



**UNIVERSIDADE FEDERAL DO CARIRI
CENTRO DE CIÊNCIAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
EM REDE NACIONAL - PROFMAT**

LUIZ CARLOS SOARES FERNANDES

**O USO DE SISTEMAS LINEARES NA TOMOGRAFIA
COMPUTADORIZADA**

**JUAZEIRO DO NORTE - CEARÁ
2018**

LUIZ CARLOS SOARES FERNANDES

O USO DE SISTEMAS LINEARES NA TOMOGRAFIA COMPUTADORIZADA

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Matemática em Rede Nacional - PROFMAT do Centro e Ciências e Tecnologia da Universidade Federal do Cariri, como requisito parcial para obtenção do Título de Mestre em Matemática. Área de concentração: Ensino de Matemática.

Orientador:
Prof. Dr. Valdir Ferreira de Paula Junior

JUAZEIRO DO NORTE - CEARÁ
2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Cariri
Sistema de Bibliotecas

- F398u Fernandes, Luiz Carlos Soares.
 O uso de sistemas lineares na tomografia computadorizada/ Luiz Carlos Soares Fernandes.
 – 2018.
 79f.: il.; color.; enc. ; 30 cm.
- Dissertação (Mestrado) – Universidade Federal do Cariri, Centro de Ciências e Tecnologia
 –Programa de Pós-graduação em Matemática em Rede Nacional, Juazeiro do Norte, 2018.
- Orientação: Prof. Dr. Valdir Ferreira de Paula Júnior.
1. Tomografia Computadorizada. 2. Sistemas Lineares. 3. Pesquisa Matemática. I. Título.

CDD 512.5

Bibliotecário: João Bosco Dumont do Nascimento – CRB 3/1355



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO CARIRI
CENTRO DE CIÊNCIAS E TECNOLOGIA
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL - PROFMAT

O Uso de Sistemas Lineares na Tomografia Computadorizada

Luiz Carlos Soares Fernandes

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Matemática em Rede Nacional – PROFMAT do Centro de Ciências e Tecnologia da Universidade Federal do Cariri, como requisito parcial para obtenção do Título de Mestre em Matemática. Área de concentração: Ensino de Matemática

Aprovada em 26 de outubro de 2018.

Banca Examinadora

Valdir Ferreira de Paula Junior

Prof. Dr. Valdir Ferreira de Paula Junior - UFCA

Orientador

Maria Silvana Alcântara Costa
Prof.^a. Dra. Maria Silvana Alcântara Costa -
UFCA

Samara Costa Lima
Prof.^a. Dra. Samara Costa Lima – UFOB

Dedico a meus pais, ao Prof. Dr. Valdir Ferreira de Paula Junior, aos meus amigos e em especial a Ramon Sales Pinheiro.

AGRADECIMENTOS

Agradeço em primeiro lugar a Deus que é senhor de tudo e de todos.

Aos meus pais.

A meu orientador, Prof. Dr. Valdir Ferreira de Paula Junior , pelo apoio, dedicação e pelos ensinamentos que foram dados na orientação deste trabalho e nas disciplinas que lecionou.

Gostaria de agradecer também, e, de forma especial, a Professora Dr^a. Maria Silvana Alcântara Costa pela condução da coordenação acadêmica institucional do PROFMAT, campus UFCA, Juazeiro do Norte-CE. Por fim, à Coordenação de Aperfeiçoamento Pessoal de Ensino Superior (CAPES) pelo apoio financeiro.

A loucura é um passo para a eternidade.

RESUMO

Tão importante quanto conhecer os procedimentos práticos, é conhecer a matemática por trás deles. Neste trabalho estudamos um pouco da matemática por trás da Tomografia Computadorizada (TC). Para isto, apresentaremos o modelo matemático da TC usando sistemas lineares. Veremos que resolver esses sistemas lineares equivale a encontrar uma parte da imagem gerada pela TC. Para resolver tais sistemas, estudaremos os algoritmos de Kaczmarz e Cimmino, mostraremos como utilizar estes algoritmos para obter soluções aproximadas de sistema lineares que possuem muitas equações e incógnitas. Vale destacar que este trabalho tem como maior objetivo servir como material para o professor do Ensino Médio introduzir seus alunos na pesquisa matemática; por isso estudaremos estes algoritmos de forma geométrica, algébrica e computacional.

Palavras-chave: Tomografia Computadorizada. Sistemas Lineares. Pesquisa Matemática.

ABSTRACT

As important as knowing the practical procedures is knowing the mathematics behind them. In this work, we present the mathematics behind computed tomography, that is, we study the mathematical fundamentals used in computed tomography. For this, we will make the mathematical model of the computed tomography using linear systems. We will see that these linear systems are equivalent to the image search generated by the CT. To obtain such systems, mark the algorithms of Kaczmarz and Cimmino, we will see that these algorithms approach a solution of these systems. It is worth pointing out that this work has as its main objective the material for the teacher of the physical, algebraic and computational form.

Keywords: Computed tomography. Linear Systems. Mathematical Research.

Lista de Figuras

1	Radiografia simples do toráx	10
2	Tomografia computadorizada	11
3	Característica do tomógrafo	12
4	Aparelho de Tomografia computadorizada	12
5	Escala de Cinzentos	13
6	Região planificada	13
7	Região planificada	22
8	Corte transversal	22
9	Corte Longitudinal	22
10	Região S	23
11	Região planificada	23
12	Região planificada	24
13	Absorção do raio L_i no pixel j	24
14	Região planificada	26
15	Solução inconsistente	28
16	Aplicação do Algoritmo de Gauss-Seidel	31
17	Método de Gauss-Seidel	32
18	Método de Jacobi	33
19	Estudo do exemplo usando o método Jacobi	34
20	Loop no algoritmo de Gauss-Seidel	35
21	Projeção de ponto	37
22	Projeção ideal	38
23	1ª projeção	40
24	2ª projeção	40
25	3ª projeção	41
26	4ª projeção	41
27	Sequência de soluções centradas em H_2	42
28	Sequência de soluções centradas em H_2 para o caso $m=2$	44
29	Método de Cimmino	45
30	Aplicação Método de Cimmino	49
31	Método de Kaczmarz para desigualdades	50

32	Reta ligando x e $P_H(x)$	52
33	Aplicação algoritmo de Gauss-Seidel	55
34	Loop no Algoritmo de Gauss-Seidel	56
35	Algoritmo de Kaczmarz	56
36	Algoritmo de Kaczmarz	57
37	Algoritmo de Kaczmarz	57
38	Algoritmo de Kaczmarz	58
39	Algoritmo de Kaczmarz	58
40	Método de Jacobi	59
41	Algoritmo de Jacobi	60
42	Algoritmo de Jacobi	60
43	Algoritmo de Jacobi	61
44	Algoritmo de Jacobi	62
45	Algoritmo de Cimmino	62
46	Algoritmo de Cimmino	63
47	Algoritmo de Cimmino	63
48	Algoritmo de Cimmino	64

SUMÁRIO

1	Introdução	10
2	Preliminares	15
2.1	Espaço Euclidiano n-dimesional	15
2.2	Matrizes e sistemas lineares	17
3	Modelo Matemático da Tomografia Computadorizada	21
3.1	Modelo Matemático da Tomografia Computadorizada Usando Sistemas Lineares	21
4	Usando Algoritmos para Resolver Sistemas Lineares	30
4.1	Os métodos de Gauss-Seidel e Jacobi	30
4.2	O operador projeção sobre hiperplanos	35
4.3	Algoritmos de Kaczmarz	39
4.4	Método de Cimmino	45
4.5	Análise das desigualdades	49
4.6	Relaxação	51
5	Ensinando os Algoritmos em Sala de Aula	54
6	Apêndice	66
6.1	O operador projeção	66
6.2	Análise de Convergência	72
7	Considerações Finais	76
	Referências	77

1 Introdução

A Tomografia Computadorizada (TC) é uma ferramenta fundamental em rotinas clínicas, de obtenção de imagem de partes internas do corpo humano, pois ela fornece um método por meio do qual é possível analisar uma parte do corpo humano sem a necessidade de intervenção cirúrgica. Além disso, em comparação com outros métodos, como a radiografia convencional ela tem a grande vantagem de fornecer uma imagem em $3D$. De fato, enquanto a radiografia convencional consiste na obtenção representativa de todas as estruturas do corpo de forma sobrepostas em uma imagem $2D$, o que é demasiado ruim para a detecção de pequenas diferenças entre os tecidos, a tomografia computadorizada (TC) nos fornece uma imagem $3D$ mais real e com mais detalhes da região em análise. A figura a seguir trata de uma radiografia convencional, onde a imagem é obtida por meio da sobreposição dos vários tecidos, conforme se observa claramente.

Figura 1: Radiografia simples do toráx

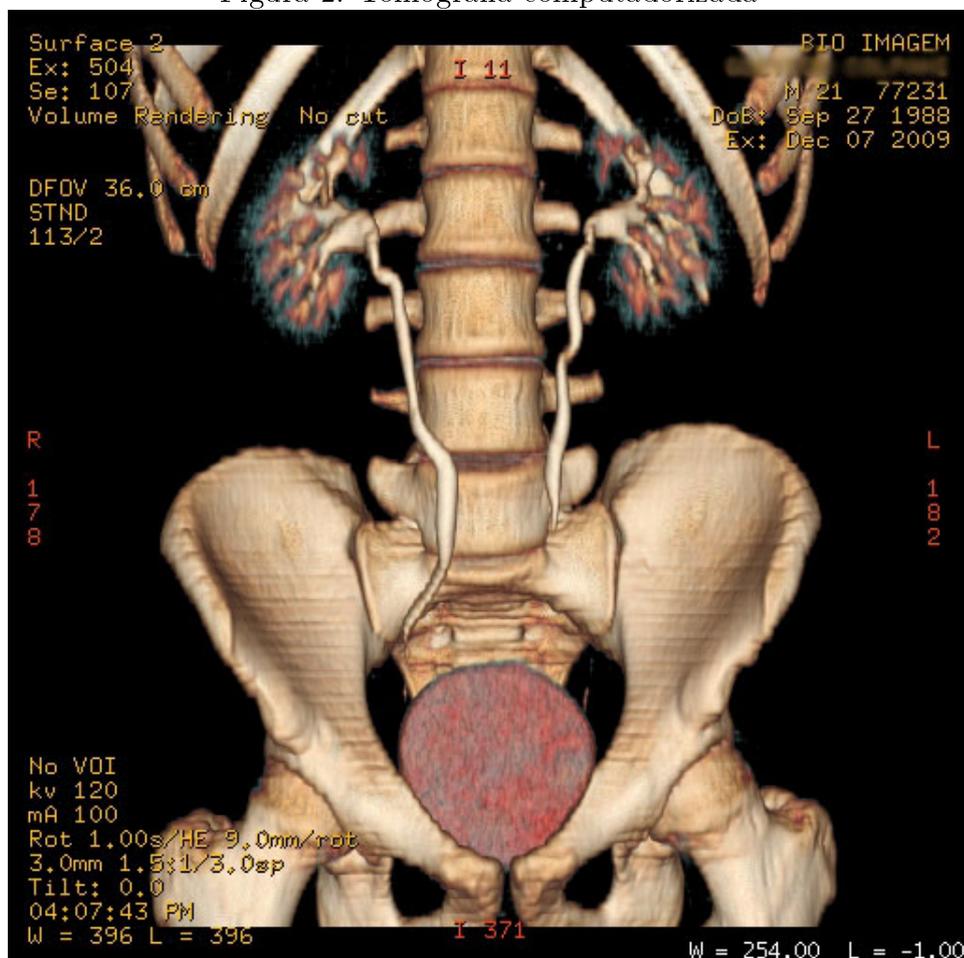


Fonte: [9]

1. INTRODUÇÃO

A Figura 2 representa uma imagem obtida através da tomografia computadorizada (TC) onde percebemos claramente a riqueza de detalhes, sendo portanto, mais adequada na elaboração de um diagnóstico clínico do que a radiografia convencional (Figura 1).

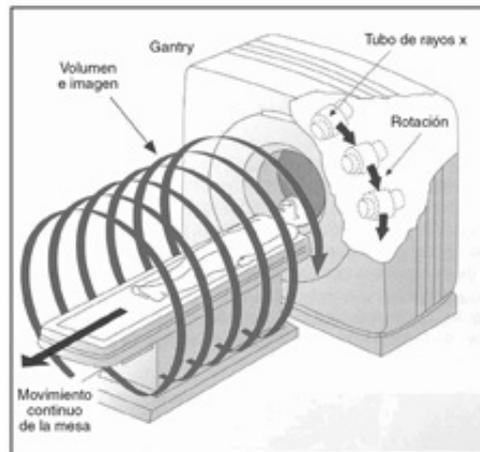
Figura 2: Tomografia computadorizada



Fonte: [10]

Para fazer a TC o paciente é colocado numa mesa que se desloca para o interior de um anel de cerca de 70 cm de diâmetro. Em volta deste encontra-se uma ampola de raios-X, num suporte circular designado gantry. Do lado oposto a ampola encontra-se o detector responsável por captar a radiação e transmitir informação ao computador ao qual está conectado. Nestas máquinas, durante o exame, um aparelho emite uma quantidade de raios X enquanto ele vai descrevendo uma volta completa (360°) em torno do paciente. Os raios, após atravessarem o corpo do paciente, são captados na outra extremidade pelo detector, conforme a figura a seguir:

Figura 3: Característica do tomógrafo



Fonte: [11]

Um computador faz a análise da variação de absorção de raios-x na região analisada e depois reconstrói a imagem por meio dos dados obtidos.

O aparelho responsável pela tomografia computadorizada recebe o nome de tomógrafo.

Figura 4: Aparelho de Tomografia computadorizada



Fonte: [12]

buir o conjunto das atenuações em linhas e colunas disposta de tal forma que determinam uma matriz em que a atenuação da primeira entrada esta associada ao primeiro quadradinho da imagem conforme Figura 6. Podemos representar o conjunto de pixels de uma figura em forma de linhas e colunas, de modo que esse conjunto determina uma matriz. Quanto maior o número de elementos desta matriz, menor será o tamanho do pixel, e logo, melhor será a qualidade da imagem obtida. Portanto, reconstruir uma imagem a partir da TC equivale a encontrar os elementos de uma matriz formada por pixels. Isto ficará claro quando formos modelar a TC , pois o processo de obtenção da imagem a partir de uma TC é modelado por uma teoria matemática consistente. Veremos no Capítulo 3 que obter uma imagem na TC equivale a resolver um sistema linear.

Esse trabalho foi baseado no artigo dos autores Iusem e De Pierro publicado na revista matemática universitária (veja [1]) tendo como objetivo estudar algoritmos matemáticos com muitas equações e incógnitas e transformá-lo numa linguagem mais simples e acessível para que o professor possa utilizá-lo em sala de aula como um exemplo prático de modelagem matemática e para motivá-los quanto a pesquisa matemática. Por isso, o texto e algumas figuras do nosso trabalho são adaptações. Também usamos [3] no Capítulo 2 e [4] nas preliminares, como referência para este trabalho. Para os professores que desejem se aprofundarem sobre os assuntos estudados aqui, recomendamos [1], [3], [7] [4], [8] e [6].

Esse trabalho esta dividido em 5 capítulos e um apêndice. No capítulo 2 faremos uma pequena revisão sobre as ferramentas necessárias para a compreensão desse trabalho. No Capítulo 3 iremos modelar a (TC) usando sistemas lineares. O Capítulo 4 é o nosso principal foco do trabalho, pois nele estudaremos como resolver os sistemas lineares, obtidos da modelagem da TC, a partir de algoritmos matemáticos. No Capítulo 5 mostraremos como o professor pode ensinar este algoritmo em sala de aula, usando apenas conhecimentos do Ensino Médio. Por fim, no Apêndice estudaremos a convergência dos algoritmos de Kaczmarz e Cimmino, mostraremos que estes algoritmos convergem para uma solução do sistema linear.

2 Preliminares

Neste Capítulo, apresentamos os conteúdos básicos que serão necessários à modelagem da tomografia computadorizada usando sistemas lineares. Inicialmente faremos uma breve introdução referente ao espaço euclidiano \mathbb{R}^n e depois relembremos os conceitos de matrizes e sistemas lineares. Mais detalhes podem ser encontrados em [4] e [5].

2.1 Espaço Euclidiano n-dimensional

Dado n um número natural. O espaço euclidiano n -dimensional, denotado por \mathbb{R}^n , é o produto cartesiano de n fatores iguais a \mathbb{R} , isto é, $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$. Seus elementos, portanto, são as sequências (ou listas) de n termos reais $x = (x_1, \dots, x_n)$. Para cada $i = 1, \dots, n$, o termo x_i chama-se a i -ésima coordenada de x . Se $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$, tem-se $x = y$ se, e somente se, $x_1 = y_1, \dots, x_n = y_n$. Assim, toda igualdade entre dois elementos de \mathbb{R}^n equivale a n igualdades entre números reais. Portanto, $\mathbb{R}^1 = \mathbb{R}$ é o conjunto dos números reais, \mathbb{R}^2 é o plano euclidiano e \mathbb{R}^3 é o espaço euclidiano tridimensional. Os elementos de \mathbb{R}^n às vezes são chamados pontos e às vezes vetores, isto vai depender muito da situação que estamos abordando. Geralmente usa-se vetores quando consideradas as operações que definiremos agora. A adição que faz corresponder a cada par de elementos $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ a soma

$$x + y = (x_1 + y_1, \dots, x_n + y_n)$$

e a multiplicação de um número real α pelo elemento $x = (x_1, \dots, x_n)$ tendo como resultado

$$\alpha x = (\alpha x_1, \alpha x_2, \dots, \alpha x_n).$$

O vetor $0 = (0, 0, \dots, 0)$, cujas coordenadas são todas zero, chama-se a origem de \mathbb{R}^n .

Dados quaisquer $x, y, z \in \mathbb{R}^n$ e $\alpha, \beta \in \mathbb{R}$ valem as seguintes igualdades:

- 1) Elemento neutro ou vetor nulo

$$x + 0 = 0 + x = x$$

2) Comutatividade

$$x + y = y + x$$

3) Simétrico

$$-x + x = x + (-x) = 0$$

4) Associatividade

$$x + (y + z) = (x + y) + z$$

5) Distributividade

$$(\alpha + \beta)x = \alpha x + \beta x, \quad \beta(x + y) = \beta x + \beta y$$

A primeira igualdade assegura que o vetor nulo é o elemento neutro da adição, já a segunda diz que $-x$ é o inverso (ou simétrico) aditivo de x .

Definição 1 *A função que associa a cada par de vetores $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ o número real*

$$\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n,$$

é chamado o produto interno canônico de x e y .

A partir da definição 1 e das operações em \mathbb{R} , pode-se verificar que o produto interno satisfaz as seguintes propriedades para quaisquer $x, y, z \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$,

a) Simétrico:

$$\langle x, y \rangle = \langle y, x \rangle;$$

b) Linear

$$\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle \quad \text{e} \quad \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle;$$

$$\langle \alpha x, y \rangle = \alpha \langle x, y \rangle \quad \text{e} \quad \langle x, \alpha y \rangle = \alpha \langle x, y \rangle;$$

c) Não-negativo:

$$\langle x, x \rangle \geq 0 \quad \text{e} \quad \langle x, x \rangle = 0 \Leftrightarrow x = 0.$$

Da definição de produto interno obtemos a seguinte definição.

Definição 2 *Diz-se que dois vetores $x, y \in \mathbb{R}^n$ são ortogonais quando*

$$\langle x, y \rangle = 0.$$

Neste caso denotamos $x \perp y$.

A próxima definição é importante, pois define tamanho de um vetor.

Definição 3 *Chama-se de norma (ou comprimento) de um vetor $x \in \mathbb{R}^n$ ao número não-negativo: $\|x\| = \sqrt{\langle x, x \rangle}$, assim*

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

A seguinte desigualdade, conhecida como Desigualdade de Cauchy-Schwarz será usada no Capítulo 4 para provar que os algoritmos de Kaczmarz e Cimmino convergirão.

Proposição 1 *(Desigualdade de Cauchy-Schwarz) Se $x, y \in \mathbb{R}^n$, então*

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Demonstração: Veja [5].

A próxima definição será muito utilizada nos capítulos seguintes, inclusive quando formos modelar a TC via sistemas lineares.

Definição 4 *Dados $v \in \mathbb{R}^n$ e $b \in \mathbb{R}$, o conjunto*

$$H = \{x \in \mathbb{R}^n | \langle x, v \rangle = b\} \tag{1}$$

é chamado de hiperplano.

Note que para $n = 3$, o conjunto H é um plano do \mathbb{R}^3 , da forma $\alpha x_1 + \beta x_2 + \gamma x_3 = b$, onde estamos tomando $v = (\alpha, \beta, \gamma)$ e $x = (x_1, x_2, x_3)$.

2.2 Matrizes e sistemas lineares

Nesta seção relembremos a definição e algumas propriedades elementares de matrizes e sistemas lineares. Para mais detalhes sobre o assunto veja [7]. Vale a pena destacar que usaremos as definições que geralmente aparecem nos livros do Ensino Médio.

A ideia geral de matriz de ordem $m \times n$ é a de um quadro retangular com mn elementos, dispostos em m linhas e n colunas. Matrizes são frequentemente utilizadas para a organização de dados, como por exemplo, as notas finais dos alunos de uma série no colégio podem formar uma matriz cujas colunas correspondem as matérias lecionadas naquela série e cujas linhas representam os alunos. Daremos agora a definição formal, de matriz, usada no Ensino Médio.

Definição 5 Uma matriz A é um conjunto (ou mais especificamente uma tabela) ou coleção que apresenta $m \cdot n$ elementos dispostos em m linhas e n colunas.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Para cada $1 \leq i \leq m$, a lista ordenada $(a_{i1}, a_{i2}, \dots, a_{in})$ chama-se a i -ésima linha ou i -ésimo vetor linha da matriz A , enquanto que para cada $1 \leq j \leq n$ $(a_{1j}, a_{2j}, \dots, a_{nj})$ é a j -ésima coluna ou j -ésimo vetor coluna de A .

A matriz A será denotada por $A = (a_{ij})_{m \times n}$, onde a_{ij} é o elemento cuja disposição na matriz ocupará, a interseção da i -ésima linha com a j -ésima coluna. Denotaremos por O a matriz cujos seus elementos são todos iguais a zero, e a chamaremos de matriz nula.

Diz-se que a matriz A é quadrada quando tem o mesmo número de linhas e colunas.

Definição 6 Duas matrizes $A = (a_{ij})_{m \times n}$ e $B = (b_{ij})_{m \times n}$ são iguais quando $a_{ij} = b_{ij}$ para todo $i (i \in \{1, 2, 3, \dots, m\})$ e todo $j (j \in \{1, 2, 3, \dots, n\})$. Isto significa que para serem iguais duas matrizes devem ser do mesmo tipo e apresentar os elementos correspondentes (elementos com índices iguais) iguais.

No conjunto de todas as matrizes $m \times n$ podemos definir a soma de matrizes e o produto por escalar da seguinte forma:

Definição 7 Dadas duas matrizes $A = (a_{ij})_{m \times n}$ e $B = (b_{ij})_{m \times n}$, chama-se soma $A + B$ a matriz $C = (c_{ij})_{m \times n}$ tal que $c_{ij} = a_{ij} + b_{ij}$, para todo i e todo j . Isto significa que a soma de duas matrizes A e B do tipo $m \times n$ é uma matriz C do mesmo tipo em que cada elemento é a soma dos elementos correspondentes em A e B .

Definição 8 Dado um número α e uma matriz $A = (a_{ij})_{m \times n}$, chama-se produto de um escalar por uma matriz, denotado por αA , a matriz $B = (b_{ij})_{m \times n}$ tal que $b_{ij} = \alpha a_{ij}$ para todo i e todo j . Isto significa que multiplicar uma matriz A por um escalar α é construir uma matriz B formada pelos elementos de A todos multiplicados por α .

Teorema 1 A adição de matrizes do tipo $m \times n$ goza das seguintes propriedades:

- 1) é associativa: $(A + B) + C = A + (B + C)$ quaisquer que sejam A, B e C do tipo $m \times n$.
- 2) é comutativa: $A + B = B + A$ quaisquer que sejam A e B , do tipo $m \times n$.
- 3) tem elemento neutro: existe M tal que $A + M = A$ qualquer que seja A do tipo $m \times n$.
- 4) todo elemento tem simétrico: para todo A do tipo $m \times n$ existe A' tal que $A + A' = M$

Demonstração:

1) Fazendo $(A + B) + C = K$ e $A + (B + C) = Z$, temos:

$$k_{ij} = (a_{ij} + b_{ij}) + c_{ij} = a_{ij} + (b_{ij} + c_{ij}) = z_{ij}$$

para todo i e todo j .

2) Fazendo $A + B = K$ e $B + A = Z$, temos:

$$k_{ij} = a_{ij} + b_{ij} = b_{ij} + a_{ij} = z_{ij}$$

3) Impondo $A + M = A$, resulta: $a_{ij} + m_{ij} = a_{ij}$ donde $m_{ij} = 0$. Daí, $M = O$, isto é, o elemento neutro da adição de matriz é a matriz nula do tipo $m \times n$.

4) Impondo $A + A' = M = O$, resulta: $a_{ij} + a'_{ij} = 0$ donde $a'_{ij} = -a_{ij}$ para todo i e todo j , ou seja, a simétrica da matriz A para a adição é a matriz A' de mesmo tipo que A .

□

Um fato interessante é que, considerando o espaço das matrizes com as operações de adição de matrizes e multiplicação de um vetor por um escalar da Definição 1 ele é um espaço vetorial, isto é, vale as propriedades comutatividade, associatividade, etc. A próxima proposição destaca algumas dessas propriedades, para mais detalhes veja ([7]).

Denotaremos o elemento neutro da adição pela matriz O , que é a matriz em que todos os seus elementos são nulos

Definição 9 Dadas duas matrizes $A = (a_{ij})_{m \times n}$ e $B = (b_{jk})_{n \times p}$, chama-se produto AB a matriz $C = (c_{ik})_{m \times p}$ tal que

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + a_{i3}b_{3k} + \dots + a_{in}b_{nk} = \sum_{j=1}^n a_{ij}b_{jk}$$

para todo $i \in \{1, 2, \dots, m\}$ e todo $k \in \{1, 2, \dots, p\}$.

Teorema 2 A multiplicação de matrizes goza das seguintes propriedades:

1. é associativa: $(AB)C = A(BC)$
2. é distributiva à direita em relação à adição: $(A + B)C = AC + BC$
3. é distributiva à esquerda: $C(A + B) = CA + CB$
4. $(kA)B = A(kB) = k(AB)$

3 Modelo Matemático da Tomografia Computadorizada

Neste capítulo, estudaremos o modelo matemático por trás da TC, usando sistemas lineares. Isto é, veremos que obter uma imagem via TC equivale a resolver um sistema de equações e inequações lineares. Para os leitores que desejem se aprofundar no assunto, recomendamos [1], no qual este capítulo foi baseado, também recomendamos [2], [3] e [8].

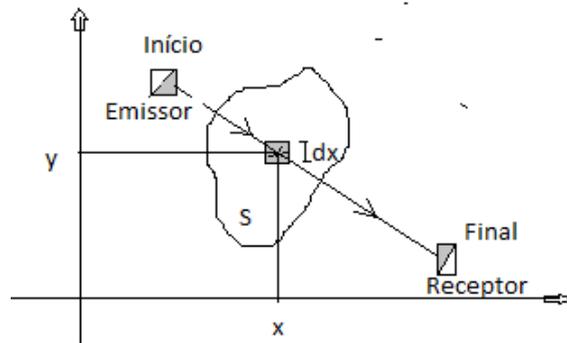
3.1 Modelo Matemático da Tomografia Computadorizada Usando Sistemas Lineares

A ideia central da construção de imagem pela TC é o de produzir uma imagem de uma parte interna do corpo humano através da medição da intensidade dos raios-X que o atravessam. Pelas propriedades do raio-X, uma parte deste raio irá passar e a outra será absorvida pelos tecidos, isso depende da densidade dos tecidos, ou seja, tecidos mais densos (como o fígado) ou com elementos mais pesados (como o cálcio presente nos ossos) absorvem mais radiação que tecidos menos densos (como o pulmão que está preenchido com ar). Assim, uma TC quantifica a radiação absorvida por cada parte do corpo analisado (radiodensidade) e traduz essas variações numa escala de cinzentos (veja a Figura 5), produzindo uma imagem.

Para estudarmos a quantidade do raio-X que irá passar, precisamos estudar a sua atenuação (perda gradual de intensidade do raio-X). Para nosso propósito iremos estudar a atenuação do raio-X em um pequeno quadradinho. Para isto, usaremos a seguinte definição que pode ser encontrada em [1] veja a página 57.

Definição 11 *Chamaremos de coeficiente de atenuação linear (ou apenas atenuação) do raio-X no ponto de coordenadas (x, y) de \mathbb{R}^2 a fração da intensidade de um raio-X que é absorvida num quadradinho de lado dx com centro em (x, y) . Indicaremos esse coeficiente por $f(x, y)$.*

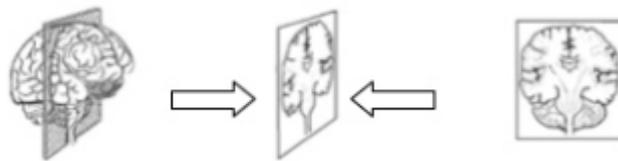
Figura 7: Região planificada



Como tecidos diferentes causem atenuações diferentes, o conhecimento do valor de f em cada ponto nos permite descrever a forma e o tamanho do órgãos que estão presentes na região em estudo. Portanto, encontrar f equivale a encontrar a imagem reconstruída pela TC.

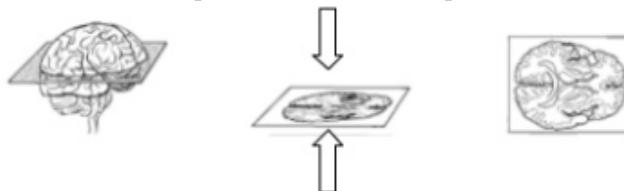
Antes de estudarmos como obter os valores da função f em cada ponto, note que uma imagem em $3D$ nada mais é que a união de várias imagem em $2D$, isto é, se tomarmos uma imagem em $3D$ e fizermos um corte transversal obteremos uma imagem em $2D$, e portanto, a imagem em $3D$ é a união de cortes transversais (imagens em $2D$), de acordo com a Figura 8 e 9. Portanto, para construirmos uma imagem em $3D$ basta construirmos imagens em $2D$ e unirmos essas imagens de forma paralela. Por isso, estudaremos como obter uma imagem em $2D$ em (TC).

Figura 8: Corte transversal



Fonte: [13]

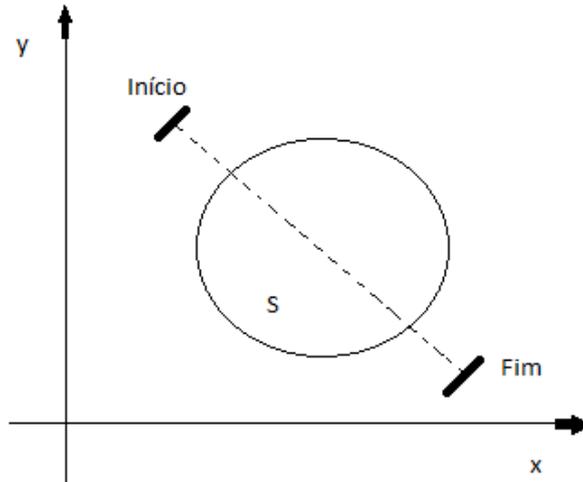
Figura 9: Corte Longitudinal



Fonte: [14]

Suponhamos que S seja um subconjunto do \mathbb{R}^2 no qual queremos obter uma imagem do seu interior (por exemplo, pode-se imaginar que S seja um corte transversal que contenha um órgão, veja Figura 8), e que sobre S disparamos raios-X, no qual medimos a intensidade inicial e final.

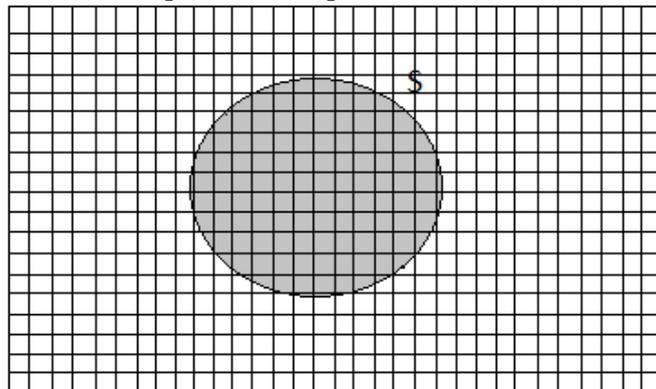
Figura 10: Região S



Fonte: [1]

Para fixar ideias, podemos supor que S esteja dividido em n quadradinhos que chamaremos de pixels de lados Δx_j ($j = 1, 2, \dots, n$). Tomando os pixels suficientemente pequenos, podemos supor que a atenuação é constante em cada pixel, isto é, a função f da Definição 11 é constante. Chamaremos de x_j ao valor de f no j -ésimo pixel. Logo, nossa incógnita não é mais a função f e sim o vetor $x = (x_1, x_2, \dots, x_n)$.

Figura 11: Região planificada

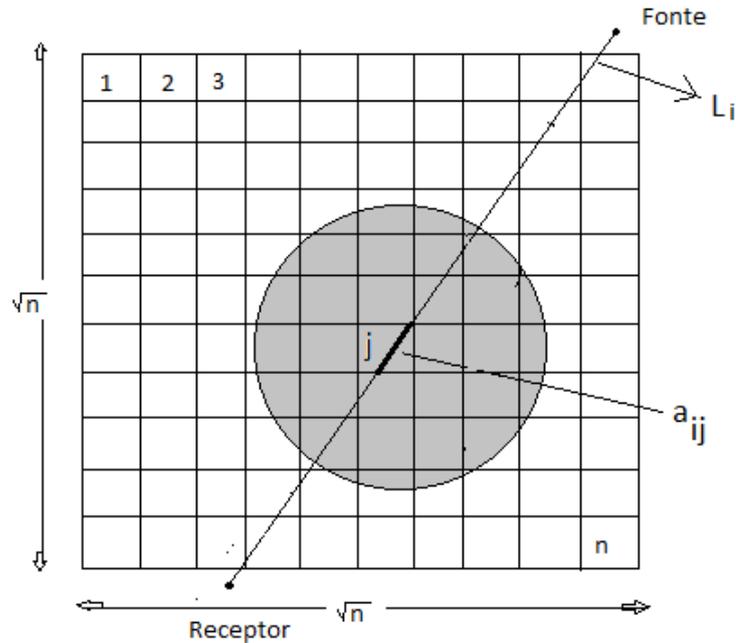


Fonte: [1]

Suponhamos que sejam disparados sobre a região S os raios-X L_1, \dots, L_m , cujas atenu-

ações medimos ao longo de todo S (por exemplo com um detector de partículas), e seja b_i a atenuação do raio-X L_i sobre todo S . Denotemos por a_{ij} o comprimento da interseção do i -ésimo pixel com j -ésimo raio. Assim a_{ij} será zero quando não existe interseção)

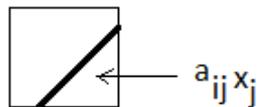
Figura 12: Região planificada



Fonte: [1]

A atenuação sofrida no pixel j pelo comprimento do raio L_i que atravessa esse pixel é dada por $a_{ij}x_j$ como mostra a figura abaixo:

Figura 13: Absorção do raio L_i no pixel j



Então, a atenuação do i -ésimo raio-X, ao longo de S é dado pela soma da atenuação em cada pixel (Figura 12), isto é, a seguinte soma dar a atenuação do i -ésimo raio:

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i.$$

representa as imagens normais para $h(x) = 0$. Porém, se o paciente está fazendo a tomografia é porque se espera (ou se teme) que $x \neq x^0$. Como h é claramente contínuo (veja [4]), o quanto menor for $h(x)$, x representará uma imagem mais próxima do normal. Essa função h é usada para excluir as soluções absurdas como pode ser visto na figura abaixo:

Figura 15: Solução inconsistente



Fonte: [16]

Para mais detalhes sobre essa função h e também para outras funções que ajudem a selecionar as melhores imagens veja [1], página 60.

Agora, nosso próximo objetivo, é estudar métodos de resolução dos sistemas de equações e inequações lineares. Existem vários métodos para determinar as soluções de um sistema de equações lineares ou de inequações lineares executando um número finito de operações; um exemplo clássico, para resolver (4), é o método de *eliminação gaussiana* (escalonamento). Já para resolver o problema (5), pode ser usado o algoritmo de Karmarkar e o algoritmo de Programação Linear (veja [1]). Na prática não se usam esses métodos clássicos para resolver tais sistemas obtidos na TC, pois a forma da matriz A , tomada no problema, vai impactar diretamente no resultado obtido, já que para termos uma boa aproximação, e que a imagem obtida seja nítida, o número n de pixels tomados tem que ser muito grande, pois a qualidade da imagem vai depender diretamente da quantidade de pixels, em que a região for dividida. Para se ter uma ideia, segundo [1], a quantidade mínima de pixels a ser tomada, deve ser de ordem superior a 10^5 e a quantidade de raios incidentes deve ser superior a 10^4 . Portanto, a matriz A de (4) é de ordem superior a $10^5 \times 10^4$, enquanto que a matriz \bar{A} de (5) é superior a $10^{10} \times 10^4$. Portanto, a ordem da matriz A é muito grande para se utilizar esses métodos.

Para a modelagem feita acima supomos que o raio-X é um raio (na verdade é um feixe), para uma modelagem em que o raio-X é fecho indicamos [8]

Os artigos [2] e [8] modelam a tomografia computadorizada usando transformada de Radon. Neste caso, supondo a função f da Definição 11 constante, obtemos o sistema (4), portanto, a modelagem que fizemos aqui é um caso particular da modelagem usando

transformada de Radon.

4 Usando Algoritmos para Resolver Sistemas Lineares

Neste capítulo estudaremos a resolução de sistemas lineares via algoritmos. Nele daremos vários exemplos para ilustrar a aplicação desses algoritmos na resolução de sistemas lineares.

4.1 Os métodos de Gauss-Seidel e Jacobi

Vamos começar pela aplicação dos Algoritmos de Gauss-Seidel e Jacobi, ambos do começo do século XIX. Tais métodos buscam a resolução de sistemas lineares que apresentam muitas equações e muitas incógnitas. O método consiste em, dados um sistema linear, gerar uma sequência de soluções aproximadas, de tal forma que essas soluções estarão tão próximas do verdadeiro resultado o quanto desejarmos, isto é, quanto mais termos dessa sequência obtivermos melhor será a aproximação.

Para uma melhor compreensão do algoritmo de Gauss-Seidel, iremos aplicá-lo para resolver um exemplo em que o sistema (4) possui a matriz A do tipo 2×2 .

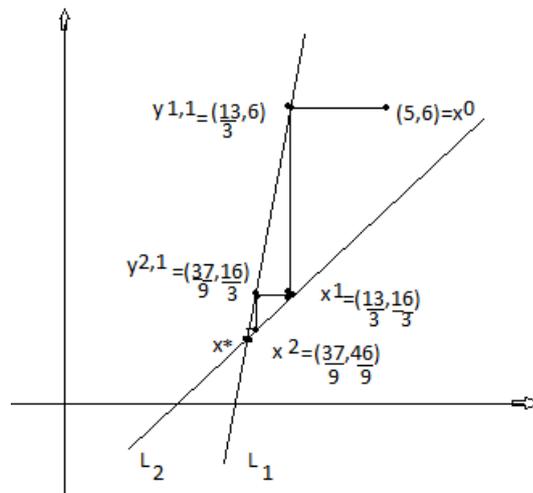
Exemplo 1 *Seja dado o sistema linear*

$$\begin{cases} 3x - y = 7 \\ x - y = -1 \end{cases} \quad (8)$$

cuja solução é o vetor $(4, 5)$. Para resolver tal sistema usando o Algoritmos de Gauss-Seidel, definimos $L_1 := \{(x, y) \in \mathbb{R}^2 \mid 3x - y = 7\}$ e $L_2 := \{(x, y) \in \mathbb{R}^2 \mid x - y = -1\}$. Note que um vetor (x, y) é solução do sistema (8) se, e somente se, $(x, y) \in L_1 \cap L_2$. O Algoritmos de Gauss-Seidel fornece um método de encontrar um ponto de $L_1 \cap L_2$. Para isso, tomemos um ponto $x^0 = (5, 6)$ (pode ser qualquer). Variando a primeira coordenada do ponto x^0 , para satisfazer a equação L_1 , e deixando a segunda coordenada fixa, obtemos o ponto $y^{1,1} = (\frac{13}{3}, 6)$. De posse do ponto $y^{1,1}$, modificamos a sua segunda coordenada para satisfazer a equação L_2 , e deixamos a primeira coordenada fixa, o que nos permite obter o ponto $x^1 = (\frac{13}{3}, \frac{16}{3})$. Modificando agora a primeira coordenada do ponto x^1 , para satisfazer

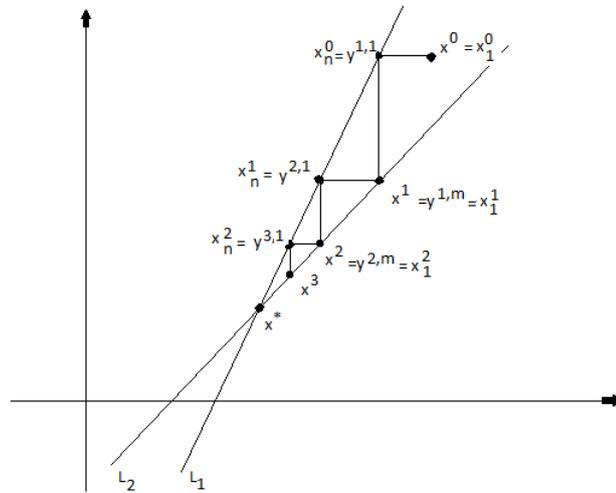
a equação L_1 , e deixamos a segunda coordenada fixa, obtemos o ponto $y^{2,1} = (\frac{37}{9}, \frac{16}{3})$. Agora, de posse do ponto $y^{2,1}$, modificamos sua segunda coordenada para satisfazer a equação L_2 , onde obtemos o ponto $x^2 = (\frac{37}{9}, \frac{46}{9})$, e deixamos a primeira coordenada fixa. Continuando, executando o processo sucessivamente, obteremos uma sequência $\{x^n\}$ que se aproximará da solução do sistema (veja a Figura 16), para ver isto, note que a primeiras coordenada da sequência x^n gera a seguinte sequência: $\{5; \frac{13}{3}; \frac{37}{9}; \dots\} = \{5; 4,33; 4,11; \dots\}$ donde observamos que ela está se aproximando de 4, enquanto que a segunda coordenada de $\{x^n\}$ gera a seguinte sequência: $\{6; \frac{16}{3}; \frac{46}{9}; \dots\} = \{6; 5,33; 5,11; \dots\}$ que se aproxima de 5.

Figura 16: Aplicação do Algoritmo de Gauss-Seidel



Descreveremos agora o algoritmo de Gauss-Seidel, para uma matriz $A \in \mathbb{R}^{n \times n}$ qualquer. Para a utilização deste algoritmo, devemos tomar um ponto inicial x^0 de \mathbb{R}^n . De posse deste ponto, modificamos a sua 1ª componente para satisfazer a 1ª equação (e deixamos o restante das equações fixas), obtemos então o ponto $y^{1,1}$. Agora, modificamos a 2ª componente de $y^{1,1}$, para satisfazer a segunda equação (e deixamos o restante das equações fixas), obtemos então o ponto $y^{1,2}$ e continuamos a iteração das equações até obtermos o ponto $y^{1,m}$, que chamamos de x^1 , isto é, $y^{1,m} = x^1$. Repetimos o procedimento com x^1 no lugar de x^0 , donde obtemos x^2 . Logo, seguindo o procedimento geramos a sequência $\{x^k\}$ com $k = 1, \dots, n$ (veja a Figura 17).

Figura 17: Método de Gauss-Seidel



Fonte: [1]

Observe que a sequência gerada por esse método depende sempre do termo anterior, ou seja, modificamos sempre o anterior para se obter o próximo. Este tipo de algoritmo recebe o nome *sequencial*, pois seus termos são obtidos a partir dos anteriores.

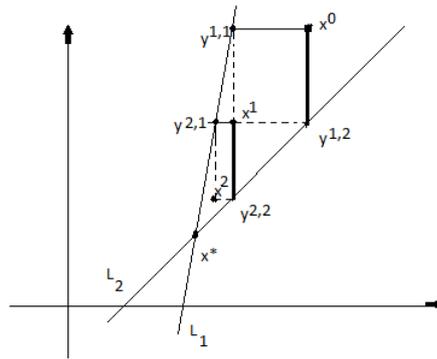
Descreveremos agora o método de Jacobi. Esse método é muito parecido com o de Gauss-Seidel, a diferença é que no método de Jacobi modificamos as coordenadas dos pontos da sequência ao mesmo tempo. Para ficar claro, vamos descrevê-lo agora.

Tomemos um ponto inicial x^0 , modificamos a sua 1ª componente para satisfazer a 1ª equação (e deixamos o restante das equações fixas), obtemos então o ponto $y^{1,1}$; modificamos a 2ª variável de x^0 , para satisfazer a segunda equação (e deixamos o restante das variáveis fixas), obtemos então o ponto $y^{1,2}$; modificamos a 3ª componente de x^0 , para satisfazer a terceira equação (e deixamos o restante das variáveis fixas), obtemos então o ponto $y^{1,3}$ e continuamos a iteração das equações até obtermos o ponto $y^{1,m}$.

Definamos o ponto $x^1 \in \mathbb{R}^m$ como sendo a primeira coordenada de $y^{1,1}$ e a segunda coordenada de $y^{1,2}$. Repetimos o procedimento com x^1 no lugar de x^0 , donde obtemos x^2 . Logo, seguindo o procedimento geramos a sequência $\{x^k\}$ com $k = 1, \dots, n$. Observe que calculamos os termos da sequência auxiliar $y^{i,j}$ de maneira paralela, todas ao mesmo tempo, para depois tomarmos o termo da sequência $\{x^k\}$, por isso, tal algoritmo é chamado de *paralelo*.

Antes de analisarmos a convergência destes métodos, observe que, do ponto de vista geométrico, conforme Figura 18, a sequência de Gauss-Seidel se aproxima de uma solução única x^* mais rápida que a sequência de Jacobi. Enquanto os vetores auxiliares $y^{k,j}$ (com k fixo) no método de Gauss-Seidel têm que ser calculados sequencialmente, ou seja, para se obter o vetor seguinte devemos conhecer um vetor anterior. Por outro lado, o algoritmo de Jacobi, permite o cálculo simultâneo dos $y^{k,j}$ (com k fixo).

Figura 18: Método de Jacobi



Fonte: [1]

No exemplo a seguir, vamos aplicar o método de Jacobi para a resolução de um sistema linear.

Exemplo 2 Considere o sistema linear:

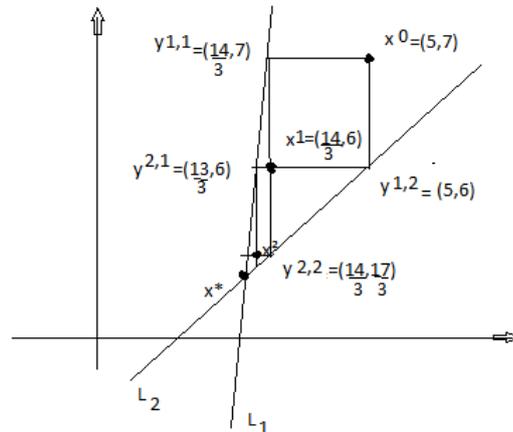
$$\begin{cases} 3x - y = 7 \\ x - y = -1 \end{cases}$$

cuja solução é $(4, 5)$. Para fixar ideias, novamente façamos $L_1 := \{(x, y) \in \mathbb{R}^2 \mid 3x - y = 7\}$, $L_2 := \{(x, y) \in \mathbb{R}^2 \mid x - y = -1\}$ e tomemos o ponto inicial $x^0 = (5, 7)$. Modifiquemos a primeira coordenada de x^0 , e deixemos a segunda coordenada dele fixa, para obtermos o ponto $y^{1,1} = (\frac{14}{3}, 7)$, que é solução da equação L_1 , ainda de posse do ponto x^0 , modifiquemos a segunda coordenada deste ponto, e deixemos a primeira fixa. Daí, obtemos o ponto $y^{1,2} = (5, 6)$ que é solução da equação L_2 . Logo, obtemos o primeiro iterado $x^1 = (\frac{14}{3}, 6)$.

Agora, de posse do ponto $x^1 = (\frac{14}{3}, 6)$, modifiquemos a primeira coordenada deste ponto, e deixemos a segunda fixa. Daí, obtemos o ponto $y^{2,1} = (\frac{13}{3}, 6)$ que é solução da equação L_1 . Ainda, utilizando o ponto x^1 , modifiquemos a segunda coordenada e deixemos a primeira fixa, donde obtemos o ponto $y^{2,2} = (\frac{14}{3}, \frac{17}{3})$, que é solução da equação L_2 . Daí, obtemos o segundo iterado $x^2 = (\frac{13}{3}, \frac{17}{3})$.

Observe que, se continuamos fazendo esse processo, é intuitivo que a sequência, $\{x^k\}$ se aproxima do ponto $(4, 5)$. Para ver isto, note que a sequência das primeiras coordenadas de $\{x^k\}$ é dado por: $\{5; \frac{14}{3}, \frac{13}{3}, \dots\} = \{5; 4,66; 4,33, \dots\}$, e portanto, está se aproximando de 4 (Figura 19). Enquanto que, a sequência das segundas coordenadas de $\{x^k\}$, é dado por: $\{7; 6; \frac{17}{3}; \dots\} = \{7; 6; 5,66; \dots\}$, onde percebemos que ela está se aproximando de 5 conforme Figura 19.

Figura 19: Estudo do exemplo usando o método Jacobi



Segundo [1], num computador com processadores paralelos o método de Jacobi pode ser mais eficiente que o método de Gauss-Seidel. Em geral, no método de Gauss-Seidel é usado apenas um vetor da forma $y^{k,j}$ com k fixo para as variáveis, enquanto que o método de Jacobi são usados dois vetores da forma $y^{j,1}$ e $y^{j,2}$. A distinção entre algoritmos sequenciais, como o de Gauss-Seidel, e paralelos, como o de Jacobi, se mantém quando usamos métodos mais sofisticados.

Discutiremos agora a convergência destes algoritmos. Para tanto, tomemos o exemplo a seguir.

Exemplo 3 Tome o sistema linear

$$\begin{cases} x + y = 0 \\ x - y = 0 \end{cases},$$

cuja solução é, claramente, $(0, 0)$. Defina $L_1 := \{(x, y) \in \mathbb{R}^2 \mid x + y = 0\}$ e $L_2 := \{(x, y) \in \mathbb{R}^2 \mid x - y = 0\}$, e tomemos um ponto $x^0 = (6, 3)$ (pode ser qualquer ponto diferente de $(0, 0)$).

Para resolver tal sistema usando o Algoritmo de Gauss-Seidel, devemos variar a primeira coordenada de x^0 , com o objetivo de satisfazer o conjunto L_1 , e deixar a segunda coordenada fixa, assim obtemos o ponto $y^{1,1} = (-3, 3)$. De posse do ponto $y^{1,1}$, modifiquemos a segunda coordenada desse ponto, e deixemos a primeira variável fixa para satisfazer o conjunto L_2 . Então, obtemos o ponto de coordenadas $y^{1,2} = (-3, -3)$ que no caso é o nosso x^1 .

De posse do ponto x^1 modifiquemos a primeira componente e deixemos a segunda fixa donde obtemos o ponto $y^{2,1} = (-3, 3)$. Tomando agora $y^{2,1}$ modifiquemos a segunda componente para satisfazer o conjunto L_2 . Daí, obtemos o ponto $y^{2,2} = (3, 3)$ que é o nosso x^2 .

Utilizando agora o ponto x^2 , modifiquemos a primeira coordenada deste ponto, para satisfazer o conjunto L_1 donde obtemos o ponto $y^{3,1} = (-3, 3)$, donde percebemos que o

Proposição 2 *Seja um hiperplano $H = \{y \in \mathbb{R}^n : \langle v, y \rangle = \beta\}$ com $v \in \mathbb{R}^n$ e $\beta \in \mathbb{R}$. Dado qualquer $x_0 \in \mathbb{R}^n$, o ponto de H mais próximo de x_0 é dado por:*

$$P_H(x_0) = x_0 - \frac{\langle v, x_0 \rangle - \beta}{\|v\|^2} v,$$

isto é, $P_H(x_0) \in H$ e $\|x_0 - P_H(x_0)\| \leq \|x_0 - y\|$, para todo $y \in H$.

Demonstração: Devemos inicialmente mostrar que o vetor v é normal ao hiperplano H . Para isso, considere $0 = (0, 0, \dots, 0) \in \mathbb{R}^n$ e $\vec{0v} = v - 0 = v$. Dados $x, y \in H$, então o vetor $\vec{xy} = y - x$. Logo

$$\langle \vec{0v}, \vec{xy} \rangle = \langle v - 0, y - x \rangle = \langle v, y - x \rangle = \langle v, y \rangle - \langle v, x \rangle = \beta - \beta = 0.$$

Logo, $\vec{0v} \perp \vec{xy}, \forall x, y \in H$. Portanto, o vetor $\vec{0v}$ é normal à H . Defina a reta $s(t) = x_0 + tv$. Como o vetor $\vec{0v}$ é normal ao plano H , segue que a reta $s(t)$ é normal ao plano H , e portanto, intersecta o plano H em um único ponto, digamos $s(t_0)$.

Logo, $s(t_0) \in H$, e portanto,

$$\langle s(t_0), v \rangle = \beta \Leftrightarrow \langle x_0 + t_0 v, v \rangle = \beta \Leftrightarrow \langle x_0, v \rangle + t_0 \langle v, v \rangle = \beta \Leftrightarrow t_0 = \frac{\beta - \langle x_0, v \rangle}{\langle v, v \rangle}.$$

Portanto, $s(t_0) = x_0 + t_0 v = x_0 + \frac{\beta - \langle x_0, v \rangle}{\langle v, v \rangle} v$. Como $\langle v, v \rangle = \|v\|^2$ então:

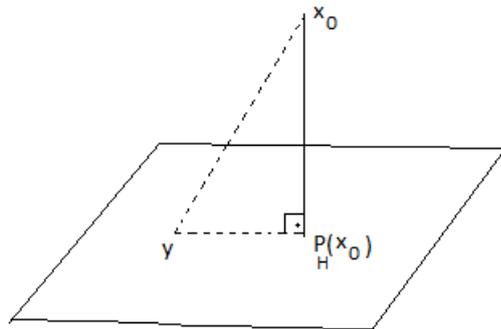
$$s(t_0) = x_0 + \frac{\beta - \langle x_0, v \rangle}{\|v\|^2} v.$$

Fazendo $s(t_0) = P_H(x_0)$, temos:

$$P_H(x_0) = x_0 - \frac{\langle x_0, v \rangle - \beta}{\|v\|^2} v \in H. \quad (9)$$

Mostremos agora que $P_H(x_0)$ é o ponto mais próximo de x_0

Figura 21: Projeção de ponto

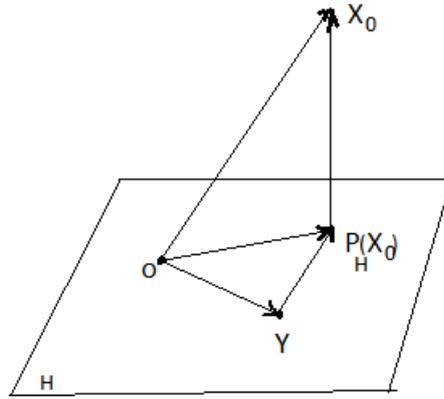


De fato, dado qualquer $y \in H$, como x_0, y e $P_H(x_0)$ são os vértices de um triângulo retângulo no ponto $P_H(x_0)$ (veja a Figura 21), pois a reta S que contém x_0 e $P_H(x_0)$ é normal à H , segue que:

$$\|P_H(x_0) - x_0\| \leq \|y - x_0\|, \quad \forall y \in H,$$

pois o segmento $\overrightarrow{yx_0}$ é a hipotenusa no triângulo retângulo de vértices $y, x_0, P_H(x_0)$. Provando a afirmação. \square

Figura 22: Projeção ideal



Provaremos agora que para cada $x_0 \in \mathbb{R}^n$, então $P_H(x_0)$ é único, isto é, ele é a aproximação ideal de x_0 a H .

Proposição 3 (*Aproximação ideal*) *Seja H um hiperplano. Se $x_0 \notin H$ e $P_H(x_0)$ a projeção de x_0 sobre H , então qualquer $y \in H$, com $y \neq P_H(x_0)$, temos:*

$$\|x_0 - P_H(x_0)\| < \|x_0 - y\|, \quad (10)$$

Demonstração: A prova consiste em mostrar que a menor distância do ponto x_0 até sua projeção $P_H(x_0)$ no hiperplano é o vetor perpendicular ao hiperplano traçado por $P_H(x_0)$. Seja $0 = (0, 0, \dots, 0) \in \mathbb{R}^n$, como $x_0 - P_H(x_0)$ é normal a H e $P_H(x_0) - y \in H$, pois $y, P_H(x_0) \in H$, segue que, $x_0 - P_H(x_0)$ é ortogonal a $P_H(x_0) - y$. Logo

$$\begin{aligned} \langle x_0 - y, x_0 - y \rangle &= \langle x_0 - y + P_H(x_0) - P_H(x_0), x_0 - y + P_H(x_0) - P_H(x_0) \rangle \\ &= \langle x_0 - P_H(x_0), x_0 - P_H(x_0) \rangle \\ &\quad + 2\langle x_0 - P_H(x_0), P_H(x_0) - y \rangle + \langle P_H(x_0) - y, P_H(x_0) - y \rangle. \end{aligned}$$

Portanto

$$\|x_0 - y\|^2 = \|x_0 - P_H(x_0)\|^2 + \|P_H(x_0) - y\|^2.$$

Como, por hipótese, $y \neq P_H(x_0)$, então $\|P_H(x_0) - y\| > 0$, o que acarreta

$$\|x_0 - y\|^2 - \|x_0 - P_H(x_0)\|^2 = \|P_H(x_0) - y\|^2 > 0,$$

Segue que,

$$\|x_0 - y\|^2 > \|x_0 - P_H(x_0)\|^2,$$

donde concluimos que

$$\|x_0 - y\| > \|x_0 - P_H(x_0)\|.$$

Portanto, a projeção ortogonal $P_H(x_0)$, de x_0 sobre H , é o *único ponto* de H mais próximo de x_0 . \square

Motivados pelas últimas duas proposições definimos o operador projeção sobre hiperplano.

Definição 12 *Seja H um hiperplano em \mathbb{R}^n . Então, defina o seguinte operador, $P_H : \mathbb{R}^n \rightarrow H$, de forma que, para cada $x \in \mathbb{R}^n$, $P_H(x)$ é o único elemento de H que satisfaz:*

$$\|x - P_H(x)\| = \min_{y \in H} \|x - y\|.$$

Ou seja, $P_H(x)$ é o ponto de H mais próximo de x .

A seguinte observação será muito importante, quando formos provar que os Algoritmos de Kaczmarz e Cimmino convergem.

Observação 1 *Segue direto da definição de operador projeção que: $P_H(x) = x$ se, e somente se, $x \in H$. Portanto, se $H_1, H_2 \subset \mathbb{R}^n$ são hiperplanos, tem-se que: se $x \in H_1 \cap H_2$, então, $P_{H_2}(P_{H_1}(x)) = x$.*

Em posse deste operador, agora estamos prontos para estudarmos os Algoritmos de Kaczmarz e Cimmino. Isto será feito nas próximas seções. Começaremos pelo Algoritmo de Kaczmarz, depois estudaremos o Algoritmo de Cimmino, ambos serão utilizados para resolver equações lineares, e posteriormente para resolver inequações.

4.3 Algoritmos de Kaczmarz

Começaremos apresentando a ideia geométrica do algoritmo de Kaczmarz, para depois descreveremos o algoritmo de forma algébrica e, por fim, o colocaremos sob ponto de vista de algoritmo computacional. Esse algoritmo consiste em projeções sequenciais sobre os hiperplanos, de forma que essas projeções criem uma sequência que convergirá para um ponto da interseção desses hiperplanos.

Para entender a ideia geométrica do Algoritmos de Kaczmarz, suponhamos dois hiperplanos H_1 e H_2 , com interseção não vazia da seguinte forma:

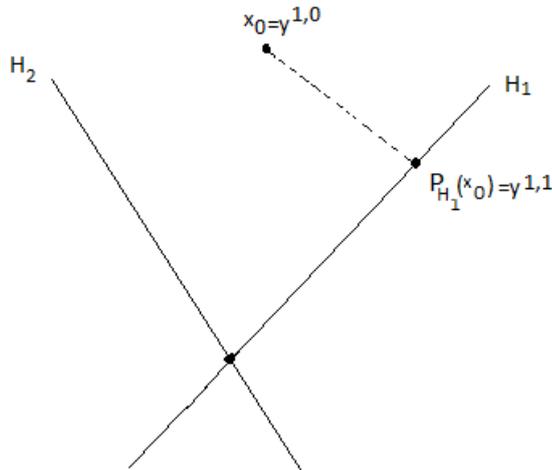
i) $H_1 = \{x \in \mathbb{R}^n | \langle x, v^1 \rangle = b_1\}$

ii) $H_2 = \{x \in \mathbb{R}^n | \langle x, v^2 \rangle = b_2\}$

onde $v^1, v^2 \in \mathbb{R}^n$ e $b_1, b_2 \in \mathbb{R}$ são fixos.

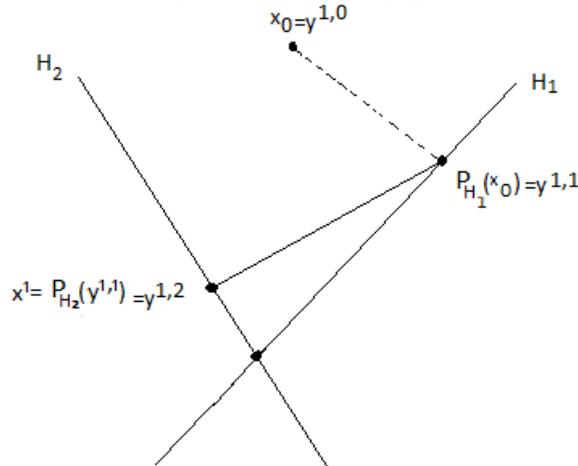
O método começa no seguinte: Dado um ponto inicial $x_0 \in \mathbb{R}^n$ que denotemos por $y^{1,0}$, ou seja, $x_0 = y^{1,0}$. Daí, projeta-se $y^{1,0} = x_0$ sobre H_1 obtendo o ponto $P_{H_1}(x_0) = y^{1,1}$. Conforme a figura a seguir:

Figura 23: 1ª projeção

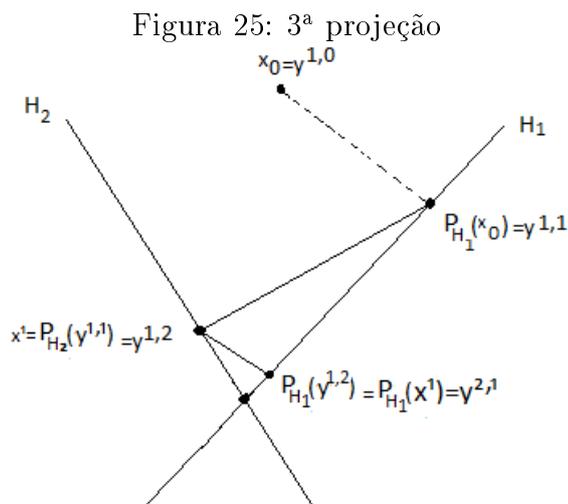


De posse do ponto $y^{1,1}$, projeta-se agora o ponto $y^{1,1}$ sobre o hiperplano H_2 , obtendo-se $P_{H_2}(y^{1,1}) = y^{1,2}$ conforme a figura a seguir:

Figura 24: 2ª projeção



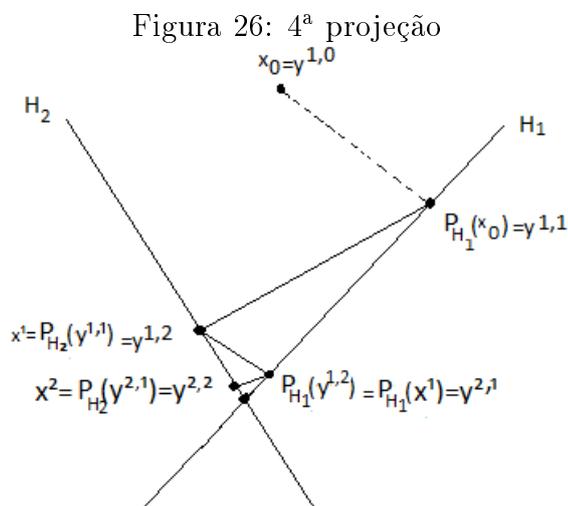
Projeta-se agora $y^{1,2} = x^1$ sobre H_1 , obtendo o ponto $P_{H_1}(y^{1,2}) = P_{H_1}(x^1) = y^{2,1}$ conforme a figura a seguir:



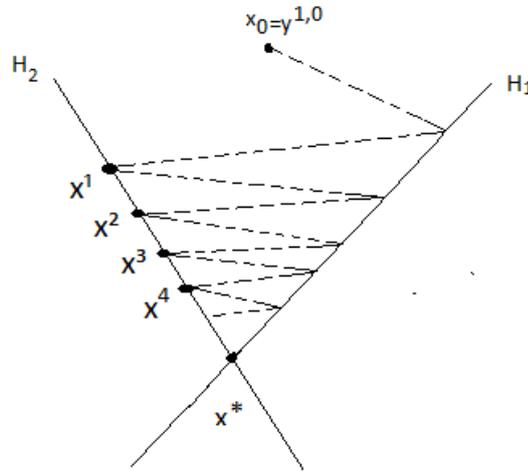
Projeta-se agora $y^{2,1} = P_{H_1}(y^{1,2}) = P_{H_1}(x^1)$ sobre H_2 , obtendo o ponto,

$$x^2 = P_{H_2}(y^{2,1}) = y^{2,2},$$

conforme a figura a



Toma-se $y^{2,2} = x^2$ e repete-se o processo acima, com x^2 em vez de x^1 , isto é, projeta-se o ponto x^2 sobre H_1 obtendo o ponto $P_{H_1}(y^{2,2}) = P_{H_1}(x^2) = y^{3,1}$, depois projeta-se $y^{3,1}$ sobre H_2 , obtendo $P_{H_2}(y^{3,1}) = y^{3,2} = x^3$.

Figura 27: Sequência de soluções centradas em H_2


Obtendo assim uma sequência da forma $\{x^n\} = \{x^0, x^1, x^2, x^3, \dots\}$.

Iremos agora descrever o método do ponto de vista computacional, para isto usaremos a Proposição 2. Suponhamos, novamente, $m = 2$. O computador deve fazer os seguintes passos:

Dado qualquer ponto $x^0 \in \mathbb{R}^n$ ele deve executar iterativamente a projeção desse ponto sobre todos os hiperplanos. Para isso devemos tomar $y^{1,0} = x^0$, e executar os seguintes passos:

1. $y^{1,1} = P_{H_1}(x^0) = y^{1,0} - \frac{\langle v^1, y^{1,0} \rangle - b_1}{\|v^1\|^2} v^1$
2. $y^{1,2} = P_{H_2}(y^{1,1}) = y^{1,1} - \frac{\langle v^2, y^{1,1} \rangle - b_2}{\|v^2\|^2} v^2$

Toma-se $y^{1,2} = y^{2,0} = x^1$, e repetimos o procedimento:

1. $y^{2,1} = P_{H_1}(x^1) = P_{H_1}(y^{2,0}) = y^{2,0} - \frac{\langle v^1, y^{2,0} \rangle - b_1}{\|v^1\|^2} v^1$
2. $y^{2,2} = P_{H_2}(y^{2,1}) = y^{2,1} - \frac{\langle v^2, y^{2,1} \rangle - b_2}{\|v^2\|^2} v^2$

Toma-se $y^{2,2} = y^{3,0} = x^2$, e repetimos o processo. Suponhamos que tenhamos repetido o processo, até o termo x^{k-1} , isto é, $y^{k-1,2} = y^{k,0} = x^{k-1}$. Para obtermos o termo x^k da sequência, assim

1. $y^{k,1} = P_{H_1}(y^{k,0}) = y^{k,0} - \frac{\langle v^1, y^{k,0} \rangle - b_1}{\|v^1\|^2} v^1$
2. $y^{k,2} = P_{H_2}(y^{k,1}) = y^{k,1} - \frac{\langle v^2, y^{k,1} \rangle - b_2}{\|v^2\|^2} v^2$

Tome $y^{k,2} = y^{k+1,0} = x^k$.

O algoritmo de Kaczmarz do ponto de vista computacional, pode ser resumido assim:

Algoritmo de Kaczmarz:

(0) **Inicialização:**

$$x^0 \in \mathbb{R}^n$$

(1) **Passo iterativo:** Dado x^k , defina: $y^{k+1,0} = x^k$. Calculemos sequencialmente as iterações:

$$y^{k+1,i} = y^{k+1,i-1} - \frac{(\langle v^i, y^{k+1,i-1} \rangle - b_i)v^i}{\|v^i\|^2} \quad (1 \leq i \leq m) \quad (11)$$

(2) Faça $x^{k+1} = y^{k+1,m}$ e **volte ao passo (1)**.

Logo, o algoritmo de Kaczmarz, consiste em projetar sequencialmente o ponto x^k sobre todos os hiperplanos $H_i = \{x \in \mathbb{R}^n | \langle v^i, x \rangle = \beta_i\}$ que determinam o sistema $Ax = b$, para obter o ponto x^{k+1} .

No nosso próximo exemplo, iremos aplicar o Algoritmo de Kaczmarz para o caso $m = 2$.

Exemplo 4 *Sejam então os hiperplanos $H_1 := \{(x, y) \in \mathbb{R}^2 | 3x - y = 7\}$ e $H_2 := \{(x, y) \in \mathbb{R}^2 | x - y = -1\}$ que representa o sistema:*

$$\begin{cases} 3x - y = 7 \\ x - y = -1 \end{cases}$$

Para a aplicação do Algoritmo de Kaczmarz tomemos $x^0 = (5, 6)$, e vamos aplicar o operador projeção para encontrar o ponto $y^{1,1}$. Pela definição do algoritmo temos:

$$y^{1,1} = P_{H_1}(5, 6) = (5, 6) - \frac{(\langle (3, -1), (5, 6) \rangle - 7)(3, -1)}{\|(3, -1)\|^2} = \left(\frac{44}{10}, \frac{62}{10}\right),$$

Agora, de posse do ponto $y^{1,1}$ projetamos sobre o H_2 para obtermos o ponto $y^{1,2}$. Pela definição de projeção ortogonal sobre hiperplano, temos:

$$y^{1,2} = P_{H_2}\left(\frac{44}{10}, \frac{62}{10}\right) = \left(\frac{44}{10}, \frac{62}{10}\right) - \frac{(\langle (1, -1), (\frac{44}{10}, \frac{62}{10}) \rangle - (-1))(1, -1)}{\|(1, -1)\|^2} = (1, -1),$$

onde obtemos o ponto $y^{1,2} = (\frac{48}{10}, \frac{58}{10})$ e façamos $x^1 = y^{1,2} = (\frac{48}{10}, \frac{58}{10})$.

De posse do ponto x^1 , projetamos sobre H_1 . Segue do algoritmo que

$$y^{2,1} = P_{H_1}\left(\frac{48}{10}, \frac{58}{10}\right) = \left(\frac{48}{10}, \frac{58}{10}\right) - \frac{(\langle (3, -1), (\frac{44}{10}, \frac{62}{10}) \rangle - 7)(3, -1)}{\|(3, -1)\|^2} = \left(\frac{432}{100}, \frac{596}{100}\right)$$

Agora, de posse do ponto $y^{2,1}$ efetuemos sua projeção sobre H_2 . Pela definição de projeção temos:

$$y^{2,2} = P_{H_2} \left(\frac{432}{100}, \frac{596}{100} \right) = \left(\frac{432}{100}, \frac{596}{100} \right) - \frac{\langle (1, -1), \left(\frac{432}{100}, \frac{596}{100} \right) \rangle - (-1)}{\|(1, -1)\|^2} (1, -1) = \left(\frac{464}{100}, \frac{564}{100} \right)$$

Assim $x^2 = \left(\frac{464}{100}, \frac{564}{100} \right)$. De maneira análoga obtemos $y^{3,1} = \left(\frac{4256}{1000}, \frac{5768}{1000} \right)$ que é a projeção de do ponto x^2 sobre H_1 . Projetamos agora o ponto $y^{3,1}$ de H_1 sobre H_2 , obtemos $x^3 = y^{3,2} = \left(\frac{4512}{1000}, \frac{5512}{1000} \right)$.

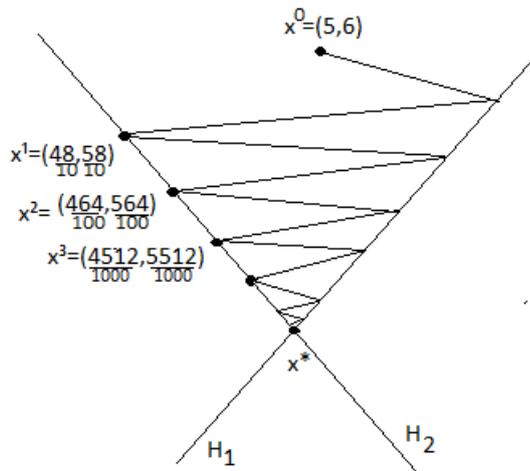
Repetindo o processo indutivamente, obtemos a sequência:

$$\begin{aligned} x^0 &= (5, 6) \\ x^1 &= \left(\frac{48}{10}, \frac{58}{10} \right) = (4, 8; 5, 8) \\ x^2 &= \left(\frac{464}{100}, \frac{564}{100} \right) = (4, 64; 5, 64) \\ x^3 &= \left(\frac{4512}{1000}, \frac{5512}{1000} \right) = (4, 512; 5, 512) \\ &\dots = \dots \end{aligned}$$

Segue daí, que a sequência formada pelas primeiras coordenadas de $\{x^n\}$ é da forma $\{5; \frac{48}{10}; \frac{464}{100}; \frac{4512}{1000}; \dots\} = \{5; 4, 8; 4, 64; 4, 512; \dots\}$, que tende a 4. Enquanto que a sequência formada pelas segundas coordenadas de $\{x^n\}$ é dado por: $\{6; \frac{58}{10}; \frac{564}{100}; \frac{5512}{1000}; \dots\} = \{6; 5, 8; 5, 64; 5, 512; \dots\}$ que tende a 5.

A Figura 28 representa o Algoritmo de Kaczmarz. Dado um ponto inicial qualquer x^0 obtemos o ponto x^1 , depois a partir de x^1 , obtemos o ponto x^2 , e assim sucessivamente.

Figura 28: Sequência de soluções centradas em H_2 para o caso $m=2$



Pelos exposto anteriormente, se denotarmos $P_1 = P_{H_1}$, $P_2 = P_{H_2}$ e $F = P_2 \circ P_1$, onde \circ significa composição de funções, então a sequência $\{x^n\}$ gerada pelo algoritmo de Kaczmarz é dado por: $x^1 = F(x^0)$, $x^2 = F \circ F(x^0)$, $x^3 = F \circ F \circ F(x^0), \dots$, isto é, $x^{n+1} = F^n(x^0)$, onde $F^n = F \circ F^{n-1}$ para $n \geq 0$ Para o caso geral, seja $P_i = P_{H_i}$, para cada $i = 1, 2, \dots, m$, e $F = P_m \circ P_{m-1} \circ \dots \circ P_1$, então a sequência $\{x^n\}$ gerada pelo algoritmo é

dado por: $x^1 = F(x^0)$, $x^2 = F \circ F(x^0)$, $x^3 = F \circ F \circ F(x^0)$, ..., isto é, $x^{n+1} = F^n(x^0)$. No apêndice provaremos que esta sequência vai convergir para um elemento de $H = \bigcap_{i=1}^m H_i$.

4.4 Algoritmo de Cimmino

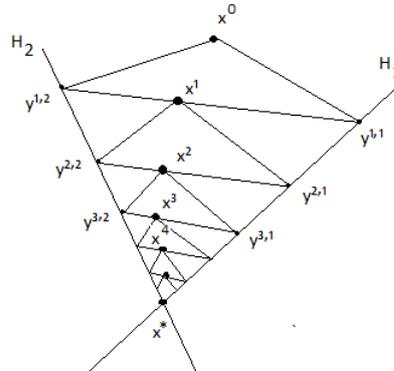
Enquanto que no Algoritmo de Kaczmarz, as projeções são efetuadas sequencialmente, no Algoritmo de Cimmino as projeções são efetuadas paralelamente, e depois pega-se a combinação convexa dessas projeções. Assim como no método de Gauss-Seidel e Kaczmarz devemos tomar um ponto inicial x^0 . Dado esse ponto devemos executar paralelamente sua projeção sobre todos os hiperplanos. Para uma melhor compreensão deste algoritmo, começaremos descrevendo como ele funciona quando $m = 2$. Para isso, consideraremos os hiperplanos:

i) $H_1 = \{x \in \mathbb{R}^n | \langle x, v^1 \rangle = b_1\}$

ii) $H_2 = \{x \in \mathbb{R}^n | \langle x, v^2 \rangle = b_2\}$

onde $v^1, v^2 \in \mathbb{R}^n$ e $b_1, b_2 \in \mathbb{R}$ são fixos.

Figura 29: Método de Cimmino



Fonte: [1]

Dado um ponto inicial x^0 e os números reais $\lambda_1, \lambda_2 > 0$ (graus de liberdade), tais que $\sum_{i=1}^2 \lambda_i = 1$. Tomando $x^0 = y^{1,0}$, e executando a projeção do ponto $x^0 = y^{1,0}$ sobre os hiperplanos, temos:

1. $P_{H_1}(x^0) = P_{H_1}(y^{1,0}) = y^{1,1}$

2. $P_{H_2}(x^0) = P_{H_2}(y^{1,0}) = y^{1,2}$.

Somando essas projeções $y^{1,1}, y^{1,2}$ com os pesos λ_1, λ_2 , obteremos o primeiro termo da nossa sequência, isto é,

$$x^1 = \lambda_1 y^{1,1} + \lambda_2 y^{1,2} = \sum_{i=1}^2 \lambda_i P_{H_i}(y^{1,0}).$$

Usando a Proposição 2 temos:

$$\begin{aligned} x^1 &= \sum_{i=1}^2 \lambda_i P_{H_i}(y^{1,0}) = \sum_{i=1}^2 \lambda_i \left[y^{1,0} - \frac{(\langle v^i, y^{1,0} \rangle - b_i)}{\|v^i\|^2} v^i \right] \\ &= \sum_{i=1}^2 \lambda_i y^{1,0} - \sum_{i=1}^2 \lambda_i \left[\frac{(\langle v^i, y^{1,0} \rangle - b_i)}{\|v^i\|^2} v^i \right]. \end{aligned}$$

Tomando-se $x^1 = y^{2,0}$, e refazendo o processo para obtemos o termo x^2 da sequência. Isto é, projeta-se x^1 , sobre os hiperplanos para obtermos:

1. $P_{H_1}(x^1) = P_{H_1}(y^{2,0}) = y^{2,1}$
2. $P_{H_2}(x^1) = P_{H_2}(y^{2,0}) = y^{2,2}$

Seja x^2 a combinação convexa das projeções $y^{2,1}, y^{2,2}$ referente aos pesos λ_1, λ_2 , isto é,

$$\begin{aligned} x^2 &= \sum_{i=1}^2 \lambda_i P_{H_i}(x^1) \\ &= \sum_{i=1}^2 \lambda_i \left[x^1 - \frac{(\langle v^i, x^1 \rangle - b_i)}{\|v^i\|^2} v^i \right] \\ &= \sum_{i=1}^2 \lambda_i x^1 - \sum_{i=1}^2 \lambda_i \left[\frac{(\langle v^i, x^1 \rangle - b_i)}{\|v^i\|^2} v^i \right] \\ &= x^1 - \sum_{i=1}^2 \lambda_i \left[\frac{(\langle v^i, x^1 \rangle - b_i)}{\|v^i\|^2} v^i \right] \end{aligned}$$

Tome $x^2 = y^{3,0}$ e repita o processo até obter o termo x^{k-1} . Para obter o termo x^k , tomemos $x^{k-1} = y^{k,0}$, e projetando-se sobre os hiperplanos obtemos:

1. $P_{H_1}(y^{k,0}) = y^{k,1}$
2. $P_{H_2}(y^{k,0}) = y^{k,2}$

Fazendo a combinação convexa, obtemos:

$$\begin{aligned}
 x^k &= \sum_{i=1}^2 \lambda_i P_{H_i}(x^{k-1}) \\
 &= \sum_{i=1}^2 \lambda_i \left[x^{k-1} - \frac{(\langle v^i, x^{k-1} \rangle - b_i)}{\|v^i\|^2} v^i \right] \\
 &= \sum_{i=1}^2 \lambda_i x^{k-1} - \sum_{i=1}^2 \lambda_i \left[\frac{(\langle v^i, x^{k-1} \rangle - b_i)}{\|v^i\|^2} v^i \right] \\
 &= x^{k-1} - \sum_{i=1}^2 \lambda_i \left[\frac{(\langle v^i, x^{k-1} \rangle - b_i)}{\|v^i\|^2} v^i \right].
 \end{aligned}$$

Portanto, obtemos:

$$x^k = x^{k-1} - \sum_{i=1}^2 \lambda_i \left[\frac{(\langle v^i, x^{k-1} \rangle - b_i)}{\|v^i\|^2} v^i \right].$$

De forma computacional, o algoritmo de Cimmino, é descrito da seguinte maneira:

Algoritmo de Cimmino:

(0) **Inicialização:**

$$x^0 \in \mathbb{R}^n \tag{12}$$

(1) **Passo iterativo:** Dado x^k , calcule:

$$x^{k+1} = x^k - \sum_{i=1}^m \lambda_i \left[\frac{(\langle v^i, x^k \rangle - b_i)}{\|v^i\|^2} v^i \right]$$

Onde $\lambda_i > 0$, para todo $i = 1, 2, \dots, m$ tal que $\sum_{i=1}^m \lambda_i = 1$.

Note que no algoritmo de Cimmino calcula-se de forma paralela (independente) todas as projeções, e depois, toma-se sua combinação convexa como sendo os termos da sequência, por isso, esse algoritmo é dito algoritmo em paralelo.

Vamos fazer um exemplo ilustrativo de aplicação do Algoritmo de Cimmino na prática, para o caso $m = 2$.

Exemplo 5 *Sejam os hiperplanos, $H_1 := \{(x, y) \in \mathbb{R}^2 \mid 3x - y = 7\}$ e $H_2 := \{(x, y) \in \mathbb{R}^2 \mid x - y = -1\}$.*

Tomemos um ponto inicial $x^0 = (5, 7)$ e os pesos $\lambda_1 = 0,6$ e $\lambda_2 = 0,4$. Daí, usando a, Proposição 2, projetamos o ponto x^0 sobre H_1 e obtemos o ponto $y^{1,1}$. Projetamos x^0

sobre H_2 , e obtemos o ponto $y^{1,2}$.

$$y^{1,1} = P_{H_1}(5, 7) = (5, 7) - \frac{\langle (3, -1), (5, 7) \rangle - 7}{\|(3, -1)\|^2}(3, -1) = \left(\frac{47}{10}, \frac{71}{10}\right),$$

$$y^{1,2} = P_{H_2}(5, 7) = (5, 7) - \frac{\langle (1, -1), (5, 7) \rangle + 1}{\|(1, -1)\|^2}(1, -1) = \left(\frac{11}{2}, \frac{13}{2}\right).$$

Tome $x^1 = \lambda_1 y^{1,1} + \lambda_2 y^{1,2}$, ou seja

$$x^1 = 0,6 \left(\frac{47}{10}, \frac{71}{10}\right) + 0,4 \left(\frac{11}{2}, \frac{13}{2}\right) = \left(\frac{502}{100}, \frac{686}{100}\right)$$

O próximo passo é obter o ponto x^2 . Para tanto, devemos executar a projeção do ponto x^1 sobre os hiperplanos H_1 e H_2 . Veja que:

$$y^{2,1} = P_{H_1}\left(\frac{502}{100}, \frac{686}{100}\right) = \left(\frac{502}{100}, \frac{686}{100}\right) - \frac{\langle (3, -1), \left(\frac{502}{100}, \frac{686}{100}\right) \rangle + 1}{\|(3, -1)\|^2}(3, -1) = \left(\frac{466}{100}, \frac{698}{100}\right);$$

$$y^{2,2} = P_{H_2}\left(\frac{502}{100}, \frac{686}{100}\right) = \left(\frac{502}{100}, \frac{686}{100}\right) - \frac{\langle (1, -1), \left(\frac{502}{100}, \frac{686}{100}\right) \rangle + 1}{\|(1, -1)\|^2}(1, -1) = \left(\frac{544}{100}, \frac{644}{100}\right).$$

Assim,

$$x^2 = \lambda_1 y^{2,1} + \lambda_2 y^{2,2} = 0,6 \left(\frac{466}{100}, \frac{698}{100}\right) + 0,4 \left(\frac{544}{100}, \frac{644}{100}\right) = \left(\frac{4972}{1000}, \frac{6764}{1000}\right).$$

Calculando x^3 . Para tanto, vamos projetar, primeiramente o ponto x^2 sobre H_1 para obtermos $y^{3,1}$ e, em sequência projetamos x^2 sobre H_2 para obtermos $y^{3,2}$.

$$\begin{aligned} y^{3,1} = P_{H_1}\left(\frac{4972}{1000}, \frac{6764}{1000}\right) &= \left(\frac{4972}{1000}, \frac{6764}{1000}\right) - \frac{\langle (3, -1), \left(\frac{4972}{1000}, \frac{6764}{1000}\right) \rangle - 7}{\|(3, -1)\|^2}(3, -1) \\ &= \left(\frac{46264}{10000}, \frac{68792}{10000}\right), \end{aligned}$$

$$\begin{aligned} y^{3,2} = P_{H_2}\left(\frac{4972}{1000}, \frac{6764}{1000}\right) &= \left(\frac{4972}{1000}, \frac{6764}{1000}\right) - \frac{\langle (1, -1), \left(\frac{4972}{1000}, \frac{6764}{1000}\right) \rangle + 1}{\|(1, -1)\|^2}(1, -1) \\ &= \left(\frac{5368}{1000}, \frac{6368}{1000}\right). \end{aligned}$$

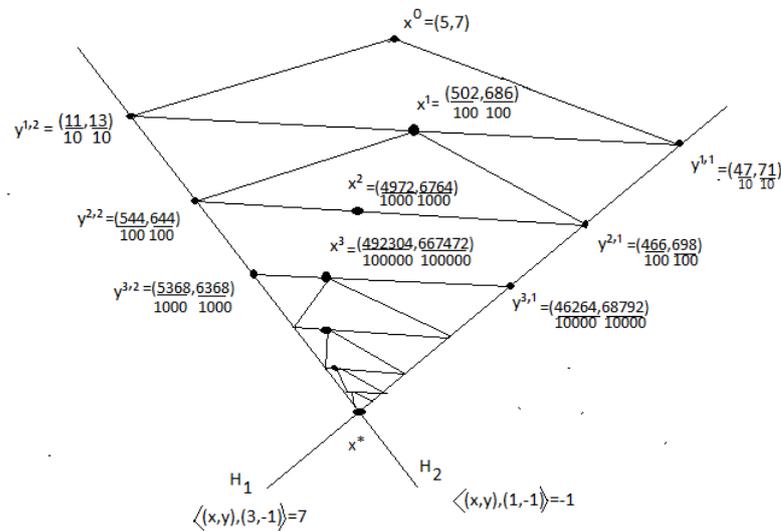
Portanto,

$$x^3 = \lambda_1 y^{3,1} + \lambda_2 y^{3,2} = 0,6 \left(\frac{46264}{10000}, \frac{68792}{10000} \right) + 0,4 \left(\frac{5368}{1000}, \frac{6368}{1000} \right) = \left(\frac{492304}{100000}, \frac{667472}{100000} \right).$$

Repetindo o processo, obtemos a sequência $\{x^n\}$. Note que as primeiras coordenadas de $\{x^n\}$ são dadas por: $\{5; \frac{502}{100}; \frac{4972}{1000}; \frac{492304}{100000}, \dots\} = \{5; 5,02; 4,972; 49,2304, \dots\}$ que converge para 4; já a segunda coordenada de $\{x^n\}$ gera a sequência $\{7; \frac{686}{100}; \frac{6764}{1000}; \frac{667472}{100000}, \dots\} = \{7; 6,86; 6,764; 6,67472; \dots\}$ que converge para 5.

Observe a figura a seguir e veja o comportamento do algoritmo de Cimmino neste exemplo.

Figura 30: Aplicação Método de Cimmino



No exemplo anterior, denotando $P_1 = P_{H_1}$, $P_2 = P_{H_2}$ e $S(x) = \lambda_1 P_1(x) + \lambda_2 P_2(x)$, onde $\lambda_1, \lambda_2 > 0$ são os pesos, com $\sum_{i=1}^2 \lambda_i = 1$, então a sequência $\{x^n\}$ gerada pelo algoritmo de Cimmino é dado por $x^1 = S(x^0)$, $x^2 = S \circ S(x^0)$, $x^3 = S \circ S \circ S(x^0), \dots$, isto é, $x^{n+1} = S^n(x)$. Para o caso geral, sejam $P_i = P_{H_i}$, para cada $i = 1, 2, \dots, m$, e $S = \sum_{i=1}^m \lambda_i P_i$, então a sequência $\{x^n\}$ converge para a solução.

No próximo capítulo provaremos que a sequência vai convergir para um elemento de $\cap_{i=1}^m H_i$.

4.5 Resolução de Inequações via algoritmos

O objetivo desta seção é estudar os análogos aos Algoritmos de Karzmarz e Cimmino, para a resolução de inequações lineares. Para isto lembremos que resolver o sistema de inequações lineares do tipo

Algoritmo de Kaczmarz para inequações:

(0) **Inicialização:**

$$x^0 \in \mathbb{R}^n$$

(1) **Passo iterativo:** Dado $y^{k+1,0} = x^k$. Calculemos sequencialmente as iterações:

$$y^{k+1,i} = y^{k+1,i-1} - \frac{1}{\|v^i\|^2} \max\{0, \langle v^i, y^{k+1,i-1} \rangle - b_i\} v^i \quad (1 \leq i \leq m) \quad (14)$$

(2) Faça $x^{k+1} = y^{k+1,m}$ e **volte ao passo (1)**.

Para Cimmino temos o algoritmos de inequações dado por:

Algoritmo de Cimmino para inequações:

(0) **Inicialização:**

$$x^0 \in \mathbb{R}^n \quad (15)$$

(1) **Passo iterativo:** Dado x^k , calcule:

$$x^{k+1} = x^k - \sum_{i=1}^m \frac{\lambda_i}{\|v^i\|^2} \max\{0, \langle v^i, x^k \rangle - b_i\} v^i$$

4.6 Relaxação

Nesta seção estudaremos o grau de liberdade que podemos tomar ao escolher as projeções sobre os hiperplanos ou sobre semi-espacos, isto é, veremos que não é obrigatório que os termos da sequência ou dos algoritmos sejam projeções, bastando tomar um elemento (ou ponto) contido na reta que passa pelos pontos x e sua projeção $P_H(x)$.

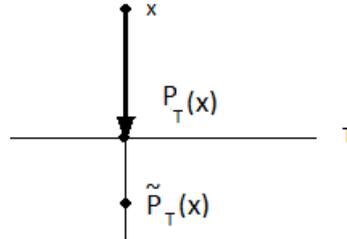
Considere x e $P_H(x)$, então a reta que passa por x e $P_H(x)$ e $(1 - \alpha)x + \alpha P_H(x)$, com $\alpha \in \mathbb{R}$. Nos algoritmos de Kaczmarz e Cimmino foram tomadas as projeções de x sobre H , ou seja, foi tomado $\alpha = 1$. A ideia de relaxação consiste em ter uma liberdade na escolha dos termos das sequências geradas pelos algoritmos. Por exemplo, para $0 < \alpha < 1$, temos os pontos x e $P_H(x)$, em vez de tomar o ponto $P_H(x)$ como feito anteriormente.

Para descrever os Algoritmos de Karzmarz e Cimmino relaxados. Seja $P_H(x)$ a projeção de x sobre o hiperplano H , definimos

$$\tilde{P}_T(x) = (1 - \alpha)x + \alpha P_H(x) \quad (16)$$

para cada $\alpha \in \mathbb{R}$.

Figura 32: Reta ligando x e $P_H(x)$



Agora, para descrever os Algoritmos de Karzmarz e Cimmino relaxados, para inequações, seja $P_T(x)$ a projeção de x sobre o semi-espaco T , definimos:

$$\begin{aligned} \tilde{P}_T(x) &= \alpha P_T(x) + (1 - \alpha)x \\ &= \alpha \left[x - \frac{1}{\|v\|^2} \max\{0, \langle v, x \rangle - \beta\}v \right] + (1 - \alpha)x \\ &= x - \frac{\alpha}{\|v\|^2} \max\{0, \langle v, x \rangle - \beta\}v. \end{aligned}$$

Assim, segue que o algoritmo de Kaczmarz relaxado

Algoritmo de Kaczmarz relaxado:

(0) **Inicialização:**

$$x^0 \in \mathbb{R}^n$$

(1) **Passo iterativo:** Dado $y^{k+1,0} = x^k$. Calculemos sequencialmente as iterações:

$$y^{k+1,i} = y^{k+1,i-1} - \frac{\alpha}{\|v^i\|^2} \max\{0, \langle v^i, y^{k+1,i-1} \rangle - b_i\}v^i \quad (1 \leq i \leq m) \quad (17)$$

(2) Faça $x^{k+1} = y^{k+1,m}$ e volte ao passo (1).

Agora a versão relaxada do Algoritmo de Cimmino é dado por:

Algoritmo de Cimmino relaxado:**(0) Inicialização:**

$$x^0 \in \mathbb{R}^n \quad (18)$$

(1) Passo iterativo: Dado x^k , calcule:

$$x^{k+1} = x^k - \alpha \sum_{i=1}^m \frac{\lambda_i}{\|v^i\|^2} \max\{0, \langle v^i, x^k \rangle - b_i\} v^i$$

Nos algoritmos de Kaczmarz e Cimmino relaxados, se $\alpha = 1$ as iterações recairão sobre o algoritmo original (não relaxado). Observe que não fizemos nenhuma restrição sobre α . Podemos passar da projeção, tomar um ponto entre a projeção e o projetado ou tomar antes do projetado.

Porém, para garantir a convergência desses métodos, veremos na Proposição 8 do apêndice, que devemos ter $0 < \alpha < 2$. Do ponto de vista computacional, segundo [1], se tomamos $\alpha \in (0, 2)$, com $\alpha \neq 1$, podemos ter uma considerável aceleração da convergência. Por isso, foi importante exibirmos os algoritmos *subrelaxado*, (quando $\alpha \in (0, 1)$), e o algoritmo *sobre-relaxado*, (quando $\alpha \in (1, 2)$).

5 Ensinando os Algoritmos em Sala de Aula

Neste capítulo descreveremos como o professor pode ensinar os algoritmos estudados aqui para resolver problemas que envolvam sistemas lineares em sala de aula. Nosso foco neste capítulo, é fugir um pouco do rigor matemático e fornecer uma ideia geométrica e analítica, de como o professor pode, ensinar estes algoritmos em sala de aula.

Inicialmente, o professor deve deixar claro para o aluno que esses algoritmos são muito úteis para resolver sistemas de equações com muita variáveis e, que a ideia deles, é obter soluções aproximadas.

Faremos alguns exemplos ilustrando os algoritmos de algoritmo de Gauss-Seidel, Jacobi, Karzmarz e Cimmino. Iniciaremos pelo Algoritmo de e Gauss-Seidel.

Exemplo 6 *Consideremos o sistema:*

$$\begin{cases} 3x - y = 8 \\ x - y = -3 \end{cases}$$

Note que encontrar pontos x e y que satisfaçam as duas equações acima, equivale a encontrar pontos (x, y) na interseção das retas: $L_1 = \{(x, y) \in R^2 | 3x - y = 8\}$ e $L_2 = \{(x, y) \in R^2 | x - y = -3\}$. Vamos aplicar o algoritmo de Gauss-Seidel para a resolução do problema. Para tanto, observe a Figura 33.

De posse do ponto $x^0 = (8, 9)$, ou qualquer outro ponto, mantendo fixa a segunda coordenada e modificando a primeira obtemos o ponto $y^{1,1} = (\frac{17}{3}, 9)$ que é solução da equação (1). De posse do ponto $y^{1,1}$ modifiquemos a segunda coordenada e mantenhamos a primeira fixa e obteremos o ponto $y^{1,2} = (\frac{17}{3}, \frac{26}{3})$ que é o nosso x^1 .

De posse agora do ponto x^1 modifiquemos a primeira coordenada e deixemos a segunda fixa. Daí obtemos o ponto $y^{2,1} = (\frac{50}{9}, \frac{26}{3})$ que é solução da equação (1). Vamos agora modificar a segunda coordenada e deixar a primeira fixa donde obtemos o ponto $y^{2,2} =$

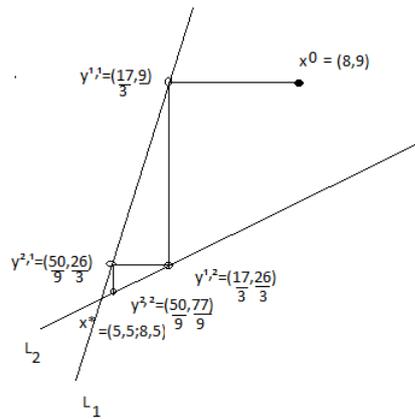
$(\frac{50}{9}, \frac{77}{3})$ que é solução da equação (2) que será o ponto x^2 . Daí, obtemos os pontos:

$$\begin{aligned} x^0 &= (8, 9) \\ x^1 &= (\frac{17}{3}, \frac{26}{3}) \\ x^2 &= (\frac{50}{9}, \frac{77}{9}) \end{aligned}$$

Perceba que, a primeira coordenada da sequência x^n gera a seguinte sequência: $\{8; \frac{17}{3}; \frac{50}{9}; \dots\} = \{8; 5,666666667; 5,555555556; \dots\}$ donde percebemos que ela converge para 5,5.

Se observamos a segunda coordenada de x^n obtemos a sequência $\{9; \frac{26}{3}; \frac{77}{9}; \dots\} = \{9; 8,666666667; 8,555555556; \dots\}$ donde percebemos que ela converge para 8,5.

Figura 33: Aplicação algoritmo de Gauss-Seidel



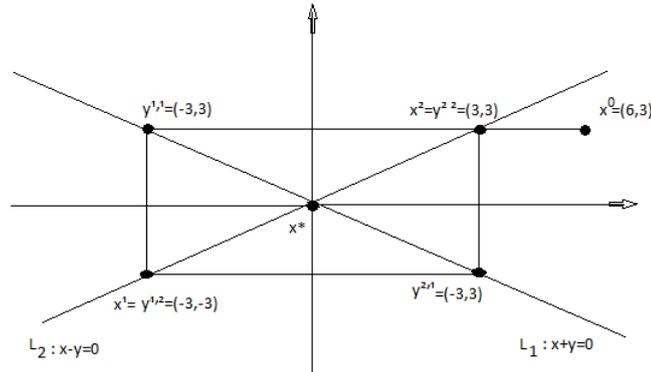
O próximo exemplo trata de um caso em que a utilização do algoritmo de Gauss-Seidel falha

Exemplo 7 Consideremos o sistema:

$$\begin{cases} x - y = 0 & (1) \\ x + y = 0 & (2) \end{cases}$$

cuja a solução é $x^* = (0, 0)$, considere as retas $L_1 = \{(x, y) \in R^2 | x - y = 0\}$ e $L_2 = \{(x, y) \in R^2 | x + y = 0\}$. Para qualquer ponto inicial x^0 , o algoritmo entrar em loop, isto é, vai ficar repetindo os seus valores. Observe a Figura a 34

Figura 34: Loop no Algoritmo de Gauss-Seidel

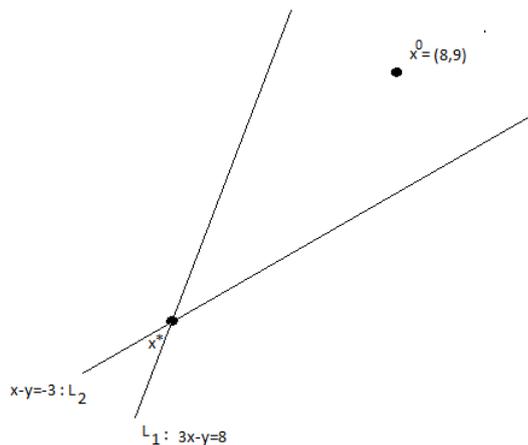


Faremos agora a utilização do algoritmo de Kaczmarz na resolução de um sistema linear.

Exemplo 8 Tomando o mesmo sistema do Exemplo 6 e considere as retas $L_1 = \{(x, y) \in \mathbb{R}^2 | 3x - y = 8\}$ e $L_2 = \{(x, y) \in \mathbb{R}^2 | x - y = -3\}$.

Na resolução desse exemplo não utilizaremos o operador projeção como foi definido na proposição 3, mas vamos utilizar conceitos intuitivos de Geometria Analítica, conteúdo que é lecionado no primeiro bimestre do 3º ano. Observe a Figura 35 onde temos a condição inicial, ou seja, o ponto que é tomado aleatório.

Figura 35: Algoritmo de Kaczmarz

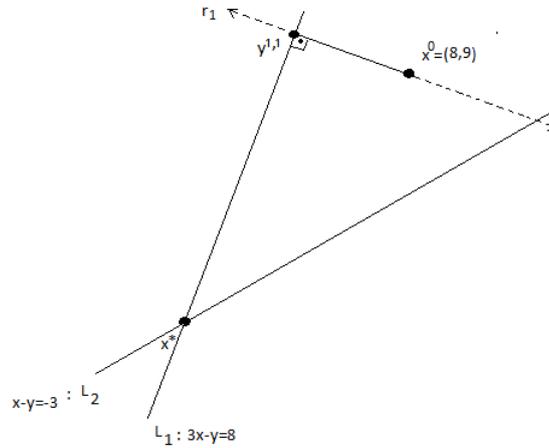


Perceba que, o coeficiente angular da reta L_1 é dado por $m = 3$ e qualquer reta perpendicular a L_1 admite coeficiente angular igual a $m'_1 = \frac{-1}{3}$. Daí, chamemos de r_1 a reta perpendicular a L_1 que contém o ponto $x^0 = (8, 9)$. Logo, a equação da reta r_1 é $r_1 : x + 3y = 35$ e, portanto determinar as coordenadas de $y^{1,1}$ consiste em resolver o sistema:

$$\begin{cases} x + 3y = 35 \\ 3x - y = 8 \end{cases}$$

Donde obtemos o ponto $y^{1,1} = (5, 9; 9, 7)$ conforme observamos na Figura 36

Figura 36: Algoritmo de Kaczmarz



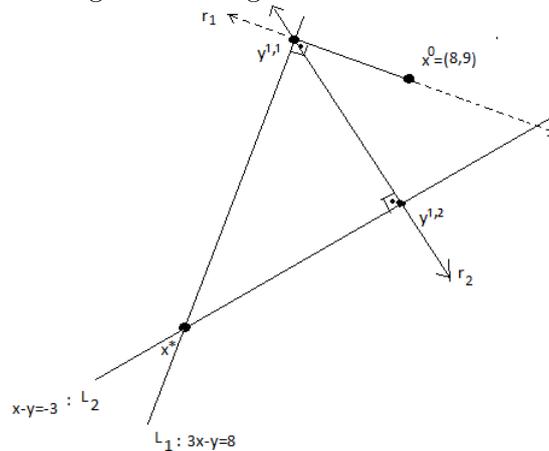
Agora percebemos que existe uma reta (única) r_2 que é perpendicular a L_2 e que passa por $y^{1,1}$. Como o coeficiente angular de L_2 é $m_2 = 1$ segue que qualquer reta perpendicular a L_2 tem coeficiente angular $m'_2 = -1$. Daí, a equação de r_2 é $r_2 : x + y = 15,6$.

Para determinarmos o ponto de coordenadas $y^{1,2}$ precisamos encontrar a solução do sistema a seguir:

$$\begin{cases} x + y = 15,6 \\ x - y = -3 \end{cases}$$

Perceba que determinar $y^{1,2}$ equivale a encontrar a solução da intersecção de $\{r_2\} \cap \{L_1\}$. Daí, resolvendo o sistema anterior obtemos o ponto $y^{1,2} = (6, 3; 9, 3)$ que no nosso caso é o x^1 conforme se observa na Figura 37

Figura 37: Algoritmo de Kaczmarz



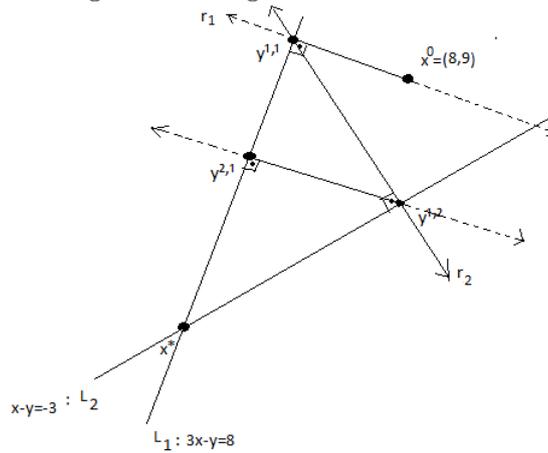
Da maneira análoga, existe uma reta r_3 perpendicular a L_1 cuja equação é dada por

$r_3 : 3y + x = 34,2$. Logo, encontrar o ponto $y^{2,1}$ equivale a resolver o sistema:

$$\begin{cases} x + 3y = 34,2 \\ 3x - y = 8 \end{cases}$$

que significa encontrar o conjunto solução de $\{r_3\} \cap \{L_1\}$ donde obtemos o ponto $y^{2,1} = (5,82; 9,46)$ conforme se observa na Figura 38

Figura 38: Algoritmo de Kaczmarz

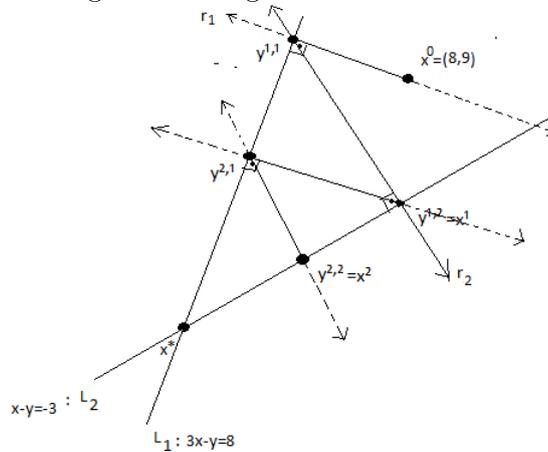


De maneira análoga escrevemos a equação de $r_4 : x + y = 15,28$ e obter o ponto $y^{2,2}$ o qual encontramos por meio da resolução do sistema:

$$\begin{cases} x + y = 15,28 \\ x - y = -3 \end{cases}$$

Dai, o ponto $y^{2,2} = (6,14; 9,14)$ donde se percebe na Figura 39.

Figura 39: Algoritmo de Kaczmarz



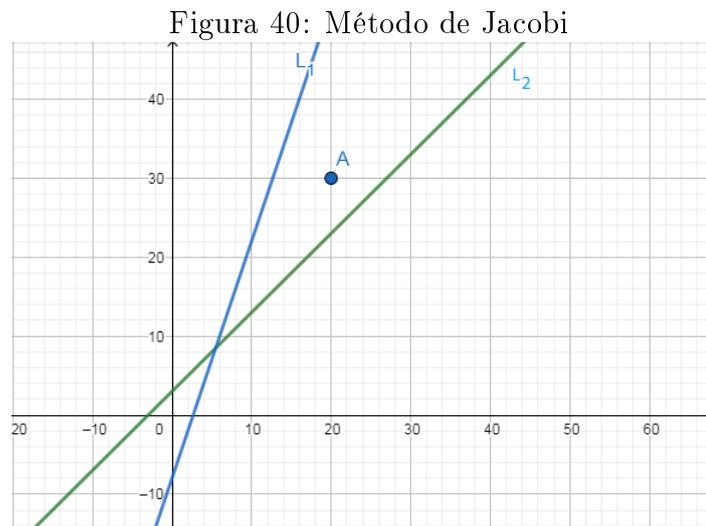
Veja que, a primeira coordenada da sequência $\{x^n\}$ gera a sequência: $\{8; 6,3; 6,14; \cdot\}$ donde percebe-se que ela se aproxima da solução 5,5. De maneira análoga, a segunda coor-

denada de x^n gera a sequência: $\{9; 9, 3; 9, 14; \dots\}$ donde percebemos que ela se aproxima de 8, 5.

Vamos agora resolver o mesmo exemplo utilizando o método de Jacobi.

Exemplo 9 Tomando o mesmo sistema do Exemplo 6 e considerando $L_1 : \{(x, y) \in \mathbb{R}^2 | 3x - y = 8\}$ e $L_2 : \{(x, y) \in \mathbb{R}^2 | x - y = -3\}$. Vamos determinar o ponto que é solução do sistema linear dado por L_1 e L_2 .

Observe a figura a seguir:



Para aplicar o método de Jacobi devemos fazer o seguinte: Dado um ponto inicial qualquer, a saber $A = (20, 30)$ conforme a Figura 40. Para encontrarmos o ponto $y^{1,1}$ devemos traçar por A a reta r_1 paralela ao eixo Ox cuja equação é dada por $r_1 : y = 30$. Daí, basta resolver o sistema:

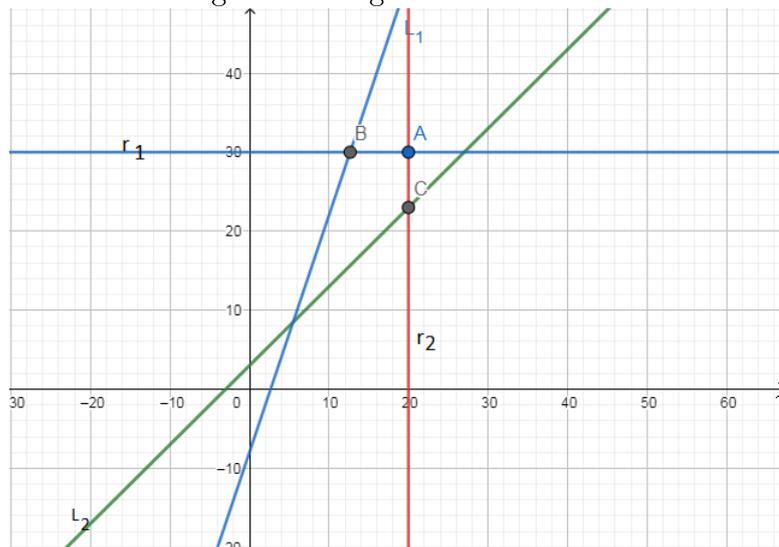
$$\begin{cases} 3x - y = 8 \\ y = 30 \end{cases}$$

donde obtemos ponto $B = y^{1,1} = (12, 66666; 30)$. Agora tracemos pelo ponto $A = x^0$ uma reta r_2 paralela ao eixo Oy de equação $r_2 : x = 20$. Logo, para determinarmos o ponto $C = y^{1,2}$ basta resolvermos o sistema a seguir:

$$\begin{cases} x - y = -3 \\ x = 20 \end{cases}$$

Observe a Figura 41 donde obtemos como solução $C = y^{1,2} = (20, 23)$.

Figura 41: Algoritmo de Jacobi



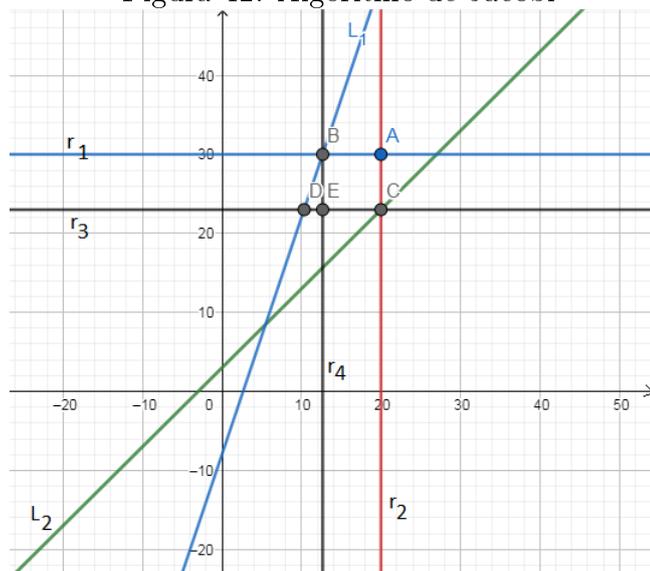
Fonte:

Agora, pelo ponto C tracemos uma reta paralela ao eixo Ox que denotaremos por $r_3 : y = 23$. Perceba que para encontrarmos o ponto $D = y^{2,1}$ devemos resolver o sistema:

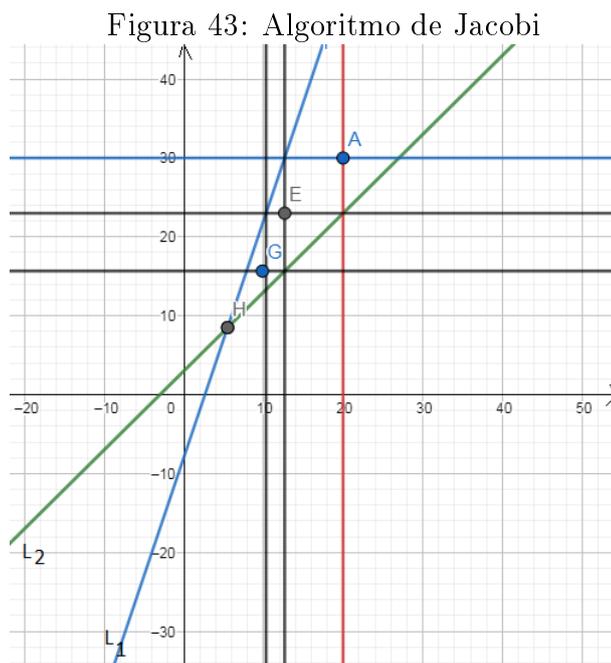
$$\begin{cases} 3x - y = 8 \\ y = 23 \end{cases}$$

onde encontramos $D = (10, 33333; 23)$. Traçando por B uma reta $r_4 : x = 12.6666666667$. Logo, o ponto $E = x^1 = (10, 3333333333; 23)$ conforme observamos na Figura 42.

Figura 42: Algoritmo de Jacobi



De maneira análoga obtemos o ponto $G = x^2 = (9, 871875; 15, 666)$ conforme se observa na Figura 43.



Observamos que intuitivamente a primeira entrada da sequência $\{x^n\}$ converge para 5,5 e a segunda entrada converge para 8,5 que é a solução procurada.

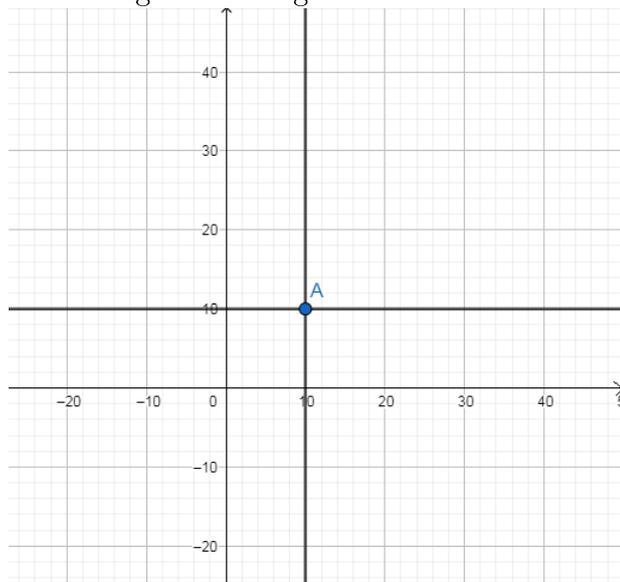
Vamos resolver agora um exemplo utilizando o algoritmo de Cimmino.

Exemplo 10 *Sejam $L_1 : x = 10$ e $L_2 : y = 10$ retas, as quais poderiam ser quaisquer. Encontrar o conjunto solução equivale a resolver o sistema:*

$$\begin{cases} x = 10 & (1) \\ y = 10 & (2) \end{cases}$$

Donde percebemos obviamente que a solução é imediata, ou seja, o ponto de coordenadas $A(10, 10)$ conforme Figura 44.

Figura 44: Algoritmo de Jacobi

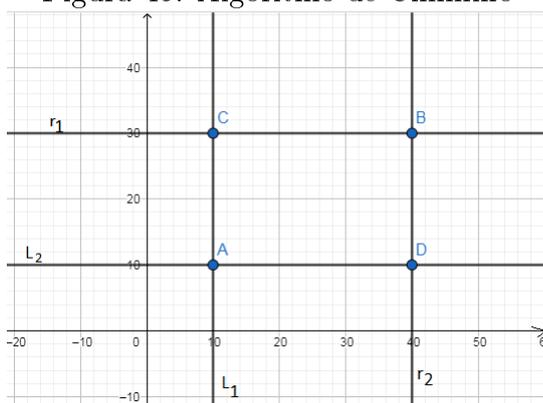


Vamos aplicar o algoritmo de Cimmino para a resolução desse problema. Primeiramente devemos tomar um ponto qualquer, tomemos então o ponto $B = (40, 30)$. Tracemos uma reta $r_1 : y = 30$ pelo ponto B . Daí, a reta r_1 e L_1 se intersectam no ponto $C = (10, 30)$ que é solução do sistema:

$$\begin{cases} x = 10 \\ y = 30 \end{cases}$$

De maneira análoga obtemos o ponto $D = (40, 10)$ conforme a Figura 45.

Figura 45: Algoritmo de Cimmino



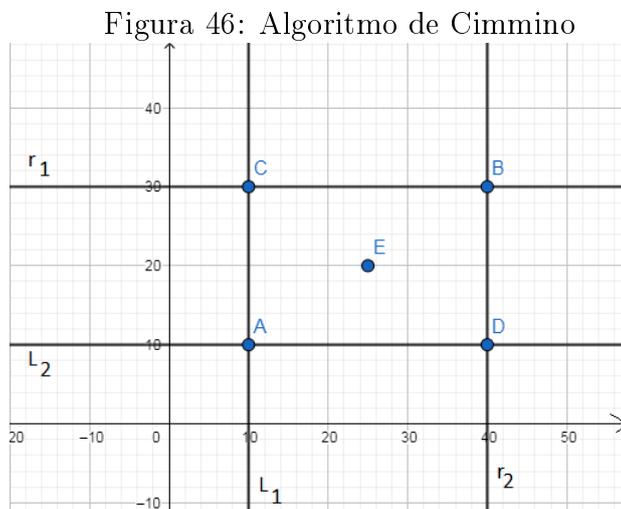
Fonte:

De posse do ponto $C = (10, 30)$ e $D = (40, 10)$, basta tomarmos agora $\lambda_1 = 0,5$ e $\lambda_2 = 0,5$ com a condição de que $\lambda_1 + \lambda_2 = 1$. Daí, para encontrarmos x^1 devemos ter:

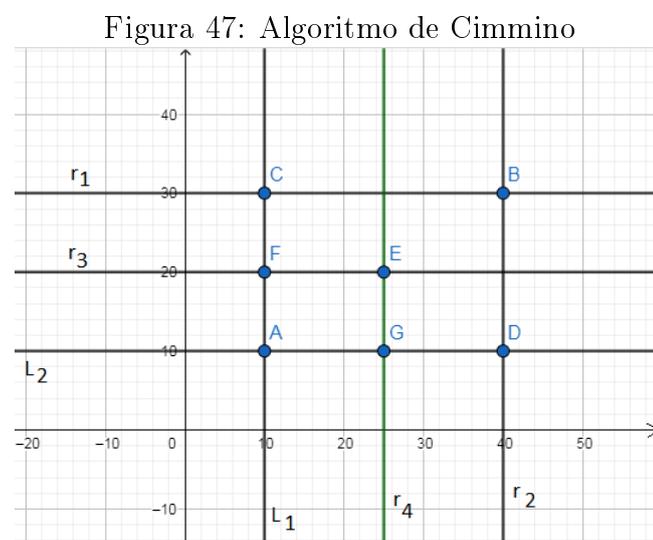
$$x^1 = \lambda_1 C + \lambda_2 D$$

$$x^1 = 0,5(10, 30) + 0,5(40, 10)$$

donde obtemos $E = x^1 = (25, 20)$ conforme vemos em Figura 46.



Analogamente traçamos as retas $r_3 : y = 20$ e $r_4 : x = 25$ donde obtemos os pontos $F = (10, 20)$ e $G = (25, 10)$.



Fonte:

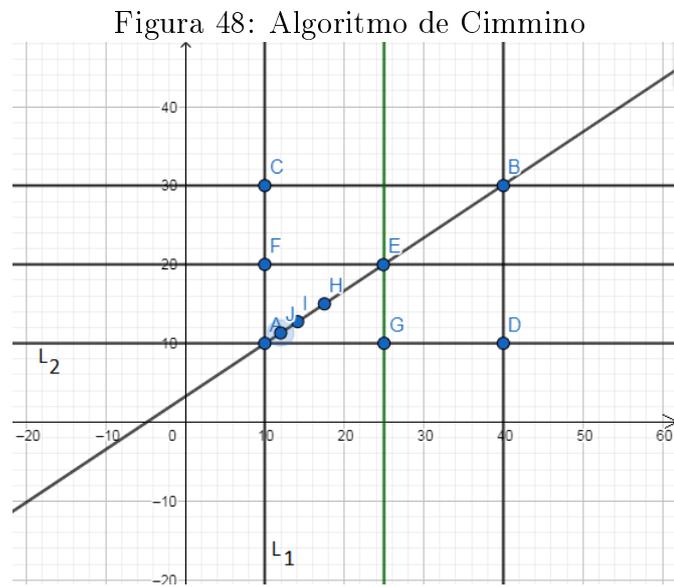
Agora devemos determinar o ponto $H = x^2$. Para tanto tomemos $\lambda_1 = 0,5$ e $\lambda_2 = 0,5$. Daí,

$$H = \lambda_1 F + \lambda_2 G$$

$$H = 0,5(10, 20) + 0,5(25, 10)$$

donde obtemos $x^2 = H = (17,5; 15)$.

Conforme observamos na Figura 48 percebemos que os pontos B, E, H, I, J se aproximam do ponto A .



6 Apêndice

Neste apêndice estudaremos a convergência dos Algoritmos de Kaczmarz e Cimmino. Isto será feito de maneira rigorosa, isto é, tentaremos provar de forma algébrica os conceitos. Para isto, iremos considerar um problema a priori mais geral, isto é, em vez de focarmos no operador projeção sobre hiperplanos, vamos provar a convergência do algoritmo para qualquer operador que satisfaça as propriedades do operador projeção sobre hiperplanos. Por isso, estudaremos uma generalização do operador projeção sobre hiperplanos, depois diremos quais propriedades esse operador satisfaz e em sequência estudaremos a convergência deste algoritmo. Vale ressaltar que vamos supor que as interseções dos hiperplanos H_1, H_2, \dots, H_m sejam não vazias.

6.1 O operador projeção

No capítulo anterior estudamos o operador projeção sobre hiperplanos, o objetivo desta seção é estudar o operador projeção sobre qualquer conjunto convexo. Para isto, vamos começar lembrando a definição de conjunto convexo.

Definição 13 (Conjunto Convexo) *Um subconjunto $H \subset \mathbb{R}^n$ é dito convexo se, para quaisquer $x, y \in H$ e $\alpha \in [0, 1]$, tivermos:*

$$(1 - \alpha)x + \alpha y \in H.$$

A próxima proposição assegura que os hiperplanos são conjuntos convexas do \mathbb{R}^n .

Proposição 4 *Dados $v \in \mathbb{R}^n$ e $b \in \mathbb{R}$. Então o conjunto $H = \{x \in \mathbb{R}^n \mid \langle v, x \rangle = b\}$ é convexo.*

Demonstração: De fato, dados $x, y \in H$ e $\alpha \in [0, 1]$, temos:

$$\begin{aligned} \langle v, \alpha x + (1 - \alpha)y \rangle &= \langle v, \alpha x \rangle + \langle v, (1 - \alpha)y \rangle \\ &= \alpha \langle v, x \rangle + (1 - \alpha) \langle v, y \rangle = \alpha b + (1 - \alpha)b \\ &= \alpha b + b - \alpha b = b. \end{aligned}$$

Logo $\alpha x + (1 - \alpha)y \in H$. Concluimos que H é convexo, provando assim a proposição.

□

A próxima proposição nos diz que a interseção de conjuntos convexos é convexo.

Proposição 5 *Sejam H_1, H_2, \dots, H_m conjuntos convexos. Então, $H = \bigcap_{i=1}^m H_i$ é um conjunto convexo.*

Demonstração: Se $\bigcap_{i=1}^m H_i = \emptyset$, então H é convexo por vacuidade. Suponhamos, agora que $\bigcap_{i=1}^m H_i \neq \emptyset$. Dados $x, y \in H$ quaisquer, então, $x, y \in H_i$ para todo $1 \leq i \leq m$. Como, por hipótese, H_i é convexo para todo $1 \leq i \leq m$, segue que $(1 - \alpha)x + \alpha y \in H_i$ para todo $1 \leq i \leq m$, e portanto, $(1 - \alpha)x + \alpha y \in \bigcap_{i=1}^m H_i$. Provando assim que H é convexo. □

O fato crucial da Proposição 2 é o fato de H ser fechado e convexo. De fato, o próximo Teorema que foi extraído do livro [6], mostra que a Definição 12 poderia ter sido feita para qualquer conjunto H que é convexo e fechado, veja [4] para a definição de conjunto fechado.

Teorema 3 (Teorema da Projeção) *Seja $D \subset \mathbb{R}^n$ um conjunto convexo, fechado e não-vazio. Então para todo $x \in \mathbb{R}^n$, existe um único ponto $\bar{x} \in D$ tal que:*

$$\|x - P_D(x)\| = \min_{y \in D} \|x - y\|.$$

Além disso, $\bar{x} = P_D(x)$ se, e somente se,

$$\bar{x} \in D, \quad \langle x - P_D(x), y - P_D(x) \rangle \leq 0, \quad \forall y \in D. \quad (19)$$

Demonstração: Veja o Teorema 3.2.4 de [6] página 101. □

Pelo Teorema da Projeção, podemos tornar a Definição 12 mais geral, como segue.

Definição 14 (Operador Projeção) *Dado $D \subset \mathbb{R}^n$ convexo e fechado. Então defina o operador $P_D : \mathbb{R}^n \rightarrow D$, de forma que, para cada $x \in \mathbb{R}^n$, $P_D(x)$ é o único elemento de D que satisfaz:*

$$\|x - P_D(x)\| = \min_{y \in D} \|x - y\|.$$

Ou seja, $P_D(x)$ é o ponto de D mais próximo de x .

A próxima proposição desta a primeira propriedade do Operador Projeção.

Proposição 6 *Seja $D \subset \mathbb{R}^n$ um conjunto convexo, fechado e não-vazio. Então para todo $x, y \in \mathbb{R}^n$,*

$$\|P_D(x) - P_D(y)\|^2 \leq \langle x - y, P_D(x) - P_D(y) \rangle.$$

Demonstração: Dados $x, y \in \mathbb{R}^n$ quaisquer. Pelo Teorema 3 os elementos $P_D(x)$ e $P_D(y)$ satisfazem:

$$\langle x - P_D(x), z - P_D(x) \rangle \leq 0$$

e

$$\langle y - P_D(y), w - P_D(y) \rangle \leq 0,$$

para todo $z, w \in \mathbb{R}^n$. Tomando $z = P_D(y)$ e $w = P_D(x)$, nas desigualdades acima, temos:

$$-\langle x - P_D(x), P_D(x) - P_D(y) \rangle \leq 0$$

e

$$\langle y - P_D(y), P_D(x) - P_D(y) \rangle \leq 0$$

Somando essas duas desigualdade, obtemos:

$$\langle y - x + P_D(x) - P_D(y), P_D(x) - P_D(y) \rangle \leq 0.$$

Como o produto interno é linear, então

$$\langle y - x, P_D(x) - P_D(y) \rangle + \langle P_D(x) - P_D(y), P_D(x) - P_D(y) \rangle \leq 0.$$

Passando o primeiro produto interno para o segundo membro obtemos

$$\|P_D(x) - P_D(y)\|^2 \leq \langle y - x, P_D(x) - P_D(y) \rangle.$$

Terminando assim a prova da Proposição. □

Como consequência desta proposição, descrevemos agora as propriedades do operador projeção que estamos interessados em estudar, tais propriedades foram extraídas de [1].

Proposição 7 (Propriedades do operador projeção) *Seja $D \subset \mathbb{R}^n$ um convexo e fechado. Então para todo $x, y \in \mathbb{R}^n$, temos:*

i) $\|P_D(x) - P_D(y)\| \leq \|x - y\|;$

ii) *se $\|P_D(x) - P_D(y)\| = \|x - y\|$, então*

a) $P_D(x) - P_D(y) = x - y;$

b) $\langle y - x, P_D(x) - x \rangle = 0.$

Demonstração: Dados $x, y \in \mathbb{R}^n$ quaisquer. Sem perda de generalidade, podemos supor $x \neq y$, pois se forem iguais a proposição é evidente.

Para provarmos o item i). Consideraremos dois casos: Se $P_D(x) = P_D(y)$, então $i)$ é evidente. O caso interessante é quando $P_D(x) \neq P_D(y)$. Neste caso, usando a Proposição 6 combinado com a desigualdade de Cauchy-Schwarz, obtemos:

$$\begin{aligned} \|P_D(x) - P_D(y)\|^2 &\leq \langle y - x, P_D(x) - P_D(y) \rangle \\ &\leq \|y - x\| \|P_D(x) - P_D(y)\|. \end{aligned} \quad (20)$$

Como por hipótese, $P_D(x) \neq P_D(y)$, podemos dividir os dois membros da desigualdade acima por $\|P_D(x) - P_D(y)\| > 0$, obtendo a desigualdade desejada.

Para provar $ii)$, se $\|P_D(x) - P_D(y)\| = \|x - y\|$ então, em (20) vale a igualdade, logo, na desigualdade de Cauchy-Schwarz vai valer a igualdade, e portanto, os vetores $y - x$ e $P_D(x) - P_D(y)$ serão colineares, isto é,

$$P_D(x) - P_D(y) = \sigma(x - y).$$

Pela hipótese $\|P_D(x) - P_D(y)\| = \|x - y\|$, segue que $\sigma = 1$ ou $\sigma = -1$.

Afirmamos que $\sigma = 1$. Suponha por contradição que $\sigma = -1$, logo, $P_D(x) - P_D(y) = y - x$. Pelo Teorema 3 os elementos $P_D(x)$ e $P_D(y)$ satisfazem:

$$\langle x - P_D(x), z - P_D(x) \rangle \leq 0 \quad (21)$$

e

$$\langle y - P_D(y), w - P_D(y) \rangle \leq 0$$

para todo $z, w \in \mathbb{R}^n$. Tomando $z = P_D(y)$ e $w = P_D(x)$, nas desigualdades acima, e somando-as obtemos:

$$\langle y - x + P_D(x) - P_D(y), P_D(x) - P_D(y) \rangle \leq 0. \quad (22)$$

Usando que $P_D(x) - P_D(y) = y - x$, resulta:

$$\langle 2P_D(x) - P_D(y), P_D(x) - P_D(y) \rangle \leq 0,$$

$$2\|P_D(x) - P_D(y)\|^2 \leq 0,$$

usando que o produto interno é bi-linear, e que ele é a norma ao quadrado, segue que,

$$2\|P_D(x) - P_D(y)\|^2 \leq 0 \rightarrow P_D(x) - P_D(y) = 0.$$

Portanto, $P_D(x) = P_D(y)$, e

$$0 = P_D(x) - P_D(y) = y - x \Rightarrow x = y.$$

O que é uma contradição com a hipótese inicial de $x \neq y$. Provando o a). Para concluir a prova desta proposição, resta provar o item b) de *ii*). Para isto, usando que o produto interno é bilinear em (22). Podemos reescrever essa desigualdade como segue:

$$\langle y - x, P_D(x) - P_D(y) \rangle + \langle P_D(x) - P_D(y), P_D(x) - P_D(y) \rangle \leq 0.$$

Passando o primeiro número para o segundo membro desta desigualdade e usando definição de norma, obtemos:

$$\begin{aligned} 0 &\leq \|P_D(x) - P_D(y)\|^2 \\ &\leq \langle y - x, P_D(x) - P_D(y) \rangle \\ &\leq \|y - x\| \|P_D(x) - P_D(y)\| \\ &\leq \|P_D(x) - P_D(y)\|^2. \end{aligned}$$

Portanto, se $\|P_D(x) - P_D(y)\| = \|x - y\|$, por (20), a desigualdade acima, na verdade é uma igualdade, e portanto, todas as contas feitas para chegar nesta fórmula vale igualdade, em particular, vale a igualdade em (21) quando tomamos $z = P_D(y)$, isto é,

$$\langle x - P_D(x), P_D(y) - P_D(x) \rangle = 0.$$

Usando agora o item a) temos:

$$\langle x - P_D(x), y - x \rangle = 0.$$

Terminando assim, a prova do item b) e conseqüentemente da proposição.

A próxima proposição justifica por que, ao estudarmos os algoritmos relaxados, temos que tomar $\alpha \in (0, 2)$.

Proposição 8 *Seja $D \subset \mathbb{R}^n$ um subconjunto convexo e fechado. Se $\alpha \in (0, 2)$ então o operador $P^\alpha = (1 - \alpha)x + \alpha P_D(x)$, satisfaz:*

$$\|P^\alpha(x) - P^\alpha(y)\| \leq \|x - y\|, \quad \forall x \in \mathbb{R}^n.$$

Demonstração: Dividiremos em dois casos:

i) Caso $0 < \alpha \leq 1$, para este caso, temos:

$$\begin{aligned}
 \|P^\alpha(x) - P^\alpha(y)\| &= \|(1 - \alpha)x + \alpha P_D(x) - (1 - \alpha)y - \alpha P_D(y)\| \\
 &= \|(1 - \alpha)(x - y) + \alpha (P_D(x) - P_D(y))\| \\
 &\leq (1 - \alpha)\|x - y\| + \alpha \|P_D(x) - P_D(y)\| \\
 &\leq (1 - \alpha)\|x - y\| + \alpha \|x - y\| \\
 &= \|x - y\|.
 \end{aligned}$$

Onde na última desigualdade utilizamos que:

$$\|P_D(x) - P_D(y)\| \leq \|x - y\|.$$

ii) Caso $1 < \alpha < 2$, temos que:

$$\begin{aligned}
 \|P^\alpha(x) - P^\alpha(y)\|^2 &= \|(1 - \alpha)x + \alpha P_D(x) - (1 - \alpha)y - \alpha P_D(y)\|^2 \\
 &= \|(1 - \alpha)(x - y) + \alpha (P_D(x) - P_D(y))\|^2 \\
 &\leq ((1 - \alpha)\|x - y\| + \alpha \|P_D(x) - P_D(y)\|)^2 \\
 &\leq (1 - \alpha)^2\|x - y\|^2 + \alpha^2 \|P_D(x) - P_D(y)\|^2 \\
 &\quad + 2(1 - \alpha)\alpha \langle x - y, P_D(x) - P_D(y) \rangle.
 \end{aligned}$$

Combinando o Proposição 6 com o fato de $1 - \alpha < 0$, temos

$$\begin{aligned}
 \|P^\alpha(x) - P^\alpha(y)\|^2 &\leq (1 - \alpha)^2\|x - y\|^2 + \alpha^2 \|P_D(x) - P_D(y)\|^2 \\
 &\quad + 2(1 - \alpha)\alpha \langle x - y, P_D(x) - P_D(y) \rangle \\
 &\leq (1 - \alpha)^2\|x - y\|^2 + \alpha^2 \|P_D(x) - P_D(y)\|^2 \\
 &\quad + 2(1 - \alpha)\alpha \|P_D(x) - P_D(y)\|^2 \\
 &= (1 - \alpha)^2\|x - y\|^2 + [\alpha^2 + 2(1 - \alpha)\alpha] \|P_D(x) - P_D(y)\|^2 \\
 &= (1 - \alpha)^2\|x - y\|^2 + \alpha(2 - \alpha) \|P_D(x) - P_D(y)\|^2.
 \end{aligned}$$

Usando a hipótese de $1 < \alpha < 2$ e também o fato do operador P_D ser não expansivo, resulta:

$$\begin{aligned}
 \|P^\alpha(x) - P^\alpha(y)\|^2 &\leq (1 - \alpha)^2\|x - y\|^2 + \alpha(2 - \alpha) \|P_D(x) - P_D(y)\|^2 \\
 &\leq (1 - \alpha)^2\|x - y\|^2 + \alpha(2 - \alpha) \|x - y\|^2 \\
 &= [(1 - \alpha)^2 + \alpha(2 - \alpha)] \|x - y\|^2,
 \end{aligned}$$

Desenvolvendo a equação em α , obtemos $(1 - \alpha)^2 + \alpha(1 - \alpha) = 1$. Portanto,

$$\|P^\alpha(x) - P^\alpha(y)\| \leq \|x - y\|$$

6.2 Análise de Convergência

Sejam $K_1, \dots, K_n \subset \mathbb{R}^n$ conjuntos convexos e fechados. Para estudar a convergência dos Algoritmos de Kaczmarz e Cimmino, usaremos a seguinte definição:

Definição 15 Para cada $i = 1, \dots, m$ e cada $\alpha \in (0, 2)$, denotemos os operadores $P_i, P_i^\alpha : \mathbb{R}^n \rightarrow \mathbb{R}^n$, por:

$$P_i(x) = P_{K_i}(x) \tag{23}$$

e

$$P_i^\alpha = (1 - \alpha)x + \alpha P_i(x) \tag{24}$$

para todo $x \in \mathbb{R}^n$, respectivamente.

Usando a definição anterior, podemos definir agora as funções $S, F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, por:

$$F(x) = (P_m^\alpha \circ P_{m-1}^\alpha \circ \dots \circ P_2^\alpha \circ P_1^\alpha)(x) \tag{25}$$

e

$$S(x) = \sum_{i=1}^m \lambda_i P_i^\alpha(x) \tag{26}$$

onde o símbolo \circ indica a composição de função.

A partir dos algoritmos estudados, podemos representar as sequências $\{x^n\}$ por:

$$x^{k+1} = F(x^k) \tag{27}$$

e para Cimmino por:

$$x^{k+1} = S(x^k) \tag{28}$$

Ou seja, agora os algoritmos estão sendo vistos como composição de operadores. Logo, basta provar que essas sequências de composições de operadores convergem para um ponto, para isto a hipótese de $K = \bigcap_i^m K_i$ não seja vazio será importante (veja Observação 8).

Motivados, pela Observação 8, temos a seguinte definição, que será usado para enunciar o Teorema de convergência dos algoritmos.

Definição 16 Dado o operador $F : K \subset \mathbb{R}^n \rightarrow K$. Diremos que $x^* \in K$ é ponto fixo do operador F , quando $F(x^*) = x^*$. Denotaremos por $Fix(F)$ o conjunto de todos os pontos fixos do operador F , ou seja, $Fix(F) = \{x^* \in K : F(x^*) = x^*\}$

Observe que encontrar ponto fixo do operador projeção equivale a encontrar um ponto na interseção dos conjuntos convexos, isto é, $F(x^*) = x^*$ se, e somente se, $x^* \in \bigcap_{i=1}^n K_i$. Portanto, para mostrar que as sequências (27) e (28) convergem, basta encontrar um ponto fixo do operador projeção.

Pelas Propriedades da Projeção Ortogonal, vistos na Proposição 7, usaremos a seguinte definição:

Definição 17 Denotaremos por \mathcal{F} ao conjunto de todas as funções contínuas $Q : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tais que:

- i) $\|Q(x) - Q(y)\| \leq \|x - y\|$
- ii) se $\|Q(x) - Q(y)\| = \|x - y\|$ então:
 - a) $Q(x) - Q(y) = x - y$
 - b) $\langle y - x, Q(x) - x \rangle = 0$

Segundo [1] o conjunto de funções \mathcal{F} é fechado para composições de funções e por combinações convexas, isto é,

$$Q_m \circ \dots \circ Q_2 \circ Q_1 \in \mathcal{F}, \quad \text{e}, \quad \sum_{i=1}^m \lambda_i Q_i \in \mathcal{F},$$

para todo $Q_1, \dots, Q_m \in \mathcal{F}$ e todo $\lambda_i \in [0, 1]$.

Como as projeções P_i sobre os conjuntos convexos fechados K_i pertencem a \mathcal{F} , veja a Proposição 7, resulta que F e S estão em \mathcal{F} , pois eles são combinações convexas ou composições de P_i .

O próximo Teorema é o principal resultado desse trabalho, pois garante que os Algoritmos de Kazmarz e Cimmino convergem sempre que $\bigcap_{i=1}^m K_i \neq \emptyset$.

Teorema 4 (Teorema do Ponto Fixo) Sejam $Q \in \mathcal{F}$ e $x^0 \in \mathbb{R}^n$. Defina a sequência $x^{k+1} = Q(x^k)$ com $k \in \mathbb{N} \cup \{0\}$.

Então a sequência $\{x^k\}$ converge se, e somente se, $\text{Fix}(Q) \neq \emptyset$. Além disso, se $\lim_{k \rightarrow \infty} x^k = x^*$, então $x^* \in \text{Fix}(Q)$.

Demonstração: Suponhamos que $\{x^k\}$ converge, isto é, $\lim_{k \rightarrow \infty} x^k = x^*$ com $x^{k+1} = Q(x^k)$. Como Q é contínua, temos:

$$\begin{aligned} x^* &= \lim_{k \rightarrow \infty} x^k \\ &= Q\left(\lim_{k \rightarrow \infty} x^{k-1}\right) \\ &= Q(x^*), \end{aligned}$$

logo, $x^* = Q(x^*)$, ou seja, $x^* \in \text{Fix}(Q)$, e portanto, $\text{Fix}(Q) \neq \emptyset$.

Para provar a recíproca, devemos mostrar que, se $\text{Fix}(Q) \neq \emptyset$, então a sequência da forma $\{x^k\}$ converge. Para isto, fixemos $z \in \text{Fix}(Q)$, então $Q(z) = z$. Usando a propriedade 1) da classe de funções \mathcal{F} , temos:

$$\|x^{k+1} - z\| = \|Q(x^k) - Q(z)\| \leq \|x^k - z\|, \quad (29)$$

Portanto, a sequência $\{\|x^k - z\|\}$ é decrescente. Repetindo (29), várias vezes, obtemos

$$\|x^{k+1} - z\| \leq \|x^k - z\| \leq \dots \leq \|x^1 - z\| \leq \|x^0 - z\|, \quad (30)$$

provando que a sequência $\{x^k\}$ está contida numa bola de centro z e raio $\|x^0 - z\|$, e portanto, é limitada. Como $\{x^k\}$ se limitada ela possui uma subsequência convergente, digamos $\{x^{p_k}\}$ a subsequência de $\{x^k\}$ para $k \in N \cup \{0\}$ que converge para um ponto x^* , isto é, $\lim_{k \rightarrow \infty} x^{p_k} = x^*$. (para mais detalhes veja corolário 1, pg 123 de [5]).

Usando (29), e que $p_k + 1 \leq p_{k+1}$ (subsequência anda mais rápido que á sequência), temos:

$$\|x^{p_{k+1}} - z\| \leq \|Q(x^{p_k}) - Q(z)\| \leq \|x^{p_k} - z\|. \quad (31)$$

Aplicando o limite quando $k \rightarrow \infty$ em (31) temos:

$$\|x^* - z\| \leq \|Q(x^*) - Q(z)\| = \|Q(x^*) - z\| \leq \|x^* - z\| \quad (32)$$

ou seja,

$$\|Q(x^*) - Q(z)\| = \|x^* - z\| \quad (33)$$

Usando a propriedade 2)a) da Definição 16, temos:

$$Q(x^*) - Q(z) = Q(x^*) - z = x^* - z \quad (34)$$

donde temos $Q(x^*) = x^*$, e portanto, $x^* \in \text{Fix}(Q)$. Provando assim a primeira parte do Teorema.

Para provar a segunda parte, tomando-se x^* no lugar de z em (29), obtemos que a sequência $\{\|x^k - x^*\|\}$ é decrescente, não negativa e possui uma subsequência $\{\|x^k - x^*\|\}$ que converge para zero, portanto, a própria sequência $\{\|x^k - x^*\|\}$ converge a zero, isto é, $\lim_{k \rightarrow \infty} x^k = x^*$ (qualquer dúvida veja [5]). Terminando assim a prova do Teorema.

Pelo Teorema 4 garantimos a convergência dos Algoritmos de Kaczmarz e Cimmino, porém a condição a ser assumida é que o conjunto de pontos fixos de F e \overline{F} sejam não vazios. No artigo do professor Iusem [1] ele estuda as condições para que tal hipótese

aconteça. Não faremos isso aqui, pois o foco é apenas introduzir a pesquisa matemática. Para o leitor que deseja, se aprofundar sobre a matemática pro trás da tomografia computadorizada, recomendamos [1], [2], [3], [8] e as referências neles contidas.

7 Considerações Finais

O professor pode usar a modelagem da tomografia computadorizada, feita no Capítulo 2, via sistemas lineares, para motivar seus alunos quanto a importância deste conteúdo na sua vida cotidiana. Em relação aos algoritmos estudados aqui, o professor pode usá-los para motivar seus alunos sobre a pesquisa em matemática, bem como fornecer métodos diferentes para resolver sistemas lineares, para isto, os exemplos podem fornecer o caminho mais rápido e natural, para que o aluno compreenda esses algoritmos. Para os alunos que tenham um conhecimento um pouco mais avançado em matemática, o professor poderá ensiná-los, como se prova a convergência de tais algoritmos. Para os professores que desejam um aprofundamento sobre os assuntos aqui estudados, recomendamos [1], no qual foi nossa principal fonte bem como as referências.

REFERÊNCIAS

- [1] Iusem, A. ; De Pierro, A. **Fundamentos Matemáticos da Tomografia Computadorizada: Método de Expansão em Séries**. Revista matemática Universitária n: 12, dezembro de 1990, pag. 57-85.
- [2] Iusem, A. ; De Pierro, A. **Fundamentos Matemáticos da Tomografia Computadorizada: Método de Transformação**. Revista matemática Universitária n: 11, junho de 1990, pag. 53-65 .
- [3] De Cezaro, A.; De Cezaro, F. T. **Problemas Inversos e a matemáticos da Tomografia Computadorizada**. V Bienal da SBM Sociedade Brasileira de Matemática UFPB-Universidade Federal da Paraíba 18 a 22 de outubro de 2010.
- [4] Lima, E. L. **Análise Real volume 2**. Coleção Matemática Universitária. Editora: IMPA Ano: 2004.
- [5] Lima, E. L. **Curso de Análise Real volume 1**. coleção Matemática Universitária. Editora: IMPA. ed. 12. Ano: 2009.
- [6] Izmailov, A.; Solodov, M. **Otimização volume 1: Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade**. Rio de Janeiro, Editora IMPA. ed. 2. Ano: 2009.
- [7] Iezzi, Gelson; Hazzan, Samoel. **Fundamentos da Matemática Elementar, Volume 4**. São Paulo, Editora Atual, ed.5. Ano 1977.
- [8] Da Silva, S. M. T. **TOMOGRAFIA COMPUTADORIZADA: Uma Proposta para a Aplicação de Matemática no Ensino Médio**. Dissertação do Profmat. Rio Grande, Rio Grande do Sul, 20 de Abril 2018.
- [9] <<https://www.vencerocancer.org.br/wp-content/uploads/2013/09/cap9-1.png> >
- [10] <<http://static.clinicabioimagem.com.br/img/tomografia-computadorizada-6.jpg>>
- [11] <http://donacion.organos.ua.es/submenu3/inf_sanitaria/proceso/imagenes_proceso/helicoidal.jpg>

-
- [12] <<http://www.integradiagnosticos.com.br/site/images/integra/exames/tomografia.jpg>>
- [13] <<https://image.slidesharecdn.com/exercicioanatomiaplanos-150921234936-lva1-app6892/95/exercicio-anatomia-planos-4-638.jpg?cb=1442879424>>
- [14] <<https://image.slidesharecdn.com/exercicioanatomiaplanos-150921234936-lva1-app6892/95/exercicio-anatomia-planos-4-638.jpg?cb=1442879424>>
- [15] <http://nedmed.info/images/img135618_99.jpg>
- [16] <<https://helioprint.com.br/blog/wp-content/uploads/2017/05/ruido-radiologico.jpg>>