^k$$

Faça $k \leftarrow k + 1$ e vá para o passo 2.

Para ilustração do algoritmo Primal-Dual será considerado os problemas (21) e (39).

Exemplo 03

Considerando o problema linear:

$$\begin{aligned} \text{minimizar } & -2x_1 + x_2 \\ \text{sujeito a } & x_1 - x_2 + x_3 = 15 \\ & x_2 + x_4 = 15 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned} \tag{90}$$

e o seu dual

$$\begin{aligned} \text{maximizar } & 15w_1 + 15w_2 \\ \text{sujeito a } & w_1 + s_1 = -2 \\ & -w_1 + w_2 + s_2 = 1 \\ & w_1 + s_3 = 0 \\ & w_2 + s_4 = 0 \\ & s_1, s_2, s_3, s_4 \geq 0 \end{aligned} \tag{91}$$

A primeira iteração será apresentada para a ilustração do algoritmo Primal-Dual na resolução dos problemas (90) e (91).

Para acelerar a convergência do algoritmo, foi utilizado um "fator de aceleração de convergência" $(2 < \gamma < 10)$, diminuindo assim o parâmetro de penalidade, que será calculado como:

$$\mu^k = \frac{(x^k)^T s^k}{n\gamma}.$$

Considerando $\alpha = 0.99$, $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 10^{-3}$ e $\gamma = 10$ e o parâmetro de penalidade inicial é: $\mu^0 = 1$, as soluções iniciais são:

$$x^0 = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T, \quad w^0 = \begin{bmatrix} 0 & 0 \end{bmatrix}^T \quad \text{e} \quad s^0 = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T$$

Seguindo os passos do algoritmo Primal-Dual, tem-se:

$$t^0 = \begin{bmatrix} 14 & 13 \end{bmatrix}^T,\ u^0 = \begin{bmatrix} -3 & 0 & -1 & -1 \end{bmatrix}^T,\ v^0 = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^T \ e\ p^0 = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^T$$

e a matriz diagonal $D_0^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

Checando a otimalidade, tem-se:

$$\mu^k = 1,\ \frac{\|t\|}{\|b\|+1} = 14\ \ e\ \ \frac{\|u\|}{\|c\|+1} = 3 \tag{92}$$

então a solução não é ótima.

Sendo assim, calculam-se as direções de busca:

$$d_w^k = \begin{bmatrix} 6.40 \\ 9.1999 \end{bmatrix},\ d_s^k = \begin{bmatrix} -9.3999 \\ -2.80 \\ -7.4 \\ -10.1999 \end{bmatrix}\ e\ d_x^k = \begin{bmatrix} 9.3999 \\ 2.80 \\ 7.4 \\ 10.1999 \end{bmatrix}$$

Embora $d_x^k > 0$ e $c^T d_x^k = -15.9999$, o problema primal não é ilimitado devido ao fato de $t^0 \neq 0$, isto é, x^0 ser infactível. Sendo assim, prossegue-se os passos do algoritmo.

Calculando o comprimento do passo tem-se: $\beta_P = 1$ e $\beta_D = 0.097549$ e atualizando a solução tem-se:

$$x^{k+1} = \begin{bmatrix} 10.39999 \\ 3.80 \\ 8.3999 \\ 11.1999 \end{bmatrix},\ w^{k+1} = \begin{bmatrix} 0.624313 \\ 0.897450 \end{bmatrix}\ e\ s^{k+1} = \begin{bmatrix} 0.083039 \\ 0.726862 \\ 0.278137 \\ 0.0050 \end{bmatrix}$$

e o valor da função objetivo é: $z^0 = -16.9999$.

Como a solução encontrada ainda não é ótima, mais iterações serão realizadas. O resultado obtido pelo algoritmo Primal-Dual e a interpretação geométrica do problema dado, com a comparação do resultado do obtido na tabela 1 e 3, podem ser vistos na tabela 3 e na figura 7, respectivamente.

iteração(k)	w_i^k	x_i^k	função objetivo
1	0.624313 0.897450	10.3999 3.80 8.3999 11.1999	-16.9999
2	0.238641 0.739639	19.33224 4.37424 0.04200 10.62576	-34.290236
3	-0.507839 0.489666	26.28840 11.288603 0.00021 3.7114	-41.288183
4	-0.547468 0.452518	29.98122 14.98144 0.0002 0.01855	-44.980989
5	-2.0000 -1.0000	29.999678 14.9999 0.000229 0.000092	-44.999449

Tabela 3: Resultados obtidos pelo algoritmo Primal-Dual

A solução ótima encontrada na 5ª iteração é: $x_1^5 = 29.999678$ e $x_2^5 = 14.9999$ e o valor da função objetivo é: $z^5 = -44.999449$.

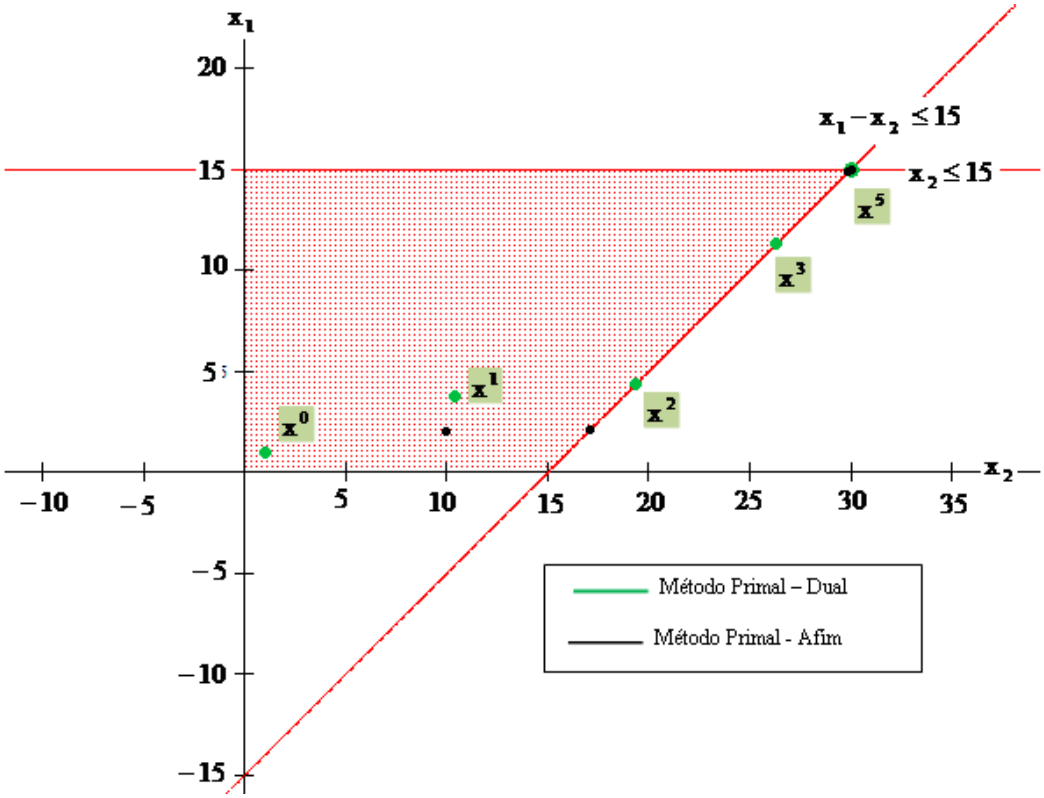

Figura 7: Comparação das soluções obtidas pelo algoritmo Primal-Afim e Primal-Dual

No próximo capítulo, o método Primal-Dual de pontos interiores é aplicado no problema de programação linear que auxilia no planejamento ótimo da radioterapia, para a verificação de sua viabilidade para a resolução deste.

4 APLICAÇÃO DA PROGRAMAÇÃO LINEAR NO PLANEJAMENTO ÓTIMO DA RADIOTERAPIA

Neste capítulo é apresentada uma aplicação da programação linear para resolução do modelo (1)-(8) que auxilia o planejamento ótimo da radioterapia.

Seguindo a metodologia descrita, inicialmente o câncer é diagnosticado e indicado o tratamento por radioterapia. Posteriormente, são realizadas imagens de tomografia computadorizada para dimensionamento do tumor e localização das regiões de interesse (tumorais, críticas e saudáveis). Assim, o médico determina a dose que será administrada no tumor (t_g). Os limites de dose no tumor são calculados por: $t_g \pm D$, onde $D = \varepsilon\, t_g$ ($t_g - D = l_t$ e $t_g + D = u_t$) e a dose máxima que as regiões crítica e saudável suportam (respectivamente u_c e u_g).

A imagem de tomografia é dividida em pixels e são obtidas as posições geométricas das localizações dos pixels referentes aos tecidos críticos, saudáveis e tumorais para serem utilizadas na construção da matriz de deposição de dose A. A partir destas informações o problema de programação linear (1)-(8) é resolvido, garantindo que o tumor receba o nível mínimo de radiação necessário para sua eliminação e que as regiões circunvizinhas sejam poupadas dos efeitos nocivos da radiação.

O fluxograma na figura 8 resume esta metodologia, baseado em Viana (2010).

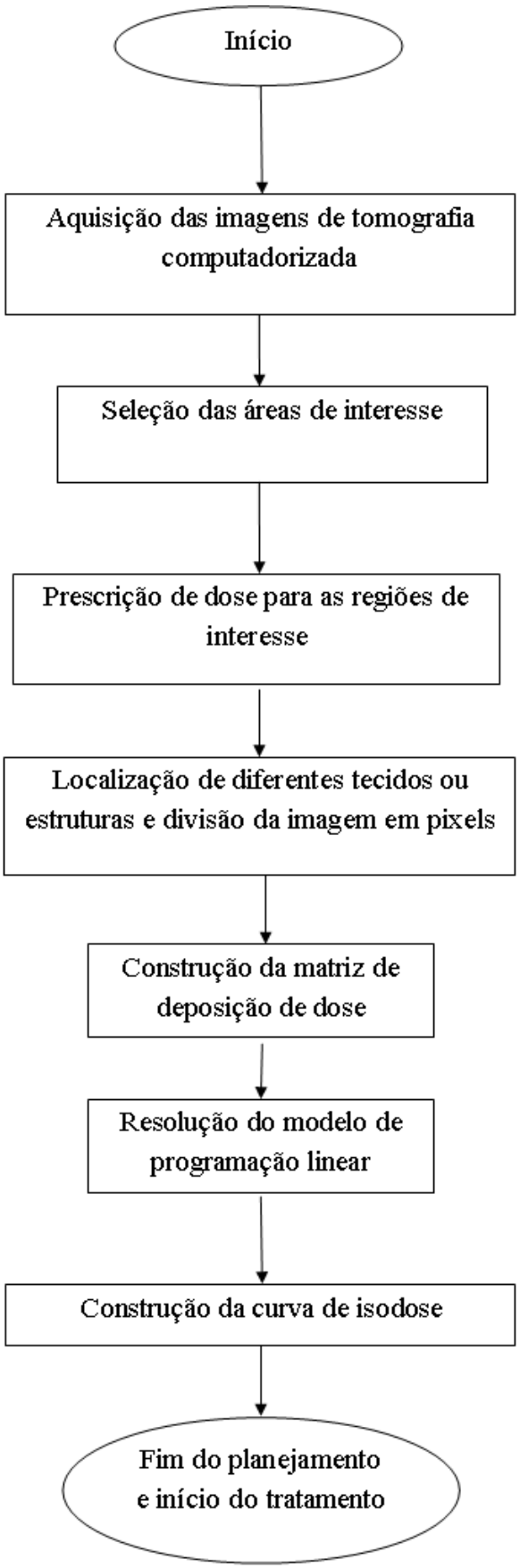

Figura 8: Fluxograma do planejamento otimizado considerando a localização das regiões de interesse

Para a aplicação desta metodologia foi utilizada uma imagem hipotética simulando um tumor no interior de uma região crítica conforme a figura 9, exemplificando um câncer de medula, que é um caso de difícil planejamento, pois o tumor está totalmente envolvido por uma estrutura crítica.

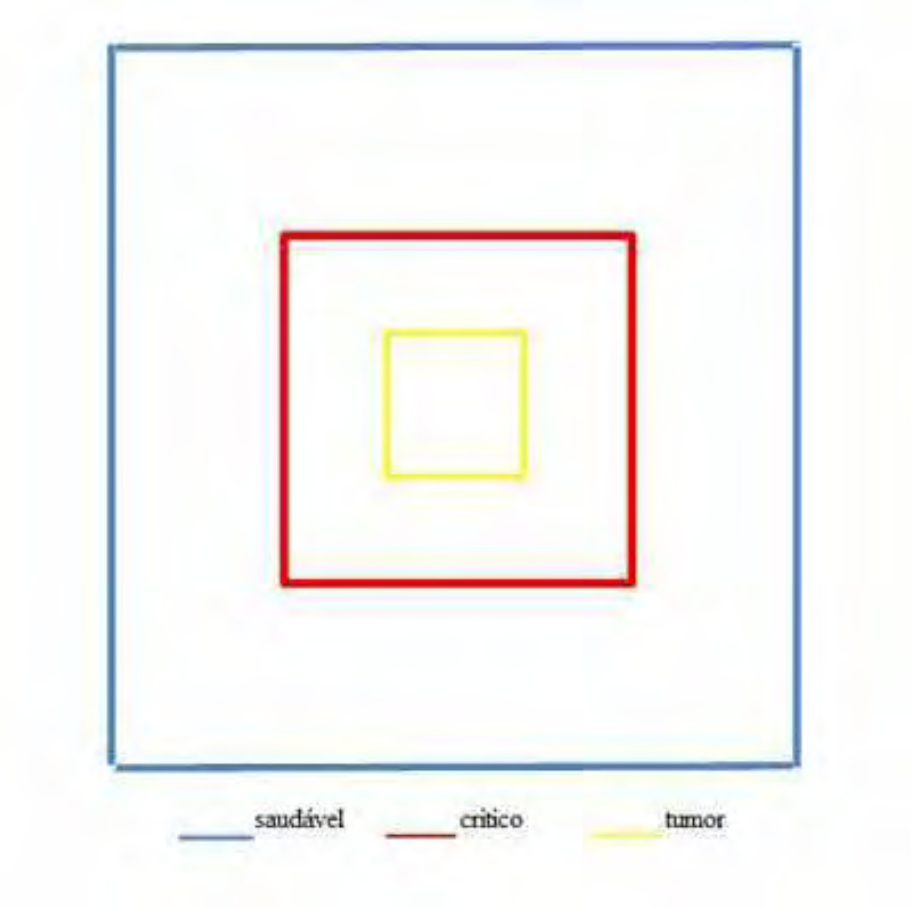

Figura 9: Imagem hipotética, simulando um tumor interior a uma região crítica

As dimensões do tecido saudável, tecido crítico e tumoral são considerados, respectivamente, 10×10 cm, 5×5 cm e 2×2 cm. A figura 10 apresenta a divisão da imagem em pixels.

Foram considerados quatro feixes incidentes ortogonalmente à face de cada lado da imagem, direcionados ao centro do quadrado mais interno, o qual representa o tumor. A largura do feixe incidente é de 6cm. A imagem foi subdividida em 100 pixels e estes foram enumerados conforme a figura 10.

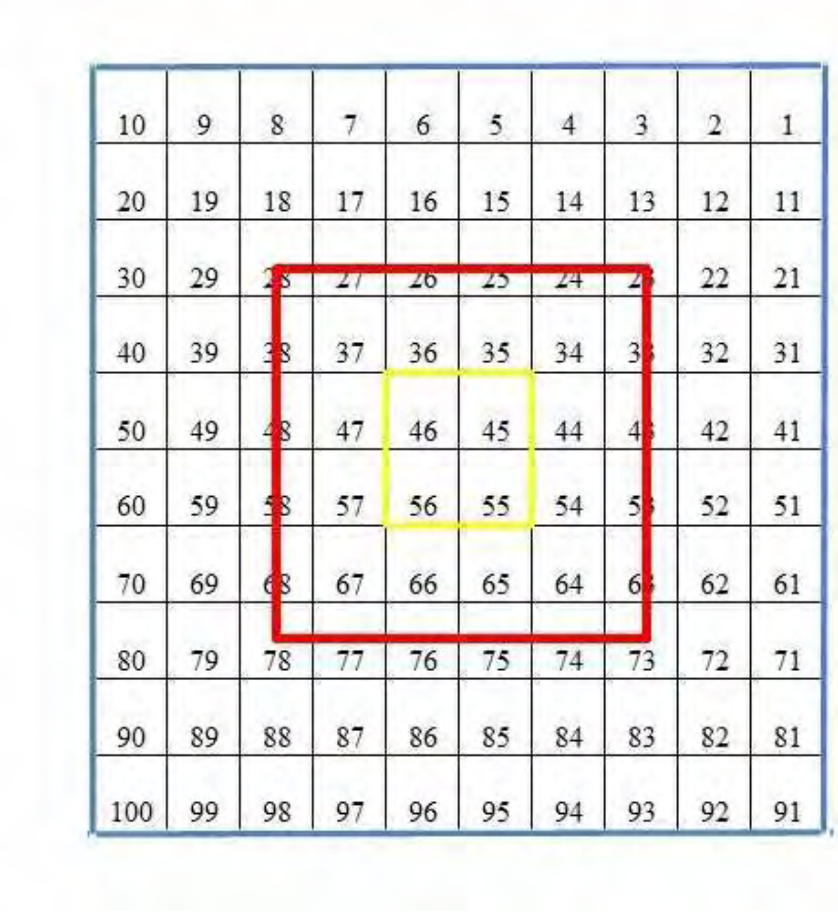

Figura 10: Divisão por pixels da imagem hipotética

A quantidade de pixels de tecidos tumoral, crítico e saudável são, respectivamente 4, 32 e 48. Conforme mostra a figura 11, existem alguns pixels que não vão receber radiação, chamados de pixels frios.

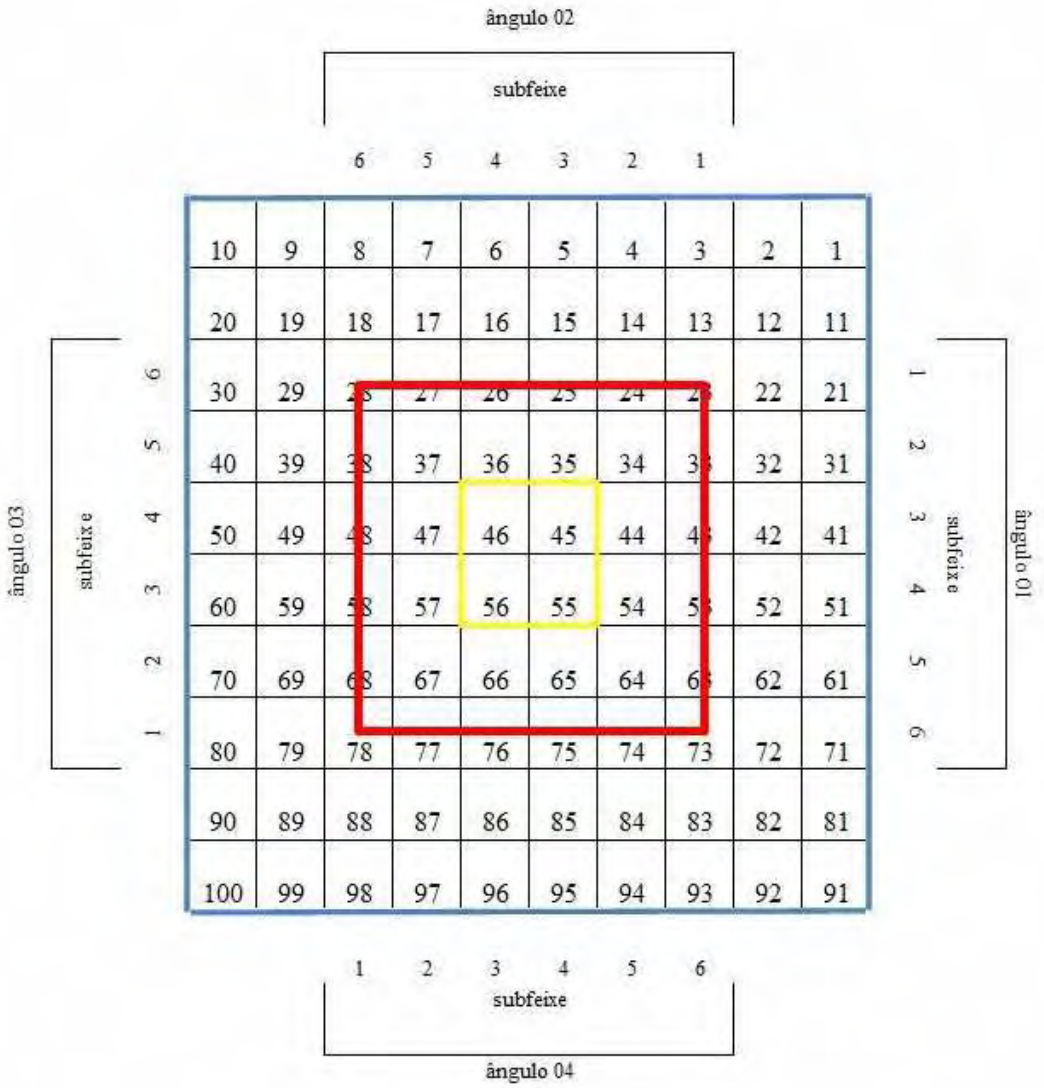

Figura 11: Representação da imagem simulada com os subfeixes de radiação

Considerando que o plano de tratamento não pode permitir que a estrutura crítica receba mais do que 40 Gy, onde Gy representa a quantidade de energia da radiação ionizante absorvida (ou dose) por unidade de massa ($1Gray(Gy) = 1J/kg$). Foi prescrito para o tumor uma dose de $t_g = 80$ Gy, com uma porcentagem de variação $\varepsilon = 2\%$, assim os limitantes adotados foram: $u_t = 81,6$ Gy, $l_t = 78,4$ Gy, $u_c = 40$ Gy, $u_g = 45$ Gy e o escalar positivo (w) igual a 1.

A composição das submatrizes da matriz de deposição de dose e a matriz

do modelo que auxilia no planejamento do tratamento do câncer por radioterapia são mostradas no Anexo A.

Para esta simulação foi utilizado o aplicativo linprog do *software* Matlab (2007a, The MathWorks, Natick, MA), que usa o método Simplex para a resolução de problemas de grande porte e uma implementação do método Primal-Dual de pontos interiores (MPDPI) visto na seção 3.4.3, em linguagem C++ utilizando-se o compilador Borland C++ Builder 6, o qual foi desenvolvido e executado com bom desempenho computacional (Homem et al., 2010). O aplicativo lingprog foi processado em um computador Dell/PC com processador Intel Core2Duo T6570 2.1GHz, memória 3GB e sistema operacional Microsoft Windows 7 Professional 32bits e o programa em linguagem C++ em microcomputador IBM/PC com processador Intel Core2Duo T5800 2GHz, memória 3GB e sistema operacional Microsoft Windows XP 32bits.

O tempo do processamento do *software* Matlab e do MPDPI foi desconsiderado pois ambos determinaram a solução ótima em tempo inferior a 10 segundos.

Utilizando os dados apresentados neste capítulo, o modelo de programação linear (1)-(8) foi testado pelo aplicativo linprog do *software* Matlab e pelo programa desenvolvido em linguagem C++ relativo ao método Primal-Dual de pontos interiores, cujos resultados das parcelas da função objetivo são apresentados na tabela 4.

softwares	Matlab	MPDPI
déficit de dose no tumor	1.839998×10^{-15}	2.67×10^{-8}
excesso de dose no tecido crítico	3.595321×10^{-16}	8.77×10^{-8}
excesso de dose no tecido saudável	8.095293×10^{-16}	2.65×10^{-7}
valor da função objetivo	3.009060×10^{-15}	3.79×10^{-7}

Tabela 4: Resultados parciais da função objetivo minimizada no planejamento otimizado

A tabela 4 mostra que o excesso de dose nas regiões críticas e saudáveis e o déficit de dose na região tumoral é praticamente zero, mostrando que o tumor recebeu toda a dose necessária para a sua eliminação e que o limite de dose permitido nas outras

regiões não foi ultrapassado.

As figuras 12 e 13 apresentam a porcentagem de dose recebida por pixel em barras e ao longo da imagem 11, respectivamente, com os resultados obtidos pelo aplicativo linprog do *software* Matlab.

As figuras 14 e 15 apresentam a porcentagem de dose recebida por pixel em barras e ao longo da imagem 11, respectivamente, com os resultados obtidos pelo programa desenvolvido em linguagem C++ relativo ao método Primal-Dual de pontos interiores.

Pode-se observar que em ambos casos, apesar do tumor estar totalmente envolvido por tecidos críticos, a região tumoral recebeu a maior quantidade de radiação, enquanto os tecidos críticos e saudáveis foram poupados.

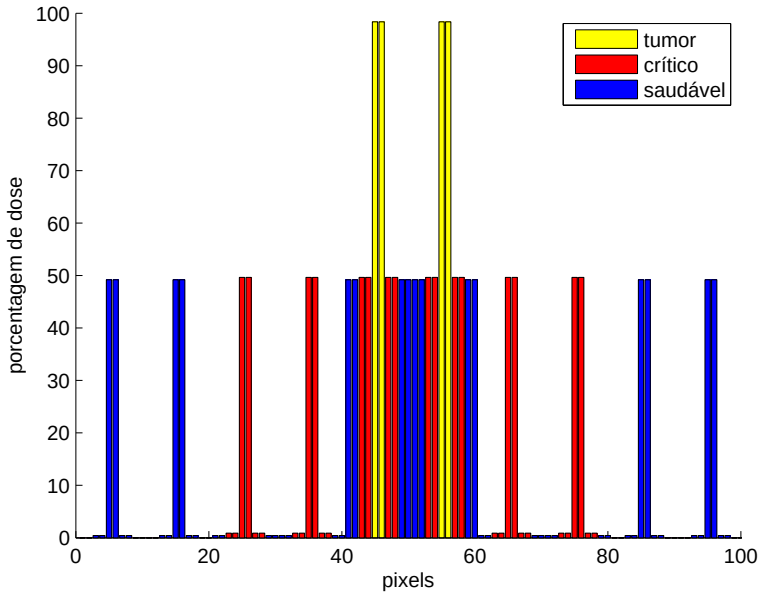

Figura 12: Distribuição da porcentagem de dose por pixel com os resultados obtidos pelo aplicativo linprog do *software* Matlab

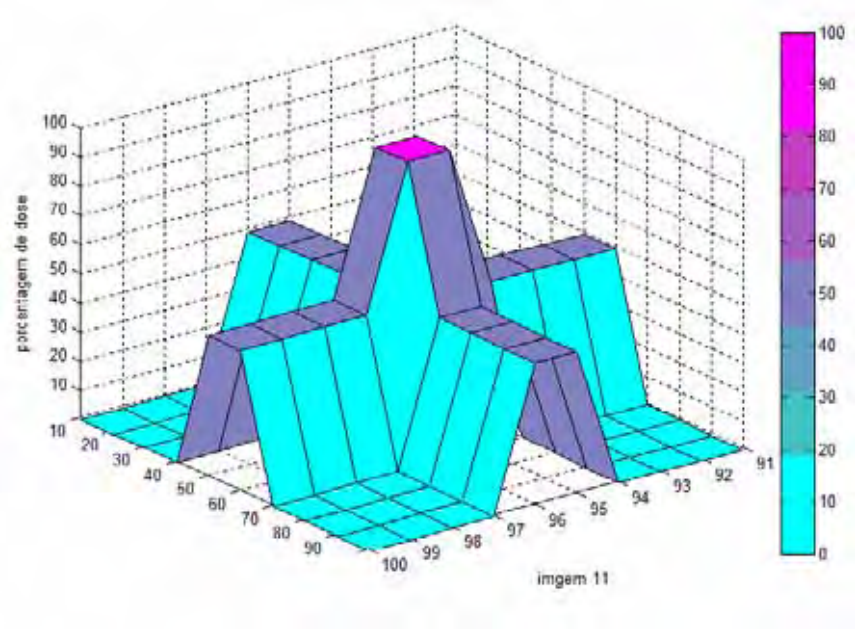

Figura 13: Distribuição da porcentagem de dose por pixel ao longo da imagem 11 com os resultados obtidos pelo aplicativo linprog do *software* Matlab

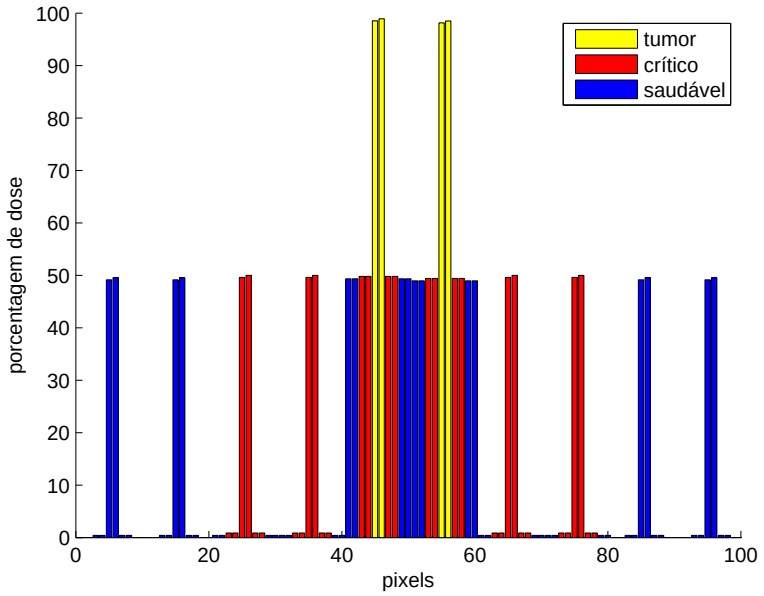

Figura 14: Distribuição da porcentagem de dose por pixel com os resultados obtidos pelo programa desenvolvido em linguagem C++ relativo ao MPDPI

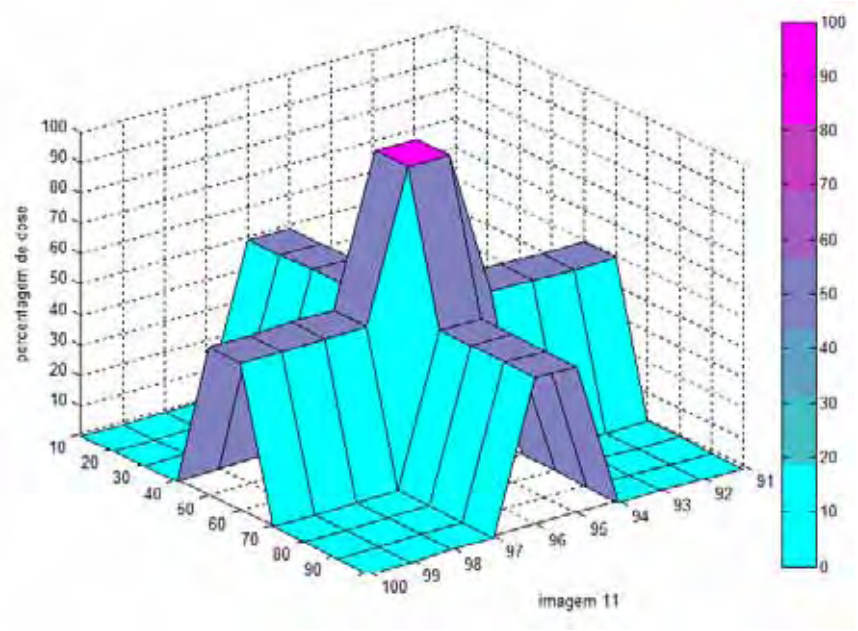

Figura 15: Distribuição da porcentagem de dose por pixel ao longo da imagem 11 com os resultados obtidos pelo programa desenvolvido em linguagem C++ relativo ao MPDPI

De acordo com os reultados da tabela 4, as simulações realizadas resolveram o modelo (1)-(8), o qual auxilia o planejamento ótimo da radioterapia. Assim, a técnica de pontos interiores mostrou-se apropriada para a resolução do problema, pois conseguiu atingir os objetivos propostos pelo modelo.

As soluções com a quantidade relativa de dose recebida por pixel determinadas pelos *softwares* Matlab e pelo método Primal-Dual de pontos interiores e os valores da porcentagem de dose recebida por pixel, apresentadas nas figuras 12 e 14 são mostradas no Anexo B.

5 CONCLUSÃO

Neste trabalho foram apresentados os principais conceitos envolvidos em um plano de tratamento do câncer por radioterapia e um modelo baseado na programação linear para o auxílio no planejamento ótimo, o qual foi investigado no capítulo 02.

Deu-se ênfase à técnica de otimização linear de pontos interiores, que foi amplamente desenvolvida no capítulo 03, discutindo-se os métodos Primal-Afim, Dual-Afim e destacando-se o método Primal-Dual de pontos interiores, o qual foi implementado em linguagem C++ e usado para a resolução do modelo (1)-(8), devido as facilidades computacionais que este método apresenta.

O modelo de programação linear proposto por Holder (2003) mostrou-se eficiente para auxiliar o planejamento ótimo, visto que mesmo em um caso complicado, como um tumor totalmente envolvido em uma região crítica, permitiu elaborar um plano de distribuição de dose onde todas as células tumorais recebem a radiação necessária para erradicação e as regiões críticas e saudáveis não recebem uma radiação superior a permitida.

A técnica de pontos interiores mostrou-se apropriada para resolução do problema de planejamento ótimo, pois conseguiu atingir os objetivos propostos pelo modelo, mostrando que o déficit de dose no tumor e o excesso de dose nas regiões críticas e saudáveis é mínimo de acordo com os resultados apresentados na tabela 4. Outro motivo pelo qual podemos afirmar que a utilização da técnica de pontos interiores foi apropriada, é o fato de que tanto na implementação do aplicativo do *software* Matlab quanto no programa em linguagem C++, as soluções encontradas são muitos próximas.

Para trabalho futuro, pretende-se estudar o método Previsor-Corretor Primal-Dual de pontos interiores para implementá-lo e aplicá-lo ao modelo de programação linear que auxilia o planejamento ótimo do tratamento do câncer por radioterapia.

Anexos

Anexo A

A matriz de deposição de dose da aplicação proposta no capítulo 04 é formada pelas submatrizes A_t, A_c e A_g, dadas por:

$$A_t = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$A_c = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A_g = \begin{bmatrix}
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&1&0&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0\\
\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\
0&0&0&0&1&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&1&0&0&0&0&0
\end{bmatrix}$$

Esta matriz é composta pelos números 0 e 1, pois foi considerado um coeficiente de atenuação nulo.

Sendo assim, a matriz do problema de programação linear que auxilia no planejamento do tratamento do câncer por radioterapia pode ser escrito como:

$$\begin{bmatrix} -L & O_1 & O_2 & -A_T \\ O_3 & O_1 & O_2 & A_t \\ O_4 & U_c & O_5 & A_c \\ O_6 & O_7 & U_g & A_G \\ -L & O_1 & O_2 & O_8 \\ L & O_1 & O_2 & O_8 \\ O_4 & -U_c & O_5 & O_9 \\ O_6 & O_7 & -U_g & O_{10} \end{bmatrix} \begin{bmatrix} t \\ c \\ g \\ X \end{bmatrix} \leq \begin{bmatrix} -l_t \\ u_t \\ u_c \\ u_g \\ o_1 \\ l_t \\ u_c \\ o_2 \end{bmatrix}$$

Onde:

$A \in \Re^{176\times108}$, $x \in \Re^{108\times1}$ e $b \in \Re^{176\times1}$.

$t = [t_1\, t_2\, t_3\, t_4]^T$: representando o número de pixels tumorais.

$c = [c_1\, c_2\, c_3 \ldots c_{32}]^T$: representando o número de pixels críticos.

$g = [g_1\, g_2\, g_3 \ldots g_{48}]^T$: representando o número de pixels saudáveis.

$X = [x_1\, x_2\, x_3 \ldots x_{24}]^T$: representando o número de subfeixes.

As matrizes L, U_c e U_g são matrizes identidade, respctivamente de ordem 4, 32 e 48, os quais correspondem os número de pixels dos tecidos tumoral, crítico e saudável.

As matrizes O_i e o_j, para $i = 1, 2.., 10$ e $j = 1, 2$ são matrizes nulas, onde:

$O_1 \in \Re^{4\times32}$ $O_5 \in \Re^{32\times48}$ $O_9 \in \Re^{32\times24}$

$O_2 \in \Re^{4\times48}$ $O_6 \in \Re^{48\times4}$ $O_{10} \in \Re^{48\times24}$

$O_3 \in \Re^{4\times4}$ $O_7 \in \Re^{48\times32}$ $o_1 \in \Re^{4\times1}$

$O_4 \in \Re^{32\times4}$ $O_8 \in \Re^{4\times24}$ $o_2 \in \Re^{48\times1}$

Anexo B

Dose de radiação recebida em cada pixel

A tabela a seguir apresenta os resultados obtidos para a dose recebida por cada pixel da figura 11, comparando os resultados obtidos pelo *software* Matlab e o método Primal-Dual de pontos interiores (MPDPI) e a porcentagem de dose administrada em cada pixel que está representado nas figuras 12 e 14.

	Matlab		MPDPI	
	dose por pixel	% de dose por pixel	dose por pixel	% de dose por pixel
1	0,00000	0,00000	0,00000	0,00000
2	0,00000	0,00000	0,00000	0,00000
3	0,36058	0,45072	0,36118	0,45147
4	0,36058	0,45072	0,36118	0,45147
5	39,34798	49,18498	39,48831	49,36039
6	39,34798	49,18498	39,16044	48,95055
7	0,36058	0,45072	0,36118	0,45147
8	0,36058	0,45072	0,36118	0,45147
9	0,00000	0,00000	0,00000	0,00000
10	0,00000	0,00000	0,00000	0,00000
11	0,00000	0,00000	0,00000	0,00000
12	0,00000	0,00000	0,00000	0,00000
13	0,36058	0,45072	0,36118	0,45147
14	0,36058	0,45072	0,36118	0,45147
15	39,34798	49,18498	39,48831	49,36039
16	39,34798	49,18498	39,16044	48,95055
17	0,36058	0,45072	0,36118	0,45147
18	0,36058	0,45072	0,36118	0,45147
19	0,00000	0,00000	0,00000	0,00000
20	0,00000	0,00000	0,00000	0,00000
21	0,36058	0,45072	0,36083	0,45104
22	0,36058	0,45072	0,36083	0,45104
23	0,72116	0,90145	0,72201	0,90251
24	0,72116	0,90145	0,72201	0,90251
25	39,70856	49,63570	39,84914	49,81143

	Matlab		MPDPI	
	dose por pixel	% de dose por pixel	dose por pixel	% de dose por pixel
26	39,70856	49,63570	39,52127	49,40159
27	0,72116	0,90145	0,72201	0,90251
28	0,72116	0,90145	0,72201	0,90251
29	0,36058	0,45072	0,36083	0,45104
30	0,36058	0,45072	0,36083	0,45104
31	0,36058	0,45072	0,36083	0,45104
32	0,36058	0,45072	0,36083	0,45104
33	0,72116	0,90145	0,72201	0,90251
34	0,72116	0,90145	0,72201	0,90251
35	39,70856	49,63570	39,84914	49,81143
36	39,70856	49,63570	39,52127	49,40159
37	0,72116	0,90145	0,72201	0,90251
38	0,72116	0,90145	0,72201	0,90251
39	0,36058	0,45072	0,36083	0,45104
40	0,36058	0,45072	0,36083	0,45104
41	39,34798	49,18498	39,33468	49,16835
42	39,34798	49,18498	39,33468	49,16835
43	39,70856	49,63570	39,69586	49,61982
44	39,70856	49,63570	39,69586	49,61982
45	78,69596	98,36995	78,82299	98,52874
46	78,69596	98,36995	78,49512	98,11890
47	39,70856	49,63570	39,69586	49,61982
48	39,70856	49,63570	39,69586	49,61982
49	39,34798	49,18498	39,33468	49,16835
50	39,34798	49,18498	39,33468	49,16835
51	39,34798	49,18498	39,63859	49,54824
52	39,34798	49,18498	39,63859	49,54824
53	39,70856	49,63570	39,99977	49,99971
54	39,70856	49,63570	39,99977	49,99971
55	78,69596	98,36995	79,12690	98,90863
56	78,69596	98,36995	78,79903	98,49879
57	39,70856	49,63570	39,99977	49,99971
58	39,70856	49,63570	39,99977	49,99971
59	39,34798	49,18498	39,63859	49,54824
60	39,34798	49,18498	39,63859	49,54824
61	0,36058	0,45072	0,36083	0,45104
62	0,36058	0,45072	0,36083	0,45104
63	0,72116	0,90145	0,72201	0,90251
64	0,72116	0,90145	0,72201	0,90251
65	39,70856	49,63570	39,84914	49,81143
66	39,70856	49,63570	39,52127	49,40159
67	0,72116	0,90145	0,72201	0,90251
68	0,72116	0,90145	0,72201	0,90251
69	0,36058	0,45072	0,36083	0,45104
70	0,36058	0,45072	0,36083	0,45104
71	0,36058	0,45072	0,36083	0,45104
72	0,36058	0,45072	0,36083	0,45104
73	0,72116	0,90145	0,72201	0,90251
74	0,72116	0,90145	0,72201	0,90251
75	39,70856	49,63570	39,84914	49,81143
76	39,70856	49,63570	39,52127	49,40159
77	0,72116	0,90145	0,72201	0,90251
78	0,72116	0,90145	0,72201	0,90251
79	0,36058	0,45072	0,36083	0,45104
80	0,36058	0,45072	0,36083	0,45104
81	0,00000	0,00000	0,00000	0,00000
82	0,00000	0,00000	0,00000	0,00000
83	0,36058	0,45072	0,36118	0,45147
84	0,36058	0,45072	0,36118	0,45147
85	39,34798	49,18498	39,48831	49,36039

	Matlab		MPDPI	
	dose por pixel	% de dose por pixel	dose por pixel	% de dose por pixel
86	39,34798	49,18498	39,16044	48,95055
87	0,36058	0,45072	0,36118	0,45147
88	0,36058	0,45072	0,36118	0,45147
89	0,00000	0,00000	0,00000	0,00000
90	0,00000	0,00000	0,00000	0,00000
91	0,00000	0,00000	0,00000	0,00000
92	0,00000	0,00000	0,00000	0,00000
93	0,36058	0,45072	0,36118	0,45147
94	0,36058	0,45072	0,36118	0,45147
95	39,34798	49,18498	39,48831	49,36039
96	39,34798	49,18498	39,16044	48,95055
97	0,36058	0,45072	0,36118	0,45147
98	0,36058	0,45072	0,36118	0,45147
99	0,00000	0,00000	0,00000	0,00000
100	0,00000	0,00000	0,00000	0,00000

pixels saudáveis
pixels críticos
pixels tumorais
pixels frios

REFERÊNCIAS BIBLIOGRÁFICAS

ADLER, I.; KARMARKAR, N.; RESENDE, M. G. C.; VEIGA, G. Data structures and programming techniques for the implementation of Karmarkar's algorithm. **ORSA Journal of Computing**, , n.1, p.84–106, 1989.

ARRUDA, W. O. 100 anos da descoberta dos raios- X. **SciELO Brasil**, v.54, n.3, p.525–531, 1996.

BARNES, E. R. A variation of Karmarkar's algorithm for solving linear programming problems. **Mathematical Programming**, , n.36, p.174–182, 1986.

BERMAN, A.; PLEMMONS, R. **Nonnegative matrices in the mathematical sciences**. New York: Academic Press, 1979.

CID, C. B. B. Planejamento do tratamento por radioterapia através de métodos de pontos interiores. Instituto de Ciências Matemáticas e de Computação (ICMC), 2003. 72p. Dissertação (Mestrado) - Universidade de São Paulo.

DIKIN, I. I. Iterative solution of problems of linear and quadratic programming. **Soviet Mathematics Doklady**, v.8, p.674–675, 1967.

FANG, S. C.; PUTHENPURA, S. Affine Scaling Algorithms. In: HALL, P. (Ed.). **Linear Optimization and Extensions: Theory and Algorithms**. New Jersey: Englewood Cliffs, 1993. p.144–200.

GILL, P. E.; MURRAY, W.; SAUNDERS, M. A.; TOMLIN, J. A.; WRIGHT, M. H. On projected barrier methods for linear programming and an equivalence to Karmarkar's projective method. **Mathematical Programming**, v.36, p.183–209, 1986.

HOLDER, A. Designing radiotherapy plans with elastic constraints and interior point methods. **HealthCare Management Science**, v.6, p.5–16, 2003.

HOMEM, T. P. D.; BALBO, A. R.; FLORENTINO, H. O. O Comportamento de Métodos Híbridos de Pontos Interiores e Branch-and-Bound em Problemas de Resíduos de Cana-de-Açucar.. In: 9TH BRAZILIAN CONFERENCE ON DYNAMICS, CONTROL AND THEIR APPLICATIONS, 9, 2010. ; resumos. Serra Negra-SP: Proceedings of the 9th Brazilian Conference on Dynamics, Control and Applications, 2010. 1.

KARMARKAR, N. A new polynomial-time algorithm for linear programming. In: , 1984. **STOC '84: Proceedings of the sixteenth annual ACM symposium on Theory of computing**; resumos. New York, NY, USA: ACM, 1984. 302–311.

KOJIMA, M.; MIZUNO, S.; YOSHISE, A. A primal-dual interior point method for linear programming. In: MEGIDDO, N. (Ed.). **Progress in Mathematical Programming: Interior-Point and Related Methods**. New York: Springer-Verlag, 1989. p.29–48.

LIM, G.; CHOI, J.; MOHAN, R. Iterative solution methods for beam angle and fluence map optimization in intensity modulated radiation therapy planning. **OR Spectrum**, v.30, n.2, p.289–309, 2008.

LORENCETTI, A.; SIMONETTI, J. P. As estratégias de enfrentamento de pacientes durante o tratamento de radioterapia. **Rev Latino-am Enfermagem**, v.13, n.6, p.944–950, 2005.

LUENBERGER, D. G.; YE, Y. The Simplex Method. In: SPRINGER (Ed.). **Linear and Nonlinear Programming**. Califórnia: Stanford University, 2008. p.33–78.

MEGIDDO, N. Pathways to the optimal set in linear programming. In: MEGIDDO, N. (Ed.). **Progress in Mathematical Programming: Interior-Point and Related Methods**. New York: Springer-Verlag, 1989. p.131–158.

MEGIDDO, N.; SHUB, M. Boundary behavior of interior point algorithm in linear programming. **Mathematics of Operations Research**, , n.14, p.97–146, 1989.

MEN, C.; ROMEIJN, H. E.; TASKIN, Z. C.; DEMPSEY, J. F. An exact approach to direct aperture optimization in IMRT treatment planning. **Physics in Medicine an Biology**, v.52, n.24, p.7333–7352, 2007.

MONTEIRO, R. C.; ADLER, I.; RESENDE, M. G. C. A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension. **Mathematics of Operations Research**, , n.15, p.191–214, 1990.

ROCHA, H.; DIAS, J. On the Optimization of Radiation Therapy Planning. In: INSTITUTO DE ENGENHARIA DE SISTEMAS E COMPUTADORES DE COIMBRA, Coimbra, 2009. ; resumos. Coimbra: , 2009.

ROSEN, I.; LANE, R.; MORRIL, S.; BELLI, J. Treatment Plan Optimization Using Linear Programming. **Medical Physics**, v.18, p.141–152, 1990.

SANTOS, C. A. D. Raios X: Descoberta casual ou criterioso experimento? **Ciência Hoje**, v.19, n.114, p.26–35, 1995.

SHEPARD, D. M.; FERRIS, M. C.; OLIVEIRA, G. H.; MACKIE, T. R. Optimization the delivery of radiation therapy to cancer patients. **SIAM Review**, v.41, n.4, p.721–744, 1999.

SONDERMAN, D.; ABRAHAMSON, P. G. Radiotherapy treatment design using mathematical programming models. **Operations Research**, v.33, n.4, p.705–725, 1985.

STOLL, B. A. Princípios gerais de radioterapia. In: **Radioterapia: Conhecimentos Gerais para médicos e estudantes de medicina**. São Paulo: Universidade de São Paulo, 1968. p.01–16.

SULTAN, A. S. A. Optimization of Beam Orientations in Intensity Modulated Radiation Therapy Planning. Kaiserslautern - Alemanha, 2006. 97p. Dissertação (Mestrado) - Universidade Técnica de Kaiserslautern.

VANDERBEI, R. J.; MEKETON, M. S.; FREEDMAN, B. A. A modification of Karmarkar's linear programming algorithm. **Algorithmica**, , n.1, p.395–407, 1986.

VIANA, R. S. S. Programa ção Linear Aplicada à criação de planejamentos otimizados em radioterapia. Instituto de Biociências de Botucatu (IBB), 2010. 73p. Dissertação (Mestrado) - Universidade Estadual Paulista.

VIANA, R. S. S.; FLORENTINO, H. O.; FILHO, P. R. F.; LIMA, E. A. B. F. Uma abordagem tridimensional na otimização do tratamento de câncer por radioterapia. In: XL SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, João Pessoa, 2008. ; resumos. João Pessoa: , 2008.

WEBB, S. The physical basis of IMRT and inverse planning. **The British Journal of Radiology**, v.76, p.678–689, 2003.

ZHANG, X.; LIU, H.; WANG, X.; DONG, L.; WU, Q.; MOHAN, R. Speed and convergence properties of gradient algorithms for optimization of IMRT. **Medical Physics**, v.31, p.1141 – 1152, 2004.

APÊNDICE

Inicialização do método Primal-Afim

De maneira semelhenate a estabelecida por Luenberger & Ye (2008), para o método Simplex, será introduzido os dois mecanismos para encontrar uma solução inicial interior factível. O método Big-M, é executado mais facilmente e apropriado para a maioria das aplicações, enquanto o método Duas-Fases é comercialmente mais utilizado devido a sua estabilidade.

Método Big-M

Neste método, adiciona-se ao programa linear original uma variável artificial x^a e um número positivo grande (M), para fazer com que $(1, 1, 1, ..., 1) \in \Re^{n+1}$ transforme-se em uma solução inicial interior factível para o seguinte problema:

$$\begin{aligned} &\text{minimizar } c^T x + Mx^a \\ &\text{sujeito a } [A \mid b - Ae] \begin{bmatrix} x \\ x^a \end{bmatrix} = b \\ &\qquad\qquad x \geq 0, \quad x^a \geq 0 \end{aligned} \tag{93}$$

onde $e = (1, 1, 1, ..., 1)^T \in \Re^n$.

No método Big-M tem-se somente *n+1* variáveis enquanto no método Simplex $n+m$ variáveis.

Quando o algoritmo Primal-Afim é aplicado ao problema (93), desde que o problema seja viável, chega-se em uma solução ótima ou conclui-se que o problema é ilimitado.

Se a variável artificial permanecer positiva na solução final (x^*, x^a) do problema Big-M, então o problema de programação linear original é infactível. Caso contrário,

ou o problema original é inferiormente ilimitado, ou x^* é a solução ótima do problema de programação linear original.

Método Duas-Fases

Escolhe-se um $x^0 > 0$ arbitrário e calcula-se $v = b - Ax^0$. Se:

1. $v = 0$, então x^0 é uma solução inicial interior factível.
2. Caso contrário, considera-se o seguinte problema de programação linear Fase-I com *n+1* variáveis.

$$\begin{aligned} &\text{minimizar } u \\ &\text{sujeito a } [A \mid v] \begin{bmatrix} x \\ \hline u \end{bmatrix} = b \\ &x \geq 0, \quad u \geq 0 \end{aligned} \tag{94}$$

O vetor $x^0 = \begin{pmatrix} x^0 \\ u^0 \end{pmatrix} = \begin{pmatrix} x^0 \\ 1 \end{pmatrix} > 0$ é uma solução interior factível ao problema Fase-I. Sendo assim, o algoritmo Primal-Afim pode ser aplicado para resolver este problema. Além disso, desde que o problema Fase-I é inferiormente limitado por zero, o algoritmo Primal-Afim terminará sempre com uma solução ótima, denominada $(x^*, u^*)^T$. A partir daí, verificamos:

1. Se $u^* > 0$, então o problema de programação linear original e infactível.
2. Se não, desde que o problema Fase-I for um problema em um espaço dimensional grande, pode-se mostrar que à exceção dos casos em que $u^* = 0$, é obtida no final da Fase-I. Então, $x^* > 0$ será uma solução interior viável inicial ao problema original.

Note que a diferença na dimensão entre o original e o problema Fase-I poderá causar cálculos extras para uma implementação simples. Primeiramente, devido às imprecisões numéricas nos computadores, a solução ótima x^* obtida na Fase-I poderia tornar

infactível para o problema original. Em outras palavras, é necessário restaurar a factibilidade primal antes de efetuar a Fase-II do método.

Além disso, a diferença da dimensão nas matrizes fundamentais $AX_k^2A^T$ (do problema original) e $\hat{A}\hat{X}_k^2\hat{A}^T$ (do problema da Fase-I) podem impedir que se utilizasse o processo de fatorização de Cholesky para o calculo rápido de suas matrizes inversas.

Conseqüentemente, seria útil operar as iterações Fase-I no espaço n-dimensional original.

Fazendo assim, suponha-se que esteja-se na k-ésima iteração da aplicação do Primal-Afim ao problema Fase-I. Denota-se:

$$\hat{A} = [A \mid v], \;\; \hat{X}_k = \begin{bmatrix} X_k & 0 \\ 0 & u^k \end{bmatrix} \;\; e \;\; \hat{A}_k = \hat{A}\hat{X}_k$$

onde a solução atual é dada por:

$$\hat{x}^k = \begin{pmatrix} x^k \\ u^k \end{pmatrix}$$

O gradiente da função objetivo é:

$$\hat{c} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

com isso, tem-se a direção de movimento no espaço original do problema Fase-I:

$$\hat{d}_x^k = -\hat{X}_k[I - \hat{A}_k^T(\hat{A}_k\hat{A}_k^T)^{-1}\hat{A}_k]\hat{c}^k$$

onde $\hat{c}^k = \hat{X}_k\hat{c}$. Definindo

$$w^k = (\hat{A}_k\hat{A}_k^T)^{-1}\hat{A}_k\hat{c}^k \tag{95}$$

então

$$\hat{d}_x^k = -\hat{X}_k(\hat{c} - \hat{A}_k^T w^k) \tag{96}$$

Efetuando cálculos resulta-se em

$$\hat{A}_k\hat{A}_k^T = [\hat{A}X_k^2A^T + (u^k)^2 v\, v^T] \tag{97}$$

e

$$\hat{A}_k\hat{c}^k = [A \mid v]\begin{bmatrix} X_k & 0 \\ 0 & u^k \end{bmatrix}\begin{bmatrix} 0 \\ \overline{u}^k \end{bmatrix} = [AX_k \mid vu^k]\begin{bmatrix} 0 \\ \overline{u}^k \end{bmatrix} = (u^k)^2 v \tag{98}$$

Combinando (95),(97) e (98), tem-se:

$$w^k = [AX_k^2A^T + (u^k)^2vv^T]^{-1}(u^k)^2v$$

Para prosseguir, é necessário a utilização do lema de Sherman-Woodbury-Morrison:

Lema: Sendo $M \in \Re^{nxn}$ uma matriz não singular e os vetores $u, v \in \Re^n$. Se $\omega = 1 + v^TMu \neq 0$, então a matriz $(M + uv^T)$ é não-singular e $(M + uv^T)^{-1} = M^{-1} - \left(\frac{1}{\omega}\right)M^{-1}uv^TM^{-1}$.

Assim:

$$w^k = \frac{1}{[(u^k)^{-2} + v^T(AX_k^2A^T)^{-1}v]}(AX_k^2A^T)^{-1}v = \frac{1}{[(u^k)^{-2} + \gamma]}(AX_k^2A^T)^{-1}v$$

onde $\gamma = v^T(AX_k^2A^T)^{-1}v$.

Adicionando os termos correspondentes encontrados em (96), vê-se que:

$$\hat{d}_x^k = -\hat{X}_k(\hat{c}^k - \hat{A}^Tw^k) = -\hat{X}_k\left(\begin{bmatrix} 0 \\ u^k \end{bmatrix} - \begin{bmatrix} X_kA^T \\ u^kv^T \end{bmatrix}w^k\right)$$

$$= -\begin{bmatrix} X_k & 0 \\ 0 & u^k \end{bmatrix}\begin{bmatrix} -X_kA^Tw^k \\ u^k(1 - v^Tw^k) \end{bmatrix} = -\begin{bmatrix} -X_k^2A^Tw^k \\ (u^k)^2(1 - v^Tw^k) \end{bmatrix}$$

Observe que:

$$(u^k)^2(1 - v^Tw^k) = (u^k)^2\left(1 - \frac{\gamma}{(u^k)^{-2} + \gamma}\right) = \frac{1}{(u^k)^{-2} + \gamma}$$

e

$$\hat{d}_x^k = \frac{1}{(u^k)^{-2} + \gamma}\begin{bmatrix} X_k^2A^T(AX_k^2A^T)^{-1}(b - Ax^0) \\ -1 \end{bmatrix} \tag{99}$$

O multiplicador escalar em (99) será considerado no comprimento do passo e o último elemento da direção do movimento é -1. Assim, sabe-se que o algoritmo sempre tenta reduzir u. Nesta expressão, tem-se o cálculo da direção $\hat{d}_x^k$ realizada no espaço original n-dimensional e o modelo para a fatoração de $AX_k^2A^T$ pode ser usada tanto na Fase-I como na Fase-II.

Para calcular o comprimento do passo, considera-se que:

$$\hat{x}^{k+1} = \hat{x}^k + \alpha_k \hat{d}_x^k$$

Como anteriormente discutido, o comprimento do passo pode ser escolhido usando:

$$\alpha_k = \min \left\{ \frac{-\alpha}{(\hat{d}_x^k)_i} / (\hat{d}_x^k)_i < 0 \right\} \quad \text{onde } 0 < \alpha < 1$$

Um ponto interessante e importante a ser observado aqui é que as iterações da Fase-I podem ser iniciadas a qualquer momento (mesmo durante as iterações Fase-II), uma vez detectado que a factibilidade de uma iteração atual é perdida devido a erros numéricos. As iterações da Fase-I podem ser aplicadas para restaurar a factibilidade, esse procedimento é chamado de "correção dinâmica de infactibilidade".

Implementações sofisticadas devem apresentar esse recurso, uma vez que o método primal é bastante sensível à truncamentos numéricos e erros de arredondamento.